

# プローブデータの取得状況と旅行時間信頼性指標の算定精度との関係

関谷 浩孝<sup>1</sup>・諸田 恵士<sup>2</sup>・高宮 進<sup>3</sup>

<sup>1</sup>正会員 国土交通省国土技術政策総合研究所 道路研究室（〒305-0804 茨城県つくば市旭 1 番地）  
E-mail: sekiya-h92tb@nilim.go.jp

<sup>2</sup>非会員 国土交通省国土技術政策総合研究所 道路研究室（〒305-0804 茨城県つくば市旭 1 番地）  
E-mail: morota-k92ta@nilim.go.jp

<sup>3</sup>正会員 国土交通省国土技術政策総合研究所 道路研究室（〒305-0804 茨城県つくば市旭 1 番地）  
E-mail: takamiya-s92tc@nilim.go.jp

国道4号の2つのOD区間（約31km及び19km）を対象に、プローブデータの取得状況（データ取得日数及びDRM区間単位の欠測状況）と、これから算定される旅行時間信頼性指標の精度との関係について分析を行った。具体的には、まずランダムサンプリングにより仮想的に欠測区間のある標本OD区間データを複数パターン作成した。これから算定される旅行時間信頼性指標と真値（欠測区間のないOD区間データから算定される旅行時間信頼性指標）との差を算定した。この結果、一定の精度を確保するために必要なデータ取得日数等に関する知見を得た。

**Key Words :** travel time reliability, variability, road traffic, probe car data

## 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・時刻に関するデータを大量かつ安価に取得できるようになり、従来の旅行時間の「平均的な値」だけでなく、旅行時間の「変動やばらつき」、いわゆる旅行時間信頼性指標（以下「指標」という。）についても比較的容易に把握することが可能となった。国土交通省では、これらの位置・時刻データから得られるDRM区間<sup>123)</sup>単位の旅行時間データを用いて、道路交通施策の実施に伴う指標の変化を算定する等、施策の効果を定量的に評価する取組みを行っている。<sup>456)</sup>

ところが、図-1に示すような8つのDRM区間で構成されるOD区間におけるデータ取得状況を見ると、1台もプローブデータが取得されていないDRM区間（以下「欠測区間」という。）がある。この場合、一定の考え方で欠測区間を補正することによりOD区間の旅行時間データを作成することができる。しかし、例えば、図-1の「2012年4月9日」のようにOD区間の延長に占める欠測区間の延長の割合（以下「欠測区間長割合」という。）が大きいと、算定される指標が実態と異なる値となってしまうことがあるだろう。また、評価対象日数（例えば60日）における全ての日（60日）に旅行時間データが取

| OD区間 延長4.8km     |     |     |     |     |     |     |     |     |         |
|------------------|-----|-----|-----|-----|-----|-----|-----|-----|---------|
| DRM区間No.         | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 欠測区間長割合 |
| DRM区間延長(m)       | 614 | 487 | 750 | 524 | 564 | 583 | 630 | 622 |         |
| 2012.4.2(月)10時台  | ○   | ×   | ○   | ×   | ○   | ○   | ○   | ○   | 21%     |
| 2012.4.3(火)10時台  | ×   | ○   | ○   | ○   | ○   | ○   | ○   | ○   | 13%     |
| 2012.4.4(水)10時台  | ○   | ○   | ○   | ○   | ○   | ○   | ○   | ○   | 0%      |
| 2012.4.5(木)10時台  | ×   | ×   | ×   | ×   | ×   | ×   | ×   | ×   | 100%    |
| 2012.4.6(金)10時台  | ○   | ○   | ○   | ×   | ○   | ○   | ×   | ○   | 24%     |
| 2012.4.9(月)10時台  | ○   | ×   | ×   | ○   | ×   | ○   | ×   | ○   | 51%     |
| 2012.4.10(火)10時台 | ○   | ○   | ○   | ○   | ○   | ×   | ○   | ○   | 12%     |
| 2012.4.11(水)10時台 | ×   | ×   | ×   | ×   | ×   | ×   | ×   | ×   | 100%    |
| 2012.4.12(木)10時台 | ○   | ○   | ×   | ○   | ○   | ×   | ○   | ○   | 28%     |
| 2012.4.13(金)10時台 | ×   | ×   | ×   | ×   | ×   | ×   | ×   | ×   | 100%    |
| 2012.4.16(月)10時台 | ×   | ×   | ×   | ×   | ×   | ×   | ×   | ×   | 100%    |
| 2012.4.17(火)10時台 | ○   | ○   | ○   | ○   | ○   | ○   | ○   | ○   | 0%      |
| ⋮                | ⋮   | ⋮   | ⋮   | ⋮   | ⋮   | ⋮   | ⋮   | ⋮   | ⋮       |

○: 10時台に1台以上のプローブデータが取得されている  
×: 10時台に1台もプローブデータが取得されていない(欠測区間)

図-1 プローブデータの取得状況のイメージ

得られている場合と、一部の日（例えば、評価対象日数60日のうち10日等）のみでしか取得されていない場合とでは、算定される指標の精度は異なるだろう。このように、データの取得状況に応じて算定される指標の精度あるいは信頼度は異なる。しかし、このことは実務で指標を算定するにはほとんど考慮されていない。

橋本ら<sup>7)</sup>は、プローブデータの取得状況と指標の算定精度との関係について分析を行っている。しかし、この研究では「旅行速度」を分析対象としており、また

DRM区間を評価単位としている。そこで筆者らは、複数のDRM区間から構成されるOD区間を対象に、プローブデータの取得状況（日数及び欠損状況）と「旅行時間信頼性指標」の算定精度との関係についての知見を得ることを目的とした分析を行った。本稿では、分析方法の概要と得られた主な知見を紹介する。

## 2. 使用データ

### a) 旅行時間データ

カーナビ等での情報提供の目的で民間会社が一般道路利用者から収集したDRM区間単位の旅行時間データ（2011年10月1日から2012年9月30日の平日234日の7時台）を使用した。ただし、特異日（年末年始、ゴールデンウィーク、お盆、休日の谷間となる平日）及び悪天候の日（大雨、降雪、台風）を除いた。

### b) OD区間

交通量が多く、旅行時間データが豊富に取得されていた国道4号上りの2つのOD区間（図-2）を分析対象とした。両区間の概要を表-1に示す。

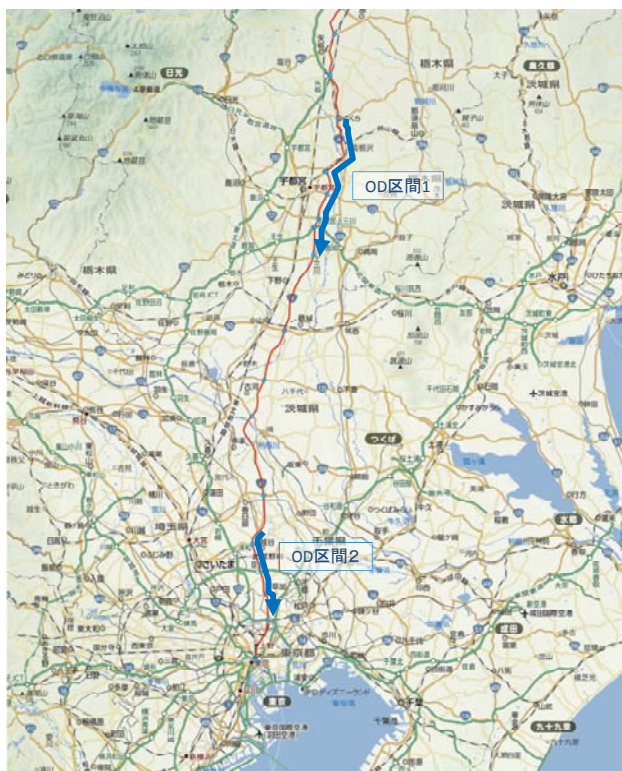


図-2 OD区間1とOD区間2の位置（国道4号上り）

表-1 OD区間1とOD区間2の概要

|         | OD区間1                    | OD区間2                   |
|---------|--------------------------|-------------------------|
| 起点      | 国道293号交差点                | 4号現道交差点                 |
| 終点      | 国道352号交差点                | 首都高中央環状                 |
| 距離 (km) | 31.0                     | 18.8                    |
| 車線数     | 4車線(約13km)<br>6車線(約17km) | 4車線(約14km)<br>6車線(約5km) |
| 信号交差点数  | 70 (2.2/km)              | 47 (2.5/km)             |
| DRM区間数  | 92 (337m/区間)             | 87 (216m/区間)            |

### c) OD区間データ

全てのDRM区間（OD区間1：92区間、OD区間2：87区間）において、同一日の7時台に旅行時間データが取得されていた日がOD区間1では215日、OD区間2では96日あった。これらの日を対象にOD区間の旅行時間（以下「OD区間データ」という）を算定した。この際、OD区間が30km程度であるため、タイムスライスを用いず、DRM区間単位の旅行時間データの単純和により算定している。

### d) OD区間データの分布

前項に示したOD区間データのうち、60日分のOD区間データについての分布及び旅行時間信頼性指標を図-3及び表-2に示す。OD区間1はOD区間2に比べ、平均旅行速度が大きく（OD区間1：38.1km/h、OD区間2：19.1km/h）、Buffer Time Index<sub>90%</sub>が小さい（OD区間1：0.09、OD区間2：0.16）。つまり、OD区間1はOD区間2より混雑の度合い及び旅行時間の変動ともに小さい。

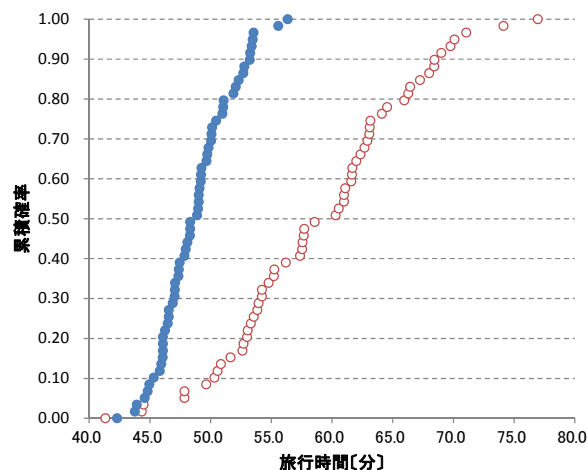


図-3 OD区間データの分布（平日7時台、60日）

表-2 旅行時間信頼性指標（平日7時台、60日）

|                                  | 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       |
| 最大値                              | 56.4分     | 77.0分       |
| 最小値                              | 42.3分     | > 41.3分     |
| 平均旅行速度                           | 38.1 km/h | > 19.1 km/h |
| Buffer Time <sub>90%</sub>       | 4.45分     | 9.49分       |
| Buffer Time Index <sub>90%</sub> | 0.09      | 0.16        |

### 3. OD区間データ取得日数と算定精度との関係

#### (1) 標本データセットの作成及び指標の算定

2章で示したOD区間データ（60日）のうち、2～59日のみでしかOD区間データが取得できなかったケースを想定し、標本データセットを作成した。この算定方法を、60日のうち10日のみでしかデータが取得できなかったケース（以下「データ取得日数10日/60日」という。）を例に示す。図4-Aに示すとおり、60日分のOD区間データの中から10日分をランダムに選定。この10個のOD区間データについての指標（■標準偏差、▲95%タイル値、□90%タイル値、×中央値及び○平均値）を算定（図4-B）。これを1,000回繰り返し（補注参照）、各指標で値の小さい順に並べると図4-Cようになる。

データ取得日数を「10日/60日」「20日/60日」及び「40日/60日」としたものを図-5に示す。ここでの真値は表-2に示す値、つまり60日分のOD区間データから算定した値とした。標本データセットから算定した指標は、データ取得日数の増加（10日→20日→40日）に伴い、ばらつきが小さくなり真値に近づいている。

#### (2) 算定精度の算定

標本データセットから算定した指標と真値との差を比較し、データ取得日数と算定精度との関係を分析した。図-6に示すように、標本データセットから算定した指標値（ $x_i$ ： $i=1\sim1,000$ ）と真値（ $X$ ）との差が真値の $\pm 5\%$ 内となる場合に「正しく算定できた」と判定し、この割合を「算定精度」とした。例えば、標本データセットから算定した90%タイル値（ $x_i^{90\%タイル値}$ ： $i=1\sim1,000$ ）1,000個のうち、真値（ $X^{90\%タイル値}=53.2$ 分）の95%～105%内（50.6分～55.9分）となるものが808個である場合、算定精度は80.8%（ $=808/1,000$ ）となる（図-6）。

データ取得日数を2～59日/60日とした場合の算定精度を図-7に示す。データ取得日数の増加（2日→60日）に伴い算定精度が向上している。

5つの指標間で算定精度を比較すると、■標準偏差の

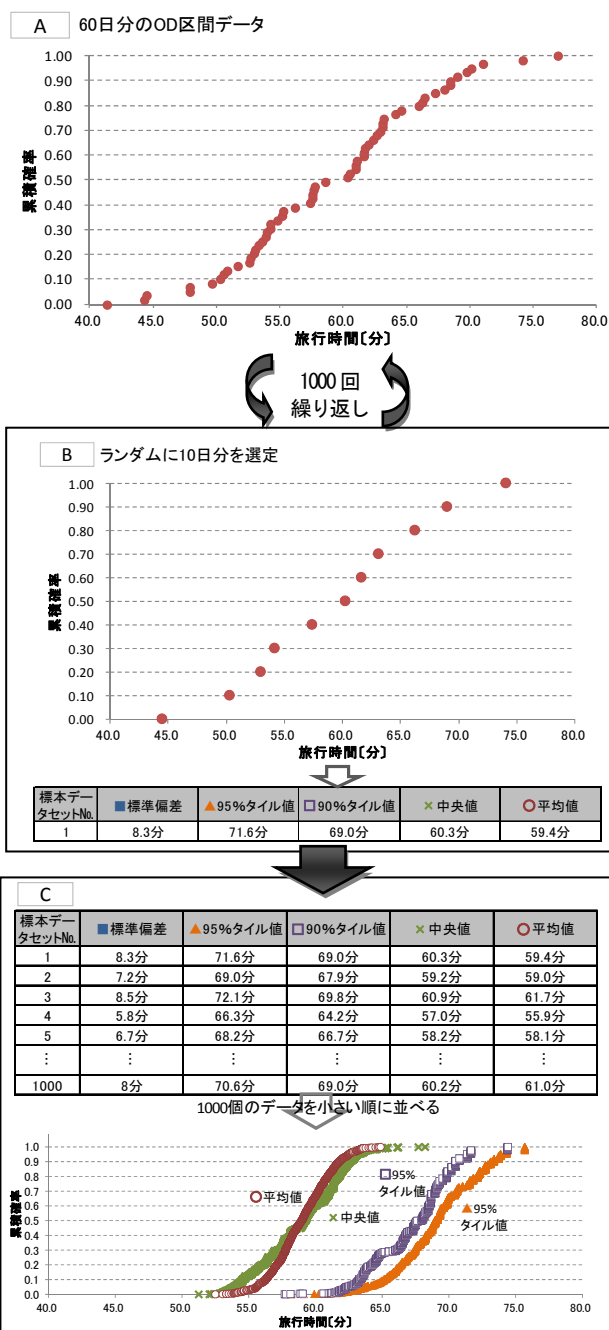


図-4 標本データセット及び指標算定のイメージ

算定精度が最も小さい。例えば、OD区間1で60日のうち30日のみでしかデータが取得できていない場合（「データ取得日数30日/60日」）でも、□90%タイル値の算定精度は99%以上である。これに対し、■標準偏差の算定精度は50%に満たない。つまり、50%以上の確率で真値との差が $\pm 5\%$ 以上となる。

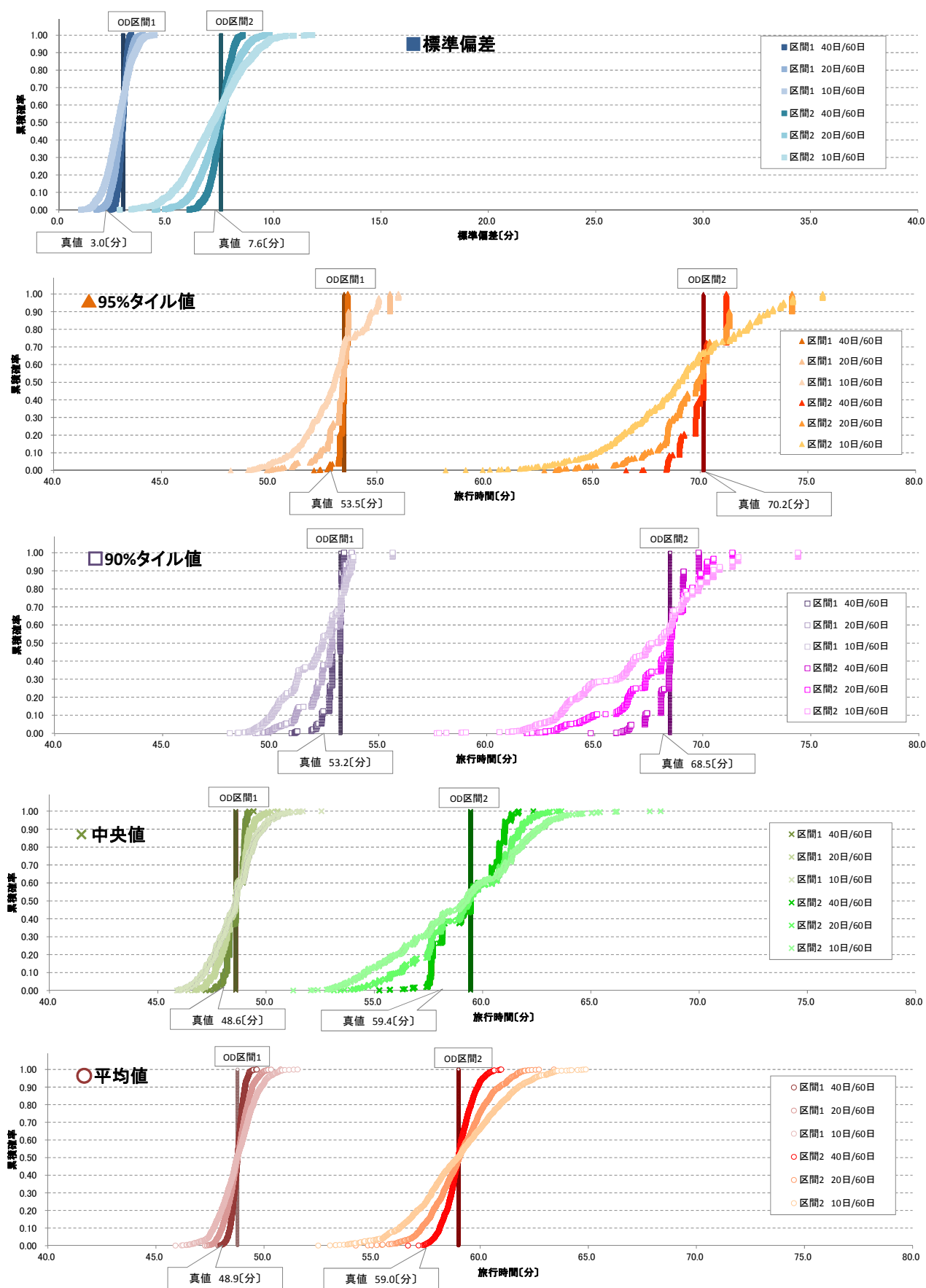


図-5 標本データセットから算定した指標の分布（欠測区間長割合 0%）

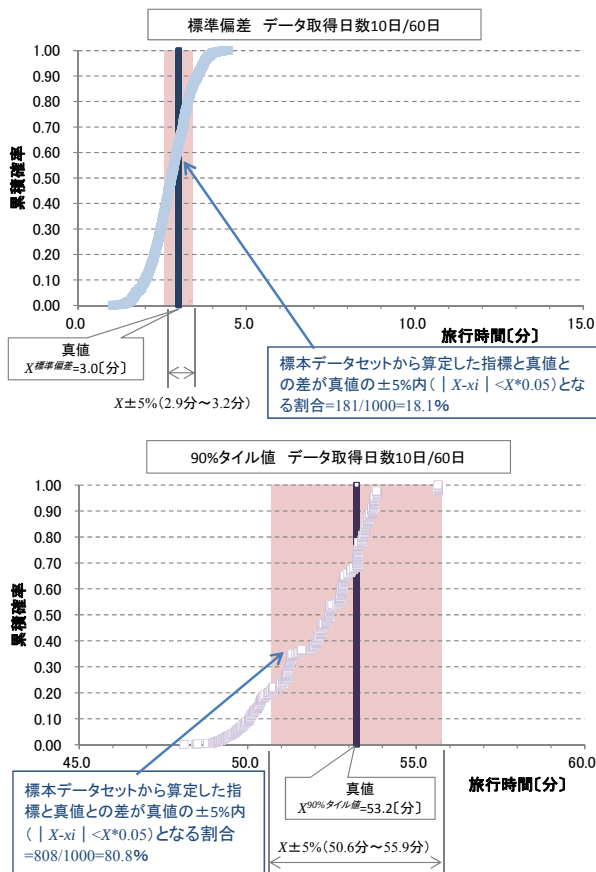


図-6 標準偏差と90%タイル値の算定精度の算定例(OD区間1)

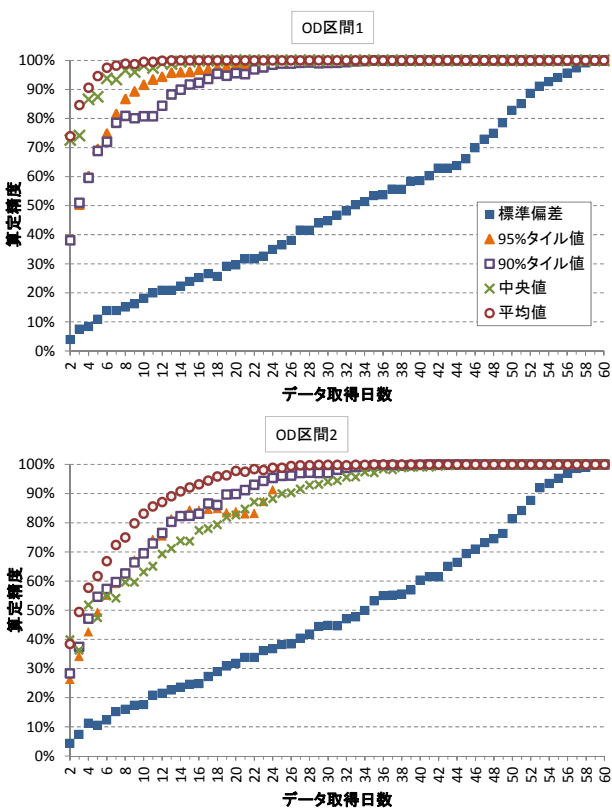


図-7 データ取得日数と算定精度との関係

表-3 算定精度を確保するために必要なデータ取得日数

| 区間    | 算定精度 | 標準偏差 | 95%タイル値 | 90%タイル値 | 中央値 | 平均値 |
|-------|------|------|---------|---------|-----|-----|
| OD区間1 | 99%  | 58日  | 22日     | 27日     | 14日 | 10日 |
|       | 95%  | 56日  | 13日     | 20日     | 8日  | 6日  |
|       | 90%  | 53日  | 10日     | 15日     | 6日  | 4日  |
| OD区間2 | 99%  | 58日  | 28日     | 33日     | 39日 | 26日 |
|       | 95%  | 55日  | 25日     | 24日     | 32日 | 18日 |
|       | 90%  | 53日  | 24日     | 21日     | 25日 | 14日 |

### (3) 一定の算定精度を確保するために必要なデータ取得日数

図-7から、99%、95%及び90%の精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。これを表-3に示す。

#### a) OD区間1と2の比較

OD区間1とOD区間2を比較すると、OD区間2の値が大きい（標準偏差を除く）。例えば、90%タイル値を90%の精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が、OD区間2では21日分必要になる。これは、2章で示したようにOD区間2の旅行時間変動がOD区間1より大きいためであると考ええる。

#### b) 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90%タイル値は、60日のうち15日（OD区間1）または21日（OD区間2）でデータが取得されていれば90%の精度で正しく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（OD区間1及びOD区間2）でデータが取得されている必要がある。

## 4. DRM区間の欠測状況を考慮した算定精度

3章では、OD区間データの取得日数と指標の算定精度との関係进行分析した。ただし、ここでのOD区間データは、OD区間内の全てのDRM区間において旅行時間データが取得されていることを想定したものであった。これに対し4章では、仮想的に欠測区間を含むOD区間データを作成し、DRM区間の欠測状況を考慮した算定精度の分析を行う。

### (1) 標本データセットの作成

3章のOD区間データに対し、データ取得日数（2～59日）毎に欠測区間長割合（10、20、30、40、50%）を設定し、データ取得日数別・欠測区間長割合別の標本データセットを作成した。図-9に示すDRM区間s1～s12を対象とした標本データセットの作成例を示す。

#### a) ランダムサンプリング

設定したデータ取得日数分のOD区間データをランダムに抽出(図-8①)。これらの抽出されたOD区間データに

対し、欠測区間長割合（10, 20, 30, 40, 50%）に応じて、仮想的に欠測とする区間をランダムに選定(図-8②)．選定した欠測区間は旅行時間を「0秒」としてOD区間データ（補正前）を算定(図-8③)．以上のランダムサンプリングをデータ取得日毎・欠測区間割合毎に1,000回繰り返す、1,000通り（補注参照）の欠測区間を含むOD区間データを作成した．

①データ取得日をランダムにサンプリング  
例えばデータ取得日数を6日とした場合、6日分をランダムに抽出( 〇 で示した日)

| データ取得日D <sub>j</sub> | s1   | s2   | s3   | s4   | s5   | s6   | s7   | s8   | s9   | s10  | s11  | s12  | 合計    |
|----------------------|------|------|------|------|------|------|------|------|------|------|------|------|-------|
| 20120402             | 27.8 | 3.7  | 13.2 | 1.3  | 4.6  | 43.0 | 15.0 | 31.2 | 1.0  | 21.6 | 5.0  | 13.8 | 181.2 |
| 20120403             | 27.0 | 4.0  | 13.2 | 2.0  | 4.0  | 30.0 | 14.8 | 31.8 | 1.0  | 12.7 | 13.2 | 13.0 | 166.7 |
| 20120404             | 29.8 | 4.0  | 13.0 | 1.5  | 3.8  | 22.3 | 15.0 | 30.4 | 1.0  | 15.6 | 7.2  | 14.2 | 157.6 |
| 20120405             | 25.3 | 3.5  | 11.3 | 1.3  | 3.8  | 24.8 | 14.3 | 30.0 | 1.0  | 26.2 | 15.7 | 14.3 | 171.2 |
| 20120406             | 25.4 | 12.2 | 15.0 | 1.3  | 4.0  | 17.2 | 13.7 | 31.3 | 2.2  | 24.4 | 5.8  | 15.1 | 167.4 |
| 20120409             | 26.4 | 4.0  | 12.5 | 2.0  | 3.8  | 23.1 | 13.7 | 30.5 | 1.0  | 20.0 | 13.8 | 19.1 | 169.8 |
| 20120410             | 25.6 | 3.6  | 11.3 | 2.2  | 3.3  | 20.8 | 13.0 | 28.9 | 1.3  | 13.5 | 23.0 | 15.2 | 161.6 |
| 20120411             | 28.4 | 5.4  | 13.4 | 2.6  | 4.3  | 23.1 | 14.6 | 34.4 | 1.5  | 19.3 | 21.6 | 17.9 | 186.4 |
| 20120412             | 28.0 | 5.0  | 14.0 | 1.6  | 4.1  | 30.3 | 15.0 | 32.7 | 1.1  | 15.4 | 6.1  | 14.6 | 168.0 |
| 20120413             | 27.5 | 12.6 | 12.4 | 2.0  | 4.0  | 30.8 | 15.2 | 30.5 | 1.0  | 20.1 | 17.8 | 16.2 | 190.0 |
| 20120416             | 26.1 | 10.6 | 12.4 | 1.0  | 3.8  | 20.9 | 16.0 | 31.9 | 1.4  | 16.2 | 10.7 | 15.1 | 166.1 |
| 20120417             | 27.6 | 4.0  | 12.7 | 1.3  | 4.2  | 27.3 | 13.8 | 30.6 | 1.5  | 21.3 | 9.4  | 15.3 | 168.8 |
| 20120418             | 26.6 | 3.7  | 12.5 | 1.3  | 4.7  | 30.7 | 15.2 | 32.4 | 1.0  | 25.1 | 22.3 | 15.3 | 190.9 |
| 20120419             | 27.4 | 4.5  | 12.7 | 1.6  | 4.3  | 20.0 | 13.6 | 29.6 | 1.0  | 19.9 | 15.6 | 15.4 | 165.4 |
| 20120420             | 27.3 | 12.2 | 12.0 | 1.0  | 3.8  | 28.8 | 14.2 | 28.2 | 2.2  | 13.5 | 13.8 | 13.5 | 170.5 |
| 20120422             | 27.5 | 8.7  | 13.9 | 1.3  | 3.9  | 20.4 | 13.9 | 31.1 | 1.0  | 17.0 | 13.1 | 16.4 | 169.0 |
| 20120424             | 28.6 | 11.8 | 13.6 | 1.7  | 4.0  | 24.0 | 13.5 | 29.2 | 1.0  | 15.0 | 21.8 | 14.8 | 179.0 |
| 20120425             | 36.7 | 9.4  | 14.1 | 1.1  | 4.4  | 31.4 | 13.9 | 28.6 | 1.0  | 17.6 | 4.1  | 13.1 | 175.4 |
| 20120426             | 28.1 | 4.0  | 12.6 | 1.8  | 4.0  | 23.8 | 14.3 | 29.8 | 1.3  | 16.6 | 17.8 | 15.4 | 169.4 |
| 20120427             | 28.3 | 4.0  | 14.9 | 1.0  | 4.4  | 32.6 | 17.0 | 32.8 | 1.0  | 14.0 | 15.1 | 15.6 | 180.6 |
| 平均旅行時間(秒)            | 27.8 | 6.5  | 13.0 | 1.5  | 4.0  | 26.3 | 14.5 | 30.8 | 1.2  | 18.2 | 13.6 | 15.2 | 172.7 |
| 時間割合                 | 0.16 | 0.04 | 0.08 | 0.01 | 0.02 | 0.15 | 0.08 | 0.18 | 0.01 | 0.11 | 0.08 | 0.09 | 1.00  |
| OD区間延長(m)            | 433  | 70   | 222  | 28   | 70   | 250  | 240  | 587  | 14   | 284  | 71   | 269  | 2538  |
| 延長割合                 | 0.17 | 0.03 | 0.09 | 0.01 | 0.03 | 0.10 | 0.09 | 0.23 | 0.01 | 0.11 | 0.03 | 0.11 | 1.00  |

②欠測区間をランダムにサンプリング  
例えば欠測区間長割合を20%とした場合、延長割合20%分の欠測区間をランダムに選定( 〇 で示した日)

| データ取得日D <sub>j</sub> | s1   | s2   | s3   | s4  | s5  | s6   | s7   | s8   | s9  | s10  | s11  | s12  | OD区間データ(補正前)T <sub>Dj</sub> (秒) |
|----------------------|------|------|------|-----|-----|------|------|------|-----|------|------|------|---------------------------------|
| 20120404             | 0.0  | 4.0  | 13.0 | 1.5 | 3.8 | 22.3 | 15.0 | 30.4 | 1.0 | 15.6 | 0.0  | 14.2 | 120.7                           |
| 20120406             | 25.4 | 12.2 | 0.0  | 0.0 | 4.0 | 0.0  | 13.7 | 31.3 | 0.0 | 24.4 | 5.8  | 15.1 | 131.9                           |
| 20120411             | 28.4 | 0.0  | 13.4 | 2.6 | 0.0 | 23.1 | 14.6 | 34.4 | 0.0 | 0.0  | 21.6 | 17.9 | 155.9                           |
| 20120417             | 27.6 | 0.0  | 12.7 | 1.3 | 0.0 | 27.3 | 0.0  | 30.6 | 1.5 | 21.3 | 0.0  | 15.3 | 137.5                           |
| 20120423             | 27.5 | 8.7  | 0.0  | 1.3 | 3.9 | 0.0  | 13.9 | 31.1 | 1.0 | 17.0 | 13.1 | 16.4 | 133.7                           |
| 20120425             | 36.7 | 0.0  | 14.1 | 0.0 | 4.4 | 31.4 | 13.9 | 28.6 | 1.0 | 17.6 | 4.1  | 0.0  | 151.7                           |

図-8 データ取得日・欠測区間のランダムサンプリングの例

## b) 欠測区間の補正

a)において作成した欠測区間を含むOD区間データについて、式(1)を用いて補正を行い、標本データセットを作成した．

$$T_{Dj}^{7:00-8:00} = \frac{\hat{T}_{Dj}^{7:00-8:00}}{1 - P\hat{T}_{Dj}^{7:00-8:00}} \quad (1)$$

D<sub>j</sub>: データ取得日 (j=1~m) (m: データ取得日数)

T<sub>Dj</sub>: 第 D<sub>j</sub> 日における OD 区間データ (補正後) (1 時間値, 7:00-8:00)

T<sub>Dj</sub><sup>ˆ</sup>: 第 D<sub>j</sub> 日における OD 区間データ (補正前) (1 時間値, 7:00-8:00)

PT<sub>Dj</sub><sup>ˆ</sup>: 第 D<sub>j</sub> 日における欠測区間時間割合 (7:00-8:00)  
第 D<sub>j</sub> 日において、欠測区間の平均旅行時間が OD 区間データに占める割合の合計値 (図-9) ．

| データ取得日D <sub>j</sub> | OD区間データの旅行時間データ(秒) |      |      |      |      |       |       |       |       |       |       |       | OD区間データ(補正前)T <sub>Dj</sub> (秒) | 欠測区間時間割合PT <sub>Dj</sub> <sup>ˆ</sup> | OD区間データ(補正後)T <sub>Dj</sub> (秒) |
|----------------------|--------------------|------|------|------|------|-------|-------|-------|-------|-------|-------|-------|---------------------------------|---------------------------------------|---------------------------------|
| 20120404             | 0.0                | 4.0  | 13.0 | 1.5  | 3.8  | 22.3  | 15.0  | 30.4  | 1.0   | 15.6  | 0.0   | 14.2  | 120.7                           | 23.9%                                 | 158.6                           |
| 20120406             | 25.4               | 12.2 | 0.0  | 0.0  | 4.0  | 0.0   | 13.7  | 31.3  | 0.0   | 24.4  | 5.8   | 15.1  | 131.9                           | 24.3%                                 | 174.3                           |
| 20120411             | 28.4               | 0.0  | 13.4 | 2.6  | 0.0  | 23.1  | 14.6  | 34.4  | 0.0   | 0.0   | 21.6  | 17.9  | 155.9                           | 17.4%                                 | 188.8                           |
| 20120417             | 27.6               | 0.0  | 12.7 | 1.3  | 0.0  | 27.3  | 0.0   | 30.6  | 1.5   | 21.3  | 0.0   | 15.3  | 137.5                           | 22.4%                                 | 177.3                           |
| 20120423             | 27.5               | 8.7  | 0.0  | 1.3  | 3.9  | 0.0   | 13.9  | 31.1  | 1.0   | 17.0  | 13.1  | 16.4  | 133.7                           | 22.7%                                 | 173.1                           |
| 20120425             | 36.7               | 0.0  | 14.1 | 0.0  | 4.4  | 31.4  | 13.9  | 28.6  | 1.0   | 17.6  | 4.1   | 0.0   | 151.7                           | 13.5%                                 | 175.3                           |
| 平均旅行時間(秒)            | 27.8               | 6.5  | 13.0 | 1.5  | 4.0  | 26.3  | 14.5  | 30.8  | 1.2   | 18.2  | 13.6  | 15.2  | 172.7                           |                                       |                                 |
| 時間割合                 | 16.1%              | 3.8% | 7.5% | 0.9% | 2.3% | 15.2% | 8.4%  | 17.8% | 0.7%  | 10.6% | 7.9%  | 8.8%  | 100%                            |                                       |                                 |
| OD区間延長(m)            | 433                | 70   | 222  | 28   | 70   | 250   | 240   | 587   | 14    | 284   | 71    | 269   | 2538                            |                                       |                                 |
| 延長割合                 | 17.1%              | 2.8% | 8.7% | 1.1% | 2.8% | 9.9%  | 23.1% | 0.6%  | 11.2% | 2.8%  | 10.6% | 10.0% |                                 |                                       |                                 |

①欠測区間時間割合を算定  
2012/4/25の欠測区間旅行時間割合PT<sub>Dj</sub><sup>ˆ</sup>:  
3.8% + 0.9% + 8.8% =13.5%  
欠測区間s2の 時間割合 欠測区間s4の 時間割合 欠測区間s12の 時間割合

②式(1)を用いて補正を行い、標本データセットを作成  
151.7 / 1-0.135 =175.3秒

図-9 欠測区間を含むOD区間データの補正の例

## (2) 算定精度

(1)で作成した標本データセットから算定した指標について、3章(2)と同様に算定精度を算定した．ここでは、「正しく算定できた」と判定する範囲を「真値の±5%内」と「真値の±2.5%内」とした場合の2ケース行った．

データ取得日数と90%タイル値の算定精度との関係を図-10及び図-11に示す．欠測区間長割合が増加すると(0%→10%→20%→30%→40%→50%)，算定精度は低下している．OD区間1とOD区間2でこの低下度合いを比較すると，OD区間2の方が大きい．これは，OD区間2の旅行時間の変動が大きい(2章参照)，前節で示した方法で欠測区間を補正すると，真値との差が大きくなってしまったためであると考えられる．特にOD区間2で欠測区間長割合が50%になると，60日全てで旅行時間データが取得できていた場合でも，算定精度(±2.5%)は70%程度である．つまり，30%の確率で真値(X<sup>90%タイル値</sup>=68.5分)の±2.5%内の値として指標を算定することができない．

また，欠測区間長割合が大きくなると，一定の精度を確保するために必要なデータ取得日数は増加する．例えば，90%タイル値を90%の精度で算定するために必要なデータ取得日数は次のとおり．欠測区間長割合 0% : 32日，10% : 37日，20% : 51日 (図-9, OD区間2) ．

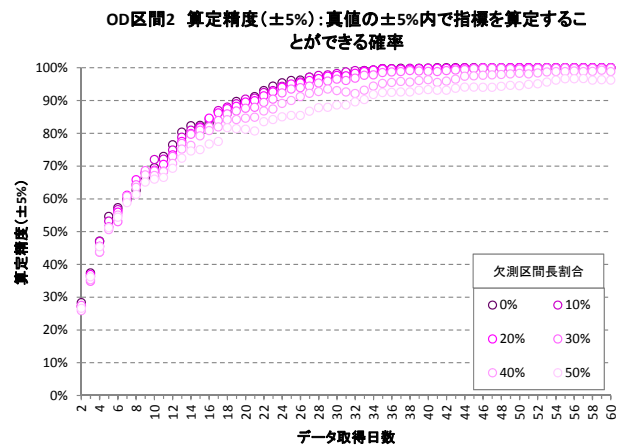
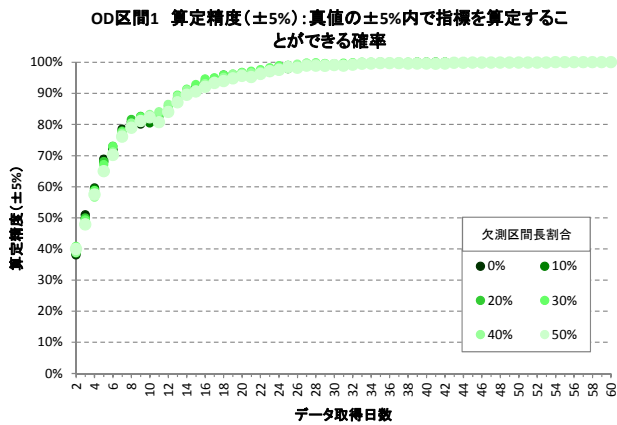


図-10 データ取得日数と90%タイル値の算定精度(±5%)との関係

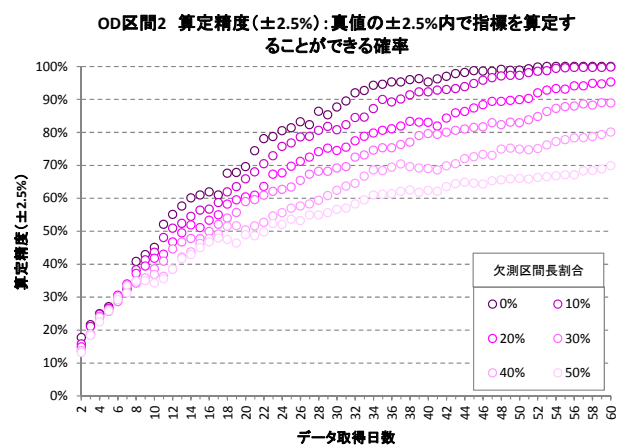
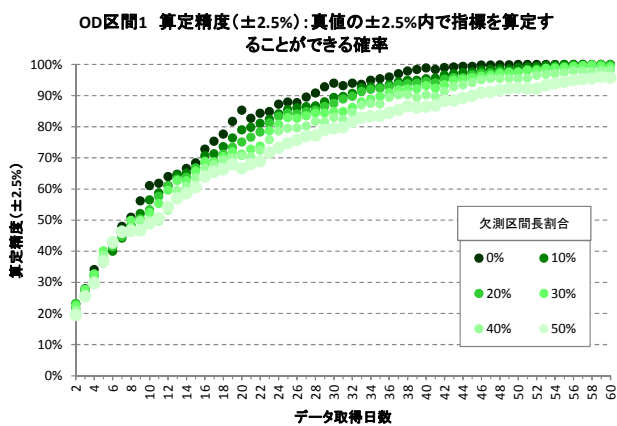


図-11 データ取得日数と90%タイル値の算定精度(±2.5%)との関係

## 5. 評価対象日数についての分析

3章及び4章では、評価対象日数を60日とした場合の分析結果を示した。本章では、評価対象日数を20日及び96日としたケースについて同様の分析を行い、結果を比較する。

### (1) 評価対象日数20日、96日のOD区間データの分布

2章で示した96日分のOD区間2のOD区間データから、評価対象日数を20日及び96日として指標を算定すると図-12及び表-4のとおりとなる。

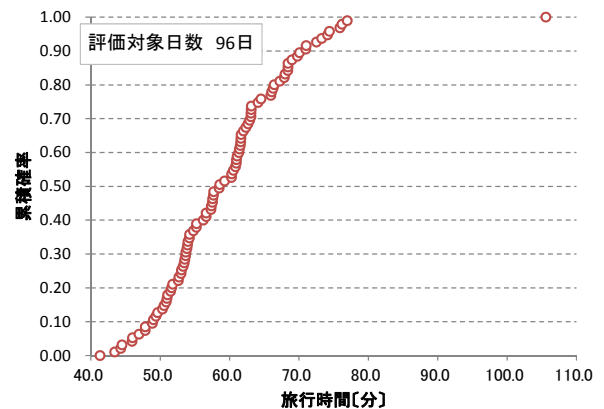
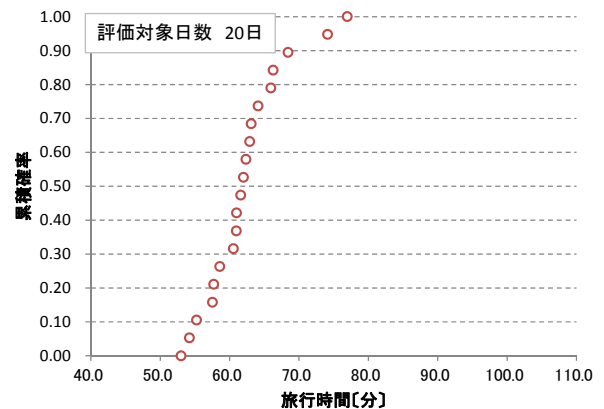


図-12 評価対象日数20日及び96日のOD区間データの分布 (OD区間2)

表-4 評価対象日数20日及び96日の概要 (OD区間2)

|                                  | 20日      | 96日      |
|----------------------------------|----------|----------|
| ■標準偏差                            | 6.0分     | 9.5分     |
| ▲95%タイル値                         | 74.3分    | 74.2分    |
| □90%タイル値                         | 69.0分    | 70.6分    |
| ×中央値                             | 61.8分    | 58.5分    |
| ○平均値                             | 62.3分    | 59.4分    |
| 最大値                              | 77.0分    | 105.6分   |
| 最小値                              | 53.0分    | 41.3分    |
| 平均旅行速度                           | 18.1km/h | 19.0km/h |
| Buffer Time <sub>90%</sub>       | 6.7分     | 11.2分    |
| Buffer Time Index <sub>90%</sub> | 0.11     | 0.19     |

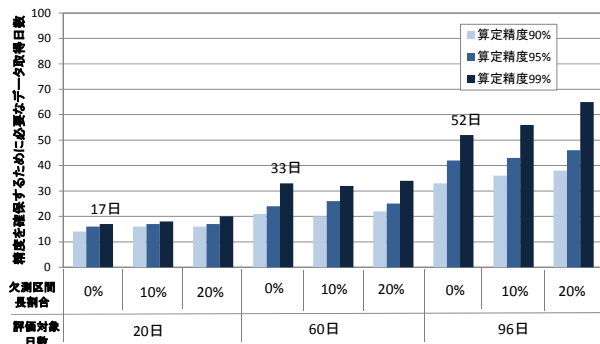


図-13 算定精度を確保するために必要なデータ取得日数

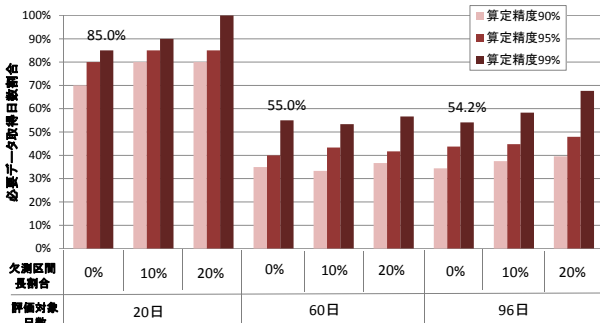


図-14 必要データ取得日数割合(算定精度を確保するために必要なデータ取得日数が評価対象日数に示す割合)(OD区間2)

## (2) 必要データ取得日数割合

(1)で示したOD区間データから、前章と同様に一定の精度を確保するために必要となるデータ取得日数を算定した。また、この日数が評価対象日数に占める割合（以下「必要データ取得日数割合」という。）を算定した。これらを図-13及び図-14に示す。

評価対象日数が増加すると（20日→60日→96日）、一定の精度を確保するために必要なデータ取得日数は増加するが、必要データ取得日数割合は減少している。例えば、欠測区間長割合が0%の場合、99%以上の算定精度を確保するための必要データ取得日数割合は、評価対象日数20日では85.0%（17日）、60日では55.0%（33日）、96日では54.2%（52日）である。

## (3) 近似式

任意の評価対象日数（20日～96日）においても、一定の算定精度を確保するために必要なデータ取得日数を把握できるよう近似式を推計した。図-15は、「評価対象日数」と「90%タイル値を99%の算定精度で算定するために必要なデータ取得日数（OD区間2、欠測区間長割合0%）」との関係を表す近似式である。前項で示したように、評価対象日数が増加すると必要なデータ取得日数の増加量は低減する。このため線形近似に加え、上に凸となる二次多項式、対数式及び累乗式を推計に用いた。全ての近似式で決定係数は0.9以上となり、高い精度で推計されている。なお、これらの推計式は、必要な

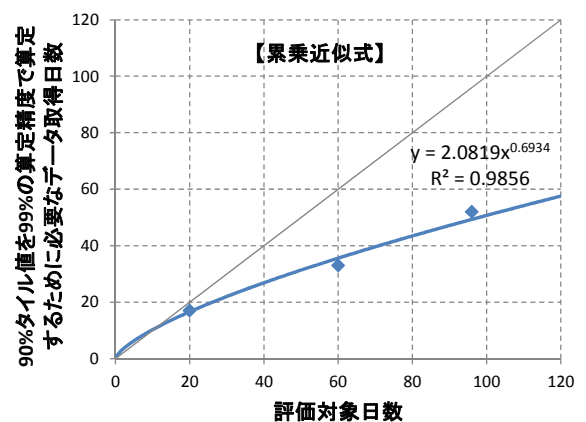
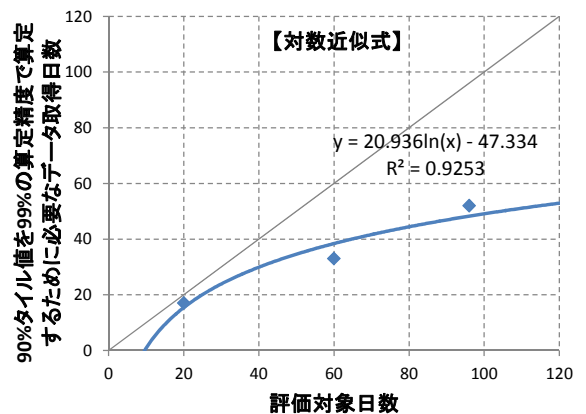
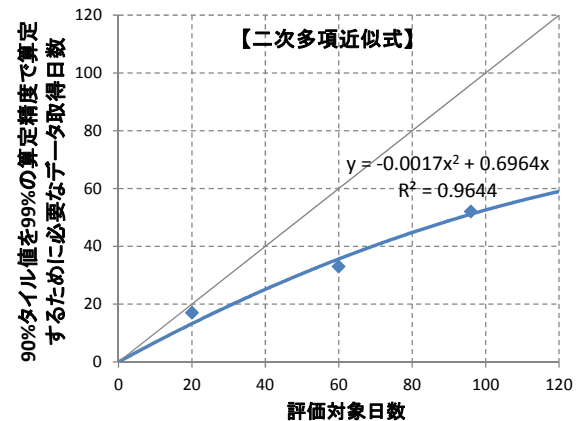
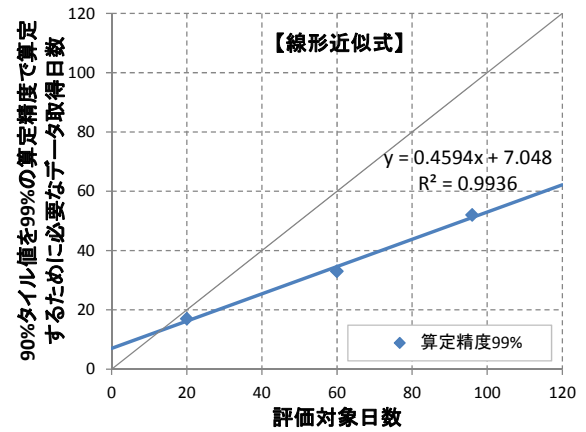


図-15 「評価対象日数」と「90%タイル値を99%の算定精度で算定するために必要なデータ取得日数（OD区間2、欠測区間長割合0%）」との関係を表す近似式

データ取得日数が評価対象日数を上回る事等があるため評価対象日数を 20 日～100 日程度とした場合に適用可能である。

## 6. 考察

(算定精度の評価方法)

本研究では「標本データセットから算定した指標値と真値との差が真値の±5%内等となる確率」をもって算定精度を表した。しかし、算定精度を評価する方法は他にも多数ある。例えば、図-16及び図-17に示す平均絶対誤差率 (MAPE) (式(2))、二乗平均誤差 (RMSE) (式(3)) がある。今後、これらの複数の評価方法を組み合わせ、多角的に算定精度を評価する方法を検討する必要がある。

平均絶対誤差率 (MAPE)

$$MAPE = \frac{\sum_{i=1}^N |X - x_i| / N}{X} \quad (2)$$

二乗平均誤差 (RMSE)

$$RMSE = \sqrt{\sum_{i=1}^N (X - x_i)^2 / N} \quad (3)$$

$X$ : 真値

$x_i$ : OD区間データ

$N$ : サンプル数 (1,000)

(算定精度の要求水準)

プローブデータの取得状況と算定精度との関係进行分析することにより、一定の精度を確保するために必要なデータ取得日数についての知見を得た。しかし、指標算定においてどの程度の精度が必要となるかといった「精度の要求水準」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精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車または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、この成果を他の区間にも適用できるわけではない。今後、汎用性のある成果（データ取得状況と算定精度との関係）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。

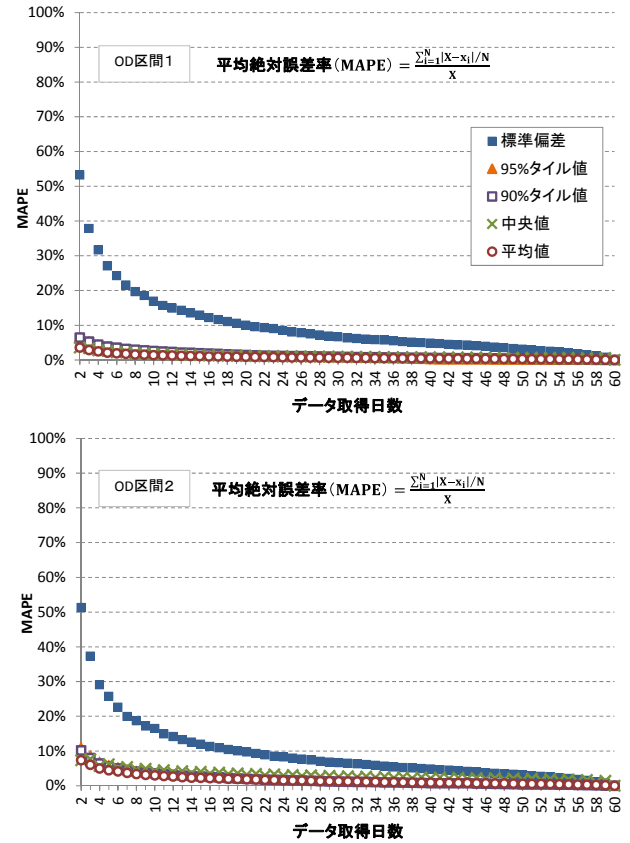


図-16 データ取得日数と平均絶対誤差率 (MAPE) との関係

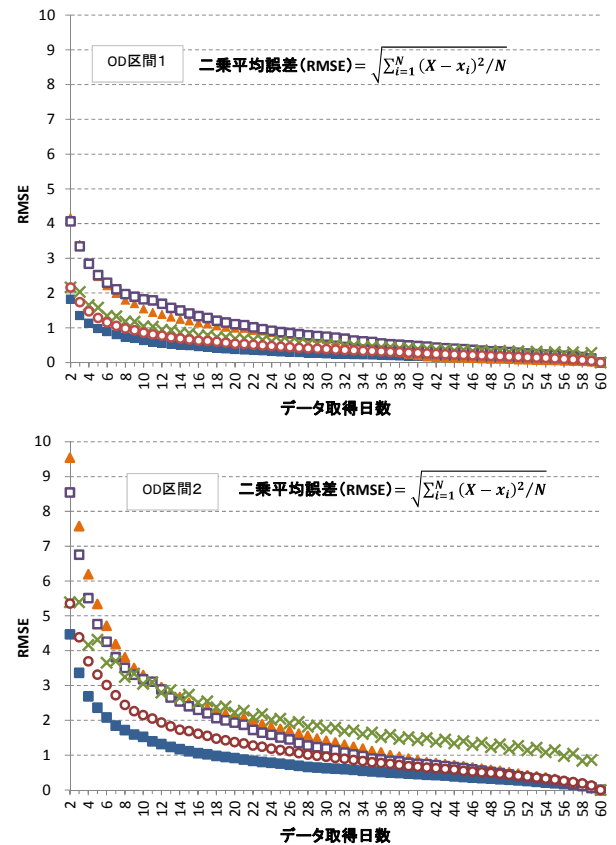


図-17 データ取得日数と二乗平均誤差 (RMSE) との関係

## 7. まとめ

プローブデータの取得状況（データ取得日数及びDRM区間単位の欠測状況）と旅行時間信頼性指標の算定精度との関係を分析した。この結果、一定の精度を確保するために必要な旅行時間データ数等について、次の知見を得た：

- ・一定の精度を確保するために必要なデータ取得日数は、旅行時間変動の大きな区間で多くなる。例えば、90%の精度で90%タイル値を算定しようとする際に必要なデータ取得日数（評価対象日数：60日）は次のとおり。  
OD区間1（旅行時間変動 小）：15日  
OD区間2（旅行時間変動 大）：21日
- ・指標によって、一定の精度を確保するために必要なデータ取得日数は異なる。例えば、90%の精度で標準偏差及び90%タイル値を算定しようとする際に必要なデータ取得日数（評価対象日数：60日）は次のとおり。  
標準偏差：53日（OD区間1及びOD区間2）  
90%タイル値：15日（OD区間1）、21日（OD区間2）
- ・欠測区間長割合（旅行時間データが取得されていないDRM区間の延長がOD区間の延長に占める割合）が大きくなると、一定の精度を確保するために必要なデータ取得日数は増加する。例えば、90%タイル値を90%の精度で算定するために必要なデータ取得日数（評価対象日数：60日、OD区間2）は次のとおり。  
欠測区間長割合 0%：32日、10%：37日、20%：51日。
- ・評価対象日数が増加すると（20日→60日→100日）、必要データ取得日数割合（一定の精度を確保するために必要なデータ取得日数が評価対象日数に占める割合）は減少する。例えば、90%タイル値を99%の精度で算定するために必要なデータ取得日数は次のとおり。  
評価対象期間 20日：85%（17日）、60日：55%（33日）、96日：54%（52日）。

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 サンプル数の考え方

算定精度の検討に用いた標本データセットのサンプル数の考え方を示す。データ取得日数の組み合わせは最大で約12京（60日のうち30日のみでしかデータが取得されていなかったケースの $6C_{30}$ ）となる。このため、次の式(4)を用いてサンプル数を設定した<sup>8)</sup>。

$$e = \alpha \sqrt{((N-n)(N-1))(p(1-p)/n)} \quad (4)$$

$n$  : 必要なサンプル数

$N$  : 母数

$p$  : 標本比率

$\alpha$  : 信頼度係数,

$e$  : 標準誤差率 $e$

算定精度の算定にあたり、標本データセットから算定した指標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、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 $N$ は、上述の12京とした。式(4)に各値を代入すると $N=790$ となる。安全側をみて、必要なサンプル数を1,000とした。

## 参考文献

- 1) 財団法人日本デジタル道路地図協会：全国デジタル道路地図データベース標準，第3.8版，2008。
- 2) 財団法人日本デジタル道路地図協会ホームページ，<http://www.drm.jp/database/structure.html>
- 3) 上坂克巳，門間俊幸，橋本浩良，松本俊輔，大脇鉄也：道路交通調査の新たな展開，土木計画学研究・論文集 Vol. 43，2011。
- 4) 関谷浩孝，上坂克巳，諸田恵士：プローブデータを用いた一般幹線道路における旅行時間信頼性指標の算定方法，土木計画学研究・論文集，Vol. 45，2012。
- 5) 諸田恵士，関谷浩孝，上坂克巳：旅行時間変動に影響を与える要因及び旅行時間信頼性指標の推計，土木計画学研究・講演集，Vol. 45，2012。
- 6) 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>
- 7) 橋本浩良，河野友彦，門間俊幸，上坂克巳：一般車プローブデータの集計対象期間と旅行速度の推計精度の関係分析，土木計画学研究・講演集，Vol. 42，2010。
- 8) 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。

(2013.5.7 受付)

## RELATION BETWEEN THE AMOUNT OF PROBE DATA AND THE ACCURACY OF TRAVEL TIME RELIABILITY INDICES -HOW DOES THE PARTIAL DEFICIT IN PROBE DATA INFLUENCE THE ACCURACY?-

Hiroataka SEKIYA, Keiji MOROTA and Susumu TAKAMIYA