

プローブ旅行時間データの取得状況と 旅行時間信頼性指標値の信頼度との関係

関谷浩孝・諸田恵士・高宮 進

1. はじめに

近年、情報通信技術の進展、カーナビゲーションの普及により、車両の走行位置・時刻情報から算定される旅行時間データ（以下「プローブ旅行時間データ」という（図-1）。）を大量に取得できるようになった。国土交通省では、一定の時間帯（7:00～8:00の1時間帯等）、一定の区間（主要交差点間等）で集計したプローブ旅行時間データを用いて「旅行時間信頼性指標」を算定し、道路交通施策の実施効果を定量的に評価する取組みを行っている¹⁾（図-2）。図-3に示すように旅行時間信頼性指標とは、旅行時間の日々のばらつきを表す指標で「90%タイル旅行時間（10日に1回程度発生する混雑時の旅行時間）」や「95%タイル旅行時間（20日に1回程度発生する混雑時の旅行時間）」等がよく用いられる。

ところが、プローブ旅行時間データは全ての時間帯で取得できているわけではない。例えば、3ヶ月間（平日60日）の7時台の旅行時間信頼性指標を算定する際、7時台に1件以上のプローブ旅行時間データが取得できた日が10日しかない場合もある。当然のことながら、この“10日分のデータから算定される旅行時間信頼性指標”は、“全日（60日）のデータから算定される旅行時間信頼性指標”と値が異なり、指標値の確からしさ、あるいは信頼度は劣ると推察される。

そこで筆者らは、プローブ旅行時間データの取得状況と算定される旅行時間信頼性指標値の信頼度との関係を明らかにすることを目的とした分析を行った。

2. 使用データ

旅行時間のばらつきの小さい「OD区間1」及びばらつきの大きい「OD区間2」（図-4）において、平日7時台に取得された60日分の旅行時間

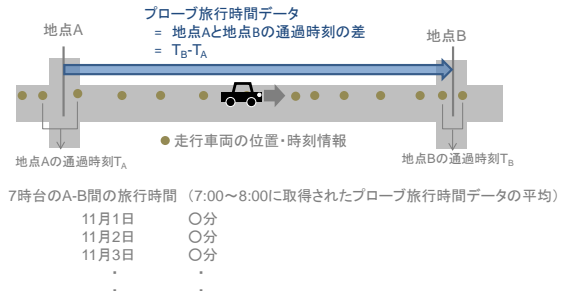


図-1 プローブ旅行時間データのイメージ

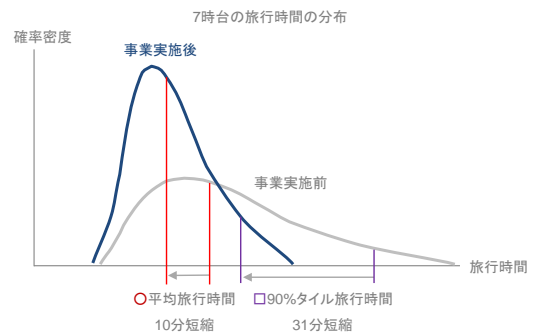


図-2 旅行時間信頼性指標を用いた施策効果の評価イメージ

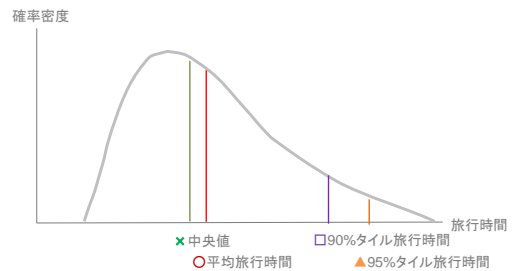


図-3 旅行時間信頼性指標の例



図-4 分析対象区間

データ（7:00～8:00に取得されたプローブ旅行時間データの平均。以下「OD区間データ」という。）を用いた。OD区間データの分布及び旅行時間信頼性指標値を図-5及び表-1に示す。

3. データ取得日数と信頼度との関係分析

60日のうち一部の日でしか旅行時間データが取得できていないケースを想定し、この場合に算定される旅行時間信頼性指標値（標準偏差、95%タイル値、90%タイル値、中央値及び平均値）が、真値とどの程度異なるかを分析した。ここでの真値は、表-1に示した「全日（60日）のOD区間データから算定した値」とした。本来「真値」とは、7時台に対象区間を走行した全ての車両の旅行時間から算定された値とすることが望ましい。しかし、一般国道では全車両の旅行時間の観測データが得られていない。このため、本研究では7時台に1件でもプローブ旅行時間データが取得できている場合、この値をこの日（7時台）の旅行時間と見なして真値を算定した。

3.1 標本データセットの作成及び指標の算定

60日のOD区間データのうち、2～59日分のOD区間データが取得されたケースを想定し、標本データセットを作成した。以下、OD区間2で60日のうち10日分のデータが取得できたケース（10日/60日）を例に示す。図-6に示すとおり、60日のOD区間データの中から10日分をランダムに選定し、旅行時間信頼性指標の標本値を算定した。これを1,000回繰り返し、1,000個の標本値を算定した〔補注1〕（表-2）。これを累積確率の形式で示すと図-7のとおりとなる。

3.2 旅行時間信頼性指標値の信頼度の算定

標本データセットから算定した旅行時間信頼性指標の標本値と真値との差が「真値の95%～105%内」となる場合に「正しく算定できた」と判定した。1,000個の標本のうち「正しく算定できた」標本の割合を「信頼度」とした。標準偏差での「信頼度」の算定例を図-8に示す。標本データセットから算定した1,000個の標本のうち、真値（7.6分）の95%～105%内（7.22～7.98分）となるものが176個あった。これより信頼度は17.6%（=176/1,000）と算定される。これは次を意味する：「全日（60日）のうち10日分の旅行時間データが取得できると、標準偏差を真値の

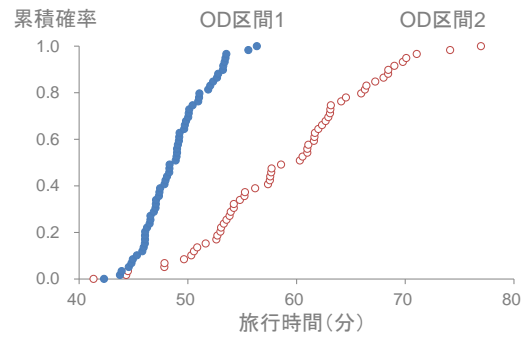
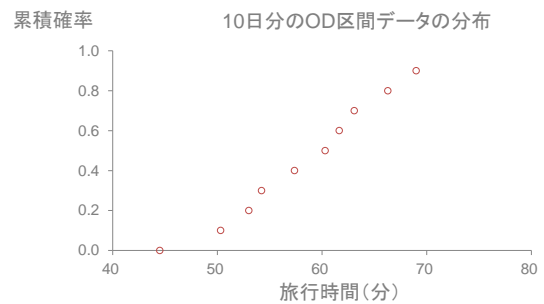


図-5 OD区間データ（60日）の分布

表-1 旅行時間信頼性指標値

	●OD区間1	○OD区間2
■標準偏差	3.0分	7.6分
▲95%タイル値	53.5分	70.1分
□90%タイル値	53.2分	68.5分
✕中央値	48.6分	59.4分
○平均値	48.8分	59.0分



標本データセットNo.	■標準偏差	▲95%T値	□90%T値	✕中央値	○平均値
1	8.3分	71.6分	69.0分	60.3分	59.4分

図-6 標本データセット（10日分）の作成及び旅行時間信頼性指標の標本値の算定

表-2 標本データセット（10日分）から算定された旅行時間信頼性指標の標本値（1,000個）

標本データセットNo.	■標準偏差	▲95%T値	□90%T値	✕中央値	○平均値
1	8.3分	71.6分	69.0分	60.3分	59.4分
2	7.2分	69.0分	67.9分	59.2分	59.0分
3	8.5分	72.1分	69.8分	60.9分	61.7分
4	5.8分	66.3分	64.2分	57.0分	55.9分
5	6.7分	68.2分	66.7分	58.2分	58.1分
⋮	⋮	⋮	⋮	⋮	⋮
1000	8.0分	70.6分	69.0分	60.2分	61.0分

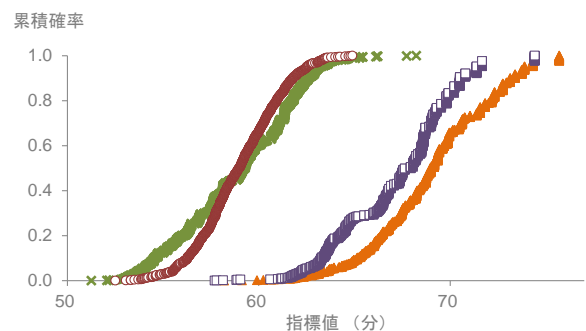


図-7 標本データセット（10日分）から算定された旅行時間信頼性指標の標本値（1,000個）の分布

95%～105%の値として正しく算定できる確率は17.6%。残りの82.4%の確率で、真値の±5%以上の誤差を持った値として算定される。」。

データ取得日数を「20日/60日」及び「40日/60日」として同様の算定を行った。これを図-9に示す。データ取得日数が増加すると（10日→20日→40日）、標本データセットから算定された1,000個の標本値の分布が真値（7.6分）に近づいている。

データ取得日数を「2日/60日」～「60日/60日」として、標準偏差、95%タイル値、90%タイル値、中央値及び平均値の信頼度を算定した。横軸を「データ取得日数」、縦軸を「信頼度」としてプロットすると図-10のとおりとなる。

4. 分析結果

図-10から、90%の信頼度で指標を正しく算定するために必要なデータ取得日数を整理した(OD区間1及びOD区間2)。これを図-11に示す。

4.1 OD区間1とOD区間2の比較

必要なデータ取得日数をOD区間1とOD区間2とで比較すると、OD区間2の値が大きい（標準偏差のみ同数）。例えば、90%タイル値を90%の信頼度で正しく算定しようとする、OD区間1では15日分のOD区間データがあればよい。これに対し、OD区間2では21日分必要になる。これは、OD区間2の旅行時間のばらつきがOD区間1より大きいためである（2章参照）。

4.2 指標間の比較

5つの指標を比較すると、指標によって必要なデータ取得日数が異なる。例えば、60日のうち15日（OD区間1）又は21日（OD区間2）分のデータが取得されていれば90%タイル値を90%の信頼度で算定することができる。一方、標準偏差を算定しようとする53日分（OD区間1及びOD区間2）のデータが必要となる。

5. まとめ

プローブ旅行時間データの取得状況（60日中のデータ取得日数）と旅行時間信頼性指標値の信頼度との関係を分析し、次の知見を得た：「10日に1回程度発生する混雑時の旅行時間（90%タイル旅行時間）」を90%の信頼度で正しく算定するためには、旅行時間のばらつきの大きい区間で

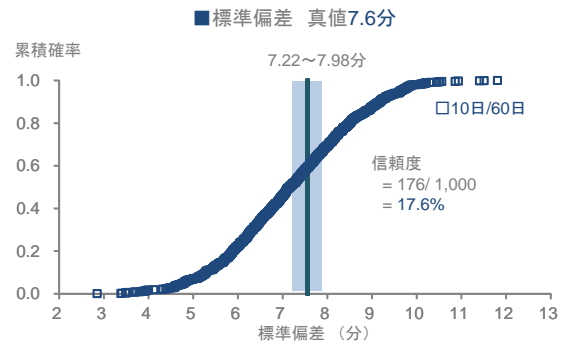


図-8 信頼度の算定（10日/60日）

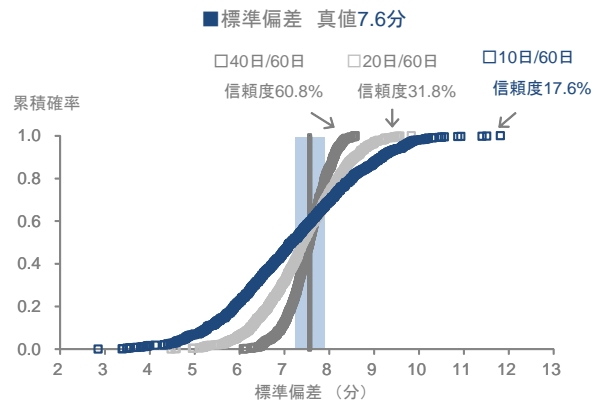


図-9 信頼度の算定（20日/60日，40日/60日）

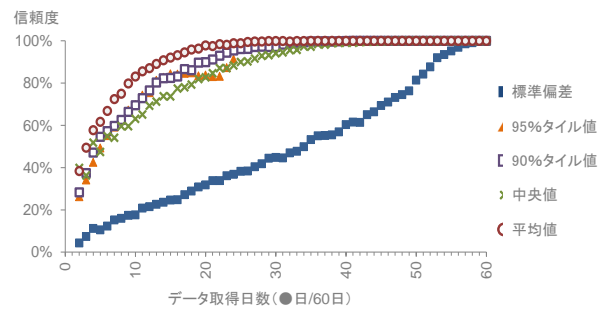


図-10 データ取得日数（2～60日）と信頼度との関係

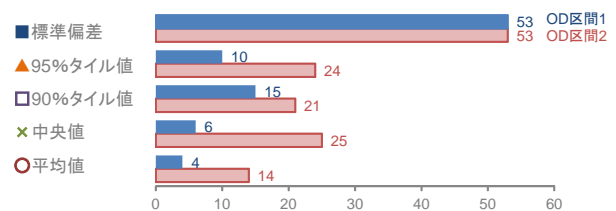


図-11 信頼度90%を確保するために必要なデータ取得日数

21日分（60日の35%）、小さい区間で15日分（60日の25%）の旅行時間データが必要である。

以下に今後の課題を示す。

（信頼度の要求水準）

指標算定においてどの程度の信頼度が必要となるかといった「要求水準」については未検討である。今後、対象とする道路の性格や指標算定の目的に応じた信頼度の要求水準についての考え方を整理する必要がある。

（分析対象のOD区間）

本研究での成果は、交通量の多い国道4号の片側2車又は3車区間を対象に行った分析結果から得られたものである。このため、必ずしも本成果を他の区間に適用できるというわけではない。今後、汎用性のある成果（データ取得状況と信頼度との関係）を得ることができるよう、異なる延長や異なる道路構造の区間（片側1車、信号交差点密度の小さい区間）においても同様の分析を行いたい。

謝 辞

本研究を進めるにあたり、「交通工学研究会『実用的な旅行時間信頼性指標に関する展開研究（代表：東京工業大学 朝倉康夫教授）』」から、丁寧かつ熱心なご指導を賜りました。ここに感謝の意を表します。

補注1 標本データセット数

3.1節では、標本データセットの数を1,000としている。以下にこの根拠を示す。データ取得日数

の組み合わせは最大で約 12×10^{16} （60日のうち30日分のデータが取得できたケースの ${}_{60}C_{30}$ ）となる。 12×10^{16} とおりの標本データセットを作成することは現実的でないことから、式(1)を用いて分析に用いる標本データセットの数を設定した²⁾。

$$e = \alpha \sqrt{((N-n)(N-1))(p(1-p)/n)} \quad \text{式(1)}$$

- n : 必要な標本データセット数
- N : 母数
- p : 標本比率
- α : 信頼度係数
- e : 標準誤差率

標本データセットから算定した旅行時間信頼性指標値が真値の一定範囲内に入るか否かを判定基準としたため、標本比率 $p=0.5$ とした。信頼度は95%や99%を用いるのが一般的であり、より厳しい信頼度99%の信頼度係数 $\alpha=2.81$ とした。標準誤差率 e は交通関連調査では20%とされることが多いが、より厳しい5%とした。母数 $N=12 \times 10^{16}$ として式(1)から標本データセット数を求めると $n=790$ となった。安全側をみて、必要な標本データセット数を1,000とした。

参考文献

- 1) 例えば、新東名(静岡県)インパクト調整会議ホームページ、
<http://www.shintomei-shizuoka.net/impact.html>.
- 2) 総務省政策統括官(統計基準担当)監修：統計実務基礎知識—平成20年度版—、(財)全国統計協会連合会、2008.

関谷浩孝



国土交通省国土技術政策総合研究所道路研究部道路研究室主任研究官
Hiroataka SEKIYA

諸田恵士



国土交通省国土技術政策総合研究所道路研究部道路研究室研究官
Keiji MOROTA

高宮 進



国土交通省国土技術政策総合研究所道路研究部道路研究室長、博士(学術)
Dr. Susumu TAKAMIYA