

期待値の信頼区間を考慮した旅行時間データの取得技術に関する研究*

A Study on Travel Time Data Collection Techniques Considering Confidence Interval of Expected Value*

有吉亮**・中村文彦***・岡村敏之****・矢部努*****

By Ryo ARIYOSHI**・Fumihiko NAKAMURA***・Toshiyuki OKAMURA****・Tsutomu YABE*****

1. はじめに

人口の減少や高齢化の進展といった社会構造の変化に伴い、交通計画の方向性にも大きな転換が求められている。すなわち、従来の大まかな計画手法に基づく需要対応型施設整備から、より細やかで局所的なニーズに対応した施設整備計画へのシフト、量から質へ、社会から個人への指向である。たとえば旅行時間という指標がある。これは道路の質的なパフォーマンスを示す指標であり、同時に個人の道路ユーザーにとって重要な情報でもある。したがって、こうした指標に対する管理者のニーズ、ユーザーのニーズは、今後さらに高まると考えられる。他方で、IT技術の目覚ましい進展に伴い、道路交通関連データの取得技術は急速に進化している。旅行時間データの取得にしても、VICS、ETC、AVI、プローブと、様々な技術で対応可能となってきた。このように、1つの分析ニーズに対して複数のデータソースが存在する場合、それぞれのツールの特性を正しく見極め、状況に応じて適当なものを選びとることが肝要である。そのような分析者の意思決定を支援するため、データの取り扱いに関する基礎的な理論と、取得技術に関する特徴点を整理した。

2. 研究の位置付け

道路の質的なパフォーマンス指標である旅行時間は、複数の主体からニーズを受け、かつ多様な取得技術が存在しているデータであるといえる。そこで本研究は、旅行時間データの取得に着目した。旅行時間の取得技術や予測手法に関する研究は数多く行われており、車両感知器データによるリンク旅行時間の推定に関する研究¹や、旅行時間データ取得技術の精度向上に関する研究²、プ

ローブデータのクレンジング処理やマップマッチング処理に関する研究³、VICSアップリンクデータに基づく旅行時間と車両感知器データに基づく旅行時間との誤差要因に関する研究⁴などがある。しかしながら、いずれも旅行時間の期待値推定に際して、推定値の信頼性と必要サンプル数についての統計的な考察はみられない。また、分析のニーズに対して旅行時間データの取得技術を横並びで評価するための理論も整備されていない。旅行時間の期待値を一定の信頼区間で推定するためのサンプルサイズは、許容誤差の大きさを設定し、サンプルの変動を仮定しなければ求めることができないが、もともと旅行時間調査は、小サンプルでの瀬踏み調査的性格が強く、推定値の統計的な有意性には取って触れられてこなかった。しかし、道路の質的な整備が求められ、道路事業の費用便益分析の評価結果がIRという形で投資者（＝国民）に広く公開されるようになった今、情報の信頼性というものから目を背け続けることはできなくなったといえる。そこで本研究では、様々な分析の場面を想定し、許容される誤差の幅に対して、旅行時間の期待値の信頼区間がどう揺らぐか、必要サンプル数がどう変動するかという点を、それぞれのデータ取得技術について検討し、横並びの評価が可能となるような理論の整備を行った。

3. 研究の方法

(1) 分析のニーズに関する整理

旅行時間の期待値を推定する際、考慮すべき分析ニーズは、どの時間帯の旅行時間を調査するか（時間のニーズ）、どの区間の旅行時間を調査するか（空間のニーズ）、期待値の信頼区間をどこまで絞るか（精度のニーズ）、の3つに大別できる。この3つのニーズについて、考慮すべき点をそれぞれ整理した。

(2) 旅行時間データ取得技術に関する特徴点の整理
旅行時間データの取得技術をシステムの特性別に分類すると、定点観測データ、フローティングカーデータ、プローブビークルデータ、バスロケーションシステムデータ、の4つに大別できる。この4タイプのデータ取得技術について、特徴点をそれぞれ整理した。

*キーワード：旅行時間、プローブ、VICS、ITS

**学生員、横浜国立大学大学院環境情報学府
(神奈川県横浜市保土ヶ谷区常盤台79-5、
TEL045-339-4039、FAX045-339-4039)

***正員、博(工)、横浜国立大学大学院環境情報研究院

****正員、博(工)、横浜国立大学大学院工学研究院

*****正員、修(工)、横浜国立大学大学院環境情報研究院

4. 分析のニーズに関する整理

時間、空間、精度の面から旅行時間データの扱い方を考察するにあたり、時間的、空間的にある程度の広がりがあり、かつ統計的な処理に耐えうるだけの密度をもった実際の旅行時間データを例にとることが有効と考えられる。そこで今回は、時間とデータ密度について有利なバスロケーションシステムデータを用いた。データの詳細については表1の通りである。

表1 分析に用いたバスロケーションデータ

区間	東京都内の国道 2.06km
データ取得時間	6:00 - 24:00
データ取得期間	11/1-12/27 2004
有効サンプル総数	3819
平均データ数	4sample/h

注) 停留所の発着時刻から旅行時間を算出 (停車時間を除いた)

(1) 時間のニーズ

この場合のニーズは、旅行時間の変動をどの程度の時間間隔で集計すればよいかということになる。時間間隔が短すぎれば、そこに存在するサンプル数も少なくなり、旅行時間の分散は大きくなってしまふ。三輪らは記述長最小原理を用いて集計時間間隔を調整しているが、この手法にはパラメータの設定に人の判断が入り、モデルもやや複雑であるため、本研究ではより単純な手法を提案する。その概念を図1に示す。

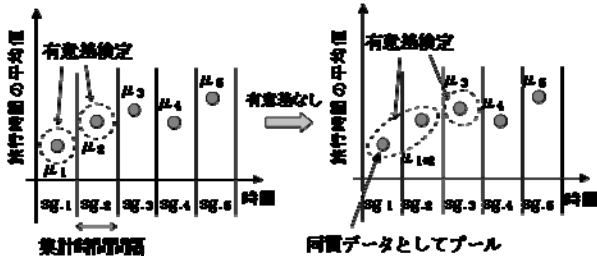
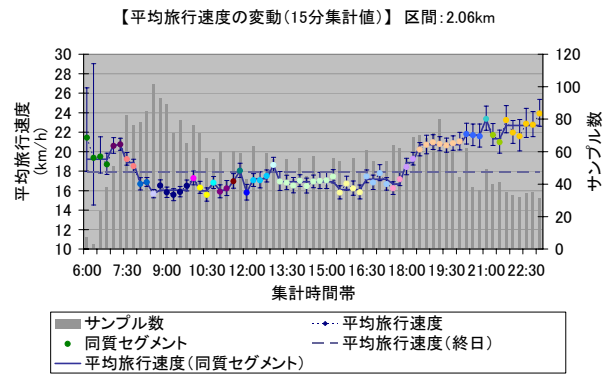


図1 同質なセグメントのまとめ上げ

まず、旅行時間の変動が時間方向について連続であると仮定する。次に、集計時間間隔を15分と決める。それぞれの時間間隔を、セグメントとして区別する。すべてのセグメントについて、旅行時間の期待値と分散を計算する。セグメント1と2のデータについて、分散の同質性検定(5%有意)を実施する。検定結果に応じて等分散または非等分散の下で期待値の差のt検定(5%有意)を行い、有意差が認められる場合は、セグメント1と2を異なる変動セグメントとして区別し、セグメント2と3についての吟味に移る。有意差が認められない場合は、セグメント1と2を同一の変動セグメントとして結合し、両者のデータをプールして期待値と分散を再計

算し、セグメント(1+2)と3について吟味に移る。以上のプロセスをすべてのセグメントに適用すると、旅行時間の終日変動をいくつかの同質なセグメントにまとめ上げることができる。続いて集計時間間隔を30分、60分、120分と変え、同様のプロセスを実行する。旅行時間の変動傾向を最も的確につかめる集計時間間隔とは、大局的な変動ポイントが有意に区別されていることと、微小な変動は大局的な変動のセグメントに内包されていることであると考え、これを満たすような集計時間間隔を採用する。図2および図3より、この調査区間の場合、60分の集計時間間隔で、旅行時間の大局的な変動傾向をつかめることがわかる。



注) 縦軸は区間距離を平均旅行時間で除した平均旅行速度表現

図2 旅行時間の変動(集計時間間隔15分)

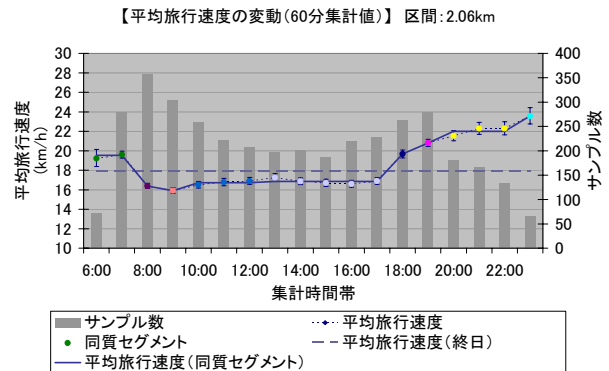


図3 旅行時間の変動(集計時間間隔60分)

(2) 空間のニーズ

この場合のニーズは、旅行時間の分布をどの程度の区間長で評価すればよいかということになる。まず、調査対象の路線(OD)をいくつかの小区間に分割することを考える。さらに、図4で示すように、i) この小区間がすべて信号交差点を含まない場合と、ii) 1つ以上の信号交差点を含む場合に分ける。

i) の場合、小区間の旅行時間は、その時間断面における交通量と、小区間内の密度分布により決まる。信号交差点から十分な距離にある小区間内の密度分布はほぼ一様であると考えられるため、この小区間における旅行

時間の分散 v_1 は相対的に小さくなると考えられる。

ii) の場合、小区間内に信号交差点が含まれているため、この区間における旅行時間は信号現示により確率的に変化する。よって、旅行時間の分散 v_2 は v_1 よりも大きくなると考えられる。また、小区間に含まれる信号交差点の数が多の場合、信号交差点での右左折がある場合、あるいは集計の時間帯が長くなる場合には、それだけ分散も大きくなる。ただし、信号交差点の数が増えても、ある程度まで小区間が長くなれば、内在する特性変動がキャンセルされて、旅行時間の変動が小さく（分散が小さく）なるという報告がある。起点を一致させた実際の旅行時間データで、その傾向が確認できることを図5に示した。

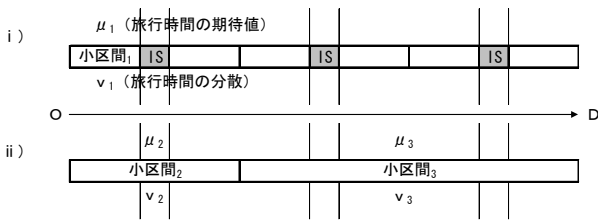


図4 対象路線 (OD) の小区間分割

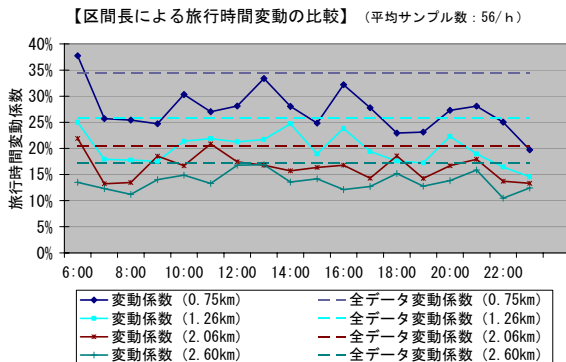


図5 区間長による旅行時間変動の比較

以上より、ある時間断面における混雑状況の分布が知りたい場合、集計区間長 (小区間長) をどのように決めるか、どの時間帯で集計するかは、サンプル数の分布を考慮しながら、分析者が調整することになる。集計区間に信号交差点を含める場合は、区間長の変化に対する旅行時間の分散の感度特に注意する。

(3) 精度のニーズ

ある区間を、ある時間帯に通過する車両の旅行時間に関する期待値を、ある信頼区間の幅 (=許容誤差の幅) を確保して求めるのにどの程度のサンプル数が必要かは、許容誤差の幅と母集団変動の仮定 (実際には推定値としてのサンプル変動の仮定) により決まる。そこで、母集団の仮定に関する理論を整理したうえで、サンプル変動と許容誤差、サンプルサイズの関係を整理した。

a) 母集団の仮定

ある時間内にある区間を通過した車両の旅行時間の期待値を推定するには、母集団、すなわち全通過車両の旅行時間の分布形を仮定しなければならない。例えばある1日のある時間帯、全車両の旅行時間が取れたとする。この全車両が母集団かといえ、当日に関してはそうである。しかし、我々が知りたいのは、同じ属性のすべての1日に関する期待値であり、ある1日の全数データは、すべての1日の全数データに関する1サンプルに過ぎない。よって、我々はより大きな母集団に対して旅行速度のサンプリング調査を行うものと考えられる。したがって、真の母集団には莫大な数の通過車両が想定される。つまり、中心極限定理により、母集団を正規分布とみなして統計処理を行ってよいことになる。以下の議論は、この仮定に基づいたものである。

b) 母集団変動と許容誤差、サンプルサイズの関係

平均 μ 、分散 σ^2 の正規母集団からとられた大きさ n の標本 (平均 \bar{x} 、分散 s_x^2) について統計量 $\frac{|\bar{x} - \mu|}{s_x / \sqrt{n-1}}$ は自由度 $n-1$ の t 分布に従う。

母平均 μ の 95% 信頼区間は次のように表せる。

$$\bar{x} - t_{0.025}(n-1) \frac{s_x}{\sqrt{n-1}} \leq \mu \leq \bar{x} + t_{0.025}(n-1) \frac{s_x}{\sqrt{n-1}}$$

サンプルの変動係数 $CV_s = \frac{n}{n-1} \left(\frac{s_x}{\mu} \right)^2$

許容誤差率 $\varepsilon = 1 - \frac{\bar{x}}{\mu}$ を用いて上式を書き直すと

$$n \geq \frac{t_{0.025}^2(n-1)}{\varepsilon^2} \frac{n}{n-1} \frac{s_x^2}{\left\{ \bar{x} \pm t_{0.025}(n-1) \cdot s_x / \sqrt{n-1} \right\}^2}$$

この式を満たす最小の n が 95% 信頼区間を与える最小サンプル数である。

したがって、許容誤差 ε と変動係数 CV_s を様々に変え、収束計算により n を求めると、図6を得る。

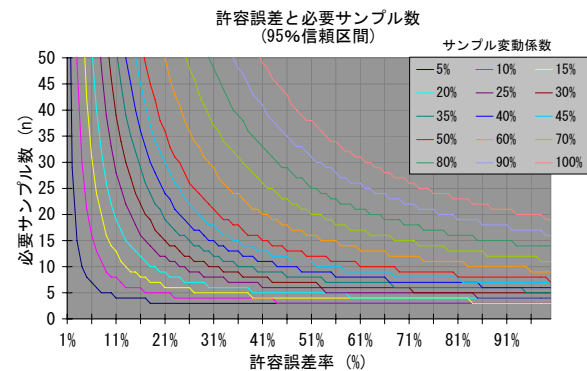


図6 分析の許容誤差と必要サンプル数

サンプルの変動 (=母集団変動の不偏推定値) が大きいほど必要サンプル数が大きくなるのがわかる。また、

必要サンプル数を求めるためには、母集団変動に関する情報が必要であるが、これは一般に未知であるため、予備的な調査を行うか、何らかの既存データを用いて推定しておく必要がある。また、推定値の誤差をどこまで許すか（許容誤差率）は、分析のニーズに応じて判断することになる。

5. 旅行時間データ取得技術に関する特徴点の整理

（１）定点観測技術

定点観測技術には、車両感知器、VICSアップリンク、ETC、AVIなどがある。車両感知器が地点通過速度を基にリンク旅行速度を推定するのに対し、VICSアップリンク、ETC、AVIは個々の車両を特定してリンク旅行速度の真値を観測できる。また、いずれも24時間の観測が可能である。VICS、ETCはデータ処理に係る手間が少なく、ストレージも相対的に小さいが、データ取得が車載機の普及率に依存している。既存研究⁵で扱われているVICSアップリンク車両の平均交通量に対する割合は、時間帯あたり1.4～2.7%と、サンプル率としてはまだ低い。また、ETCでは一般道路のデータ取得ができない。一方、AVIは全車両を観測可能であるが、画像処理によって車両を特定するため、システムの計算負荷が大きく、ストレージもかなり大きくなる。

（２）フローティングカー

フローティングカーとは、主に旅行時間の観測を目的として、交通流の中に追加的に投入される調査車両を指す。調査員が乗車するため、状況に応じた柔軟で能動的なデータ取得が可能である。GPSによる位置情報の取得に加え、人の手による交通状況の判断および記述、ビデオ撮影などが可能である。一般に、これらの調査で得られるデータはオフラインでの処理となる。また、データの取得時間や調査サンプル数は予算的な制約を受ける。

（３）プローブビークル

プローブビークルは、位置情報などのデータが、受動的にオンラインで取得される車両を指す。プローブビークルには、一般車の他に、商用車やバス、タクシーなどがある。このシステムの利点としては、24時間のデータ取得が可能であること、データの分布が空間的な広がりをもつこと、リアルタイムで情報が処理されることなどが挙げられる。反面、システムの構築に係るイニシャルコストが大きい、車両の走行特性や属性によってデータの分布が偏る、特定の路線について統計的な処理に耐えるだけのデータ密度が得られないといった問題がある。

（４）バスロケーションシステム

バスロケーションシステムからは、停留所の発着時刻情報が得られる。バスの位置データには、車両の属性によるバイアスが存在し、データサンプル数・データ取得時間・データ取得空間が事業者のサービスレベルに依存するといった問題があるが、データ処理に係る手間が少ない、ストレージが小さい、長期的なデータの蓄積があるといった利点を活用できれば、データの利用価値は高い。実際、交通量の多い都心部の幹線道路、すなわち旅行時間に対する分析ニーズが存在するような路線では、バスのサービスレベルも高く、統計的な処理に耐えうるだけの十分な密度のデータが得られることが多い。

6. まとめと今後の展開

実際の旅行時間データの分析結果を用いて理論の整理を行ったことで、旅行時間データの取り扱いに関していくつかの知見が得られた。時間に関する分析ニーズについては、同質な変動のセグメントを結合していくことにより、旅行時間の大局的な変動が明確となる集計時間間隔が得られることがわかった。空間に関する分析ニーズについては、集計区間長と旅行時間変動の関係ふまえ、分析者が適切な集計区間長を調整することが重要であるとわかった。精度に関するニーズについては、様々な路線の変動条件のもと、分析者が許容する誤差の幅と必要な調査サンプル数についての関係が明らかになった。旅行時間の取得技術については、それぞれのシステムの特徴点を整理することができた。

今後は、より具体的に詳細な分析のシナリオを設定し、その場面ごとに時間・空間・精度・コストの面から、データ取得技術の適用可能性を評価したい。また、複数のデータ取得技術の組合せによって対応可能な分析ニーズの存在についても検討していきたい。

¹ 高橋ら：車両感知器を用いた市街路旅行時間の推定方式，土木学会年次学術講演会講演概要集第4部，Vol. 34，1979年

² 上杉ら：断片的なプローブ軌跡の接合による区間旅行時間の期待値と分散の推定，土木計画学研究・講演集 Vol. 26，2002年

³ 堀口ら：プローブデータのクレンジング処理と車種別の運行特性分析土木計画学研究・講演集，Vol. 26，2002年

⁴ 堀口ら：アップリンクを利用した旅行時間と地点速度データに基づく旅行時間の誤差要因に関する理論的考察，第23回交通工学発表会論文報告書集，pp. 117-120，2003年

⁵ 森田ら：光ビーコンで収集した交通情報の解析とその応用の検討，SEIテクニカルレビュー，第164号，2004年