

# 混合円周分布のパラメータを用いた 交通量の時間変動の分析

長崎 滉大<sup>1</sup>・中西 航<sup>2</sup>・朝倉 康夫<sup>3</sup>

<sup>1</sup>学生会員 東京工業大学 環境・社会理工学院 (〒152-8552 目黒区大岡山 2-12-1)  
E-mail: k.nagasaki@plan.cv.titech.ac.jp

<sup>2</sup>正会員 東京工業大学助教 環境・社会理工学院 (〒152-8552 目黒区大岡山 2-12-1)  
E-mail: nakanishi@plan.cv.titech.ac.jp

<sup>3</sup>正会員 東京工業大学教授 環境・社会理工学院 (〒152-8552 目黒区大岡山 2-12-1)  
E-mail: asakura@plan.cv.titech.ac.jp

道路の一日の中の交通量は時間によって変動する。その変動をノンパラメトリックなモデルを用いて分析する研究は多いが、分布を仮定して確率密度関数のパラメータの推定をし、それによって時間変動を分析する研究は少ない。また、交通量の時間変動のように24時間単位で繰り返されているデータに対しては、通常の統計処理を行うのではなく、角度を扱う方向統計学を適用し、0時と24時を同一と見なして分析する必要がある。本研究では、方向統計学における確率分布である円周分布を混合した分布を用いて、一日の交通量の分布のパラメータ推定を行った。このことにより、朝の出勤と夕方の帰宅のピークを分離することに成功した。その結果、追越車線、走行車線にかかわらず、前者のピークは左に歪んだ立ち上がりの早いピークであり、後者のピークは緩やかなピークであることがわかった。

**Key Words :** *traffic volume, parameter estimation, directional statistics, circular distribution*

## 1. はじめに

道路の交通量は最も観測しやすい交通状態のひとつである。その交通量は季節や曜日、一日の中の時間などの様々な要因で変動している<sup>1)</sup>。例えば、観光地周辺の道路ではその観光地が人気な期間の交通量が多く、都市部の道路では通勤の影響で平日の交通量が多い。一日の中の時間変動という点では、朝の通勤と夜の帰宅の交通量が多く、深夜の交通量は少ないといった特徴がある(以後、観測された単位時間あたりの通過台数の一日の中での変動を交通量の時間変動と呼ぶ)。このような事実を踏まえて交通量の予想や分析をする研究は多く存在する<sup>2,3,4)</sup>。しかし、交通量の時間変動について、確率分布を仮定してパラメータを推定する研究は少ない。

交通量の時間変動のように24時間単位で繰り返されているデータに対しては、通常の統計処理では不具合が生じてしまう。例えば、24時間全体で23:59の2分後を0:01として処理するのは、両端が接続していない通常の統計処理では不可能である。この問題を解決するために、角度を扱う方向統計学を適用して、時刻で記述されるデータを角度に変換し処理する必要がある。しかし、交通量

の時間変動の分析に方向統計学の知見を用いている研究も非常に少ない。さらに、都市部の高速道路のような朝の交通と夕方の交通が卓越していると考えられる交通量の時間変動に対して、混合分布を用いてパラメータ推定をしている研究は著者らの知る限り存在しない。

以上のことを踏まえて、本研究では交通量の時間変動に対して確率分布のパラメータ推定を行う。この時、角度を扱う方向統計学を適用して分析する。さらに、複数のピークを持つと考えられる交通量の時間変動に対して、混合分布を用いてパラメータ推定を行う。このことにより、方向統計学の確率分布である円周分布(後述)を混合した分布を用いて交通量の時間変動を分析することを本研究の目的とする。

## 2. データ

本研究で用いるデータは2016年6月6日から7月7日、8月22日から10月10日までに阪神高速3号神戸線の上り線の魚崎-尼崎東間の複数箇所のトラフィックカウンターで観測されたデータである。該当区間は走行車線と追越

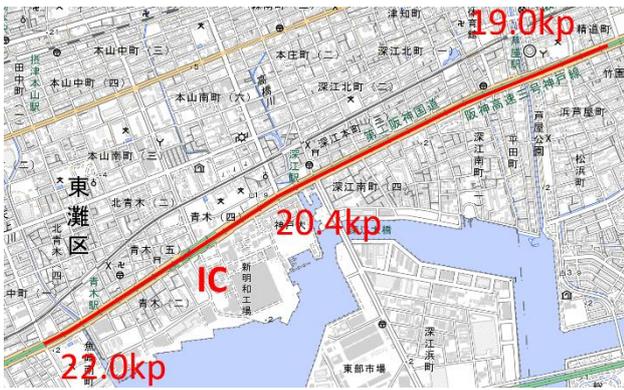


図-1 交通量データの取得範囲(地理院地図(国土地理院)に著者が加筆して作成)

車線の二車線であり、18.0 kpから22.3 kpの区間にある10地点計17車線のデータを取得した。本研究では、同地点で2車線同時に観測している19.0 kpから22.0 kpの区間にある両車線のデータを用いる。21.3 kpに、深江出口が走行車線側に存在する。対象領域の地図を図-1に示す。この区間の7地点の観測点において、前述の期間全ての車両の通過時刻データを取得できている。

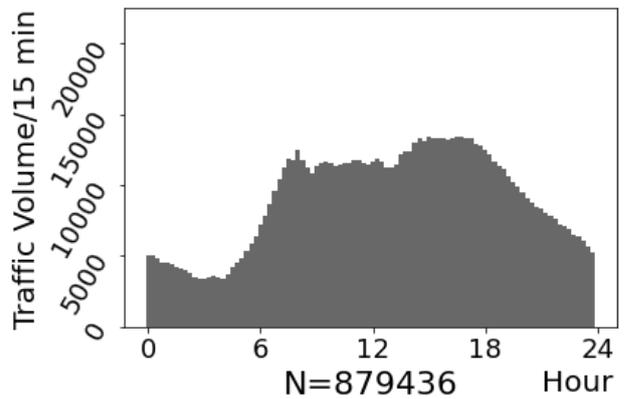
一例として、20.4 kpの走行車線、追越車線において取得した期間の平日全46日間を合計した交通量のヒストグラムを図-2に示す。横軸の左端が0時、右端が24時を示し、ヒストグラムの階級の幅は15分である。走行車線の46日間の総通過台数は879,436台、追越車線の46日間の総通過台数は1,121,262台であり、それぞれ一日あたり19,118台、24,375台である。走行車線について、7:00付近から18:00付近まではほぼ均等に12000台\*46日/15分程度の交通量が流れており、これは1043台/時に相当する。追越車線について、7:00付近にかなり立ち上がりの速い急峻なピークがあり、交通量は22000台\*46日/15分程度である。これは、1913台/時に相当する。また、16:00付近に右側の裾の厚いピークが確認できる。これらは朝の出勤にあたるピークと夕方の帰宅にあたるピークであると考えられる。追越車線について、前者のピークは左に歪み急峻であり、後者のピークは緩やかであるため、確率分布の当てはめを行う上で柔軟な形状の表現を必要としている。

### 3. 分析手法

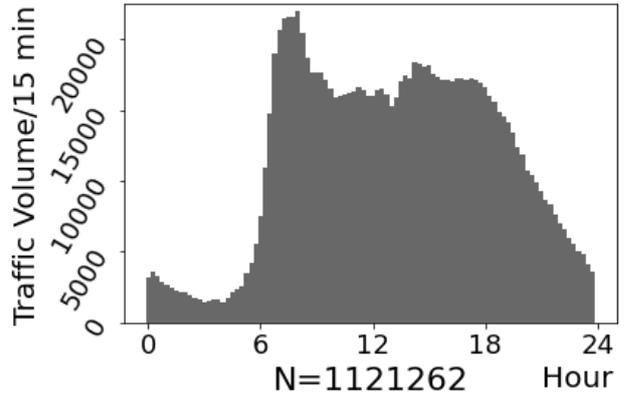
#### (1) 方向統計学

##### a) 方向統計学の概要

方向統計学 (Directional Statistics)<sup>9)</sup> とは角度データを扱う統計学である。角度データは $2\pi$ の周期を持つため、観測値が角度である場合、定義域が $[0, 2\pi)$ や $[-\pi, \pi)$ のように幅 $2\pi$ に制限される。さらに、定義域の両端は本来連



(a) 走行車線の交通量



(b) 追越車線の交通量

図-2 平日46日の合計の15分ごとの交通量

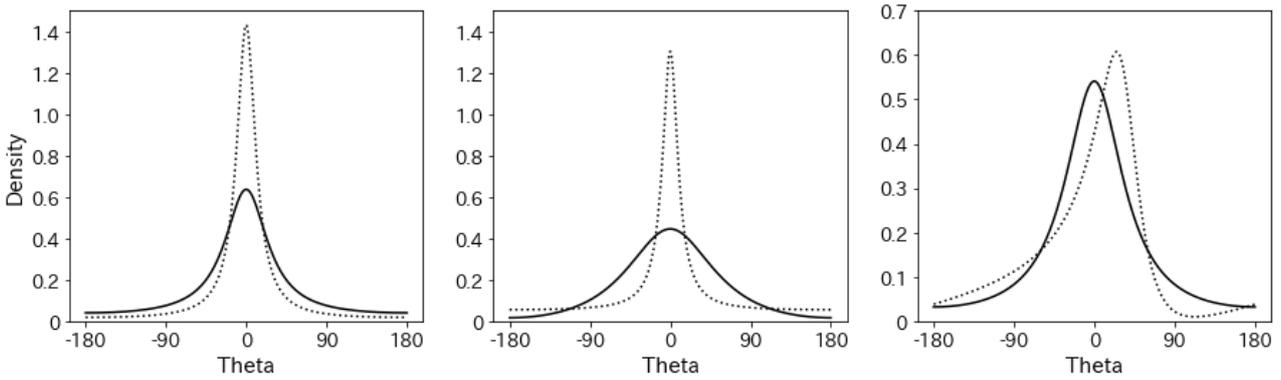
続している。本研究で用いる車両の通過時刻のデータも一日、すなわち24時間の周期で繰り返されている。したがって、このようなデータに対しても方向統計学を適用することができる。例えば、時刻 $t$ を角度 $\theta$ で $\theta = 2\pi(t - 12)/24$ と置き換えることで、0時を $-\pi$ 、12時を0、24時を $\pi$ という $[-\pi, \pi)$ の範囲の角度データとして用いることができる。

##### b) 円周分布

方向統計学における確率密度関数も $2\pi$ の周期を持つ必要がある。すなわち、角度上の確率分布は以下の三つの制約を満たす分布でなければならない。

- 1) 全ての実数 $\theta$ について、 $f(\theta) \geq 0$
- 2) 確率の和は1となる。すなわち、 $\int_a^{a+2\pi} f(\theta) d\theta = 1$
- 3)  $f(\theta)$ は周期 $2\pi$ の関数となる。  
すなわち、 $f(\theta + 2k\pi) = f(\theta)$

条件1)は通常の実軸の分布と共通する条件だが、2), 3)は $2\pi$ の周期を持つ角度データ独特の条件である。このような角度データに対する分布を円周分布と呼ぶ。



左の図は $\gamma$ による形状の違い( $\mu=0, \bar{\alpha}_2 = \gamma^2, \bar{\beta}_2 = 0$ のもとでの $\gamma = 0.6$ (実線),  $\gamma = 0.8$ (点線))を, 中央の図は $\bar{\alpha}_2$ による形状の違い( $\mu=0, \gamma = 0.6, \bar{\beta}_2 = 0$ のもとでの $\bar{\alpha}_2 = 0.2$ (実線),  $\bar{\alpha}_2 = 0.5$ (点線))を, 右の図は $\bar{\beta}_2$ による形状の違い( $\mu=0, \gamma = 0.6, \bar{\alpha}_2 = 0.2$ のもとでの $\bar{\beta}_2 = 0$ (実線),  $\bar{\beta}_2 = 0.2$ (点線))をそれぞれ示している.

図-3 Kato and Jones分布の形状

(2) Kato and Jones分布とその混合分布

a) Kato and Jones分布

本研究では, 様々な分布形状を表現できる柔軟性の高いKato and Jones分布<sup>6)</sup>を用いて交通量の時間変動のパラメータ推定を行う. 確率密度関数を以下の式(1)に示す.

$$g(\theta) = \frac{1}{2\pi} \left[ \frac{1 + 2\gamma^2}{\gamma^2 + \bar{\alpha}_2^2 + \bar{\beta}_2^2 - 2\gamma\{\bar{\alpha}_2 \cos(\theta - \mu) + \bar{\beta}_2 \sin(\theta - \mu)\}} \right] \quad (1)$$

ここで,  $-\pi \leq \mu < \pi, 0 \leq \gamma < 1, (\bar{\alpha}_2, \bar{\beta}_2) \neq (\gamma, 0)$ であり,  $(\bar{\alpha}_2 - \gamma^2)^2 + \bar{\beta}_2^2 \leq \gamma^2(1 - \gamma^2)$ である. 四つのパラメータ $\mu, \gamma, \bar{\alpha}_2, \bar{\beta}_2$ は, それぞれ位置, 集中度, 尖度, 歪度を表す. これら四つのパラメータにより, 図-3のように様々な分布形状を表現することができる. さらに, 確率密度関数が初等関数で構成されているため, 数値的な扱いが比較的簡単である. なお, 図-3の左図は $\bar{\alpha}_2 = \gamma^2, \bar{\beta}_2 = 0$ のときの確率分布であり, この時Kato and Jones分布はWrapped Cauchy分布という特殊な場合に帰着する. 以後 $\bar{\alpha}_2$ による尖度の大小について,  $\gamma^2$ を基準にして行う.

b) 混合Kato and Jones分布 (MoKJ)

Kato and Jones分布を混合し, 多峰の確率分布を表現することを考える.  $m$ 個のKato and Jones分布を各分布 $h$ の比率 $\pi_h$ で混合した時の混合Kato and Jones分布の確率密度関数を式(2)に示す.

$$f(\theta) = \sum_{h=1}^m \pi_h g_h(\theta | \mu_h, \gamma_h, \bar{\alpha}_{2h}, \bar{\beta}_{2h}) \quad (2)$$

ここで,  $\pi_h \geq 0$ であり, 確率の和は1であるため $\sum \pi_h =$

1となる. 便宜上, この分布をMoKJ (Mixture of Kato and Jones)と呼ぶ.

(3) パラメータ推定

Kato and Jones分布は円周分布のなかでは比較的簡潔な形式だが, パラメータの最尤推定には数値計算が必要である. 一方, 混合分布のパラメータ推定に用いられるEMアルゴリズム<sup>7)</sup>は, 最尤推定を解析的に行えない場合には多大な計算時間が必要となる. したがって, MoKJのパラメータ推定にEMアルゴリズムを適用することは現実的でない. そこで, 著者らはEMアルゴリズムの一部をモーメント推定法で代用したEMMアルゴリズム<sup>8)</sup>を提案した. さらに, 解の特定可能性および計算時間の面でより優れた推定法を開発した<sup>9)</sup>. これは, (通常の局所解を有する非凸問題の最尤推定と同様に)多数のパラメータの初期値から尤度の最大化を試みる方法のひとつである. 以降の推定結果は, この最新の推定法を用いたものである.

4. 分析結果

(1) パラメータ推定の結果

前述のパラメータ推定手法を用いて, 対象範囲の交通量の時間変動のパラメータを推定する. 追越車線において左右に二つの顕著なピークが見られるため, MoKJの混合数は2とした. 同様に, 走行車線についても混合数は2とした. なお, 3以上の混合数についても推定を試したが, 2つのピークが卓越しているため3つ目のピークの比率が著しく小さくなり, 推定結果は混合数2のものほとんど変わらなかった.

まず, 7地点の走行車線について, 推定した確率密度

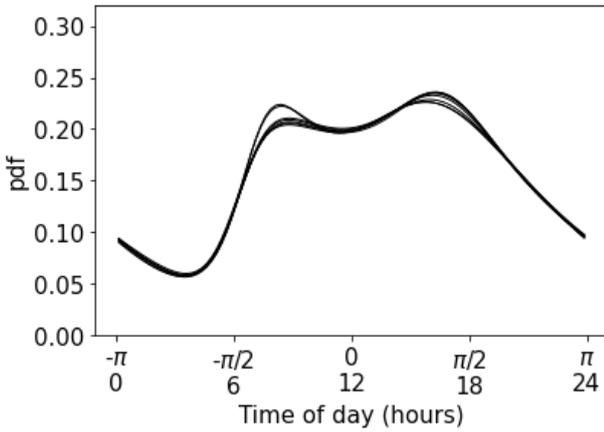


図-4 7地点の走行車線の推定結果

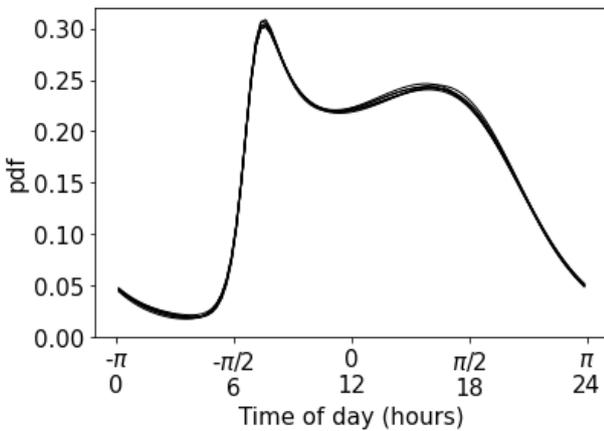


図-5 7地点の追越車線の推定結果

関数を重ね合わせたものを図-4に示す。6:00過ぎのピークについて、緩やかなピークのもの、少し盛り上がったピークの二種類が見られる。緩やかなピークは深江出口の下流側の5地点のものであり、少し盛り上がったピークは出口の上流側の2地点のものである。18:00前のピークについても出口の前後で多少の差異が見られる。

次に、7地点の追越車線について、推定した確率密度関数を重ね合わせたものを図-5に示す。これらの推定結果は7地点全てにおいてほぼ変わらず、ほとんど重なっているように見える。これは、対象領域は最大でも3 kmほどしか離れておらず、さらには追越車線側には出入口も存在しないため、ほぼ全ての車両が数分間に通過するためだと考えられる。

深江出口より上流側に位置する22.0 kpの走行車線の推定結果を図-6、表-1に示す。また、下流側に位置する20.4 kpの走行車線の推定結果を図-7、表-2に示す。以下、図-6や図-7の左側の分布( $h = 1$ )を第一の分布、右側の分布( $h = 2$ )を第二の分布と呼ぶ。22.0 kpの結果について、第一の分布は平均が11:40程度に位置し、最頻値は8:20程度に位置する。第二の分布は平均が17:10程度に位置し、

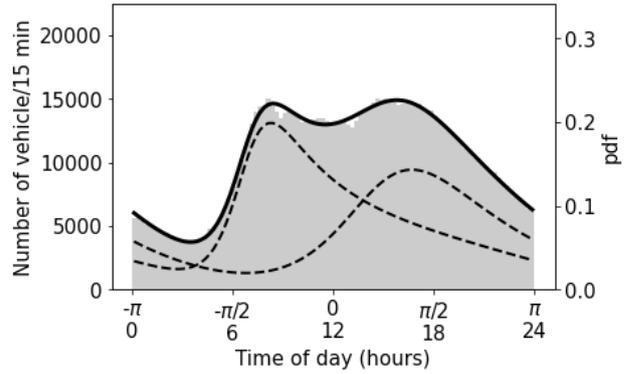


図-6 22.0kp地点の走行車線の推定結果

表-1 22.0kp地点の走行車線の推定パラメータ

$h$	$\mu_h$	$\gamma_h$	$\bar{\alpha}_{2h}$	$\bar{\beta}_{2h}$	$\pi_h$
1	-0.335	0.371	0.102	-0.156	0.553
2	1.456	0.413	0.070	-0.056	0.447

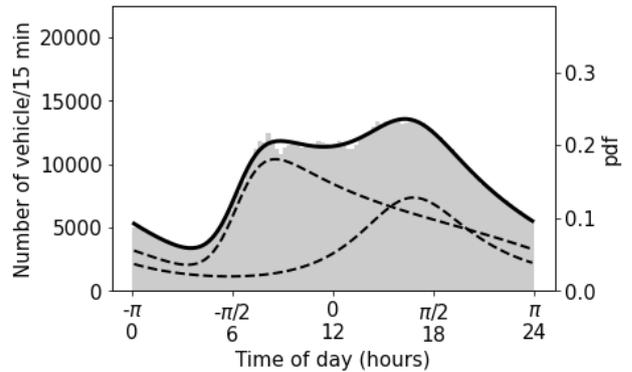


図-7 20.4kp地点の走行車線の推定結果

表-2 20.4kp地点の走行車線の推定パラメータ

$h$	$\mu_h$	$\gamma_h$	$\bar{\alpha}_{2h}$	$\bar{\beta}_{2h}$	$\pi_h$
1	-0.085	0.273	0.023	-0.129	0.642
2	1.346	0.425	0.134	-0.026	0.358

最頻値は16:40程度に位置する。一方、20.4 kpの結果について、第一の分布は平均が10:40程度に位置し、最頻値は8:30程度に位置する。第二の分布は平均が17:30程度に位置し、最頻値は16:50程度に位置する。また、図-6に示される22.0 kpの交通量は1,007,078台であり、図-7に示される20.4 kpの交通量よりも127,646台多い。これは1日あたり約2,775台に相当する。この交通量は深江出口から流出する交通量におおむね等しい。各時間帯の交通量について、20.4 kpよりも全体的に22.0 kpの方が多いが、朝のピークについてその差は特に顕著である。そのことが双方の $\gamma_1$ と $\bar{\alpha}_{2_1}$ に表現されており、22.0 kpの $\gamma_1$ と $\bar{\alpha}_{2_1}$ が20.4 kpのものよりも大きくなっている。さらに、22.0 kp、

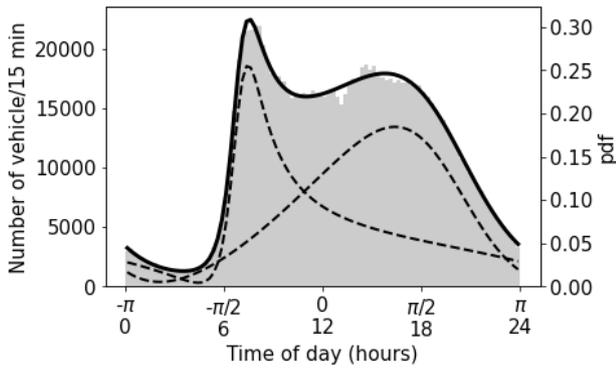


図-8 22.0kp地点の追越車線の推定結果

表-3 22.0kp地点の追越車線の推定パラメータ

$h$	$\mu_h$	$\gamma_h$	$\bar{\alpha}_{2h}$	$\bar{\beta}_{2h}$	$\pi_h$
1	-0.457	0.411	0.192	-0.231	0.448
2	0.873	0.495	0.039	0.074	0.552

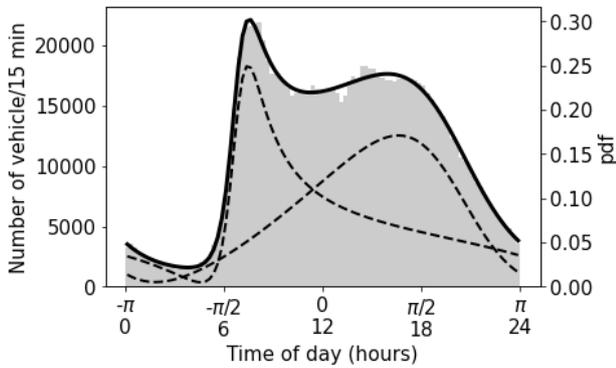


図-9 20.4kp地点の追越車線の推定結果

表-4 20.4kp地点の追越車線の推定パラメータ

$h$	$\mu_h$	$\gamma_h$	$\bar{\alpha}_{2h}$	$\bar{\beta}_{2h}$	$\pi_h$
1	-0.384	0.375	0.154	-0.225	0.485
2	0.869	0.485	0.037	0.088	0.515

20.4 kpの双方について、 $\bar{\beta}_{21}$ が絶対値が大きい負の値になっている。これにより朝の立ち上がりの早いピークを表現している。第二の分布については、両者ともに緩やかなやや左に歪んだ形をしている。これにより夕方の緩やかなピークを表現している。各分布の比率を決める $\pi_h$ について、22.0 kpのものは顕著な差はないものの、20.4 kpのものは第一の分布の方がかなり大きくなっている。これは、第一の分布が小さい $\bar{\alpha}_{21}$ によって裾の広い分布となっているため、夕方のピークについても第一の分布が担っているためであると考えられる。

次に、深江出口より上流側に位置する22.0 kpの追越車線の推定結果を図-8、表-3に示す。また、下流側に位置

する20.4 kpの追越車線の推定結果を図-9、表-4に示す。22.0 kpの結果について、第一の分布は平均が10:20程度に位置し、最頻値は7:30程度に位置する。第二の分布は平均が15:20程度に位置し、最頻値は16:20程度に位置する。一方、20.4 kpの結果について、第一の分布は平均が10:30程度に位置し、最頻値は7:30程度に位置する。第二の分布は平均が15:20程度に位置し、最頻値は16:40程度に位置する。また、図-8に示される22.0 kpの交通量は1,111,781台であり、図-9に示される20.4 kpの交通量よりも9,481台少ない。これは1日あたり約206台に相当する。地点間の交通量の差はほとんどなく、パラメータの推定結果にもほとんど差はない。双方ともに、 $\gamma_1$ より $\gamma_2$ の方が大きい、 $\bar{\alpha}_{21}$ が $\bar{\alpha}_{22}$ よりも大きいため、第一の分布のピークの方が急峻になっている。さらに、 $\bar{\beta}_{21}$ は絶対値のかなり大きい負の値であるため、第一の分布はかなり左に歪んでいる。そのことにより7:30辺りの立ち上がりの早いピークを表現している。第二の分布について、緩やかなやや右に歪んだ形状をしている。それらが小さい $\bar{\alpha}_{22}$ と、絶対値の小さい正の $\bar{\beta}_{22}$ で表現されている。また、各分布の比率について、両者ともに $\pi_h$ は0.5程度だが、若干 $\pi_2$ の方が大きくなっている。

(2) パラメータの解釈

全ての結果について、第一の分布は朝の出勤の分布、第二の分布は夕方の帰宅の分布と捉えることができる。走行車線について、朝の出勤時間周辺交通の一部が深江出口で流出し交通量が減っているため、推定結果についても深江出口前後で大きな違いが見られる。特に、朝の出勤の分布の形状は大きく違う。そのことが第一の分布のパラメータに表れており、22.0 kpの第一の分布のパラメータは、比較的大きい $\bar{\alpha}_{21}$ と、絶対値が大きい負の $\bar{\beta}_{21}$ を持つ。これに対して、下流側の20.4 kpの $\bar{\alpha}_{21}$ は22.0 kpのものよりもかなり小さい。第二の分布についても、形状は出口の上流側の方が下流側に比べて大きくなっている。しかし、これは下流側の20.4 kpの第一の分布の右の裾が分厚くなっているためでもあると考えられる。また、走行車線では深夜帯の交通量が追越車線に比べて多い。ピーク時との比較では、追越車線が1割程度であるのに対して、走行車線は3割程度もある。走行車線における深夜帯の交通量について、出口より上流側の22.0 kp地点では第二の分布が、下流側の20.4kp地点では第一の分布が過半数を占めていることが図-6、図-7からわかる。出口の上流と下流にかかわらず、深夜帯の交通は共通して第一、もしくは第二の分布のどちらかの影響を強く受けていると予想していたが、その予想とは反する結果が得られた。しかし、上流、下流両方について左に歪んでおり、右に裾が厚い第一、第二の二つの分布の右の裾で表現されているという共通した特徴がみられた。

追越車線について、第一の分布は出勤の時間が8:30や9:00などの時間に規定されており、通過交通の多い追越車線においては、その時間に間に合うように多くの交通が集中することが予想される。その結果、分布の最頻値はその時間よりも一時間程度前の時間になり、推定した $\bar{\alpha}_{2_1}$ は比較的大きい値に、 $\bar{\beta}_{2_1}$ は絶対値がかなり大きい負の値になっている。一方、帰宅の時間は出勤の時間よりも比較的自由度が高いため、交通量は裾の厚い緩やかな分布になっていることが予想される。そのことが小さい $\bar{\alpha}_{2_2}$ によって表現されている。また、第二の分布が右に歪んでいる理由として、帰宅の時間を極端に遅くしたくないという利用者の心理が交通量に反映されていると考えられる。

## 5. 結論

本研究では、円周分布のひとつで、柔軟な分布形状を表現できるKato and Jones分布を混合したMoKJを用いて、交通量の時間変動のパラメータ推定を行った。このことにより、二つのピークを持つ交通量分布を二つの確率分布に分離することに成功した。第一のピークは朝の出勤の交通と考えられうる、急峻で左に歪んだ分布であった。第二のピークは夕方の帰宅のピークと考えられうる、緩やかな分布であった。

本研究で用いたデータはトラフィックカウンターで観測した交通量のデータであり、そもそもの交通需要とは必ずしも一致しないと考えられる。高速道路の道中ではなく、インターチェンジの流入交通量などを用いて、本手法を通して交通需要の特性を分析することは今後の課題の一つであると考えている。

謝辞：本研究で用いた交通量データは阪神高速道路株式会社からご提供頂いた。統計数理研究所の加藤昇吾准教授からパラメータ推定手法についてご教示いただいた。また、本研究の一部は科学研究費(若手研究:18K13846)の助成により行われた。この場を借りて厚く御礼申し上げます。

## 参考文献

- 1) 中山晶一郎：道路の時間信頼性に関する研究レビュー，土木学会論文集D3（土木計画学），Vol. 67, No. 1, pp. 95-114, 2011.
- 2) 高山純一，飯田恭敬：周期変動を考慮した実測交通量による交通需要推計法，土木計画学研究・論文集，No. 1, pp. 43-50, 1984.
- 3) 佐々木邦明，名取優太，井口均，西川啓幸：周期変動および自己回帰を考慮した高速道路の時間交通量の変動特性の分析，交通工学論文集，Vol. 3, No. 2, pp. A\_178-A\_186, 2017.
- 4) Kumar, S. V., and Vanajakshi, L.: Short-term traffic flow prediction using seasonal ARIMA model with limited input data, European Transport Research Review, Vol. 7, No. 3, pp. 1-9, 2015.
- 5) 清水邦夫：角度データのモデリング，近代科学社，2018.
- 6) Kato, S., and Jones, M. C.: A tractable and interpretable four-parameter family of unimodal distributions on the circle, Biometrika, Vol. 102, No. 1, pp. 181-190, 2015.
- 7) Ng, S. K., Krishnan, T., and McLachlan, G. J.: The EM algorithm. In Handbook of computational statistics, Springer, Berlin, Heidelberg. pp. 139-172, 2012.
- 8) 長崎滉大，中西航，朝倉康夫：方向統計学を用いた道路ネットワークの分析，土木計画学論文集 D3（土木計画学），Vol. 75, No. 6, pp. I\_119- I\_205, 2020.
- 9) Nagasaki, K., Nakanishi, W. and Kato, S: Private communication, unpublished.

(2021.?? 受付)

## ANALYSIS OF TRAFFIC VOLUME BY THE ESTIMATED PARAMETER OF THE MIXTURE OF CIRCULAR DISTRIBUTION

Kota NAGASAKI, Wataru NAKANISHI and Yasuo ASAKURA

The traffic volume varies by time of a day. For example, traffic volume of an expressway between cities has two typical peaks for the morning rush hour and evening rush hour. Many studies analyze the fluctuation of traffic volume by applying the non-parametric model. However, little studies estimate the parameters for probability density function by assuming the parametric model. Moreover, it is necessary to process the periodic data such as traffic volume in a day by applying directional statistics. In this study, we estimate the parameter for the mixture of circular distribution, which is the probability density function for angular data. Finally, the two peaks of morning rush peak and evening rush peak are divided. It is clarified that the earlier peak is sharp and left skewed peak, and the latter peak is gentle peak for both passing and driving lanes.