

주행거리 및 연비를 이용한 연료소비량 산정방법

- 한국교통연구원 (2012년 5월) -

한국데이터정보과학회지 3월

『도로교통부문에서 주행거리를 이용한 연료소비량 산정방법: - 몬테카를로 기법 중심으로』

경기대학교 수학과 박천건

에너지경제연구원 연구위원 소진영

동국대학교 통계학과 이영섭

목차

I . 연구목적 및 연구계획

II . 활동자료

III . 연료소비량 추정 방법

IV . 모의실험

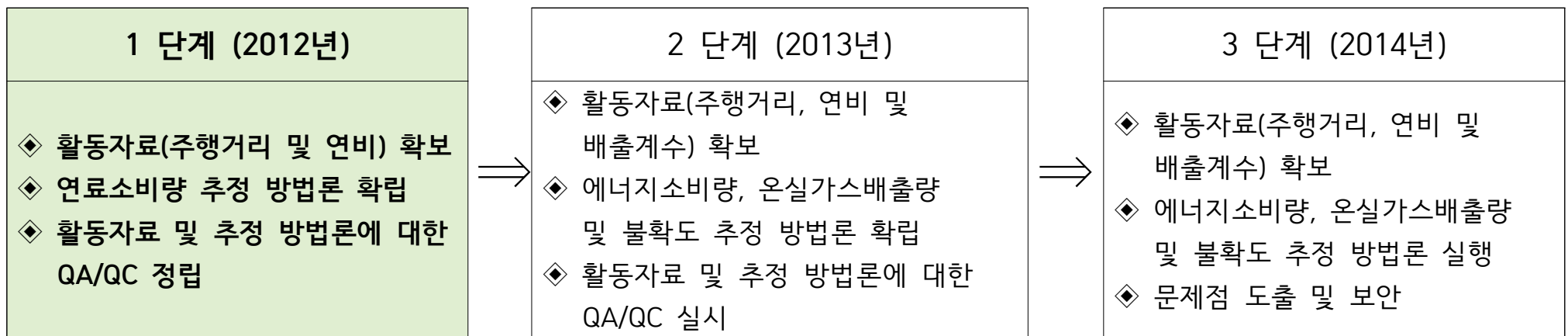
V . 토의

I . 연구목적 및 연구계획 - 도로교통부문 중심

■ 도로교통부문에서 활동자료의 종합적인 활용의 방법론 확립

- ◆ 에너지소비량
- ◆ 주행거리
- ◆ 연비
- ◆ 배출계수 등

■ 도로교통부문에서 연구추진 계획



Ⅱ. 활동자료 - 도로교통부문

■ 에너지소비량

- ◆ 에너지통계연보 - 에너지경제연구원
- ◆ 에너지총조사보고서(5년 주기) - 에너지경제연구원

■ 주행거리

- ◆ 연간 주행거리 실태조사 - 교통안전공단
- ◆ 기타

■ 연비

- ◆ 자동차연비조사 - 지식경제부
- ◆ 각종 단체

■ 교통량

- ◆ 도로교통량 통계연보
- ◆ 지자체 교통량 통계 - 서울특별시 교통량 조사, 광역시 교통량 조사

Ⅲ. 연료소비량 추정 방법론 - 활동자료

■ 연료소비량에 관련된 작성기관 및 보고서

- ◆ 에너지통계연보 - 에너지경제연구원
- ◆ 에너지총조사보고서(5년 주기) - 에너지경제연구원

■ 활용 문제점 (IPCC에 부합하는 결과 도출)

- ◆ 에너지통계연보 - 연간 자료이지만 차종에 대한 구분이 불가능
- ◆ 에너지총조사보고서(5년 주기) - 5년 주기의 자료이고 원지자료의 신뢰성 결여

■ 보조자료로 활용방안 모색

Ⅲ. 연료소비량 추정 방법론 - 기호

※ 총연료소비량 추정(T)

■ 자료의 기호

	내용	기호
자료	모집단 (모든 차량)의 수	N
	표본 (일부 차량)의 수	n
	i 번째 차량 주행거리 (km)	d_i
	i 번째 차량 연료소비량 (ℓ)	$c_i = d_i / e_i$
	i 번째 차량 연비 (km/ ℓ) = 연료 ℓ 당 주행거리 (km)	$e_i = d_i / c_i$

Ⅲ. 연료소비량 추정 방법론 - 방법 1

<방법 1> 확률표본 - 각 차량의 연료소비량(c_i)

$$\hat{T}_1 = N \cdot \bar{c}$$

- 모든 차량수 N
- 표본 차량의 연료소비량 $c_j = d_j/e_j, j = 1, \dots, n (< N)$
- 연료소비량의 표본평균 $\bar{c} = \sum_{j=1}^n c_j/n$

Ⅲ. 연료소비량 추정 방법론 - 방법 2

<방법 2> 확률표본 - 모든 차량의 주행거리(d_i) 및 표본차량의 연비(e_i)

$$\hat{T}_2 = \frac{1}{\bar{e}} \sum_{i=1}^N d_i$$

- 모든 차량수 N
- 모든 차량의 주행거리 $d_i, i = 1, \dots, N$
- 표본평균연비 $\bar{e} = \sum_{j=1}^n d_j / \sum_{j=1}^n e_j$

Ⅲ. 연료소비량 추정 방법론 - 방법 3

<방법 3> 표본 차량의 주행거리(d_i) 및 연비(e_i)

(단계 1) 표본 차량 n 개의 주행거리와 연비로부터 연료소비량 (c_i) 계산:

$$c_i = d_i / e_i, \quad i = 1, \dots, n$$

(단계 2) 연료소비량의 확률밀도함수를 추정: $f(c)$

반복 (단계 3)~(단계 4): $k = 1, \dots, m$ {

(단계 3) 연료소비량의 확률밀도함수로부터 모든 차량수 (N)대한 난수생성:

$$c_{(j,k)} \sim f(c), \quad j = 1, \dots, N$$

(단계 4) 생성된 연료소비량의 합을 계산: $\tilde{T}_k = \sum_{j=1}^N c_{(j,k)}$ }

(단계 5) 생성된 총연료소비량의 평균을 계산: $\hat{T}_3 = \frac{\sum_{k=1}^m \tilde{T}_k}{m}$

Ⅲ. 연료소비량 추정 방법론 - 방법 4

<방법 4> 모든 차량의 주행거리(d_i) 및 표본 연비(e_i)

(단계 1) 표본 n 개 차량의 연비로부터 확률밀도함수를 추정: $f(e)$

반복 (단계 2)~(단계 4): $k = 1, \dots, m$ {

(단계 2) 연비의 확률밀도함수로부터 모든 차량수 (N)에 대한 난수생성:

$$e_{(i,k)} \sim f(e), \quad i = 1, \dots, N$$

(단계 3) 생성된 연비를 무작위로 각 차량의 주행거리에 할당:

$$(d_i, e_{(i,k)}), \quad i = 1, \dots, N$$

(단계 4) 모든 차량수 (N)의 연료소비량을 계산: $c_i = d_i / e_{(i,k)}, \quad i = 1, \dots, N$

(단계 5) 총연료소비량을 계산: $\tilde{T}_k = \sum_{i=1}^N c_{(i,k)} \}$

(단계 6) 생성된 총연료소비량의 평균을 계산: $\hat{T}_4 = \frac{\sum_{k=1}^m \tilde{T}_k}{m}$

Ⅳ. 모의실험 - 가정

<가정>

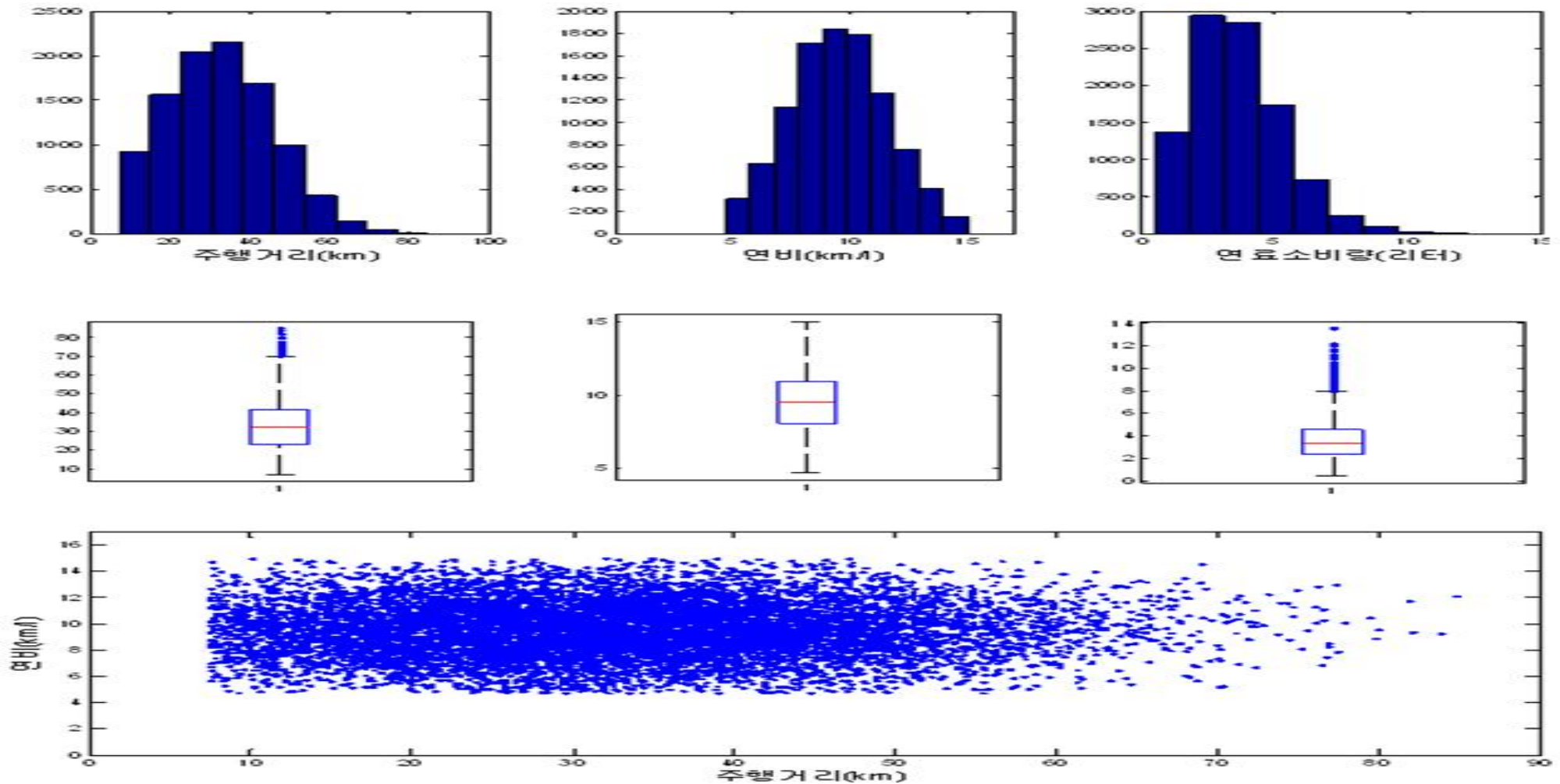
주행거리와 연비에 대한 정규분포의 모수 및 조건

	평균	표준편차	최솟값	최댓값
주행거리 (d_i , km)	31.5	15.5	7.6	327
연비 (e_i , km/ℓ)	9.5	2.2	4.7	15

* 모집단(모든 차량수), $N = 10,000$, $i = 1, \dots, N$

* 표본 차량수, $n = 30(0.3\%), 100(1\%), 500(5\%), 1000(10\%)$

Ⅳ. 모의실험 - 가정의 시각적 표현



모집단에 대한 주행거리, 연비 그리고 연료소비량의 형태

Ⅳ. 모의실험 - 가정의 요약통계값

총연료소비량($T=36,146 \ell$)

100번 반복에 대한 표본 연비

표본수 (표본비율)	평균					표준편차				
	최솟값	평균	중앙값	최댓값	표준 편차	최솟값	평균	중앙값	최댓값	표준 편차
30 (0.3%)	8.58	9.55	9.59	10.49	0.39	1.45	2.01	1.99	2.60	0.23
100 (1%)	9.06	9.55	9.56	9.94	0.19	1.79	2.06	2.05	2.34	0.12
500 (5%)	9.34	9.52	9.52	9.81	0.09	1.88	2.05	2.05	2.18	0.05
1,000 (10%)	9.38	9.53	9.52	9.69	0.06	1.95	2.05	2.05	2.18	0.04

Ⅳ. 모의실험 - 결과의 요약통계값

총연료소비량($T=36,146 \ell$)

100번 반복에서 총연료소비량의 추정에 대한 요약통계값

		표본수 (전체 차량의 수 10,000)			
총연료소비량		30	100	500	1,000
방법 1	최 소 값	29,025	31,597	34,117	35,088
	평 균	36,047	35,998	36,223	36,189
	중 앙 값	35,878	36,073	36,273	36,133
	최 대 값	45,084	40,016	37,864	38,082
	표준편차	3,475	1,670	733	589
방법 2	최 소 값	31,384	34,484	35,322	35,567
	평 균	36,064	36,049	36,193	36,185
	중 앙 값	35,995	36,056	36,183	36,167
	최 대 값	40,925	38,141	37,159	36,982
	표준편차	1,714	854	388	272

		표본수 (전체 차량의 수 10,000)			
총연료소비량		30	100	500	1,000
방법 3	최 소 값	30,146	32,953	35,413	36,284
	평 균	37,787	37,490	37,545	37,491
	중 앙 값	37,430	37,533	37,550	37,416
	최 대 값	47,903	41,270	39,572	39,460
	표준편차	3,727	1,756	794	613
방법 4	최 소 값	32,404	34,377	34,963	35,374
	평 균	35,869	35,908	35,989	35,959
	중 앙 값	35,783	35,935	35,979	35,970
	최 대 값	39,567	37,704	36,692	36,486
	표준편차	1,439	690	326	227

IV. 모의실험 - 결과의 요약통계값

총연료소비량($T=36,146 \text{ l}$)

$$MSE = \frac{1}{100} \sum_{j=1}^{100} (\hat{T}_{i_j} - T)^2, \text{ } i \text{ 는 제시된 방법, } j \text{ 는 반복수}$$

범위와 변동계수

		표본수 (전체 차량의 수 10,000)			
		30	100	500	1,000
범위	방법 1	16,059	8,419	3,747	2,994
	방법 2	9,541	3,657	1,837	1,415
	방법 3	17,757	8,317	4,159	3,176
	방법 4	7,163	3,327	1,729	1,112
변동계수	방법 1	0.0964	0.0464	0.0202	0.0163
	방법 2	0.0475	0.0237	0.0107	0.0075
	방법 3	0.0986	0.0468	0.0211	0.0164
	방법 4	0.0401	0.0192	0.0091	0.0063

평균제곱오차 (MSE)

		표본수 (전체 차량의 수 10,000)			
		30	100	500	1,000
MSE	방법 1	7.47e+6	3.00e+6	5.14e+5	2.48e+5
	방법 2	2.23e+6	9.23e+5	1.72e+5	6.50e+4
	방법 3	1.37e+7	5.36e+6	2.45e+6	2.04e+6
	방법 4	1.52e+6	6.67e+5	1.53e+5	8.00e+4

V. 토의

■ 활동자료에 대한 심도있는 분석이 필요

- ◆ 주행거리 - 추정, 차종, 연식 등
- ◆ 연비(연료소비량) - 추정, 차종, 연식 등

■ 확률밀도함수 추정

- ◆ 비모수적인 방법
- ◆ 모수적인 방법

■ 연료소비량 추정의 방법론에 대한 객관적인 비교 연구 필요

■ 기타

필요한 활동자료

■ 주행거리

차량번호	용도별	차종1	차종2	차종3	출시연 도	전 검사일	전 검사일 주행거리	현 검사일	현 검사일 주행거리
	자가용자동차	승용	일반형	1,000cc 미만					
				1,500cc 미만					
				2,000cc 미만					
				2,000cc 이상					

■ 실제연비(?)

