

# カーネル密度推定を適用した鉄道通勤者の 到着分布の推定

宮内 弘太<sup>1</sup>, 高田 和幸<sup>2</sup>, 藤生 慎<sup>3</sup>

<sup>1</sup>学生会員 学士 (工学), 東京電機大学理工学研究科

<sup>2</sup>正会員 博士 (工学), 東京電機大学理工学部 (〒350-0394 埼玉県比企郡鳩山町石坂)

e-mail: takada@g.dendai.ac.jp Phone: 049-296-5708

<sup>3</sup>正会員 博士 (学際情報学), 金沢大学理工研究域

近年, 首都圏では鉄道の遅延が慢性的に発生しており, 所要時間の信頼性の低下が社会問題となっている. 所要時間信頼性の向上に向けて様々な観点から調査・分析・研究が行われているが, 鉄道利用者個人の到着状況のモデル化を行った先行研究は多くない. そこで本研究では, インターネット調査を行い, 到着状況 (早着・定時着・遅着) に関するデータを収集した. そしてカーネル密度推定を適用して, 鉄道通勤者個人の到着状況の分布を特定した.

**Key Words :** railway commuters, arrival distribution, kernel density estimation,

## 1. 背景・目的

近年, 首都圏においては列車の遅延本数が増加傾向にある. 著者らの調査<sup>1)</sup>によると, 首都圏の鉄道利用者は「新規路線の整備」「所要時間の短縮」よりも, 「定時運行」「運行停止時間の短縮」といった所要時間の信頼性の向上を求めていることが明らかとなった.

軌道系交通機関の特徴である時間厳守性が守られなければ, 利用者は一定時間の遅延を見込んで出発時刻を早めざるを得ず, 不便を強いられる. このようなことから, 鉄道の所要時間の信頼性を向上させる取り組みが必要である.

2012年7月に改定された「鉄道プロジェクトの評価手法マニュアル 2012 (国土交通省鉄道局) <sup>2)</sup>」においても, 列車の遅延時間短縮便益の計測手法の開発の必要性が記されている. また2014年4月に出された交通政策審議会への諮問<sup>3)</sup>においても, 鉄道遅延への対応が挙げられている.

所要時間信頼性に関しては様々な観点から調査・研究が行われているが, 鉄道利用者個人の到着状況のモデル化を行った先行研究は多くない.

そこで本研究では, 「鉄道の利用者の到着状況が, 特定の分布では表現しきれない」という筆者らのこれまで

の研究成果<sup>4)</sup>を踏まえ, 個々の鉄道利用者の到着状況をカーネル密度推定を用いて特定することを目的とする.

なおカーネル密度推定は, ビッグデータを用いた機械学習による将来予測の分野で活用が広がっているが, 土木計画学分野における適用は, 松田ら<sup>5)</sup>に見られる程度でまだ多くない.

## 2. 既往研究

### (1) 異種混合モデル推定についての研究

道路, 鉄道の旅行時間信頼性の評価を行った研究は数多くあり, 利用者の到着状況について様々な設定がなされている. 中でも, 利用者の到着時刻は正規分布や対数正規分布, 指数分布に従うと仮定した研究が多い.

中山<sup>6)</sup>は, 道路交通利用時の旅行時間の分布に関するレビューを行い, 旅行時間の減少は限界があるが, 増大には限界が無いことから, 旅行時間には対数正規分布を仮定するのが自然であるとしている. 一方, 高浪ら<sup>4)</sup>は鉄道利用者の到着分布は特定の分布に従うものではなく, 複数の分布が混在していること示した上で, 異種混合分布 (2つの正規分布と1つの指数分布) に表現できることを示している (図-1)

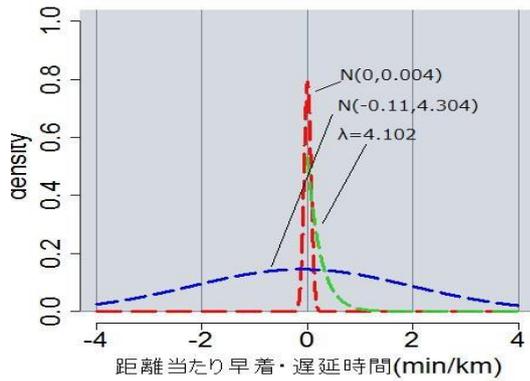


図-1 異種混合分布を適用した到着分布<sup>4)</sup>

表-1 分布形の組み合わせとBICの対応値<sup>7)</sup>

分布波形	BIC
正規分布 1つ	244983
正規分布 2つ	12689
正規分布 1つ+指数分布 1つ	125537
正規分布 2つ+指数分布 1つ	8607
正規分布 3つ+指数分布 1つ	26903

一方、宮内ら<sup>7)</sup>はk-means法を用いて、到着分布の形状に基づいたセグメンテーションを行っている。クラスタ数を2~5の4種を設定し、設定数による影響を確認した結果、クラスタ数が3の時が最も適している事を示している(表-1)

### 3. 分析データ

#### (1) アンケート調査

本研究では、株式会社マクロミルの調査モニターを活用して、鉄道の利用状況に関するアンケート調査を実施した。調査対象者は、首都圏の1都3県(東京・神奈川・千葉・埼玉)に居住する15歳以上の有職者(パート・アルバイトを含む)で、かつ週に5日以上鉄道を利用する者とした。調査は平成27年2月23、24日の2日間で行い、計1000名からの回答を得た。有効回答数は1000である。

図1は、アンケートで示した到着状況に関する設問である。この回答データをもとにして、希望到着時刻と実際の到着時刻との乖離に関するデータを作成した。

また、乗車・降車・乗換えに利用した駅、利用路線に関する情報に基づき、経路検索サイト(Yahoo!経路, 2015年4月)を参照して移動距離、到着までの所要時間、乗車料金、乗り換え回数を算出した。

#### (2) 標本の属性

回答者の個人属性(性別、県別居住地、年代、業種)の内訳を図-2に示す。性別、年代、居住地について示す。大都市交通センサスと比較すると、男性の割合と20代の割合において差が生じているが、他の個人属性は特定の属性に偏ることなく標本が抽出されており、本調査は大都市交通センサスと概ね同じ条件でデータが抽出できたと考えられる。

Q1. 時間内到着状況について	
到着時間が前後1分以内に着くのは100回中何回ですか?	
Q2. 早着時間の到着状況について	
・1分以上早く着くのは100回中何回着きますか?	
・2分以上早く着くのは100回中何回着きますか?	
・3分以上早く着くのは100回中何回着きますか?	
・4分以上早く着くのは100回中何回着きますか?	
・5分以上早く着くのは100回中何回着きますか?	
・10分以上早く着くのは100回中何回着きますか?	
・20分以上早く着くのは100回中何回着きますか?	
・30分以上早く着くのは100回中何回着きますか?	
Q3. 遅着時間の到着状況について	
・1分以上遅く着くのは100回中何回着きますか?	
・2分以上遅く着くのは100回中何回着きますか?	
・3分以上遅く着くのは100回中何回着きますか?	
・4分以上遅く着くのは100回中何回着きますか?	
・5分以上遅く着くのは100回中何回着きますか?	
・10分以上遅く着くのは100回中何回着きますか?	
・20分以上遅く着くのは100回中何回着きますか?	
・30分以上遅く着くのは100回中何回着きますか?	
Q4. あなたが自宅の最寄り駅から乗車する時間は何時何分ですか?	
Q5. あなたが会社の最寄り駅に到着する理想の時間は何時何分ですか?	

図-2 アンケート調査内容(一部抜粋)

表-2 回答者の個人属性

		第11回大都市交通センサス	本調査
男女比	男性	65.2%	78.3%
	女性	37.7%	21.3%
年齢層	10代	0.4%	0.0%
	20代	22.3%	5.6%
	30代	28.5%	28.5%
	40代	22.7%	26.9%
	50代	15.4%	24.3%
	60代以上	8.5%	14.7%
居住地	埼玉県	19.0%	14.5%
	千葉県	17.0%	13.1%
	東京都	42.0%	49.2%
	神奈川県	22.0%	23.2%

### 4. 到着状況に関する分析

#### (1) 所要時間の認知状況

アンケートでは鉄道利用者が自宅の最寄り駅を出発する出発時刻と、理想とする到着時刻をを聞いている。次に理想到着時刻を経路検索サイトに入力し、表示さ

れた出発時刻を調べ、理想到着時刻に着くための出発時刻を決定する。すると実際の出発時刻と、路線検索サイトで調べた出発時刻との間には差が生じる。この差が正で多い場合は、余裕時間を大きく設定しており、また負で大きい場合には理想の到着時刻よりも遅れて到着することを許容していることを示している。図-3はこれらの時間の関係を示した概念図である。

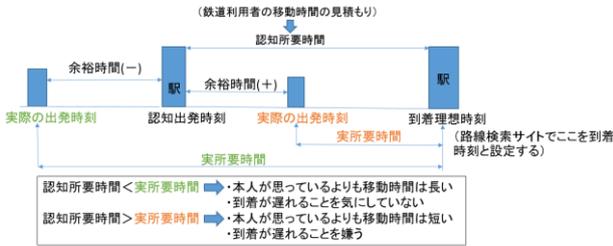


図-3 余裕時間の概念図

この余裕時間を通勤時の移動距離で割ることにより、単位kmあたりの余裕時間が算出される。この時の階級と頻度の関係を図-4に示す。プラス側であれば、単位kmあたりの時間分だけ早着をしている。一方、マイナス側であれば、単位kmあたりの時間分だけ遅着をしている。鉄道利用者の認知所要時間と実所要時間との時間差は大きくばらついていることが分かる。全体的に見てみると比較的マイナス側の方が多いことが分かる。このことから多くの鉄道利用者は事前に余裕時間を持った行動をしていることが分かる。

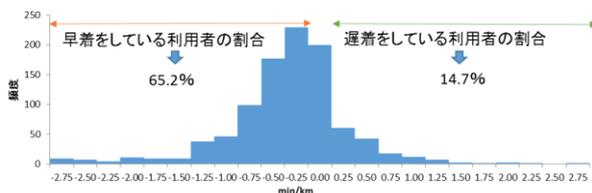


図-4 認知所要時間と実所要時間の乖離

## 5. カーネル密度推定を用いた到着分布

### (1) カーネル密度推定について

先行研究より、鉄道利用者の到着分布については、複数の分布が混在していることが明らかとなっている。そこで本研究では、カーネル密度推定を用いて、鉄道利用者一人一人の到着分布の特性を分析する。

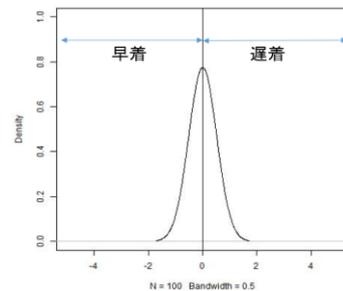
なおカーネル密度推定とは、高次元式で表すことが可能でありサンプル数が多いほど複雑な関数で表せられる。本研究ではガウスクーネルを用いて平滑化した到着分布を求めた。

### (2) データ

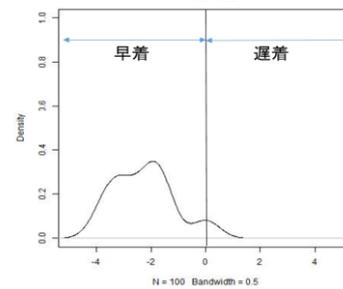
到着状況について17区分された回答データを用いた。回答者の100回分の到着状況を元にしてデータを作成した。得られた早着時間・遅着時間を個々人の乗車距離(km)で割り、単位距離あたりの早着時間・遅着時間に変換した。またカーネル推定を行う際には、バンド幅を0.5(min/km)とした。

### (3) 到着分布の特定結果

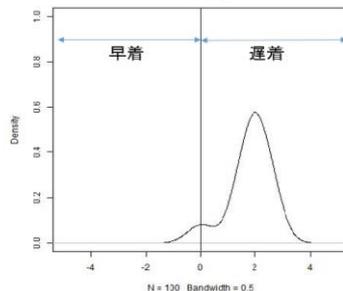
カーネル密度推定を行った結果を図-5に示す。大きく分けて3つの到着分布が出てきた。最も多かったのが期待値0の正規分布の形をした到着分布が多かった。次いで多かったのが早着型の形をした到着分布が多かった。最も少なかったのが遅着型の形をした到着分布となった。この事から鉄道利用者の到着分布というのは大きく分けて3つ存在している。



(a)



(b)



(c)

図-5 カーネル密度推定による到着分布の推定

## 6. まとめと今後の課題

鉄道利用者の目的地への到着状況を知るために早遅着の有無をアンケート調査よりデータを得た。鉄道利用者の到着状況をカーネル密度推定を行うことにより3つの到着分布が得られた。得られた到着分布と異種混合分布を比較してみると、異種混合分布で推定された到着分布と特性は近似している。またカーネル密度推定により3つに分けられた分配数と異種混合分布で推定された分配率を比較してみると、こちらも近似した結果となった。カーネル密度推定を行うにあたり、バンド幅の設定が重要となってくるが、今回の推定の場合バンド幅は0.5(km/h)が妥当である。鉄道利用者一人一人の到着分布を推定することにより、到着状況と個人変数の関係性が説明できる。

今後の展望について記す。

得られた個々人の到着分布より、期待値 $\mu$ 、標準偏差 $\sigma$ 、また微分係数等を算定し、これらの変数を用いて分布形状に基づくクラス分類を行い、クラスごとの分布の特徴を明らかにした上で、異種混合分布を適用して出発時刻選択行動のモデル化を行う予定である。

### 謝辞:

本研究は、科学研究費補助金基盤研究(C)(26420520)の助成を受けたものである。

(2016.7.31受付)

### 参考文献:

1) 高田和幸, 鈴木孝典, 藤生慎: 鉄道の遅延時間を考

- 慮した出発時刻決定行動に関するモデル分析, 土木学会論文集D部門, Vol.68, 2012
- 2) 国土交通省鉄道局監修: 鉄道プロジェクトの評価手法マニュアル2012年改訂版
  - 3) 国土交通省 交通政策審議会 陸上交通分科会鉄道部会: 諮問書(諮問第198号)「東京圏における今後の都市鉄道のあり方について」, 平成26年度第1回(第10回)
  - 4) 高浪裕三, 杉山茂樹, 藤生慎, 高田和幸: 鉄道利用者の所要時間分布の特性分析, 土木計画学研究発表会・講演集, Vol.46, 2012
  - 5) 松田耕治, 向井智哉, 藤生慎, 高山純一, 中山晶一郎: クルーズ客の観光行動分析~カーネル密度推定を用いた検討~, 第53回土木計画学研究発表会・講演集, CD-ROM, 2016
  - 6) 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, 2011
  - 7) 宮内弘太, 高田和幸, 高浪裕三, 藤生慎: 混合モデル分布を適用した鉄道通勤者の到着分布の推定, 交通工学研究会, 2016

## ESTIMATION OF THE ARRIVAL DISTRIBUTION OF THE RAILWAY COMMUTERS BY KERNEL DENSITY ESTIMATION

Kota MIYAUCHI, Kazuyuki TAKADA and Makoto FUJII

Recently, the delay of railway operation chronically occurs in the Tokyo metropolitan area, and decrease in reliability of travel time has become one of the social problems. There are many previous studies focusing on the reliability of travel time. However, there are few previous studies modeling the individual arrival situation of railway users. Therefore, in this study, internet questionnaire survey was conducted in order to collect the data about the arrival situation of the railway commuters in the Tokyo metropolitan area. The kernel density estimation was applied to model the individual arrival situation in this study. As the results of the estimation, density of the arrival situation for each railway commuter was identified.