



도로종류별 교통사고 추세분석 및 시계열 분석모형 개발

Road Accident Trends Analysis with Time Series Models for Various Road Types

한상진* 김근정**
Han, Sang Jin Kim, Kewn Jung

Abstract

Roads in Korea can be classified into four types according to their responsible authorities. For example, Motorway is constructed, managed, and operated by the Korea Highway Corporation. Ministry of Construction and Transportation is in charge of National Highway, and Province Roads are run by each province government. Urban/County Roads are run by corresponding local government. This study analyses the trends of road accidents for each road type. For this purpose, the numbers of accidents, fatalities, and injuries are compared for each road type for last 15 years. The result shows that Urban/County Roads are the most dangerous, while Motorways are the safest, when we simply compare the numbers of accidents, fatalities, and injuries. However, when we compare these numbers by dividing by total road length, National Highway becomes the most dangerous while Province Roads becomes the safest. In the case of road accidents, fatalities, and injuries per vehicle·km, which is known as the most objective comparison measure, it turns out that National Highway is the most dangerous roads again. This study also developed time series models to estimate trends of fatalities for each road type. These models will be useful when we set up or evaluate targets of national road safety.

Keywords : road accidents, times series model, number of accident, number of fatalities, number of injuries

요 지

우리나라의 도로는 종류별로 관리주체가 서로 다르다. 고속국도는 한국도로공사, 일반국도는 건설교통부, 지방도는 광역자치단체, 시·군도는 기초자치단체에서 건설·관리 및 운영·책임을 맡고 있다. 본 연구에서는 이렇듯 관리주체가 다른 우리나라의 도로에서 발생하고 있는 교통사고발생 추세를 도로종류별로 분석하는데 목적이 있다. 이를 위해 과거로부터 축적된 도로종류별 교통사고건수, 사망자수, 부상자수의 자료를 기반으로 사고위험을 단순건수, 도로연장 당 건수, 그리고 대·km 당 건수로 비교하였다. 그 결과 단순 건수에 의한 비교에서는 도시부 도로가 전부분에서 가장 위험하고 고속국도는 전 부분에서 가장 안전한 것으로 분석되었다. 그러나 1km당 발생건수는 일반국도가 가장 위험하고 지방도가 가장 안전한 것으로 분석되었다. 교통사고 발생률을 가장 객관적으로 비교할 수 있는 지표인 10만 차량 대·km 당 비교에서는 일반국도가 가장 위험하고, 고속국도가 상대적으로 안전한 것으로 분석되었다. 한편, 도로종류별 사망자수 추세를 설명하는 시계열 분석모형을 개발한 결과 고속국도, 지방도, 도시부도로의 경우 AR(1)모형이 추세를 잘 설명할 수 있는 것으로 분석되었고, 일반국도의 경우 ARIMA(2, 3, 0)모형이 추세를 설명할 수 있는 것으로 분석되었다 이러한 모형은 장래 교통안전계획의 목표치를 수립하거나 평가하는데 근거자료로 활용될 수 있을 것으로 보인다.

핵심용어 : 교통사고, 시계열 모형, 교통사고건수, 사망자수, 부상자수

* 정회원 · 한국교통연구원 도로교통연구실 연구위원
** 정회원 · 한국교통연구원 도로교통연구실 연구위원



1. 서론

우리나라의 도로는 종류별로 관리주체가 서로 다르다. 즉, 고속국도는 한국도로공사, 국도는 건설교통부, 지방도는 광역자치단체, 시·군도는 기초자치단체에서 건설·관리 및 운영 책임을 맡고 있다. 본 연구에서는 이렇듯 관리주체가 다른 우리나라의 도로에서 발생하고 있는 교통사고발생 추세를 도로 종류별로 분석하는데 목적이 있다. 이를 위해 과거로부터 축적된 도로종류별 교통사고건수, 사망자수, 부상자수의 자료를 도로종류별 연장 혹은 10만 차량 대·km 자료를 서로 연관시켜 분석하고자 한다. 이러한 분석은 어떤 종류의 도로에서 상대적으로 사고감소를 위한 노력을 더 펼치고 있는지를 파악하는데 도움이 될 것으로 보인다. 또한 본 연구에서는 시계열 분석기법으로 교통사고 추세모형을 구축하고자 한다. 시계열 분석기법은 자료의 전반적 추세를 이해하거나 단기간의 예측에 있어 유용한 분석방법으로 알려져 있다.

2. 교통사고 추세분석

본 연구는 도로종류별로 교통사고발생의 추세가 어떻게 다른지를 파악하기 위해 교통사고 발생건수, 사망자수, 부상자수를 단순히 비교한 후에 이들을 도로 종류별 연장과 도로 종류별 10만 차량 대·km로 나누었을 때 그 결과가 어떻게 달라지는지를 파악한다. 모든 자료는 『교통사고 통계분석, 1990~2005, 도로교통안전관리공단』에 근거하여 구축되었다. 도로종류별 10만 차량 대·km는 도로 종류별 평균 일교통량(ADT: Average Daily Traffic)에 도로종류별 연장거리를 곱하여 계산하였다.

도로는 도로법 제11조(도로의 종류와 등급)에 근거하여 고속국도, 일반국도, 특별시도·광역시도, 지방도, 시도, 군도, 구도 등으로 구분된다. 그러나 본 연구에서는 특별시도·광역시도, 시도, 군도를 묶어

도시부 도로로 분류하였다. 따라서 본 연구에 사용된 도로의 종류에는 고속국도, 일반국도, 지방도, 도시부도로가 있다.

2.1 단순건수에 의한 비교

도로종류별(고속국도, 일반국도, 지방도, 도시부도로) 단순사고건수, 사망자수 및 부상자수 추이는 그림 1~그림 3과 같은 추세로 나타났다. 도로종류별 사고발생건수와 부상자수의 순위는 도시부도로, 일반국도, 지방도, 고속국도 순으로 나타났다. 사망자수의 순위는 1989년부터 2004년까지 계속해서 도시부도로가 가장 높은 수치를 나타냈고 그 다음으로 일반국도, 지방도, 고속국도 순위이다. 그러나, 사망자수는 모든 도로에서 1996년 이후 감소하는 경향을 보이고 있고, 발생건수, 부상자수의 경우 대체로 감소하는 경향을 보이나 도시부도로의 경우는 증감이 반복되고 있다.

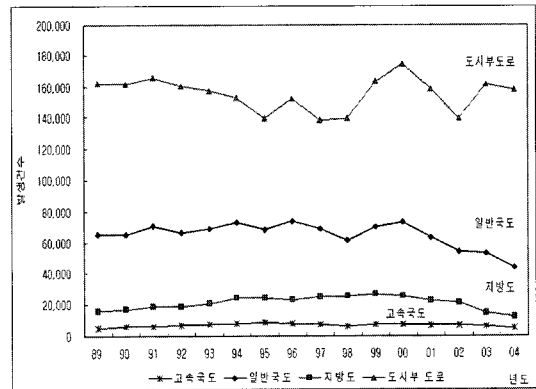


그림 1. 도로종류별 사고발생건수 추이

설재훈(2005)은 1996년 이후 사망자수 감소경향과 정부정책과의 관계를 표 1과 같이 정리하고 있다. 이중 교통사고 감소에 큰 영향을 미친 대표적인 정책으로 1997년 어린이통학버스 특별보호규정 도입, 무인단속카메라 도입, 위반차량 과태료 부과제도 도입을 꼽을 수 있다.



표 1. 정부의 연도별 주요 교통안전정책

연도	교통안전정책
1996	· 대형버스, 대형화물차 속도제한장치 부착 의무화
1997	· 어린이통학버스 특별보호규정 도입 · 무인단속카메라 도입, 위반차량 과태료 부과제도 도입, 자동차보유대수 1,000만대 돌파
1998	· IMF경제위기로 사망자수 대폭감소(-21.9%)
1999	· 일반도로 및 자동차전용도로 최고속도 각 10km/시 상향 · 운전면허 정지처분 벌점기준 완화(30점→40점)
2000	· 교통사고 사망자수 기준을 사고발생후 30일 이내로 변경 · 국무총리실 안전관리개선기획단 설치·운영
2001	· 교통위반 신고보상금제 도입 · 안전띠 착용 생활화 및 강력단속 전개 · 운전중 휴대전화사용 금지 · 음주운전 면허취소기간 연장(1년→2년)
2002	· 운행기록계 미작동 · 고장방치차량 교통법칙금 신설 · 신고보상금제 시행 중단
2003	· 자특회계법 개정안 시행(교통법칙금·과태료를 교통안전사업에 투자) · 어린이 교통안전 원년 선포(어린이 사망자수 5년내 1/2감소 효과)
2004	· 교차로 교통사고 자동기록장치 도입

자료) 우리나라의 도로교통사고 특성과 안전정책 현황, 한국도로학회지 제7권4호, 한국도로학회, 2005

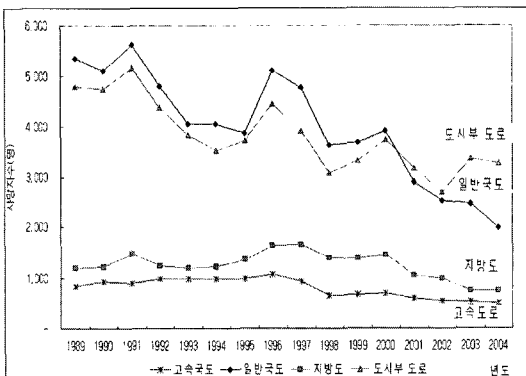


그림 2. 도로종류별 사망자수 추이

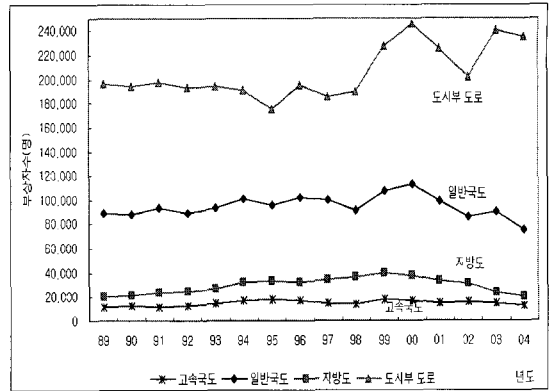


그림 3. 도로종류별 부상자수 추이

2.2 도로연장별 건수 비교

도로종류별(고속국도, 일반국도, 지방도, 도시부도로) km당 사고발생건수, 사망자수 및 부상자수 추이는 그림 4~그림 6과 같은 추세로 나타났다. 그림 4에 제시된 도로종류별 km당 사고발생건수 순위는 1993년에 일반국도, 도시부도로, 고속국도, 지방도 순에서 일반국도, 고속국도, 도시부도로, 지방도 순으로 변경되었으며, 10년만인 2003년을 기점으로 도시부도로의 사고발생건수가 증가 추세를 보이면서 다시 고속국도와 순위를 바꿔 일반국도, 도시부도로, 고속국도, 지방도의 순위를 보이고 있다.

그림 5에 제시된 도로종류별 km당 사망자수 순위는 고속국도, 일반국도, 도시부도로, 지방도로 나타났으며, 고속국도와 일반국도의 사망자수 차이는 점점 감소추이를 보이고 있고, 1993년부터 2002년 사이에는 지방도가 도시부도로 보다 높은 수치를 보이고 있다.

그림 6에서 제시된 도로종류별 km당 부상자수 순위는 1991년 근소한 차이로 고속국도, 일반국도, 도시부도로, 지방도로 나타나며, 1999년부터 급격한 증감으로 고속국도와 일반국도의 순위가 바뀌었다.

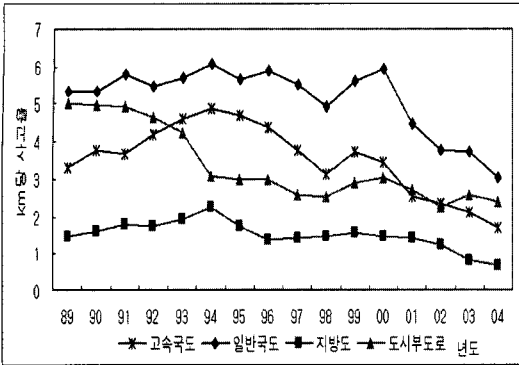


그림 4. 도로종류별 km당 교통사고 발생건수 추이

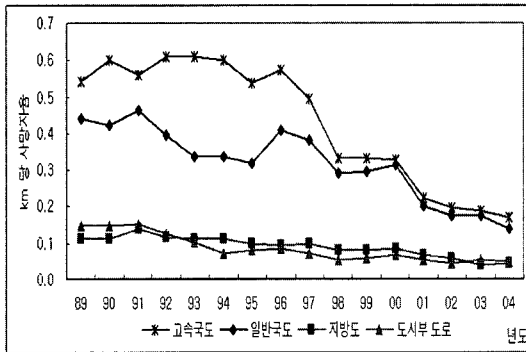


그림 5. 도로종류별 km당 사망자수 추이

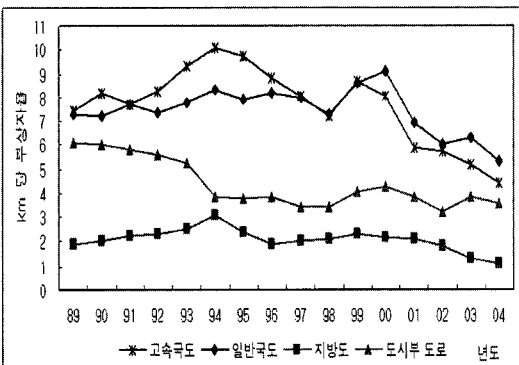


그림 6. 도로종류별 km당 부상자수 추이

2.3 대·km별 건수 비교

도로종류별(고속국도, 일반국도, 지방도) 10만 대·km당 사고발생건수, 사망자수 및 부상자수 추이는

그림 7~그림 9에 제시되어있다. 도로종류별 평균일 교통량(ADT)은 『도로교통량 통계연보, 1990~2004, 건설교통부』에서 고속국도, 일반국도, 지방도에 대해 추출하였으나, 도시부도로의 경우에는 제시되어 있지 않아 대·km별 건수 비교에서 제외되었다. 그림 7에 제시된 도로종류별 10만대·km당 사고발생건수 순위는 1989년에서 1992년까지 일반국도와 지방도에서 거의 동일한 감소패턴을 보이나, 1993~1994년까지 일반국도의 감소추세가 더 급격하게 줄어 지방도의 사고발생건수가 더 크게 나타났다. 그러나 1995년을 기점으로 일반국도, 지방도, 고속국도의 순위를 보이고 있다. 그림 8에 제시된 도로종류별 10만대·km당 사망자수 순위는 사고발생건수와 거의 동일한 추세로 따라가며 1995년을 기점으로 일반국도, 지방도, 고속국도의 순위를 보이고 있다.

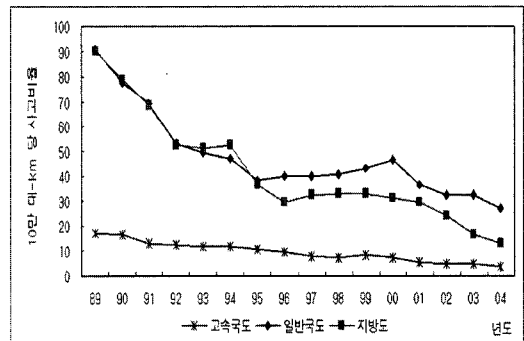


그림 7. 도로종류별 10만대·km당 사고건수 추이

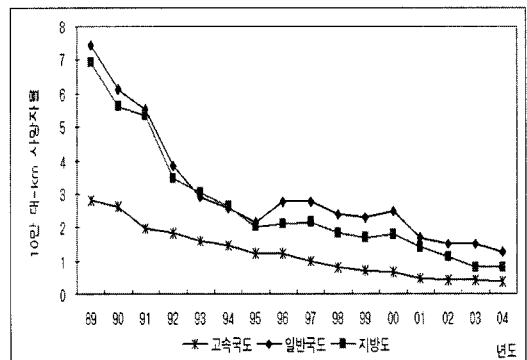


그림 8. 도로종류별 10만대·km당 사망자수 추이

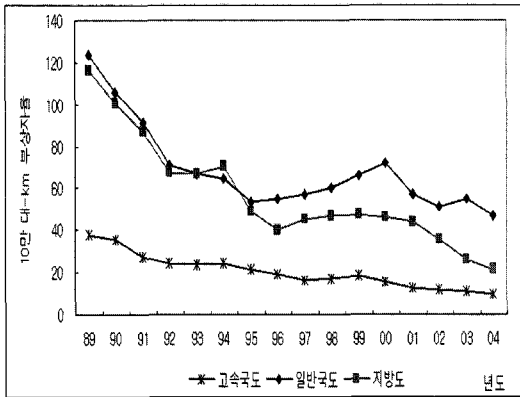


그림 9. 도로종류별 10만대·km당 부상자수 추이

그림 9 도로종류별 10만대·km당 부상자수는 1998년 이후부터 대체로 감소추세를 나타내고 있다. 도로종류별 10만대·km당 부상자수 순위는 일반국도, 지방도, 고속국도의 순으로 나타난다.

2.4 소결

앞서 제시한 도로종류별 사고건수, 사망자수, 부상자수의 순위를 단순건수, km당 지표, 10만대·km당 지표로 정리하면 표 2와 같다. 표 2를 살펴보면, 단순건수에 의한 순위에서는 고속국도가 가장 안전하고, 도시부도로가 가장 위험한 것으로 보이나, km당 지표 및 10만대·km당 지표에서는 대체로 일반국도가 가장 위험한 것으로 분석되고 있다. 일반적으로 도로의 안전도를 평가할 때 단순건수에 의한 비교보다는 교통량과 통행거리가 동시에 반영된 대·km당 지표를 선호하는 경향이 있다. 이런 측면에서 일반국도에 대한 교통안전 노력이 다른 도로보다 더 요구된다고 볼 수 있다. 일반국도의 10만대·km당 교통사고 발생건수, 사망자수, 부상자수가 1990년대 중반 이후 다른 도로보다 항상 높게 나타나고 있기 때문이다.

표 2. 교통사고 지표의 종합 순위표

구 분	단순건수			km당 건수			10만대·km당 건수		
	발생건수	사망자수	부상자수	발생건수	사망자수	부상자수	발생건수	사망자수	부상자수
고속국도	4	4	4	3(2) ^b	1	1(2) ^d	3	3	3
일반국도	2	1(2) ^a	2	1	2	2(1) ^d	1(2) ^e	1(2) ^e	1(2) ^f
지방도	3	3	3	4(4) ^b	4(3) ^e	4	2(1) ^e	2(1) ^e	2(1) ^f
도시부도로	1	2(1) ^a	1	2(3) ^b	3(4) ^e	3(2) ^d	N/A [*]	N/A [*]	N/A [*]

각주: ()^a 2001년 - 2004년 순위임 ()^b 1993년 - 2000년 순위임
 ()^c 1993년 - 2002년 순위임 ()^d 1998년, 2000년-2004년 순위임
 ()^e 1993년 - 1994년 순위임 ()^f 1994년 순위임

* ADT가 제시되지 않음

* 도시부도로는 단순합계를 적용하였음

출처: 1990년~2005년 교통사고 통계분석, 도로교통안전관리공단, 1990~2005

표 3. 교통사고 지표값(2004년)

구 분	단순건수			km당 건수			10만대·km당 건수		
	발생건수	사망자수	부상자수	발생건수	사망자수	부상자수	발생건수	사망자수	부상자수
고속국도	4,941	502	12,855	1.690	0.172	4.938	3.741	0.380	3.741
일반국도	43,613	1,990	75,399	3.061	0.140	5.293	27.324	1.247	27.324
지방도	12,382	745	19,980	0.709	0.043	1.143	13.169	0.792	13.169
도시부도로	157,657	3,261	234,765	2.402	0.050	3.577	N/A [*]	N/A [*]	N/A [*]

3. 교통사고 발생추세 모형

본 장에서는 앞서 제시한 교통사고지표 중에서 도로종류별 사망자수의 추세를 모형화하여 향후 5년 동안의 교통사고 사망자수를 예측하고자 한다. 이런 경우 일반적으로 사용할 수 있는 예측기법에는 회귀분석기법과 시계열분석기법이 있으나 본 연구에서는 시계열분석기법으로 향후 5년의 교통사고 사망자수를 도로 종류별로 예측하고자 한다.

3.1 회귀분석

교통사고 발생추세 모형 중 지금까지 가장 많이 사용된 모형이 회귀분석모형이다. 회귀분석에서는 시간에 따른 사고추세를 직선 혹은 곡선(log 혹은 지수 곡선)식으로 추정한다. 시간 t 에 대한 단순선형회귀모형(simple linear regression model)의 경우 다음 식과 같이 나타낼 수 있다.

$$y_i = \beta_0 + \beta_1 t + \epsilon_i$$

여기서, $y_i = i$ 연도의 사고건수(비율)

$\beta_0 =$ 직선의 절편

$\beta_1 =$ 직선의 기울기

$\epsilon_i = i$ 연도 값의 오차

회귀분석모형을 이용한 거시적 사고예측모형에는 임헌연·이일병, 박병호, 이수범·김정현·홍다희·유창남 등 다수가 있다. 우선 임헌연·이일병 (1990)은 교통사고건수, 부상자수, 사망자수를 종속변수로 하고, 인구수, 자동차보유대수, 운전면허 소지자수, 도로연장거리, 국민총생산과의 상관성을 검증한 후 자동차보유대수를 독립변수로 하는 회귀분석모형을 개발하였다. 박병호(1995)는 충청권(대전, 충북, 충남)의 사고예측모형을 개발하기 위해 자동차화의 수준과 개인의 사고 위험성 간을 멱함수식과 지수함수식으로 나타냈다. 이수범·김정현·홍다희·유창남 (2003)은 도로를 규모와 등급에 따라 고속도로, 2차

로 도로, 4차로 도로(중앙분리대 유/무)로 구분하여 사고건수와 교통량, 교차로수, 연결로수, 횡단신호등수, 설계속도와의 관계를 다중회귀분석모형으로 밝혀냈다.

위 선행연구들은 국가 전체나 지역, 특정된 도로구간의 교통사고특성모형을 구축하기 위해 회귀모형을 사용하였으며, 설명변수로는 경제, 교통 및 도로특성을 반영할 수 있는 요소를 사용하였다. 회귀분석모형은 사고발생의 전반적 추세를 설명하는 데는 유용하지만 단기간의 미래예측에 있어서는 시계열분석보다 설명력이 다소 떨어질 수 있다.

3.2 시계열분석

국민총생산, 물가지수, 주가지수 등과 같이 시간에 따라 지표가 산출되는 자료를 시계열자료라 부르며 이러한 자료에 대한 분석기법을 시계열분석기법이라 부른다. 이 기법은 Box와 Jenkins (1976)에 의해 ARIMA 모형이 개발된 이후 비약적인 발전을 이루고 있다.

시계열모형 중 대표적인 모형으로는 AR(자기회귀, Auto-Regressive)모형과 MA(이동평균, Moving Average)모형 그리고 이들을 결합시킨 ARMA(자기회귀 이동평균)모형이 있다. 또한 비정상 시계열자료를 정상시계열자료로 전환시키는 차분 과정이 포함된 ARIMA(통합자기회귀 이동평균, AutoRegressive Intereated Moving Average)모형이 있다.

평균이 μ 인 AR(p), MA(q), ARMA(p, q), ARIMA(p, d, q)모형은 다음과 같이 표현된다.

- 평균이 μ 인 AR(p): $Z_t - \mu = \phi_1(Z_{t-1} - \mu) + \phi_2(Z_{t-2} - \mu) + \dots + \phi_p(Z_{t-p} - \mu) + a_t$
- 평균이 μ 인 MA(q): $Z_t - \mu = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$
- 평균이 μ 인 ARMA(p, q): $Z_t - \mu = \phi_1(Z_{t-1} - \mu) + \dots + \phi_p(Z_{t-p} - \mu) + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$



- 평균이 μ 인 ARIMA(p, d, q): $\phi(B)(W_t - \mu) = \theta(B)a_t, W_t = (1 - B)^d Z_t$

Z_t : 시점 t 에서의 관측값

μ : 평균

θ : 백색잡음($a_{t-1}, a_{t-2}, a_{t-2}, \dots$)이 Z_t 에 미치는 영향을 설명하는 파라미터

ϕ : 계수, Z_t 에 대한 $Z_{t-1}, Z_{t-2}, Z_{t-2}, \dots$ 의 기여도

a_t : 시점 t 에서의 백색잡음

B : 후향연산자, d : 차분의 차수, p : AR모형의 차수,

q : MA모형의 차수

일반적으로 시계열분석모형을 수립하기 위해 3단계인 모형식별→모수추정→모형진단과정을 거친다.

1단계인 모형식별단계(Model Identification)에서는 자료의 산점도와 표본자기상관함수(SACF)를 통해 시계열의 비정상성 여부를 찾아내고, 만약 비정상성 자료라면 분산안정화 변환과 차분을 통해 정상성 자료로 변환을 한다.

일반적으로 시계열모형은 AR(p), MA(q), ARMA(p, q), ARIMA(p, d, q)모형 등 여러 가지 대안으로 요약될 수 있는데 이중 가장 적합한 모형을 선택하기 위해서는 대개 AIC(Akaike Information Criteria: Akaike 정보량 기준)값이 최소가 되는 모형을 선택한다. 일반적인 모형식별과정은 그림 10과 같이 정리된다.

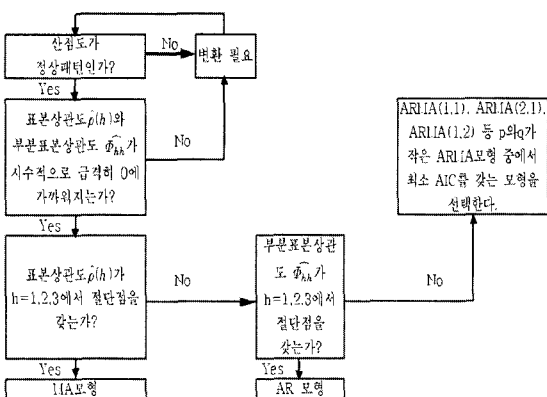


그림 10. 모형식별과정

2단계는 모수추정단계(Parameter Estimation)로 주로 최우추정법(Maximum Likelihood:ML)과 조건부최소제곱추정법(Conditional Least Squares:CLS) 등을 통해 추정하며 모수추정치가 유의한지를 판단하기 위해 t -검정을 실시한다. t -검정은 귀무가설($H_0: \theta$ or $\phi=0$)로 추정된 모수가 0일 가능성이 높은지의 여부를 판단하는 것이다.

3단계 모형진단단계(Diagnostics)에서는 과다적합진단(Overfitting Diagnostics)과 잔차분석(Residual analysis)을 통해 시계열 모형의 차수 (p 또는 q)를 줄일 수 있는 가능성이 있는지 혹은 잔차에 뚜렷한 경향성이 발견되는지의 여부를 판단한 후 최종 예측모형을 결정한다. 시계열분석기법에 대한 보다 자세한 내용은 조신섭·황선영(1997)을 참고하면 된다.

4. 도로종류별 사망자수 추세모형

4.1 모형의 식별

고속국도, 일반국도, 지방도, 도시부도로의 사망자수에 대한 시계열분석모형을 식별하기 위해 표본자기상관함수를 살펴보면 대체로 시차 2(lag 2)에서 급격히 0으로 떨어지는 경향이 나타난다. 이는 AR(1)모형으로 시계열자료를 적합시키는 것이 합리적임을 나타낸다. 보다 다양한 시계열모형을 검토하기 위해 AR(1), MA(1), ARMA(1, 1), ARIMA(1, 1, 0), ARIMA(0, 1, 1), ARIMA(1, 1, 1)모형으로 시계열자료를 적합시키고 이들의 AIC 값 및 추정된 모수들의 t 값을 분석하였다.

표 4~표 7은 도로종류별 교통사고 사망자수 자료를 기반으로 형성된 시계열모형들의 AIC값과 t 값을 정리한 것이다. 고속국도 사망자수에 대한 모형은 AIC값을 기준으로 할 때 ARIMA 모형이 더 좋게 나타났으나 t 값이 좋지 않아 AR(1)모형이 더 적절한 것으로 판단되었다. 일반국도의 경우에는



ARIMA(2, 3, 0)모형이 t 값도 통계적으로 유의하고 AIC값도 가장 낮아 적절한 모형으로 선택하였다. 지방도와 도시부도로의 경우도 AR(1)모형이 t 값 및 AIC값을 고려할 때 가장 적절한 것으로 판단되었다.

표 4. 고속국도의 모형별 AIC값 & t 값

구분	AR (1)	MA (1)	ARMA (1,1)	ARIMA (1,1,0)	ARIMA (0,1,1)	ARIMA (1,1,1)
$\hat{\mu}$	8.44	14.15	9.12	-0.80	-0.73	-0.70
$\hat{\phi}$ or $\hat{\theta}$	6.73	-5.78	3.96, -1.76	0.24	-0.53	-0.35, -0.56
AIC	194.2612	202.6957	194.8238	182.4608	182.3836	183.9999

표 5. 일반국도의 모형별 AIC값 & t 값

구분	AR (1)	MA (1)	ARMA (1,1,0)	ARIMA (0,1,1)	ARIMA (2,3,0)
$\hat{\mu}$	7.64	11.53	-1.53	-6.61	-3.23
$\hat{\phi}_1$	7.80	-4.69	-0.46	3.92	3.04
$\hat{\phi}_2$	-	-	-	-	-2.52
AIC	254.9583	262.7607	237.7348	233.9084	206.149

표 6. 지방도의 모형별 AIC값 & t 값

구분	AR (1)	MA (1)	ARMA (1,1)	ARIMA (1,1,0)	ARIMA (0,1,1)	ARIMA (1,1,1)
$\hat{\mu}$	7.28	14.88	7.72	-0.60	-0.60	-0.59
$\hat{\phi}$ or $\hat{\theta}$	4.49	-5.84	247, -0.56	0.11	-0.12	-1.11, -1.35
AIC	214.062	216.9708	215.59	202.1941	202.1926	203.7484

표 7. 도시부도로의 모형별 AIC값 & t 값

구분	AR (1)	MA (1)	ARIMA (1,1,0)	ARIMA (0,1,1)	ARIMA (2,3,0)
$\hat{\mu}$	8.94	17.30	-0.78	-9.14	-0.79
$\hat{\phi}$ or $\hat{\theta}$	7.17	-3.03	-0.21	3.36	-1.07, -2.47
AIC	247.8328	250.1051	233.0589	226.1036	233.4054

4.2 모수의 추정

조건부최소제곱법으로 추정한 모수 추정치 및 유의확률이 표 8에 정리되어 있다.

표 8. 도로종류별 사망자수 시계열모형의 모수추정치 및 통계량

구분		estimate	standard error	T value	Pr > T
고속국도	$\hat{\mu}$	827.6637	98.02535	8.44	< 0.0001
	$\hat{\phi}_1$	0.96701	0.14378	6.73	< 0.0001
일반국도	$\hat{\mu}$	-626.04882	658.7036	7.64	< 0.0001
	$\hat{\phi}_1$	0.75422	0.24834	3.04	0.0125
	$\hat{\phi}_2$	-0.65315	0.25913	-2.54	0.0304
지방도	$\hat{\mu}$	1,173.6	161.56069	7.28	< 0.0001
	$\hat{\phi}_1$	0.844	0.18802	4.49	0.0005
도시부도로	$\hat{\mu}$	4,658.1	521.21855	8.94	< 0.0001
	$\hat{\phi}_1$	0.93998	0.1311	7.17	< 0.0001

4.3 모형의 진단

표 8에 기반한 시계열모형은 표 9와 같이 정리된다.

표 9. 도로종류별 사망자수 시계열모형

구분	모형종류	모형식
고속국도	AR(1)	$Z_t = 27.30868 + 0.96701Z_{t-1} + a_t$
일반국도	ARIMA (2, 3, 0)	$(1 - 0.75422B + 0.65315B^2)W_t - \mu = a_t, W_t = (1 - B)^3 Z_t$
지방도	AR(1)	$Z_t = 183.0874 + 0.8440Z_{t-1} + a_t$
도시부도로	AR(1)	$Z_t = 279.6018 + 0.93998Z_{t-1} + a_t$

표 9에 제시된 도로종류별 사망자수 시계열모형들에 대한 잔차분석 결과 잔차들이 특별한 경향을 띄지 않았고, 이들 모형이 과다적합된 것으로도 보이지 않아 최종모형으로 선택하였다.



4.4 예측

표 9에 제시된 도로종류별 사망자수 시계열모형들을 이용하여 향후 5년간 예측한 사망자수 추정치가 표 10에 제시되어있다. 그 결과 AR(1)모형으로 예측된 고속국도, 지방도, 도시부도로의 경우 사망자수가 앞으로 증가할 것으로 나타났으며, ARIMA(2, 3, 0)모형으로 예측된 일반국도는 2006년까지 증가하다 2007년 이후 감소하는 추세로 나타났다.

표 10. 도로종류별의 향후 5년간(2005년~2009년) 사망자수 예측치

구분	연도	Forecast	Std Error	95% Confidence Limits	
고속국도	2005	513	98.83	319.0247	706.4658
	2006	523	137.4926	253.6555	792.6164
	2007	533	165.6696	208.4773	857.8903
	2008	543	188.2398	173.9569	911.8434
	2009	552	207.1319	146.3248	958.2669
일반국도	2005	2,208	608.0645	1016.5582	3400.1274
	2006	2,261	761.6237	768.5368	3754.0468
	2007	1,480	763.3467	-16.1720	2976.0918
	2008	1,393	809.6434	-194.2392	2979.5047
	2009	1,416	843.6220	-237.0433	3069.8941
지방부	2005	812	183.5905	452.0331	1171.6946
	2006	868	240.2391	397.4367	1339.1566
	2007	916	273.5276	379.8214	1452.0299
	2008	956	294.9576	378.0181	1534.2305
	2009	990	309.3186	383.7985	1596.3051
도시부도로	2005	3,345	527.2123	2311.5442	4378.1784
	2006	3,424	723.5600	005.5374	4841.8403
	2007	3,498	860.5391	1811.1591	5184.4105
	2008	3,567	965.5302	1675.0287	5459.8376
	2009	3,633	1049.5925	1575.7374	5690.0644

그림 11~그림 14는 도로종류별 실제 사망자수와 모형에 의한 사망자수 추정치 및 향후 5년간의 예측치를 보여주고 있다.

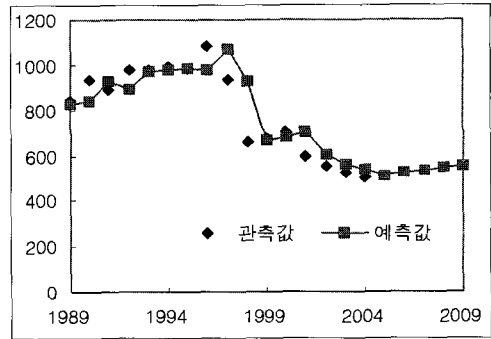


그림 11. 고속국도의 시계열모형에 의한 예측치

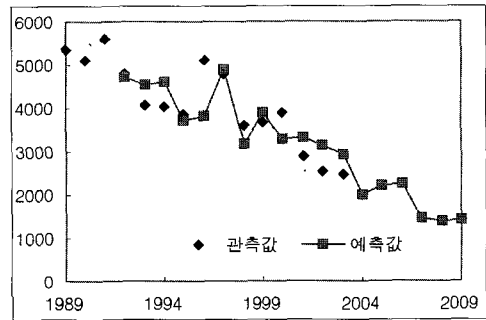


그림 12. 일반국도의 시계열모형에 의한 예측치

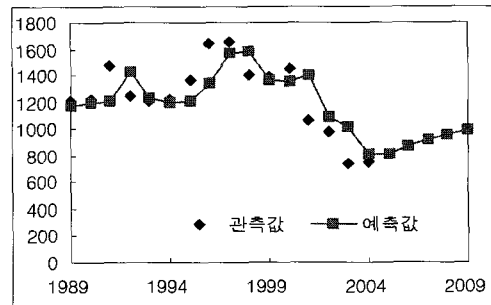


그림 13. 지방도의 시계열모형에 의한 예측치

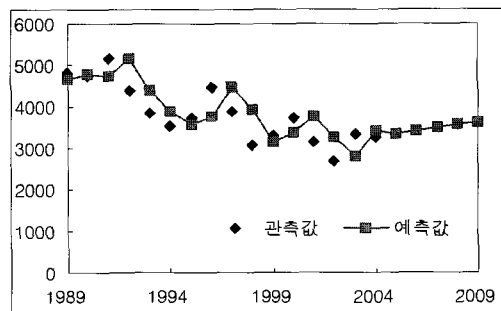


그림 14. 도시부도로의 시계열모형에 의한 예측치

4.5 시계열예측모형의 한계

도로 종류별 사고예측모형 중에서 AR(1)모형으로 추정된 고속국도, 지방도, 도시부도로의 경우 사망자수는 향후 5년간 계속 증가하는 것으로 예측된다. 하지만 도로별 사망자수가 5년간 지속적으로 증가할 것이라고 예측하는 것은 합리적이지 않다. 도로교통안전관리의 책임을 지는 도로관리청 혹은 경찰청이 교통사고 감소를 위해 적극적으로 노력하는 한 교통사고 사망자수는 증가하다가도 다시 낮아지는 것이 일반적이기 때문이다. 따라서 도로의 사망자수 예측모형으로써 AR(1) 모형은 통계적으로는 타당하나 실질적으로는 타당성이 결여되었다고 판단할 수 있다.

AR(1) 모형이 5년 동안 지속적으로 증가하는 예측값을 제시한 이유는 AR 모형이 지니는 근원적인 특성 때문이다. 가령, 평균이 μ 인 AR(1)모형 $Z_t - \mu = \phi(Z_{t-1} - \mu) + a_t$, $|\phi| < 1$ 가 주어질 때, 미래시차 l 의 예측값 $\hat{Z}_t(l) = \mu + \phi^l(Z_t - \mu)$ 이 된다. 즉, 미래시차 l 이 커질수록 ϕ^l 값은 점점 작아져 $\phi^l(Z_t - \mu)$ 은 0에 근사하게 되고, $\hat{Z}_t(l)$ 은 평균(μ)에 근사하게 된다. 이러한 경향은 일반적인 AR(P)의 경우에도 나타난다. 본 연구에서 고속국도, 지방도, 도시부도로의 사망자수가 계속 증가한 것도 AR(1) 모형의 예측치가 점점 평균으로 가까워지려는 경향 때문에 나타난 것이다.

이에 반해 비정상 시계열모형이 ARIMA(2, 3, 0) 모형으로 추정된 일반국도의 경우는 향후 5년간 사망자수가 증가와 감소를 반복할 것으로 예측되었다.

만약 MA모형으로 최종모형이 결정된다 하여도 AR모형처럼 예측치의 활용도는 높지 않다. MA(q) 모형은 미래시차 l 이 q 보다 크면 예측치는 항상 평균 μ 이기 때문이다. 가령 평균이 μ 인 MA(1)모형이 $Z_t - \mu = a_t - \theta a_{t-1}$, $|\theta| < 1$ 으로 주어질 때, 미래시차 l 이 1인 경우 예측치는 $\hat{Z}_t(1) = \mu - \theta a_t$ 이다. 이는 Z_t, Z_{t-1}, \dots, Z_1 에 대한 조건부 기대값으로 구할 수 있다. l 이 2인 경우에는 $Z_{t+2} = \mu + a_{t+2} - \theta a_{t+1}$, $|\theta| < 1$

로 조건부 기대값을 취하면 $E(a_{t+1}|Z_t, Z_{t-1}, \dots, Z_1)$ 과 $E(a_{t+2}|Z_t, Z_{t-1}, \dots, Z_1)$ 값은 0이 된다. a_{t+1} 혹은 a_{t+2} 와 Z_t, Z_{t-1}, \dots, Z_1 이 서로 독립이기 때문이다. 따라서 $\hat{Z}_t(2) = \mu$ 이 된다. 일반적인 MA(q)의 경우로 확장하면 미래시차 l 이 $q+1$ 이상일 때 예측값은 $\hat{Z}_t(l) = \mu$ 이 된다.

이와 같은 시계열예측모형의 특성을 정리하면, 사고자료가 정상시계열모형인 AR 혹은 MA 모형으로 적합된다면 모형의 타당성은 인정된다 하더라도 실질적으로는 예측의 신빙성이 낮아진다고 볼 수 있다. 그러나 사고자료가 비정상 시계열모형인 ARIMA모형으로 적합된다면 모형의 타당성뿐만 아니라 실질적인 예측의 신빙성도 높아진다. 따라서, 사고자료가 정상시계열의 특성을 띠는지 혹은 비정상시계열의 특성을 띠는지가 시계열분석모형의 형태 및 그 유용성에 큰 영향을 미치는 것으로 결론내릴 수 있다.

교통사고자료는 대체로 추세를 지니기 때문에 정상시계열 자료로 분류하기는 힘들다. 교통사고는 어느나라 건 최고치까지 증가하다가 다시 감소하는 추세를 지니기 때문이다. 하지만 본 연구에서처럼 연단위 자료를 활용하면서 동시에 자료의 수가 적은 경우에는 사고자료를 정상시계열 자료로 분류할 가능성도 있다. 따라서, 향후 연구에서는 더 많은 자료를 활용하여 시계열모형을 다시 추정할 필요가 있을 것으로 보인다. 특히, 월단위 자료를 통해 모형을 구축한다면 더 유용한 모형을 얻을 수 있을 것이다. 자료의 수도 많아질 뿐만 아니라 상당한 시차(l)을 두고 사망자수가 늘어나더라도 큰 문제가 되지 않기 때문이다.

5. 결 론

5.1 결 론

본 연구에서는 도로종류별 교통사고 추세를 비교하기 위해 교통사고 발생건수, 사망자수, 부상자수를 단순비교와 동시에 이들을 도로종류별 연장, 그



리고 10만 대·km로 나누어 비교하여 보았다. 그 결과 단순건수에 의한 비교에서는 도시부도로가 발생건수, 사망자수, 부상자수 등 전부분에서 가장 위험한 것으로 나타났고, 반면 고속국도는 전부분에서 가장 안전한 것으로 분석되었다. 1km 당 발생건수는 일반국도가 가장 위험하고 지방도가 가장 안전한 것으로 분석되었다. 교통사고발생률을 가장 객관적으로 비교할 수 있는 지표인 10만대·km 당 교통사고 발생건수, 사망자수, 부상자수에서는 일반국도가 가장 위험한 것으로 나타났으며 고속국도는 상대적으로 안전한 것으로 분석되었다. 단, 도시부도로의 경우는 평균교통량 자료를 구할 수 없어 분석에서 제외되었다.

도로종류별 사망자수 추세를 시계열분석모형으로 추정된 결과 고속국도, 지방도, 도시부도로의 경우 AR(1)모형으로 구축되었고, 향후 5년간 계속 증가하는 추세로 나타났다. 일반국도의 경우는 ARIMA(2, 3, 0)모형으로 구축되어 예측치가 향후 5년 동안 증감을 반복할 것으로 예측되었다. 고속국도, 지방도, 도시부도로의 사망자수가 향후 5년간 증가할 것으로 예측된 이유는 AR모형이 지니는 본질적 특성, 즉 예측치가 평균으로 회귀하는 특성 때문이다. 따라서 사고자료가 정상시계열모형인 AR모형으로 적합되었을 경우에는 예측치가 실질적으로 의미있는 숫자가 될 가능성이 적다고 볼 수 있다. 연단위 사고자료는 일반적으로 추세를 갖는 점을 고려하여 가급적 ARIMA 등 비정상시계열 모형으로 적합시키는 편이 바람직하다.

5.2 연구의 한계 및 향후 연구

본 연구에서는 16년 동안의 교통사고통계를 이용하여 시계열분석모형을 구축하였으나 일반적으로 30개 이상의 자료를 사용해야 보다 신뢰성 있는 시계열분석모형을 구축할 수 있는 것으로 알려져 있다. (조신섭, 황선영, 1997). 따라서 자료수를 대폭 늘리기 위해 월별 교통사고자료로 계절적 변동 및 추세

변동을 파악하여 예측하는 것이 필요할 것으로 보인다. 월별 교통사고자료에 대한 시계열분석모형 구축은 또한 사고자료를 정상시계열모형으로 적합시키는 오류를 줄일 수 있을 것으로 보인다. 설령, 정상시계열모형으로 적합된다 하더라도 모형에 의한 예측치가 실질적으로 의미있는 수치가 될 가능성이 높아진다. 단기적으로 예측치가 평균으로 회귀한다 하더라도 그 기간이 그리 길지 않기 때문이다. 5년 동안 사망자수가 계속 증가할 것이라고 예측하는 것은 합리적이지 않지만, 5개월 동안 사망자수가 계속 증가할 것이라는 예측은 충분히 가능하기 때문이다.

한편, 본 연구에서는 교통사고 사망자수에 대한 시계열분석모형만을 개발하였으나, 이러한 모형은 교통사고 원인별 사고통계, 차종별 사고통계, 사고관련 운전자 연령대별 사고통계 등 다양한 자료를 기반으로 구축이 가능하다. 따라서 보다 다양한 사고통계 자료를 기반으로 시계열분석모형을 구축하여 교통사고의 전반적인 추세를 이해하고 장래를 예측한다면, 보다 실효성 높은 교통정책을 수립하거나 평가하는데 큰 도움이 될 것으로 판단된다.

감사의 글

본 논문은 건설교통부가 출연하고 한국건설교통기술평가원에서 위탁시행한 2005년도 건설기술기반구축사업(05기반구축 D02: 도로교통안전진단 및 관리를 위한 통합정보시스템 구축)의 지원으로 이루어졌습니다.

참고문헌

- 권혁제(2000), SAS 활용 표준통계학, 청목출판사
- 도로교통안전관리공단(1990~2005), 1990년~2005년 교통사고 통계분석
- 박병호(1995), 충청권의 교통사고예측모형에 관한 연구, 대한교통학회지, 제15권
- 설재훈(2005), 우리나라의 도로교통사고 특성과 안전정책 현황, 한국도로학회지 제7권4호, 한국도로학회



설재훈(2004), 최근 교통사고의 감소 특징과 향후 교통 안전 대책 방향, 월간교통, 한국교통연구원

송문섭(1993), SAS를 이용한 통계자료분석, 자유아카데미

송근원(2005), 아리마 시계열분석, 금정

이성덕(2003), SAS를 이용한 시계열자료분석, 탐진

이수범(2003), 김정현, 김태희, 도로 및 교통특성에 따른 계획단계의 도시부 도로 교통사고 예측 모형개발, 대한교통학회지, 제21권 제4호

임현연, 이일병(1990), 한국의 교통사고예측모델 개발에 관한 연구, 대한교통학회지, 제12권

조신섭, 황선영(2003), 시계열 분석, 한국방송통신대학교 출판부

하태준, 강정규, 박제진(2001), 신호교차로 교통사고 예측 모델의 개발 및 적용, 대한교통학회지, 제19권 제6호

Box, G.E.P. and G.M. Jenkins (1976) Time Series Analysis: Forecasting and Control, 2nd ed. San Francisco: Holden-Day

접 수 일: 2006. 7. 6
심 사 일: 2006. 7. 22
심사완료일: 2007. 7. 19