■ 論 文 ■

통행시간가치의 신뢰구간 추정(II)

Estimating Confidence Interval of Value of Travel Time(II)

조 중 래
(명지대학교 교통공학과 교수)

박 철 규
(수원대학교 응용통계학과 교수)

목 차

요 약

통행시간가치는 교통계획 및 교통투자정책 평가에 있어서 매우 중요한 의미를 갖는다. 특히 교통시설투자와 관련된 타당성분석에서 통행시간가치는 해당사업의 경제적 타당성을 판단함에 있어 결정적인 영향을 미치게 된다.

지금까지 교통시설투자와 관련된 많은 타당성조사사업에서 시간절약 편익산정 결과에 대한 신뢰성문제가 빈번하게 제기되어 왔고, 심지어는 시간절약편익을 편익항목에서 제외하여야 할 것이라는 극단적인 주장도 있었다. 'IMF시대'로 이야기되는 최근의 경제상황을 고려할 때 투자사업평가에 대한 분석결과의 정확성에 대한 사회적 및 행정적 요구는 더욱 커질 것이며, 따라서 시간절약편익에 대한 신뢰성의 문제는 앞으로 더욱 강조될 것으로 예상된다.

본고에서는 통행시간과 통행비용의 비율에 대한 점근분포함수를 추정하고 이를 통하여 통행시간가치의 신뢰구간을 추정하는 방법을 제시하였다. 점근분포함수를 이용한 AD법은 파라메타의 신뢰구역을 이용한 기존의 CR법에 비하여 통계이론적 기초가 탄탄하며, 또한 사례연구의 결과를 통해서 볼 때, CR법에 비하여 그 분석결과가 더 정교한 것으로 분석되었다. 그러나 AD법은 선택모형을 정산하기 위하여 사용되는 표본자료의 수가 많아야 한다 한계를 갖으며, 따라서, 모형정산을 위한 표본자료의 수가 충분하지 못한 경우에 있어서는 CR법이 사용되어야 할 것이다.

AD법을 이용하여 분석된 서울시 출근통행자 시간가치의 95%신뢰구간은 7341.25±1945.05(원/시간)로 추정되었다.

Ⅰ. 서론

통행시간가치는 교통계획 및 교통투자정책 평가에 있어서 매우 중요한 의미를 갖는다. 교통시설투자와 관련된 타당성분석에서 통행시간가치는 해당사업의 경제적 타당성을 판단함에 있어 결정적인 영향을 미치게 되며, 혼잡통행료 등 교통수요관리와 관련된 가격정책에 있어서는 가격기준설정을 위한 기본척도가 된다.

'IMF시대'로 이야기되는 최근의 경제상황을 고려할 때 투자사업평가에 대한 분석결과의 정확성에 대한 사회적 및 행정적 요구는 더욱 커질 것이며, 최근 시행되기 시작한 '예비타당성조사'는 이러한 경향을 반증하고 있다고 하겠다.

교통수요관리를 위한 가격정책의 내용은, 집중관리 대상이 되는 특정교통수단, 특정가로구간 혹은 특정지역에 대한 사용자비용을 조절함으로써 정책목표가 되는 교통수단 혹은 지역의 수요를 감축하고, 그를 통하여 도로교통의 혼잡을 완화하고자 하는 것이다. 개인교통수단에 대한 주차요금정책과 주행세는 특정교통수단에 대한 관리정책으로, 혼잡통행료정책은 특정가로구간의 관리를 위하여, 그리고 도심진입세는 특정지역의 교통혼잡관리를 위하여 적용된다.

이러한 각종 가격정책의 이론적 배경은 통행시간과 통행비용의 '상대가격(Relative Price)'의 조절을 통하여 교통수단 혹은 통행경로에 대한 개별적 선택행위에 사회적 공공성을 부여하고자 하는 것이다. 여기에서 상대가격이라 함은 동일한 단위로 환산된 통행시간과 통행비용의 비율을 의미하는 것으로, 단위의 단일화를 위해서는 일반적으로 통행시간가치를 사용한다. 이것은 통행자들의 통행시간가치가 각종 도로가격정책의 '적정가격수준'을 결정하는 핵심적인 요소가 됨을 의미한다. 따라서, 시간가치의 신뢰성 문제는 혼잡통행료 등의 가격정책에 있어 설정된 가격수준의 신뢰성뿐만 아니라 관련정책의 사회적 공감대를 확보함에 있어 매우 중요한 요인으로 작용하게 된다.

조중래(1998)는 이산형 선택모형의 효용함수를 구성하는 통행시간변수와 통행비용변수에 대한 매개상수들의 신뢰구역(Confidence Region)을 이용하여 통행시간가치의 점근신뢰구간(Asymptotic Confidence Interval) 을 추정하는 방법을 제시한 바 있다. 본 연구는 앞의 연구에 대한 후속 연구로서, 본고에서는 앞의 연구에서 제안된 CR(Confidence Region)법의 한계를 검토하고 그 극복방안을 제시한다. 2장에서는 CR법의 내용과 한계점을 검토하고, 3장에서는 분포함수를 통한 시간가치 신뢰구간추정에 관한 방법론을 제시한다. 사례연구를 통한 두 방법간의 비교분석결과는 4장에서 설명된다.

Ⅱ. VOT 신뢰구간 추정법(I) : CR법

통행시간가치(VOT : Value of Travel Time)는 통행시간에 대한 한계효용과 통행비용에 대한 한계효용의 비율에 관한 함수가 된다. 따라서, 이산형 선택이론(Discrete Choice Theory)을 기본으로 한 한계대체율법을 이용하여 통행자의 통행시간가치를 구할 경우, u_m을 선택모형에서 사용된 선택대안(Choice Alternative) m의 효용이라 할 때, 선택대안 m을 이용하는 통행자의 통행시간가치는

$$VOT = \frac{\partial u_m / \partial T_m}{\partial u_m / \partial C_m} = f(\frac{\widehat{\beta_T}}{\widehat{\beta_C}}) \tag{1}$$

로 표현된다. 식(1)에서 $\widehat{\beta_T}$ 및 $\widehat{\beta_C}$는 각각 효용함수에 포함된 통행시간 T_m 및 통행비용 C_m의 파라메타를 나타낸다.

조중래(1998)는 통행시간 파라메타와 통행비용 파라메타에 대한 신뢰구역(Confidence Region)을 이용하여 통행시간가치의 점근신뢰구간(Asymptotic Confidence Interval)을 추정하는 방법을 제시한 바 있다. 일반적으로 $\widehat{\boldsymbol{\beta}}$를 최우추정법에 의해 추정된 모형내의 파라메타 벡타라고 할 때, 파라메타 벡타 $\boldsymbol{\beta}$에 대한 유의수준 $(1-\alpha)$에서의 신뢰구역(Confidence Region)은

$$(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \widehat{\Sigma}_{\widehat{\beta}}^{-1} (\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})^{T} \le \chi^2_{k,\alpha} \tag{2}$$

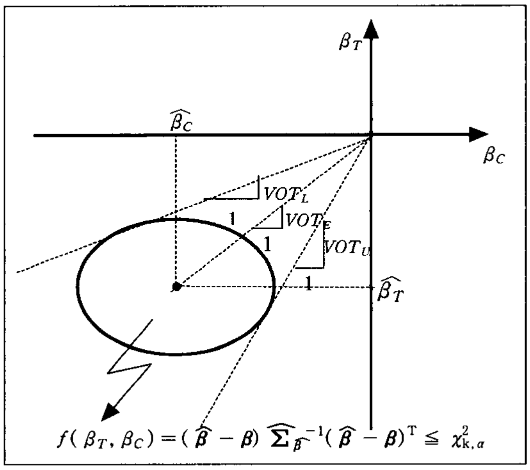
로 표현된다. 여기서, k는 χ^2분포의 자유도를 의미하며 $\widehat{\Sigma}_{\widehat{\beta}}$ 는 $\widehat{\boldsymbol{\beta}}$의 공분산행렬에 대한 추정치를 나타낸다. 특히, 파라메터 벡타가 $\boldsymbol{\beta} = (\beta_T, \beta_C)$일 때 자유도는 $k=2$가 되며, $\widehat{\boldsymbol{\beta}}$의 점근 공분산행렬은

$$\Sigma_{\widehat{\beta}} = \begin{bmatrix} Var(\widehat{\beta_T}) & Cov(\widehat{\beta_T}, \widehat{\beta_C}) \\ Cov(\widehat{\beta_C}, \widehat{\beta_T}) & Var(\widehat{\beta_C}) \end{bmatrix} \qquad (3)$$

과 같다.

조중래가 제시한 파라메타의 신뢰구역을 이용한 방법(Confidence Region Method:CR법)은 통행시간가치가 식(2)로 표현된 $\beta_C - \beta_T$평면상의 임의의 한 점에서 원점에 이르는 직선의 기울기로 표현되며, 따라서 신뢰구간의 상한치(Upper Bound)와 하한치(Lower Bound)는 원점을 지나면서 신뢰구역에 접하는 접선의 기울기로 표현된다는 점에 기초한다. 〈그림 1〉은 CR법의 개념도를 나타낸 것으로, 점($\widehat{\beta_C}$, $\widehat{\beta_T}$)에서 원점에 이르는 직선의 기울기 VOT_E는 시간가치 기대값의 추정치를 나타내며, 같은 논리로 원점에서 신뢰구역에 접하는 두 개의 직선의 기울기 VOT_U 및 VOT_L은 각각 유의수준 α에서의 시간가치의 상한치와 하한치를 의미한다(조중래, 1998).

그러나, 이 방법은 시간가치를 규정하는 파라메터들의 추정량에 대한 이차변형(Quadratic Transformation)을 통하여 신뢰구간을 제시하므로, 신뢰상한과 하한이 추정치에 대하여 대칭을 이루지 못 한다. 또한, 이차변형에 따른 분산의 왜곡은 신뢰구간의 크기에 영향을 주며 따라서 실제로 주어진 신뢰계수($1-\alpha$)의 정확성을 저하시키게 된다.



자료 : 조중래, "통행시간가치의 신뢰구간추정", 대한교통학회지 제16권 제4호, 1998, p.222.

〈그림 1〉 통행시간가치 신뢰구간 추정을 위한 CR법 개념도

CR법의 이러한 이론적 한계를 극복하기 위하여, 다음절에서는 통행비용과 통행시간의 비율(β_T/β_C)에 대한 추정량의 점근분포(Asymptotic Distribution)를 도출하고, 이에 근거한 시간가치의 신뢰구간을 추정하는 방법을 제시한다.

Ⅲ. VOT 신뢰구간 추정법(II) : AD법

점근분포함수를 통하여 시간가치의 신뢰구간을 추정하는 AD법(Asymptotic Distribution Method)을 설명하기 위하여 본고에서는 로짓형태의 수단선택모형을 가정한다. 로짓모형 효용함수의 파라메타에 대한 최우추정량은 점근적으로 정규분포를 따른다(Ben-Akiva and Lerman, 1985). 시간가치의 신뢰구간을 추정하려면 우선 로짓모형 효용함수의 속성변수로 포함되는 통행시간 파라메타 β_T와 통행비용 파라메타 β_C의 비율(β_T/β_C)의 추정량에 대한 확율분포를 구하여야 한다. 그러나, 유한표본하에서는 이들 개별 추정량의 분포를 전혀 알 수 없기 때문에, 시간가치의 추정량에 대한 분포를 도출하는 것이 불가능하다. 또한, 통행시간과 통행비용의 추정량 벡타가 점근적으로 이변량 정규분포를 따른다고 할지라도, 통행시간가치의 추정량에 대한 점근분포가 이변량 정규분포로부터 변수변환을 통해 얻어지는 것이 아니다. 따라서, 통행시간가치에 대한 신뢰구간의 구축은 근사적으로 이루어질 수밖에 없으며, 이를 위해 변수변환에 적용될 수 있는 일반적 점근이론(Asymptotic Theory)이 필요하다. 차후의 논의를 위해 확률벡타 $\mathbf{X}_n = (X_{n1}, X_{n2}, \dots, X_{nk})$에 대하여 $\sqrt{n}(\mathbf{X}_n - \mathbf{u})$의 분포가 n이 무한대로 커짐에따라 다변량정규분포 $N(\mathbf{u}, \widehat{\Sigma})$로 수렴하면 $\mathbf{X}_n \sim AN(\mathbf{u}, \frac{\Sigma}{n})$로 표시한다. 다음의 정리는 점근적으로 다변량 정규분포를 따르는 확율변수들의 변환함수가 다시 정규의 점근분포를 갖게 됨을 보여 준다.

정리 1(Serfling, 1980)

함수 $g(\mathbf{t})$가 $\mathbf{t} = \mathbf{u}$에서 0 이 아닌 미분치를 가질 때, $\mathbf{X}_n$이 $\mathbf{X}_n \sim AN(\mathbf{u}, \frac{\Sigma}{n})$을 만족하면 $g(\mathbf{X}_n) \sim AN(g(\mathbf{u}), \frac{D(\mathbf{u})\Sigma D(\mathbf{u})^T}{n})$이다. 여기에서,

$$\mathbf{D}(t) = \left(\frac{\partial g(t)}{\partial t_1}, \frac{\partial g(t)}{\partial t_2}, \cdots, \frac{\partial g(t)}{\partial t_k} \right).$$

이 정리를 이용하면 통행시간가치에 대한 근사적 신뢰구간을 제시할 수 있다. 표본의 크기가 n일 때, 로짓모형의 파라메타 $\boldsymbol{\beta} = (\beta_T, \beta_C)$에 대한 최우추정량 $\hat{\boldsymbol{\beta}}$은

$$\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \longrightarrow^{d} N(\mathbf{0}, \boldsymbol{\Sigma}) \quad (4)$$

를 만족한다. 다시 말해, 식(4)는 $\hat{\boldsymbol{\beta}} \sim AN(\boldsymbol{\beta}, \frac{\boldsymbol{\Sigma}}{n})$와 같으며, $\hat{\boldsymbol{\beta}}$의 점근적 평균벡타와 공분산 행열이 각각 $\boldsymbol{\beta}$와 $\boldsymbol{\Sigma}/n$임을 보여준다. 여기서, 편의상 $\hat{\boldsymbol{\beta}}$의 점근적 공분산행열 $\boldsymbol{\Sigma}/n$을 $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}}$로 표시하기로 한다. 효용함수가 통행시간과 통행비용에 대하여 선형일 경우, 식(1)에서 보는 바와 같이 통행시간가치는 단순히 두 속성변수의 파라메타의 비율로 표현된다. 따라서, 함수 g를

$$g(\boldsymbol{\beta}) = \frac{\beta_T}{\beta_C} \quad (5)$$

라고 정의하면, 통행시간가치가 $VOT = g(\boldsymbol{\beta})$와 같고 그 추정량은 최우추정량의 불변성(invariance property)에 의해 $VOT_E = g(\hat{\boldsymbol{\beta}})$로 표현된다. 따라서, 함수 $g(\boldsymbol{\beta})$를 1차 미분하여 얻어지는 벡타를,

$$\mathbf{D}(\boldsymbol{\beta}) = [1/\beta_C, \ -\beta_T/\beta_C^2] \quad (6\text{-}1)$$

라 하면 정리 1에 의해

$$g(\hat{\boldsymbol{\beta}}) \sim AN(g(\boldsymbol{\beta}), \mathbf{D}(\boldsymbol{\beta})\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}}\mathbf{D}(\boldsymbol{\beta})^T) \quad (6\text{-}2)$$

이 성립한다. 이 사실은 표본의 크기 n이 충분히 클 때

$$\Pr[-\Phi \le g(\hat{\boldsymbol{\beta}}) - g(\boldsymbol{\beta}) \le \Phi] \doteq 1 - \alpha \quad (7)$$

임을 의미하며, 여기서 $\Phi = Z_{\frac{\alpha}{2}}\sqrt{\mathbf{D}(\boldsymbol{\beta})\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}}\mathbf{D}(\boldsymbol{\beta})^T}$를, 그리고 $Z_{\alpha/2}$는 표준정규분포에서의 $100(1-\alpha/2)$ 백분위수를 나타낸다.

그러나, 아직은 식(7)을 이용하여 신뢰구간을 제시할 수 없다. 그것은 $g(\hat{\boldsymbol{\beta}})$의 점근분산인 $\mathbf{D}(\boldsymbol{\beta})\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}}\mathbf{D}(\boldsymbol{\beta})^T$에 미지의 파라메터 $\boldsymbol{\beta}$가 포함되어 있기 때문이다. 하지만, $g(\hat{\boldsymbol{\beta}})$의 점근분산 $\mathbf{D}(\boldsymbol{\beta})\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}}\mathbf{D}(\boldsymbol{\beta})^T$에 대하여 그 추정량인 $\mathbf{D}(\hat{\boldsymbol{\beta}})\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}}\mathbf{D}(\hat{\boldsymbol{\beta}})^T$이 강한 일치성을 갖는다는 성질을 이용하면, 식(7)에서 $g(\hat{\boldsymbol{\beta}})$의 점근분산에 그 추정량을 대체함으로써 근사적인 신뢰구간의 구축이 가능하다. 따라서, 신뢰구간의 정의에 따라 식(5)와 식(7)로부터 통행시간가치 VOT의 $100(1-\alpha)\%$ 근사 신뢰구간은

$$VOT_E \pm Z_{\frac{\alpha}{2}}\sqrt{\mathbf{D}(\hat{\boldsymbol{\beta}})\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}}\mathbf{D}(\hat{\boldsymbol{\beta}})^T} \quad (8)$$

로 표현된다.

대부분의 수단선택분석 프로그램에서 $\hat{\boldsymbol{\beta}}$의 공분산행열에 대한 추정치 $\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}}$가 제공된다. 또한, 식(6-1)에서 정의된 바와 같이

$$\mathbf{D}(\hat{\boldsymbol{\beta}}) = [1/\hat{\beta}_C, \ -\hat{\beta}_T/\hat{\beta}_C^2] \quad (9)$$

이므로, 식(8)에 의한 통행시간가치의 신뢰구간을 쉽게 구축할 수 있다.

IV. 사례연구 : CR법과 AD법의 신뢰구간 추정결과 비교

조중래는 전게서에서 1996년에 실시한 서울의 가구통행실태 조사자료로부터 추출된 3,659샘플의 출근통행 자료를 이용하여 수단선택모형을 구축하고, 그로부터 CR법을 이용하여 서울시 출근통행자의 통행시간가치에 대한 근사신뢰구간을 추정한 바 있다.

〈표 1〉의 수단선택모형으로부터 CR법을 이용하여 추정된 서울의 출근 통행자의 평균시간가치는 7,341(원/시)였고, 95%신뢰구간은,

$$5,454(\text{원/시간}) \le VOT \le 10,806(\text{원/시}) \quad (10)$$

였다(조중래, 1998).

〈표 1〉 서울시 출근통행 수단선택모형

변 수 명		Coefficient(t-value)	
1.승용차더미	(d_a)	1.7660	(10.694)
2.버스더미	(d_b)	8.5631	(20.514)
3.지하철더미	(d_s)	8.3579	(19.931)
4.통행시간	(TIME)	-0.055818	(-21.782)
5.통행비용	(COST)	-0.0³45620	(-7.851)
6.승용차유무	(CAR)	5.1849	(23.776)
7.면허소지여부	(LIC)	0.79207	(2.734)
8.소득	(INC)	0.0²14676	(1.957)
9.가족수	(FS)	0.042828	(1.000)
10.연령	(AGE)	0.012552	(1.948)

Statistics :

Number of observations	: 3,659
Log-Likelihood	: -2581.488
Restricted(Slope=0)	: -5072.451
Chi-Squared	: 4981.926
ρ^2	: 0.49108

Covariance Matrix :

	TIME	COST
TIME	$0.0^5 65668$	
COST	$0.0^9 74016$	$0.0^8 33764$

주 : 0.0^n1234는 0.1234×10^{-n}을 의미한다.
자료 : 조중래, "통행시간가치의 신뢰구간추정", 대한교통학회지 제16권 제4호, 1998, p.222.

본 절에서는 〈표 1〉에서 제시된 바와 동일한 모형에 대하여, 앞에서 제안된 AD법을 이용하여 통행시간가치를 추정하고 이를 CR법으로부터 추정된 결과와 비교한다.

우선, 〈표 1〉의 모형정산을 위하여 사용된 자료의 표본수는 3,659로 AD법을 이용하기에 충분하다고 판단된다. 이제, 〈표 1〉에 나타난 정산결과로부터 AD법을 적용하기 위해 필요한 통계치들을 요약하면 다음과 같다.

$$\widehat{\beta_T} = -0.055818 \tag{11-1}$$

$$\widehat{\beta_C} = -0.0^3 45620 \tag{11-2}$$

$$\widehat{\Sigma_{\beta}} = \begin{bmatrix} 0.0^5 65668 & 0.0^9 74016 \\ 0.0^9 74016 & 0.0^8 33764 \end{bmatrix} \tag{11-3}$$

식(11-1)과 식(11-2)의 값을 이용하면, 통행시간가치의 기대값은

$$VOT_E = \frac{\widehat{\beta_T}}{\widehat{\beta_C}} = \frac{-0.055818}{-0.0^3 45620} = 122.35(\text{원/분}) \tag{12}$$

이 되고, 이를 다시 시간당 시간가치로 환산하면 $VOT_E = 7341.25$(원/시간) 이 된다. 한편, 식(9)에서

$$\mathrm{D}(\widehat{\beta}) = \left[\frac{1}{-0.0^3 45620},\ -\frac{-0.055818}{(-0.0^3 45620)^2}\right] = [-2192.02,\ \ 268203.05]$$

를 얻는다. 따라서, 식(11)로부터

$$\mathrm{D}(\widehat{\beta})\widehat{\Sigma_{\beta}}\mathrm{D}(\widehat{\beta})^T = 256.94 \tag{13}$$

이 된다. 식(12), 식(13) 및 표준정규분포의 97.5%의 백분위수 $Z_{0.025}=1.96$을 식(8)에 대입하여 95%신뢰구간을 구하면 122.35±31.42(원/분)이 되며, 이를 시간당 가치로 환산하여 한시간에 대한 통행시간가치를 구하면 7341.25±1945.05(원/시간)이 된다. 따라서, 통행시간가치의 95%신뢰구간은

$$5,396(\text{원/시간}) \le VOT \le 9,287(\text{원/시간}) \tag{14}$$

의 범위를 갖는다.

위의 AD법을 이용하여 추정된 결과를 식(10)의 CR법을 이용하여 추정된 결과와 비교할 때, 신뢰구간의 상한치와 하한치 모두에 있어서 CR법에 비하여 AD법의 추정결과가 작은 것으로 나타났다. 즉, CR법에 비하여 AD법의 추정결과가 하한치에 있어서는 58(원/시), 그리고 상한치에 있어서는 1,519(원/시) 작은 것으로 분석되었다. 이는 CR법의 추정결과를 기준으로할 때, 추정방법의 차이에 따라 상한치에 있어서는 14.0%, 그리고 하한치에 있어서는 1.1%의 편차가 있음을 의미한다. CR법의 또 다른 문제는 신뢰구간의 중심(8175.5원/시)과 시간가치의 기대값(7341.3원/시)이 일치하지 않는다는 점이다. 또한, AD법에 의한 시간당 시간가치의 신뢰구간의 길이가 3,891(원/시)로서 CR법의 5,352(원/시)보다 짧아 AD법에 의한 신뢰구간이 더 정교함을 알 수 있다.

V. 결론

본고에서는 통행시간과 통행비용의 비율에 대한 분포함수를 추정하고 이를 통하여 통행시간가치를 추정하는 방법을 제시하였다. 본고에서 제시된 AD법은 파라메타의 신뢰구역을 이용한 기존의 CR법에 비하여 통계이론적 기초가 탄탄하며, 또한 사례연구의 결과를 통해서 볼 때, CR법에 비하여 그 분석결과가 더 정교하다. 그러나 AD법은 선택모형을 정산하기 위하여 사용되는 표본자료의 수가 많아야 한다는 한계를 갖는다. 따라서, 모형정산을 위한 표본자료의 수가 충분하지 못한 경우에 있어서는 CR법이 사용되어야 할 것이다.

참고문헌

1. 조중래, '통행시간가치의 신뢰구간 추정', 대한교통학회지 제16권 제4호, 1998.
2. Ben-Akiva M. and S. Lerman, Discrete Choice Analysis: Theory and Application to Travel Demand, MIT Press, Cambridge, 1985.
3. Serfling R. J., Approximation Theorems of Mathematical Statistics, John Wiley & Sons, 1980.