

■ 論 文 ■

부스트랩 기법을 이용한 TCS 데이터로부터 차종별 교통량 추정모형 구축

Construction of vehicle classification estimation model
from the TCS data by using bootstrap Algorithm

노 정 현

(한양대학교 도시대학원 교수)

김 태 군

(한양대학교 도시공학과 박사과정)

차 경 준

(한양대학교 수학과 교수)

박 영 선

(한양대학교 수학과 이학박사)

남궁 성

(한국도로공사 도로연구소 책임연구원)

황 부 연

(한국도로공사 도로연구소 책임연구원)

목 차

- I. 서론
 - 1. 연구의 배경 및 목적
 - 2. 연구의 방법
- II. 표준 차종분류 설정
 - 1. 각 조사별 차종분류 현황
 - 2. 표준차종 분류 설정
- III. 톨게이트별 카테고리 설정
 - 1. 카테고리설정의 필요성
 - 2. 주제별 카테고리 설정
 - 3. 카테고리의 유의성 검정
 - 4. 카테고리분류 최종결과
- IV. 표본조사
 - 1. 조사표본 톨게이트 선정
 - 2. 조사방법
- V. 톨게이트 차종별 교통량 추정모형 구축
 - 1. 모형의 이론적 고찰
 - 2. 모형의 정립
 - 3. 차종별 구성비 작성
 - 4. 대안별 모형결과
 - 5. 모형의 검정
 - 6. 모형의 적용
- VI. 결론
참고문헌

Key Words : TCS자료, 차종분류, 표준차종, 톨게이트 유출입교통량, 부스트랩

요 약

차종별 교통량자료는 자료의 출처별로 차종이 동일하지 않아 자료간 호환이 어려우며 이들 자료의 활용도 또한 매우 낮다. 특히, 고속도로의 경우에는 전수자료인 TCS 자료가 있음에도 불구하고 TCS의 차종분류는 차종 내에 승용, 승합, 화물차가 혼재 되어있어 실질적으로 활용도가 매우 낮다.

이에 본 연구에서는 각 출처별 자료들의 차종구분과 호환할 수 있도록 차종구분을 표준화하고 고속도로 톨게이트 유출입 차종별 교통량을 표준화된 차종별로 추정하기 위한 모형을 개발하였다. 즉, 톨게이트를 그 특성에 따라 몇 개의 카테고리로 분류하였고, 각 카테고리별로 각 차종의 구성비를 점추정량을 이용한 기법(산술평균, 기하평균, 조화평균)과 비모수적 통계기법인 부스트랩을 이용하여 표준화 분류별 교통량을 추정하는 모형을 개발하였다. 그 결과 두 방법 모두 비교적 유의한 수준의 결과가 도출되었으나, 표본의 크기에 따라 발생할 수 있는 극단치에 대한 오추정 문제를 감안할 수 있는 부스트랩기법이 우수한 것으로 나타났다.

본 연구의 결과로 향후 TCS 자료의 활용성 증대와 TCS 자료를 이용한 고속도로 구간교통량 추정과 고속도로 정기교통량 조사자료의 좀더 구체적인 비교가 가능할 것으로 기대된다.

1. 서론

1. 연구의 배경 및 목적

교통량자료는 교통분야전반의 각종 연구에 필요한 기초자료로서 이용되나, 기관이나 자료 출처별로 교통량자료는 차종구분이 통일되어 있지 않아 연구자간 공용이 어려우며, 자료의 활용도 또한 낮은 실정이다.

고속도로의 경우 요금징수 시스템인 TCS(Toll Collection System)를 이용하여 얻어진 차종별 톨게이트 유출입 교통량 자료는 전수자료임에도 불구하고 동종내에도 승용, 승합, 화물차가 혼재되어있는 이유로 사람통행 즉, 승용·승합의 통행량은 전체자료의 일정비율을 할당하여 산출하는 방법을 이용하고 있어 교통계획 및 정책 등에 신뢰성 있는 자료로 활용되지 못하고 있다.

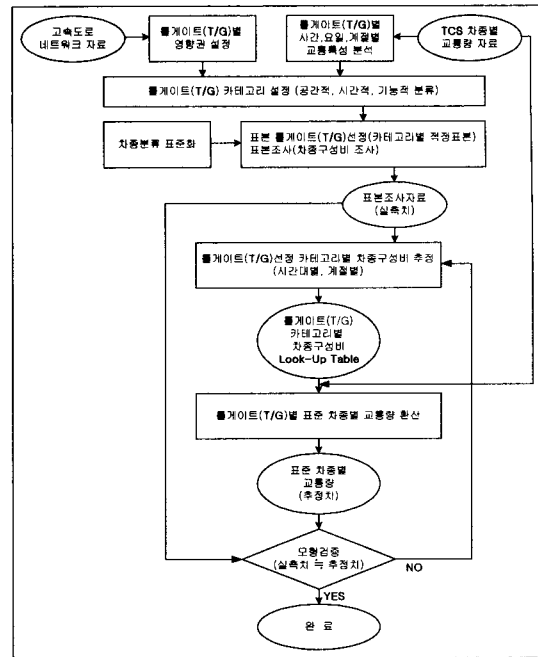
또한, 일반적으로 톨게이트는 복합적인 특성을 가지고 있으며 그 특성에 따라 이용교통량의 규모 및 차종구성비도 달라 톨게이트의 차종구성비는 각 톨게이트마다 고유특성을 가지고 있으며, 향후 톨게이트가 계속 증가될 경우에도 또 다른 복합적인 특성을 가진 톨게이트가 등장할 것이다.

따라서, 본 연구에서는 각 조사에서 사용되고있는 분류기준을 표준화하고, 각 톨게이트의 복합적인 특성을 고려할 수 있는 특성별 카테고리를 설정하여 톨게이트의 복합적인 특성을 카테고리간의 연계를 통해 설명하고(Wootton & Pick, 1967), 통계적 추정기법인 붓스트랩(bootstrap)기법(Jose Banos-Pino et al., 1999 ; Urban Hjorth, 2002)을 이용하여 각 톨게이트에서 취합된 유출입 차종별 TCS 자료를 사람통행과 화물통행을 분류하여 집계할 수 있도록 표준화된 차종분류별 교통량으로 전환할 수 있는 모형을 개발하였다.

2. 연구의 방법

본 연구의 연구방법은 <그림 1>에서 나타낸 바와 같이 다음의 몇 가지 단계로 정리할 수 있다.

첫째, 기존의 연구에서 이용된 교통량자료와 TCS 자료가 서로 호환될 수 있도록 새로운 차종분류기준을 표준화하는 단계로서 현재의 차종분류는 타 자료와



<그림 1> 연구진행 과정

호환이 어려운 상태이므로 현재 각 연구 및 조사에서 사용되고 있는 교통량조사자료와 호환될 수 있도록 차종분류를 표준화한다.

둘째, 톨게이트별로 그 특성을 고려하여 몇 개의 카테고리로 구분하고, 각 카테고리간 차종구성비의 독립성검정을 통해 카테고리분류의 유의성을 검토한다.

셋째, 상기단계의 결과를 이용하여 톨게이트별로 차종구성비를 추정한다. 추정과정은 두 부분으로 나누어 설정하였는데 먼저, 각 카테고리내의 그룹 대표치를 추정한 다음 카테고리의 대표치를 이용하여 톨게이트별 차종구성비를 추정한다.

II. 차종분류의 표준화

현재 여러 기관 및 연구에서 사용되고 있는 차종분류는 TCS자료를 제외하고, 일반적으로 크게 승용차, 승합차, 화물차로 구분하고 있다. 일반적으로 차종분류기준은 승합차는 탑승가능인원으로, 화물차는 차축구성과 적재가능용량을 병산하여 분류하고있다(<표 1> 참조)(건설교통부, 2000).

TCS자료는 고속도로 이용차량을 대상으로 통행요금 징수된 차량에 대해서 분류하고 있으며, 요금면제 차량을 포함하여 7종으로 분류하고 있다.

〈표 1〉 조사자료별 차종분류 기준

구분	분류	차종분류의 특성
TCS자료	5종 (6종)	요금을 기준으로 5종으로 나눔. 경차는 6종으로 분류하여 할인혜택을 줌.
고속도로 정기교통량 조사	3종 (9종)	대분류 : 승용차, 승합차, 화물차 소분류 : 9종으로 나누며, 군차량은 기타로 처리.
교통 영향평가	3종 (5종)	승용차, 버스 화물차로 나눔. 버스와 화물차를 대형, 소형으로 나눔.

〈표 2〉 TCS자료상의 차종분류 기준

TCS 자료	분류기준
7종	1종 승용차, 16인승이하의 승합차 적재중량 2.5톤미만 화물차
	2종 17인이상 23인승이하 승합차 2.5톤이상 5.5톤이하 화물차
	3종 33인승이상 승합차 5.5톤초과 10톤미만 화물차
	4종 10톤이상 20톤미만 화물차
	5종 20톤이상 화물차
	6종 경차
	기타 면제차량

전술한 바와 같이 조사자료별로 차종분류기준이 다르므로 국내 각 용도별 교통량자료를 검토하여 국내 차량체원조사를 기준으로 TCS 데이터의 5개¹⁾ 차종분류를 〈표 3〉과 같이 11개 차종으로 세분화하여 정리하였으며, 본 연구에서는 새로 정리된 11개 차종분류를 '표준차종분류'로 정의하기로 한다.

〈표 3〉 표준차종분류

분류	분류기준	분류	분류기준
1	승용차	7	화물차 (5톤초과~10톤미만)
2	승합차(16인승이하)	8	화물차 (10톤이상~20톤미만)
3	승합차 (17인승이상~25인승이하)	9	화물차(20톤이상)
4	승합차(25인승초과)	10	세미트레일러
5	화물차(2.5톤미만)	11	플트레일러
6	화물차 (2.5톤이상~5톤이하)		

1) TCS 상에서의 6종인 경차는 승용차이므로, 1종과 같은 차종내에 포함시켜 정의한다.
2) Urban Hjorth(2002)는 Bootstrap 사용에 있어서 kernel regression filter를 사용하였다.

〈표 4〉 조사별 분류체계 비교

표준 차종	TCS 분류	고속도로 정기 교통량조사	교통영향 평가관련조사		
1	(1종)	승용차	승용차		
2	(1종)	승 합 차	버스		
3	(2종)			소형(16인승이하)	소형(15인승이하)
4	(3종)		대형(16인승이상)		
5	(1종)	화 물 차	화 물 차		
6	(2종)			소형(1톤미만)	소형(5톤미만)
7	(3종)			보통(8톤미만)	
8	(4종)			대형(8톤이상)	대형 (5톤이상, 트레일러)
9	(5종)				
10	(5종)	플트레일러			
11	(5종)	기타			

표준차종분류와 기존의 차종분류체계인 'TCS자료', '고속도로 정기교통량조사자료', '교통영향평가 관련자료'에서 제시한 차종분류체계와 비교해보면 〈표 4〉와 같이 나타낼 수 있다.

III. 톨게이트별 카테고리 설정

1. 카테고리설정의 필요성

차종분류 표준화를 위해서는 각 톨게이트별 실측자료가 필요하나 전체 고속도로 톨게이트를 전수조사하기에는 많은 시간과 인력 및 비용이 소요된다. 또한 전수조사를 하더라도 특정날짜가 조사자료를 통해 산출된 표준차종구성비가 톨게이트의 차종구성비를 대표한다고 보기에는 무리가 따른다. 그러므로 본 연구에서는 톨게이트의 특성에 따라 유사한 차종구성형태를 가질 것으로 판단하고, 각 특성별로 유사성(Similarity)을 갖는 톨게이트를 그룹화 하여 차종구성비의 유사성을 찾고자 하였다(Stoper & McDonald, 1983). 더욱이 주제별로 설정된 그룹을 카테고리별로 연계시킴으로서 톨게이트가 갖는 다중특성을 반영할 수 있을 뿐만 아니라 자료의 신뢰성 및 대표성을 확보함과 동시에 최종모형구축을 위한 필터링(filtering)²⁾역할로서 그 필요성을 들 수 있겠다(Urban Hjorth, 2002).

2. 주제별 카테고리 설정

각 톨게이트의 다중특성을 반영하기 위한 카테고리의 주제를 크게 다음의 네 가지로 분류하였다.

- 영향권특성에 따른 분류
- 이용 교통량규모에 따른 분류
- 교통량의 시간적 변동에 따른 분류
- 차종구성비에 따른 분류

1) 영향권특성에 따른 분류

한국도로공사(1998)는 톨게이트를 이용하는 통행자의 기·종점지역을 중심으로 이용권역을 설정하여 이를 영향권으로 하였던 바 본 연구에서는 톨게이트의 영향권에 의한 특성이 차종구성비에 영향을 줄 것으로 보여 영향권의 규모 및 기능에 따라 분류하였다 (한국개발연구원, 1999).

(1) 영향권의 규모에 따른 분류

영향권규모에 따라 톨게이트의 이용차량의 특성은 해당 톨게이트영향권의 인구규모에 따라 다르므로 영향권의 인구규모에 따라 톨게이트를 구분하였다. 그 결과 <표 5>에서 보는 바와 같이 분류하여 분석에 포함시켰다.

<표 5> 영향권의 규모에 따른 분류기준

분류		기준
광역 도시권역	대도시 권역	행정구역상으로 서울특별시를 포함한 6대 광역시에 위치한 톨게이트
	위성도시 권역	지리적으로 대도시의 인접한 도시로써 대도시의 영향권에 위치한 톨게이트
중소도시권역		행정구역상 시에 위치하고 있는 톨게이트
농어촌권역		행정구역상 군에 위치하고 있는 톨게이트

(2) 영향권의 기능에 따른 분류

영향권의 기능에 따라 <표 6>과 같이 공업지역, 관광지역, 일반지역으로 차종구성비를 구분하였다.

<표 6> 지역의 기능별 분류기준

분류	대상지역
공업지역	도시계획법, 산업입지 및 개발에 관한 법률에 의거한 공업지역 및 산업단지 배치현황을 바탕으로 한 공업지역
관광지역	도시계획법 및 관광권역 분류에 따른 관광단지를 중심으로 한 지역
일반지역	공업 및 관광도시기능을 제외한 지역

2) 이용교통량규모에 따른 분류

일반적으로 대도시의 경우에는 교통량의 규모가 매우 높은편이며, 승용·승합차의 비율이 중소도시 및 농어촌지역에 비해 상대적으로 높은 편이다. 이렇듯이 이용교통량은 그 지역의 규모 및 특성과 연관되며, 이러한 교통량규모에 따라 차종구성비에서도 차이를 보이기 때문에 이를 반영하기 위하여 연평균 톨게이트 이용교통량을 기준으로 <표 7>과 같이 톨게이트를 분류하였다. 또한, 분류기준을 설정하기 위하여 톨게이트의 교통량자료를 근거로 군집분석(평균연결법)을 통해 분류하였다.

<표 7> 교통량 규모분류 기준

구분	교통량
I 그룹	25000대 이상
II 그룹	10000대 이상~25000대 미만
III 그룹	10000대 미만

3) 교통량의 시간변동에 따른 분류

광역도시권역의 톨게이트일수록 차종구성비는 그 이용시간이 출·퇴근시간과 그 외 시간에 따라 차이를 가지게 되며, 평일과 휴일에 따라 교통량과 구성비는 달라지게 된다. 따라서 계절, 요일, 침두시 등 시간적 변화에 따라 고속도로 이용차량의 구성이 다소간의 차이를 보여 시간변동에 따른 차종구성비 변화를 분석하였다.

그 결과, 계절별 변동의 경우 이용차량의 증감에는 영향을 주었으나, 구성비에서의 변동은 크지 않았다. 그러나 요일별, 시간대별 변동은 존재하는 것으로 분석되었다.

(1) 요일별 변동에 따른 분류

요일별 차종구성비분포를 분석한 결과, 요일별로 연평균 차종구성비가 평일에는 유사한 분포를 보였으나, 휴일 및 토요일에는 톨게이트별로 상이한 결과를 보여 <표 8>과 같이 평일과 휴일(토,일,연휴) 2개의 그룹으로 나누어 분석하였다.

<표 8> 그룹별 분류기준

구분	카테고리
1 그룹	평일(월요일~금요일)
2 그룹	휴일(토,일,연휴기간)

〈표 9〉 요일별 변동에 따른 톨게이트 분류

유형	구분		톨게이트수
	평일	휴일	
1	M	L, H	99
2	L	H	27
3	H	L	26

또한, 평일과 휴일의 총 교통량 및 각 톨게이트의 평균 교통량구성비를 산정하였다. 각 톨게이트의 경우 평균 교통량구성비는 총 교통량의 평균 교통량구성비(μ)와 각 톨게이트의 평균 교통량구성비와의 편차($\mu - \sigma, \mu + \sigma$)를 기준으로 하여 톨게이트의 평균 교통량 구성비가 총 교통량의 평균과 유사한 형태이면 M, 총 교통량의 평균보다 높은 경우 H, 총 교통량의 평균보다 낮은 경우를 L로 설정하였다.

분류결과, 톨게이트의 평일과 휴일의 평균 교통량구성비는 〈표 9〉와 같이 3가지 유형으로 분류되었다(단 본 연구의 표본조사는 평일을 기준으로 시행되어 분류를 평일중심으로 설정하였다).

(2) 시간대별 변동에 따른 분류

고속도로의 시간대별 교통량은 특별한 침두시 없이 거의 일정한 교통량을 보인 경우와 오전·오후 침두시간이 나타나는 2가지 형태로 나타났다. 또한 시간에 따른 차종별 교통량 분포의 경우 2~5종은 일정한 반면, 1종은 다른 차종에 비해서 확연히 다른 분포를 나타내었다. 따라서, 본 연구에서는 일일 침두율을 사용하여 톨게이트를 분류하였던 바, 톨게이트별 일 침두율의 분포가 0.4 와 0.6 부근에서 변곡점이 발생하는 것을 볼 수 있어 〈표 10〉과 같이 세 그룹으로 나누어 분류하였다.

〈표 10〉 시간변동별 톨게이트 분류결과

구분	구분	톨게이트	
		유입	유출
1	0.2 ≤ 침두율 < 0.4	7	10
2	0.4 ≤ 침두율 < 0.6	125	111
3	0.6 ≤ 침두율 < 0.8	20	31

4) 차종구성비에 따른 분류

TCS 5종 데이터는 여러 요인으로 각 차종간의 구

〈표 11〉 1종 차종구성비에 따른 분류

구분(%)	톨게이트	
	유입	유출
1 80 ≤ 1종 구성비 < 90	7	6
2 70 ≤ 1종 구성비 < 80	60	70
3 60 ≤ 1종 구성비 < 70	85	76

성비는 동일하지 않다. 그러나 대부분의 톨게이트에서 1종이 차지하는 비중이 가장 크기 때문에 표준차종분류에서도 마찬가지로 승용차의 분포가 차종구성비의 상당부분을 차지할 것으로 판단된다. 따라서, 본 연구에서는 TCS 1종의 구성비에 따라 톨게이트를 분류하였다(단, 이 경우 유입, 유출을 별도로 분류하였다).

2. 카테고리간의 독립성 검증

본 연구에서 제시한 각 카테고리의 유의성을 살펴 보기 위하여 동질성의 가설을 갖는 카이제곱 검증 (Chi square test)으로 검증하였다.

카이제곱 통계량은 피어슨의 카이제곱 통계량(Pearson's Chi square statistics:Chi-s)이 가장 일반적이고 (Bishop et al., 1975), 이를 보정한 우도비 카이제곱 통계량(likelihood ratio Chi square:L.R.)의 형태가 보편화되어 있다. 또한 카이제곱 통계량을 수정 보완한 Mantel-Hanenszel:M-H)방법 등이 제안되어 활용되고 있다(Conte & deBor, 1972 ; Cook et al, 1972). 또한, 연관성 척도로서 피어슨(Phi Coefficient :P.C.), 분할계수(Contingency Coefficient:C.C.) 등이 적용되고 있다. 따라서 카테고리간의 독립성검정의 기준을 유의수준 $\alpha = 0.05$ 로 하고, Chi-s, L.R., M-H 통계량 모두 성립하는 기각역으로 독립성유무를 판단하였다(Cook et al., 1972 ; Devroye et al., 1986) 그 결과, 영향권규모와 기능에 따른 카테고리간 독립성에 대한 검증통계량의 유의성정도는 〈표 13〉에서 보는 바와 같이 p-value가 Chi-s인 경우 0.077, M-H은 0.529로서 0.024인 L.R.의 경우를 제외하면 유의수준 $\alpha = 0.05$ 에서 독립성을 갖는다고 판단되며, 세 가지 검증통계량을 모두 적용하면 그 기각역은 작다고 하겠다. 따라서, 영향권규모에 따른 카테고리과 기능에 따른 카테고리는 상호 독립적 성향을 갖고 있다고 할 수 있다(Edward et al., 1977).

영향권의 규모-이용 교통량규모에 따른 카테고리의

경우 Chi-s, L.R 검정통계량이 각각 77.868, 80.792 (p-value : 0.001)로 파이계수와 분할계수(각각 0.716, 0.582)에서 보듯이 종속적인 성향이 나타났다.

그러나 M-H 검정에서는 이를 뒷받침하고 있지 않아 종속적이라기 보다는 상호연계정도는 있으나 독립 성향을 띠고 있다고 결론지을 수 있다.

영향권의 규모-요일변동에 따른 카테고리, 영향권의 기능-요일의 변동에 따른 카테고리, 그리고 영향권의 기능-시간의 변동(유출)에 의한 카테고리의 경우 유사한 독립성 결과를 보였으나 적합성 또는 연관성(각각, 0.291, 0.284, 0.221, 0.314)이 높지 않았다. 또한, 이용 교통량규모-시간변동(유입) 카테고리의 경우 세 가지 검정통계량이 각각 21.880, 19.436, 18.728(p-value : 0.001)로서 독립적이라는 결론은

어려웠으나 연관성정도는 0.268~0.379정도로 미미하였다.

따라서, 톨게이트의 특성에 따른 각각의 카테고리의 독립성검정결과는 사회적, 경제적발전과 더불어 톨게이트의 고유의 기능과 시·공간적 변화양상을 고려한다면 다소 그룹간(카테고리영역별) 상관관계가 있지만, 그 정량적인 값은 미미하고 카테고리의 탐색적인 그룹화 결과와 부합되는 결과로 판단되어 설정된 카테고리를 모두 사용하기로 하였다.

3. 카테고리 분류 최종결과

각 카테고리의 독립성검정을 통해 총 6개 유형의 유의성있는 카테고리를 설정하였다.

<표 12> 각 카테고리간 독립성 검정결과

		통계량(p-value)			계수		
		Chi-S	L. R.	M-H	P. C.	C. C.	C. Vx
영향권의 규모	지역의 기능	11.391(0.077)	14.537(0.024)	0.397(0.529)	0.274	0.264	0.194
	이용 교통량규모	77.868(0.001)	8(0.792(0.001)	0.997(0.318)	0.716	0.582	0.506
	요일변동	25.747(0.001)	31.368(0.001)	0.004(0.948)	0.412	0.381	0.291
	시간변동(유입)	8.081(0.232)	8.682(0.192)	1.358(0.244)	0.231	0.225	0.163
	1종 차종구성비(유입)	5.125(0.528)	4.917(0.555)	0.322(0.571)	0.184	0.181	0.13
	시간변동(유출)	10.243(0.115)	9.207(0.162)	0.098(0.754)	0.26	0.251	0.184
	1종 차종구성비(유출)	6.53(0.367)	7.491(0.278)	0.003(0.958)	0.207	0.203	0.147
영향권의 기능	이용 교통량규모	5.965(0.202)	8.308(0.081)	2.402(0.121)	0.198	0.194	0.14
	요일변동	9.227(0.056)	8.53(0.074)	0.749(0.387)	0.246	0.239	0.174
	시간변동(유입)	24.459(0.001)	19.858(0.001)	1.707(0.191)	0.401	0.372	0.284
	1종 차종구성비(유입)	12.689(0.013)	15.421(0.004)	6.67(0.01)	0.289	0.278	0.204
	시간변동(유출)	14.807(0.005)	12.961(0.011)	0.062(0.803)	0.312	0.298	0.221
	1종 차종구성비(유출)	11.865(0.018)	13.826(0.008)	3.488(0.062)	0.279	0.269	0.198
이용 교통량규모	요일변동	12.718(0.013)	17.176(0.002)	0.031(0.861)	0.289	0.278	0.205
	시간변동(유입)	21.88(0.001)	19.436(0.001)	18.728(0.001)	0.379	0.355	0.268
	1종 차종구성비(유입)	4.522(0.34)	4.403(0.354)	2.818(0.093)	0.172	0.17	0.122
	시간변동(유출)	9.007(0.061)	9.422(0.051)	8.704(0.003)	0.243	0.237	0.172
	1종 차종구성비(유출)	4.153(0.386)	4.684(0.321)	0.13(0.718)	0.165	0.163	0.117
요일변동	시간변동(유입)	6.942(0.139)	7.251(0.123)	2.232(0.135)	0.214	0.209	0.151
	1종 차종구성비(유입)	1.836(0.766)	1.753(0.781)	0.98(0.322)	0.11	0.109	0.078
	시간변동(유출)	30.04(0.001)	24.867(0.001)	3.399(0.065)	0.445	0.406	0.314
	1종 차종구성비(유출)	4.631(0.327)	4.514(0.341)	0.753(0.386)	0.175	0.172	0.123
시간변동(유입)	1종 차종구성비(유입)	4.631(0.327)	4.514(0.341)	0.753(0.386)	0.175	0.172	0.123
시간변동(유출)	1종 차종구성비(유출)	3.779(0.437)	3.433(0.488)	2.103(0.147)	0.158	0.156	0.111

주) Chi-s : Chi-statistic, L. R. : likelihood ratio Chi-statistic, M-H : Mantel-Hanenszel Chi-statistic P. C. : Phi Coefficient, C. C. : Contingency Coefficient, C. Vx : Cramer's V

〈표 13〉 카테고리별 톨게이트수

		1그룹	2그룹	3그룹	4그룹	합계	
영향권 특성	영향권의 규모	19	23	63	47	152	
	영향권의 기능	115	19	18	-	152	
이용 교통량규모		22	24	106	-	152	
시간적 변화	요일별	99	27	26	-	152	
	시간 대별	유입	7	125	20	-	152
		유출	10	111	31	-	152
차종구성비 변화	1종차종 구성비	7	60	85	-	152	
	유출	6	70	76	-	152	

즉, 영향권의 규모 4개, 영향권의 기능 3개, 이용교통량 규모 3개, 요일별 변동 3개, 시간적 변동 3개, 그리고 1종 차종구성비의 변동 3개 그룹으로 총 972가지의 경우로 분류하였다(Puget Sound, 1994).

그러나 현재 존재하는 톨게이트의 수는 152개로 경우의 수에 비해 매우 적을 뿐 아니라 카테고리분류의 경우의 수가 유입방향의 경우의 수는 71가지, 유출방향의 경우의 수는 76가지에 불과하다.

하지만 본 연구에서 제시한 경우의 수는 향후 건설되는 톨게이트를 고려하여 설정한 것이며, 또한, 기존 톨게이트의 토지이용 변화로 인해 그 특성이 변경되었을 때, 능동적으로 대처할 수 있도록 하여, 카테고리의 분류가 향후에도 지속적으로 사용될 수 있도록 고려하였다.

IV. 표본조사

1. 조사 표본 톨게이트 설정

조사표본 톨게이트는 난수생성에 의한 무작위 추출법을 통해 계층별로 균형을 맞춰 최저 표본율을 선택하는 방법을 이용하여 전국 톨게이트를 대상으로 카테고리별로 각 그룹당 20% 내외를 추출하였다.

그 결과 최종 조사대상 톨게이트는 32개소로 설정되었으며, 조사대상 톨게이트의 그룹현황을 살펴보면 〈표 14〉와 같다.

〈표 14〉 표본톨게이트의 카테고리별 그룹현황

톨게이트	영향권 특성		톨게이트 이용 교통량 규모	요일별	시간별		1종 차종구성비	
	지역의 규모	지역의 기능			유입	유출	유입	유출
인천	1	1	1	3	3	3	1	1
성남	2	1	1	1	2	3	2	2
구리	2	1	1	1	3	3	2	2
토평	2	1	3	1	2	2	1	1
김포	2	1	1	1	2	2	2	2
서울	1	1	1	3	2	2	2	2
동수원	1	2	1	1	3	3	3	3
수원	1	1	1	3	2	3	2	2
기흥	3	1	2	3	2	1	2	2
오산	3	1	2	3	3	3	3	3
안성	3	1	2	3	2	3	3	3
천안	3	1	2	1	2	3	3	3
대전	1	1	2	1	2	2	2	2
서대구	1	1	1	1	3	2	3	3
경주	3	3	2	2	2	2	3	3
부산	1	1	1	1	2	2	3	3
울산	1	2	2	1	3	3	3	3
유성	1	3	3	1	2	2	2	2
백양사	4	3	3	2	2	2	3	2
광주	1	1	2	1	2	2	3	3
용인	3	3	2	3	3	2	2	2
이천	3	1	2	1	2	2	3	3
여주	4	1	3	1	2	2	3	3
새말	3	3	3	1	2	3	3	3
둔내	4	3	3	1	2	3	2	2
월정	4	3	3	2	2	3	2	2
동서울	1	1	1	1	2	2	2	2
일죽	3	1	3	1	2	2	3	3
증평	4	1	3	1	2	2	3	3
동광양	3	2	3	3	1	1	2	2
진월	3	2	3	3	1	1	3	3
발안	3	1	3	3	2	2	3	3

2. 조사방법

각 톨게이트를 입·출구로 나누어 기본적으로는 조사원에 의한 방법을 원칙으로 조사를 실시하였으나, 일부 지역³⁾

3) 각 톨게이트별로 시간대에 따라 유출·입 게이트가 운영되는 수가 다르나 24시간 통행권 조사시에 통과차량들의 영향을 받지 않는 유성, 백양사, 동광양, 진월, 둔내에 한해서는 톨게이트 이용시에 사용되고 있는 통행권을 이용하여 조사하였다.

에 한해서는 출구방향에 대해 통행권을 이용한 조사방법을 사용하였다. 조사원을 동원하는 경우, 시간당 45분간 조사하도록 하여 조사원이 피로로 인해 발생할 수 있는 비표본오차(non-sampling error)를 최소화하도록 하였다. 차종분류는 본 연구에서 제시하고 있는 표준차종으로 분류하였다.

V. 톨게이트 차종별 교통량 추정모형 구축

1. 모형의 이론적 고찰

본 연구에서 설정된 모형은 표본 톨게이트의 표준차종구성비를 바탕으로 차종별 대표치를 구하였으며, 그 방법으로 점 추정량을 산출하는 방법과 비모수적 추정기법중의 하나인 붓스트랩(bootstrap)과의 비교를 통하여 극단적 표본치로 인해 발생될 수 있는 위험률(risk)을 최소화할 수 있도록 하였다(Efron, 1982).

1) 평균치를 이용한 방법

점 추정량을 산출하는 방법으로서 평균치를 이용할 수 있는데 그 대표적인 방법으로 산술평균, 기하평균, 조화평균 등이 있다.

산술평균은 모든 자료의 총합을 총 도수로 나눈 값으로 통계적으로 가장 일반화되어 있는 방법이다. 그러나, 산술평균은 극단값에 민감해서 과대 또는 과소 추정되는 경향이 있다.

$$\text{산술평균} = \frac{1}{n} \sum_{i=1}^n x_i$$

기하평균은 관측자료의 편차가 큰 경우를 고려하여 각 자료의 로그 변환한 산술평균으로 정의된다. 그러므로 어떤 시점을 기준으로 할 때의 물가지수, 인구증가율 등의 비율로서의 평균값의 의미로 이용된다.

$$\text{기하평균} = \sqrt[n]{x_1 \times x_2 \times \dots \times x_n}$$

조화평균은 각 구간의 거리가 일정한 경우와 단위 시간 개념의 평균속도 및 평균가격 등을 나타낼 때 이용되며, 산술평균보다는 안정된 대표값으로 활용되고 있다.

$$\text{조화평균} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$$

2) 붓스트랩을 이용한 방법

(1) 경험 분포와 플러그-인 원리

통계적인 추정문제에 있어서 모집단의 분포에 대한 정보가 전혀 없으며, 단지 표본 (x_1, x_2, \dots, x_n) 만이 주어져 있을 때, 이 자료를 가지고 모집단의 확률분포를 알아보려고 하는 경우가 있다. 분포함수 F 에 대한 추정치로 일반적으로 다음과 같이 정의되는 경험 분포함수(empirical distribution function) \hat{F}_n 를 사용한다.

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n I(x_i \leq x),$$

여기서, $I(x_i \leq x) = \begin{cases} 1, & x_i \leq x \\ 0, & x_i > x \end{cases}$

경험분포함수의 정의를 이용하여 확률밀도 함수와 어느 사상의 확률에 대한 추정치도 다음과 같이 표현할 수 있다.

$$\hat{f}_n(x) = \frac{1}{n} \sum_{i=1}^n I(x_i = x),$$

$$\hat{P}_n(A) = \frac{1}{n} \sum_{i=1}^n I(x_i \in A)$$

또한, 표본으로부터 모수를 추정하기 위해 경험분포함수를 이용할 때, 자료만을 사용하기 때문에 이를 '플러그-인 원리' 라고 부른다.

(2) 붓스트랩의 용어정립

분포 F 에서 독립적으로 추출된 표본 (x_1, x_2, \dots, x_n) 으로 관심있는 모수 θ 를 추정하고자 할 때, 분포 F 를 모르는 경우에 플러그-인 추정량 $\hat{\theta}$ 으로 모수 θ 를 점 추정할 수 있다. 이때, $\hat{\theta}$ 의 모수 θ 에 대한 정밀도(precision)는 구간추정이 요구되는데 분포를 모르는 상황에서 $\hat{\theta}$ 의 분산 또는 표준편차를 알 수 있는 방법으로서 붓스트랩 기법이 제안되어 활용되고 있다.

특히, 이 기법은 경제적 측면에서 적은수의 표본을 가지고 유효성 있는 추정을 하는 경우와 미지의 분포 특성의 충분성(sufficiency)을 위한 추정기법이라 할

수 있다(Gelfund et al., 1984). 붓스트랩기법은 추정량의 오차를 줄일 수 있는 비모수적인 방법으로 그 응용과 활용에 있어서 검증된 고급의 추정기법으로 알려져 있다(Law et al., 1982 ; Marriot, 1971). 또한 점추정량의 단점인 신뢰구간을 제공함으로써 실용성(performance)문제에서도 능률적으로 대처할 수 있는 장점을 지니고 있다(물론, 점추정량이 신뢰구간을 갖는 경우에는 데이터의 분포를 아는 경우 가능하나 실제로, TCS 자료의 경우 카테고리별로 고유의 특성은 있으나, 특정한 분포가정은 무리가 따른다).

붓스트랩 방법은 분포를 모르는 경우에 유용하며, 특정분포를 가정하지 않는 비모수적 접근방식으로 여러 분야에서 활용되고 있는데, 이는 전산알고리즘을 통하여 표본을 재생성(resampling)하는 방법이므로 일반적인 경우(산술, 기하, 조화평균 등)와의 비교를 <표 15>와 같이 정리할 수 있다.

<표 15> 실제 경우와 붓스트랩 방법의 비교

구분	실제 경우	붓스트랩 방법
사용 분포	모르는 확률분포 F	경험분포함수 F_n
자료	관측된 표본 (x_1, x_2, \dots, x_n)	붓스트랩 표본 $(x_1^*, x_2^*, \dots, x_n^*)$
실험의 반복	반복이 없음	반복수를 임의로 설정
추정량	$\hat{\theta} = T(x)$	$\hat{\theta}^* = T(x^*)$
추정방법	점추정	점추정, 구간추정

(3) 붓스트랩 알고리즘

주어진 표본 $x=(x_1, x_2, \dots, x_n)$ 으로 모수 θ 를 추정하고자 할 때, 붓스트랩 알고리즘을 이용한 차종 구성비(%) 추정은 먼저, 카테고리내 그룹별 차종구성비를 산출하고 그 결과를 통해 카테고리간 차종구성비 산출한다.

① 카테고리내 그룹별 차종구성비 산출단계

[단계1]

주어진 TCS 표본 $x=(x_1, x_2, \dots, x_n)$ 으로 경험분포함수를 만든다.

여기서, x_n 은 각 카테고리의 그룹내 차종 구성비(%)

이고 먼저, 표본조사 32개중 그룹내의 톨게이트로서 순서통계량을 산출한다.

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$$

그리고, 각 순서통계치에 해당되는 경험분포함수 값을 다음과 같이 대응시키고, 같은값을 갖는 표본에 대하여는 확률을 합하여 1종~11종 차종구성비를 누적분포함수(cumulative distribution function)화한다.

$$\frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1$$

n : 카테고리내 그룹별 톨게이트 수

[단계 2]

재 표본추출 및 추정과정을 100번 반복한다((단계 1)에서 구한 경험분포함수를 이용하여 그룹 내 차종비를 무작위추출과정으로 $B=100$ 개를 모의 실험하여 붓스트랩 표본을 생성하고, 표본오차를 추정한다).

그 과정은 먼저, [단계 1]의 경험분포함수와 0과 1사이의 난수를 이용하여 새로운 표본을 반복 생성한다(이를 붓스트랩 표본이라고 부른다).

$$x^{*k} = (x_1^{*k}, x_2^{*k}, \dots, x_n^{*k})$$

$$k = 1, 2, \dots, B \quad (B=100)$$

n : 카테고리내 그룹별 톨게이트 수

그리고, 재생성된 붓스트랩 표본으로 추정치 $\hat{\theta}(*k)$ 를 구한다. 이것은 새로운 표본을 반복생성과정에서 생성된 $B=100$ 개의 재 표본 추출과정에 의한 붓스트랩 표본의 표본오차를 계산하여 그룹내 각 차종구성비 $N=40$ 개⁴⁾를 산출하기 위한 전 단계이다.

[단계 3]

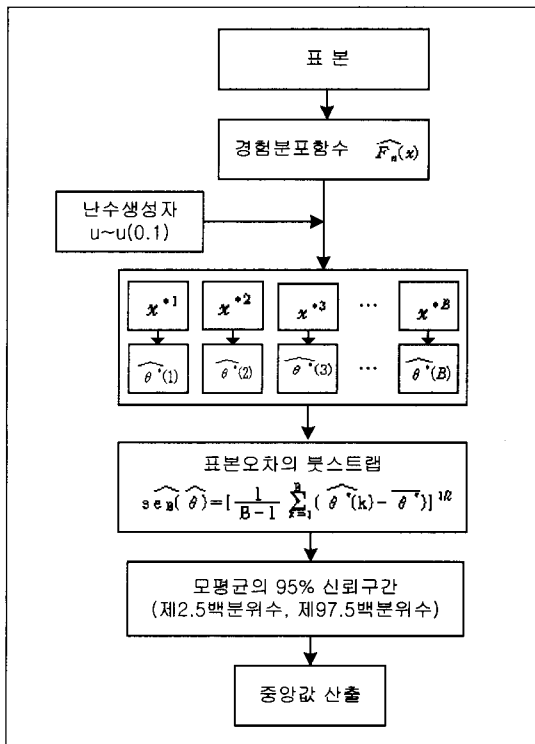
반복단계에서 구한 붓스트랩 추정치들을 이용하여 추정량의 분포를 추정한 후, 최종적인 그룹별 1~11개 차종구성비($N=40$)를 산출한다.

4) 붓스트랩과 같은 비모수적 방법에서의 신뢰구간은 모수적 방법과는 달리 기각역의 수는 자연수가 되어야 한다. 즉, 95% 신뢰구간(유의수준 5%)인 경우에 $N=40$ 일 때, 양쪽 기각역 2.5%는 1개씩의 표본에 해당되므로 양측 표본을 제외시킴으로써 신뢰구간을 구하게 된다. 그러므로 95% 신뢰구간은 통상 40의 배수로 설정한다. 따라서 본 연구에서는 표본을 최소 $N=40$ 으로 설정하였다.

산출된 그룹별 40개의 차종구성비의 순서 통계량을 구하여 95% 신뢰구간 추정치를 출력한다. 여기서 점추정치, 즉 그룹별 11개 차종구성비는 중앙값으로 최종 출력에 포함시킨다(구간 추정값에서 점추정량 산출은 일반적으로 중앙값이 안정적으로 평가받고 있다).

② 카테고리간 차종구성비 산출단계

입/출력에 따른 6개의 카테고리간 11개 차종구성비추정의 붓스트랩 알고리즘은 위 ① 과정의 붓스트랩 추정 알고리즘과 동일한 방법으로 산출한다.



<그림 2> 붓스트랩 알고리즘 순서도

[단계 1]

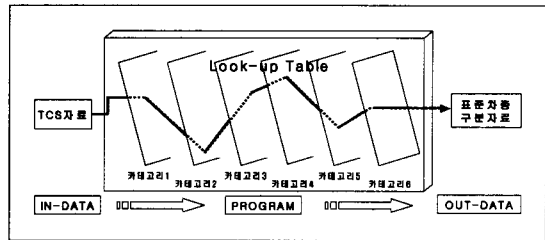
앞에서 카테고리내 그룹별 대표치(95% 신뢰구간의 중앙값)를 6개의 데이터표본을 설정한다.

[단계 2]

설정된 6개의 표본을 앞에서 제안한 붓스트랩 추정 알고리즘을 동일한 방법으로 적용하여 틀게이트별 11개 차종구성비 95%신뢰구간과 점추정치(중앙값)를 출력한다.

2. 모형의 정립

각 카테고리별 차종구분 구성비를 이용하여 표준차종별 교통량을 산출하는 모형의 기본적인 구조는 <그림 3>과 같다.



<그림 3> 모형의 기본개념도

$$T_m = T_n \times f(P_{nm}^1, \dots, P_{nm}^5)$$

- T_m : i 틀게이트의 m 수단(표준차종별) 유출입량
- T_n : i 틀게이트의 n 수단(5종) 유출입량
- $f(\cdot)$: 통계적 대표치 추정함수
- P_{nm}^i : i 번 카테고리의 5종(n)별 표준차종(m) 구성비

여기서 말하는 대표치추정함수는 전술한 방법인 평균치와 붓스트랩을 이용한 방법을 통하여 카테고리의 그룹별 대표치를 산출하는 방법중의 최적의 방법을 의미한다.

각 그룹의 표준차종 구성비의 대표치를 구하는 방법은 각 카테고리에 대한 통계적 대표치를 이용하는 방법과 동일하게 적용하여 산출하였다. 이때, 유입과 유출을 구분하여 산출하며, 산출결과를 실측치와 비교하여 오차가 적으면서 논리적이고 합리적으로 설정할 수 있는 방법을 최종모형에 적용하였다.

3. 차종별 구성비 작성

카테고리의 그룹별로 작성되어질 각각의 차종구성비는 전체 교통량, 즉 틀게이트별 교통량에 대한 백분율을 환산하였을 경우 차종별 교통량이 적은 경우에 발생될 수 있는 오류를 최소화하기 위하여 개개의 TCS차종별로 표준차종구성비를 추정하였다. 그러나 TCS자료상에서의 기타차량(면제차량 등)에 대해서는

고려하기가 어렵고, 그 양이 매우 적어 무시하기로 하였다.

4. 대안별 모형결과

1) 카테고리내 그룹별 대표치 산정

평균치를 이용한 방법과 붓스트랩 방법을 통하여 카테고리 내 그룹별 대표치를 각각 산출하였다.

붓스트랩을 이용한 경우에는 무작위추출 반복횟수를 100회⁵⁾로 하였으며, 구간 추정치에서 대표값은 극단치를 고려하여 중앙값(median)으로 하였다.

2) 카테고리간 연결 대표치 산정

그룹별 대표치에서 얻은 결과를 통하여 6개의 카테고리별 대표치를 재정산하여 하나의 새로운 대표치로 설정하였다.

정산방법은 그룹별 대표치를 산출할 때와 동일한 방법을 적용하였고, 산술, 기하, 조화평균인 경우에는 점추정량을 붓스트랩인 경우에는 중앙값으로 설정하였다.

5. 모형의 검정(Test)

이상에서 제시된 통계적 추정방법 중에서 실측치와 비교하고, 오차범위를 설정한 후 가장 좋은 추정치를 도출하여 이를 최종 표준차종분류에 적용하였다. 또한, 실험오차 및 극단값으로 발생한 값을 제거한 후 2가지 기법을 비교·선택하였다.

검정방법으로는 재조사(re-sampling)방법이 가장 안정적이라 할 수 있으나 유입표본을 기준으로 모형을 구축하여 유출을 비교 판단할 수 있으며, 152개의 틀게이트 중에서 본 연구에서 사용된 적지 않은 표본수, 그리고 조사비용 등을 고려하여 모형치를 실측치에 직접비교하는 방법을 사용하였다.

1) 검정방법

(1) Theil의 부등계수

Theil의 부등계수(Theil's inequality coefficient)

는 H. Theil(1961)에 의해 소개되었으며 RMSE(Root Mean Square Error)의 개념으로 정의되고 있다.

이는 실측치와 관측치와의 차이를 비교하여 관측치가 실측치를 얼마나 잘 표현하고 있는가를 설명하는 이론으로서 Theil의 부등계수 U⁶⁾값의 특성을 살펴보면 다음과 같다.

$$U = \frac{\sqrt{\frac{1}{T} \sum_{i=1}^T (Y_i^s - Y_i^a)^2}}{\sqrt{\frac{1}{T} \sum_{i=1}^T (Y_i^s)^2} + \sqrt{\frac{1}{T} \sum_{i=1}^T (Y_i^a)^2}}$$

여기서,

Y_i^s : 추정값

Y_i^a : 실측값

T : 관측수

RMSE의 경우 관측치와 추정치사이의 중앙값에 대한 분산정도를 설명하는 반면, Theil의 부등계수는 관측자료가 정규분포를 따르는지에 관계없이 오차범위를 결정할 수 있는 방법으로서 RMSE와 달리 절대적인 평가가 가능하며, U값이 0에 가까울수록 모형의 설명력이 매우 높다고 볼 수 있다(Theil, 1961).

(2) 검정방법

실측치에 대한 모형치의 검정은 표본조사를 통한 11개의 표준차종을 실측치로 가정하고 표본조사를 통해 산출된 카테고리별 차종구성비를 이용한 산출결과를 모형치로 가정하여 이들간의 오차를 비교하는 방법으로 Theil의 부등계수를 사용하였다.

비교대상은 앞에서 제시한 그룹의 대표치 산출방법과 각 카테고리의 연결을 통한 대표치를 조합한 16가지의 방법을 유입·출 방향에 대해 각각 산출하여 최소오차이면서 이론적인 합리성이 있는 방법을 선택하였다. 또한, 표준차종 중 TCS자료와 일치하는 8종(TCS자료의 차종분류상 4종)에 대해서는 검정대상에서 제외시켰다.

2) 검정결과

검정결과를 살펴보면, 유입방향의 경우 기하평균-

5) 기 연구된 바에 의하면 반복횟수가 30회 이상이면 유의성을 가진다고 발표되었다.

6) 이하부터는 Theil의 부등계수를 U값으로 칭하기로 한다.

〈표 16〉 주요 톨게이트의 표준차종에 의한 교통량 추정치 결과 예 (고속도로 정기교통량 조사시점)

ID	톨게이트	년	월	일	유출입	1종	2종	3종	4종	5종	6종	7종	8종	9종	10종	11종
101	서울	1999	10	22	유입	49,765	13,902	210	3,644	1,2840	6,392	3,498	1,294	486	541	72
101	서울	1999	10	22	유출	53,516	14,204	654	3,677	13,900	6,283	3,480	1,479	589	633	93
102	동수원	1999	10	22	유입	22,082	6,458	200	1,230	5,841	5,095	1,243	1,961	1,478	1,971	246
102	동수원	1999	10	22	유출	20,276	5,487	722	1,205	5,480	4,343	1,209	1,889	1,433	1,790	208
103	수원	1999	10	22	유입	20,189	5,638	79	807	5,215	2,380	775	541	265	294	40
103	수원	1999	10	22	유출	20,614	5,489	229	819	5,438	2,200	784	572	288	295	48
140	부산	1999	10	22	유입	13,461	3,953	48	750	3,741	1,629	761	893	828	974	122
140	부산	1999	10	22	유출	13,871	3,935	173	870	3,964	1,590	806	1,072	910	1,055	125

산술평균, 조화평균-산술평균의 Theil의 부등계수 U 값이 0.04894로서 가장 실측치에 유사한 것으로 나타났다. 그러나 전체적으로 볼 때, 조화평균-기하평균, 조화평균-조화평균의 경우를 제외하고 모두 U 값이 0.05내외로서 상당히 실측치에 유사한 수준으로 정산되었음을 판단할 수 있었다.

유출방향의 모형대안별 U값을 살펴보면 붓스트랩-기하평균의 U값이 0.04303으로 가장 작은 값을 보여 실측치에 가장 유사한 값으로 나타났다. 또한 전체적인 분포를 살펴보면 유입의 경우와 마찬가지로 조화평균-기하평균, 조화평균-조화평균의 경우를 제

외하고는 모두 U값이 0.05이내로 상당히 실측치에 유사한 수준으로 정산되었다. 유입의 경우와 비교해 볼 때, 유출의 경우 U값이 작아 실측치에 다소 더 근사한 것으로 나타났다.

이러한 분석결과는 도시의 규모, 도시의 기능, 톨게이트 교통량규모 그리고 요일별 변동 카테고리의 분류가 유출데이터 기준으로 되어있기 때문으로 보여진다.

따라서, 유입방향의 오차를 줄이기 위해 4개의 카테고리의 데이터가 유출방향의 속성이 들어가므로 유입방향의 데이터를 적용할 경우에는 좀더 오차가 적어질 것으로 예측되어진다.

결과적으로 평균치와 붓스트랩을 이용한 방법 등은 몇 가지조합을 제외하고는 모두 유의한 수준으로 실측치에 가깝게 추정되었다. 따라서 실험오차와 극단치를 고려하여 붓스트랩-붓스트랩방법을 최종적으로 설정하였다.

6. 모형의 적용

본 연구에서 설정된 모형은 1999년 고속도로 정기교통량 조사시점인 10월 22일~23일 TCS자료 중 일일교통량을 이용하여 당시의 전체톨게이트인 152개를 대상으로 적용하였다.

그 중에서 서울톨게이트를 예로 살펴보면, TCS의 1종에 해당하는 표준차종의 1종인 승용차의 경우 TCS 1종의 65% 정도에 해당하며 표준차종 2종인 승합차의 경우 18%, 화물차의 경우 17% 수준으로 승합차와 승용차의 비율이 타지역에 비해 높게 나타났다.

〈표 17〉 모형대안별 Theil의 부등계수 결과

카테고리내 그룹 대표치	카테고리간 그룹 대표치	부등계수(U값)	
		유입	유출
기하평균	기하평균	0.05009	0.04558
	붓스트랩	0.04892	0.04483
	산술평균	0.04894	0.04476
	조화평균	0.05358	0.04685
붓스트랩	기하평균	0.05177	0.04303
	붓스트랩	0.05246	0.04337
	산술평균	0.05215	0.04331
	조화평균	0.05150	0.04290
산술평균	기하평균	0.05088	0.04321
	붓스트랩	0.05159	0.04348
	산술평균	0.05123	0.04344
	조화평균	0.05065	0.04316
조화평균	기하평균	0.06306	0.06627
	붓스트랩	0.04856	0.04769
	산술평균	0.04894	0.04759
	조화평균	0.08960	0.09685

〈표 18〉 주요 톨게이트의 표준차종에 의한 교통량 추정치의 유의수준 예 (고속도로 정기교통량 조사시점)

ID	톨게이트	년	월	일	유출입	통계값	1종	2종	3종	4종	5종	6종	7종	8종	9종	10종	11종
101	서울	1999	10	22	유입	상한값	49,957	13,972	220	3,672	12,949	6,397	3,522	1,294	492	547	74
101	서울	1999	10	22	유입	하한값	49,468	13,789	202	3,591	12,687	6,381	3,456	1,294	481	534	70
101	서울	1999	10	22	유출	상한값	53,853	14,357	757	3,720	14,088	6,365	3,528	1,479	593	638	96
101	서울	1999	10	22	유출	하한값	53,081	14,035	567	3,631	13,711	6,201	3,444	1,479	578	620	90
102	동수원	1999	10	22	유입	상한값	22,140	6,494	225	1,235	5,941	5,111	1,247	1,961	1,529	2,030	253
102	동수원	1999	10	22	유입	하한값	21,949	6,410	174	1,217	5,738	5,074	1,232	1,961	1,418	1,901	238
102	동수원	1999	10	22	유출	상한값	20,410	5,554	777	1,220	5,579	4,426	1,225	1,889	1,472	1,821	214
102	동수원	1999	10	22	유출	하한값	20,138	5,430	656	1,191	5,387	4,294	1,195	1,889	1,398	1,748	202
103	수원	1999	10	22	유입	상한값	20,267	5,665	81	812	5,250	2,383	781	541	268	297	40
103	수원	1999	10	22	유입	하한값	20,106	5,612	76	800	5,165	2,377	768	541	263	292	39
103	수원	1999	10	22	유출	상한값	20,732	5,544	246	828	5,524	2,242	795	572	290	295	48
103	수원	1999	10	22	유출	하한값	20,522	5,446	206	812	5,369	2,193	777	572	285	292	46
140	부산	1999	10	22	유입	상한값	13,489	3,964	49	753	3,762	1,629	761	893	839	987	124
140	부산	1999	10	22	유입	하한값	13,375	3,923	47	744	3,691	1,627	754	893	813	956	118
140	부산	1999	10	22	유출	상한값	13,885	3,941	185	878	3,987	1,600	808	1,072	915	1,065	126
140	부산	1999	10	22	유출	하한값	13,830	3,918	150	861	3,930	1,564	794	1,072	891	1,030	120

VI. 결론

본 연구에서는 고속도로전수자료인 TCS 데이터를 이용하여 조사별로 일치되어있지 않은 분류차종을 표준화함으로써 상호간 호환이 가능하도록 기존의 5종 체계의 TCS자료를 11종으로 전환하였다.

이 전환과정에서 영향권특성, 톨게이트 이용교통량 특성, 시간적 특성, 차종구성비에 따른 특성을 고려한 카테고리방법을 사용함으로써 최종적인 대표치를 비모수적인 방법인 붓스트랩을 이용하여 설정하였다.

그러나 연구결과에 대한 재 표본추출(resampling) 작업을 수행하지 않은 관계로 각 대안별로 우위를 판단하기에는 다소무리가 있으나, 평균치를 이용한 방법과 비교하여 붓스트랩기법의 추정오차가 크지 않았으며, 평균치추정기법의 표본수에 따른 영향도 특히, 향후 국가시책 및 지역별 발전향방에 따라 톨게이트 수 및 교통량변동성을 고려할 때, 평균치 추정기법은 대안별 검증결과가 다르게 나타나 오차의 변동폭이 매우 클 것으로 판단된다.

결론적으로 교통량의 경우와 같이 분포의 복잡성 (complexity), 주변의 여러 가지 여건에 의해 발생되는 불규칙성(irregularity) 그리고 표본의 추출과정에서 발생하는 오차 등을 고려할 때, 이러한 문제에 보다 유연한(flexible) 방법으로서 또한, 이론적

근거가 좀더 명확하고 추정치에 대한 신뢰구간을 제시할 수 있는 붓스트랩기법이 보다 합리적이라고 할 수 있겠다.

본 연구결과를 통하여 향후 TCS자료의 활용성이 증대되고, TCS자료를 이용한 고속도로 구간교통량추정과 고속도로 정기교통량 조사자료를 비교할 수 있는 계기를 마련할 수 있을 것으로 기대된다.

참고문헌

1. 한국도로공사(1988), "고속도로 Network Analysis 모형적용을 위한 자료구축 및 실용화 연구".
2. 건설교통부(2000), "도로교통량통계연보".
3. 한국개발연구원(1999), "무안-광양간 고속도로 예비타당성 분석".
4. Bishop, YMM., Fienberg, SE., Holland, PW. (1975), "Discrete Multivariate Analysis: Theory and Practice," MIT Press, MA: Cambridge.
5. Conte, S. D. and deBor, C.(1972), "Elementary numerical analysis," Mcgraw Hill, New York.
6. Cook, D., Craven, A. H. and Clarke, G. M. (1982), "Basic statistical computing," Edward Arnold.
7. Devroye, L.(1986), "Non-uniform random

- variate generation." Springer-Verlag, New York.
8. Edward, K. Morlak(1977), "Introduction to Transportation Engineering and Planning," pp.425~442.
 9. Efron, B.(1982), "Jackknife, the bootstrap and other resampling plans," SIAM.
 10. Gelfund, A. E. and A. F. Smith, M.(1984), "Sampling-based approaches to calculating marginal densities," JASA, 85, pp.398~409.
 11. Jose Banos-Pino, Victor Fernandez-Blanco, Ana Rodriguez-Alvarez(2002), "The allocative efficiency measure by means of distance function : The case of Spanish public railways," EJOR, 610, pp.191~205.
 12. Law, A. M. and Kelton, W. D.(1982), "Simulation modeling and analysis," McGraw-Hill, New York.
 13. Marriot, F. H. C.(1971), "Practical problems in a method of cluster analysis," Biometrics, 27, pp.501~514.
 14. Puget Sound Regional Council(1994), "Travel Model Development And Refinement trip Generation Final Report".
 15. Stoper. P. R. & K, G, McDonald.(1983), "Trip generation by cross-classification : An alternative methodology," Transportation Research Record, No. 2, pp.84~91.
 16. Theil, H.(1961), "Economic Forecasts and Policy," pp.30~37.
 17. Urban Hjorth(2002), "Traffic subflow estimation and bootstrap analysis from filtered counts," TRB, 36, pp.345~359.
 18. Wootton, H. J. and G. W. Pick(1967), "A Model for Trip Generated by Household," Journal of Transport Economics and Policy, 1, pp.137~153.

✉ 주 작 성 자 : 노정현

✉ 논문투고일 : 2001. 10. 26

논문심사일 : 2001. 11. 19 (1차)

2001. 12. 5 (2차)

2001. 12. 31 (3차)

2002. 1. 30 (4차)

심사판정일 : 2002. 1. 30