

SNS が鉄道需要創出に与える影響－四国を対象に－

芝浦工業大学大学院 学生会員 ○楠田 早紀子
芝浦工業大学 正会員 岩倉 成志

1. 背景と目的

鉄道プロジェクトの評価基準には、費用便益分析が用いられてきた。プロジェクトの評価を向上するためには、主に一般化費用の改善が行われるが、B/Cは整備前の鉄道需要を育てることによっても上昇する。そこで、近年急激に普及・発展している SNS マーケティングを行うことにより鉄道需要の創出につながるのではないかと考えた。

本研究では四国の観光地に関する SNS 投稿の増加に伴い、付近の公共交通機関の利用者が増加するという仮説を立て、観光地に関する SNS 投稿と観光地付近の駅利用者数の相関性を分析することを目的とする。

2. 対象地及びデータ概要

本研究では四国の観光地から徒歩またはバスで移動できる最寄り駅を対象とし、丸亀駅・栗林公園北口駅・大歩危駅・観音寺駅・詫間駅・琴平駅・鳴門駅・下灘駅・伊野駅の9地点の day-to-day 定期外乗降者数データを用いた。対象期間は2019年4月3日から2022年7月31日である。

近年の四国県庁駅（高松駅、徳島駅、松山駅、高知駅）の日別駅乗降者数を図-1に示す。土休日や年末年始、お盆休み、連休といった期間に鉄道需要が高まることが分かる。また、2020年4月には新型コロナウイルス感染拡大に伴う緊急事態宣言が発出されたことによる鉄道需要の減少が起きている。

SNS 投稿数は表-1に示すキーワードを含む投稿を Twitter API より抽出し、週別に集計したものを用いる。なお、本研究では投稿内容について解析していないため、投稿内容のネガティブ、ポジティブに左右されてしまう可能性のあるリツイートやリプライによる投稿は SNS 投稿数に含まず分析することとした。

3. 分析手法

駅乗降者数と SNS 投稿数の相関性を分析するため、2つの手法で行った。1つ目は SNS 投稿以外の影響要因を排除した乗降者数を時系列モデルにより推定し、その時系列モデルと実利用者数との残差と、SNS 投稿数との相関性を分析する異常検知技術の手法である（以下、手法①）。

2つ目は乗降者数を目的変数とした時系列モデルを推定する際、外生変数に SNS 投稿数を入れ、パラメータの統計的有意性を確認する手法である（以下手法②）。

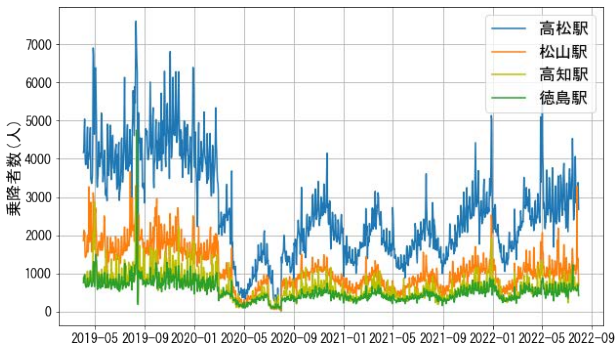


図-1 四国県庁所在地駅の日別乗降者数

表-1 観光地に関するキーワード

駅名	丸亀駅	栗林公園北口	大歩危駅	観音寺駅	
キーワード	丸亀城	栗林公園	祖谷のかずら橋	高屋神社	
			かずら橋	観音寺	
駅名	詫間駅	琴平駅	鳴門駅	下灘駅	伊野駅
キーワード	父母ヶ浜	金刀比羅宮	大塚国際美術館	下灘駅	仁淀川
					仁淀ブルー

4. 時系列モデル（SARIMAXモデル）の概要

駅乗降者数データのような時系列を分析するためには1年を周期とする季節変動を考慮したモデルを扱う必要がある。そのため、時系列モデルの中でも季節変動成分やトレンド成分、外部変数Xを考慮できる次式の SARIMAX モデルによってモデル推定を行う。

$$y_n = \beta exog_n + \sum_{i=1}^p a_i x_{n-i} + \sum_{i=0}^q b_i \varepsilon_{n-i} + c + \sum_{j=1}^P A_j \sum_{i=1}^p a_i x_{n-sj-i} + \sum_{j=1}^Q B_j \sum_{i=1}^q b_i \varepsilon_{n-sj-i}$$

ここで p は自己回帰モデルの次数、 q は移動平均モデルの次数、 P, Q はそれぞれ季節自己回帰モデル、季節移動平均モデルの次数である。また、 β は外生変数のパラメータ、 $exog_n$ は外生変数、 A, a, B, b は内生変数のパラメータ、 ε_n は誤差、 c は定数項を示す。本研究では週別データ系列を扱うので、周期 s は 52 となる。

モデルの選定方法は、各次数の組み合わせ当てはめたモデルを複数作り、赤池情報量基準 (AIC) の値が最小となるモデルを選んだ。

また、日照時間、緊急事態宣言、第1回蔓延防止策、第2回蔓延防止策が SNS 以外の乗降者数に寄与する要因として挙げられるため、外生変数としてパラメータを推定する。

表-2 各駅の推定された SARIMAX モデルの係数

	丸亀駅	栗林公園北口	大歩危駅	観音寺駅	詫間駅	琴平駅	鳴門駅	下灘駅	伊野駅
外 生 変 数	日照時間	0.047	0.033	0.038	-0.004	0.032	0.022	0.005	0.029
	緊急事態宣言	-0.158	-0.101	0.003	-0.092	-0.183	-0.182	-0.111	-0.229
	蔓延防止策①	0.015	-0.056	-0.027	-0.036	-0.046	-0.019	-0.061	-0.053
	蔓延防止策②	-0.01	-0.186	-0.043	-0.015	-0.033	-0.05	-0.043	-0.108
	SNS投稿数(P値)	0.152(0.001)	0.245(0.001)	0.166(0.000)	0.103(0.008)	0.222(0.000)	0.065(0.235)	0.178(0.002)	0.227(0.232)
内 生 変 数	ar.L1	-1.047	—	—	—	—	-0.644	0.237	—
	ar.L2	-0.347	—	—	—	—	-0.383	—	—
	ma.L1	0.843	-0.64	-0.496	-0.267	-0.354	—	-0.848	-0.746
	ma.L2	—	—	—	—	-0.309	—	0.196	—
	ar.S.L52	-0.253	-0.506	-0.239	-0.268	-0.365	-0.986	—	-0.626
	ma.S.L52	—	—	—	—	—	0.886	—	—
	sigma	0.213	0.623	0.234	0.233	0.37	0.437	0.65	0.768
	R ²	0.741	0.247	0.683	0.711	0.526	0.437	0.021	-0.063

内生変数は全て 5% 有意

5. モデルの推定結果

手法①では推定モデルの残差と SNS 投稿数の相関性が確認できなかったため、以下では手法②の推定結果を説明する。手法②で推定された SARIMAX モデルの係数を表-2 に、丸亀駅の推定モデルを例にモデルの再現性を図-2 に示す。

丸亀駅、大歩危駅、観音寺駅、詫間駅の 4 駅は決定係数が 0.5 を超え、モデルの精度は悪くない。さらに SNS 投稿数の p 値は 4 駅とも 0.05 を下回り、有意な値となったことから、この 4 駅は駅乗降者数に SNS 投稿数が寄与していると言える。

一方で、栗林公園北口駅、琴平駅、鳴門駅、下灘駅、伊野駅は決定係数が 0.5 を下回りモデル精度が良くない結果となった。考えられる要因として、SNS 投稿数データにはネットニュースに伴う投稿も含まれているため、ニュースの内容や投稿内容によって鉄道需要に負の影響を与えている可能性がある。

図-3 に丸亀駅 SARIMAX モデルの SNS 投稿数を 0 から 500 まで増加させた際の乗降者数の感度分析結果を示す(上図 SNS 投稿数 0 件, 下図 SNS 投稿数 500 件)。SNS 投稿数の増加に伴い、乗降者数が増加していることが分かる。さらに、SNS 投稿数が 1 件増加した際の丸亀駅駅降車客数は 2.29 人、大歩危駅、観音寺駅、詫間駅はそれぞれ 0.28 人、0.27 人、0.13 人増加することが分かった。

6. 結論

観光地名を投稿内容に含む SNS 投稿数と観光最寄り駅の乗降者数に相関性があると認められた一方で、課題も残った。

1 つ目は、競合路線の影響についてである。推定結果で SNS 投稿数のパラメータが有意とならなかった琴平駅や栗林公園北口駅では JR の他に琴電やその他の路線が競合している。そのため、競合路線数も考慮する必要があると考える。

2 つ目は、本研究で SNS 投稿数と乗降者数の相関性について分析を行ったが、SNS の投稿内容に

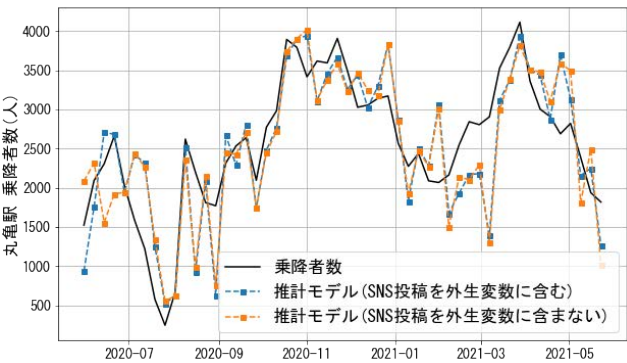


図-2 丸亀駅の推定モデル

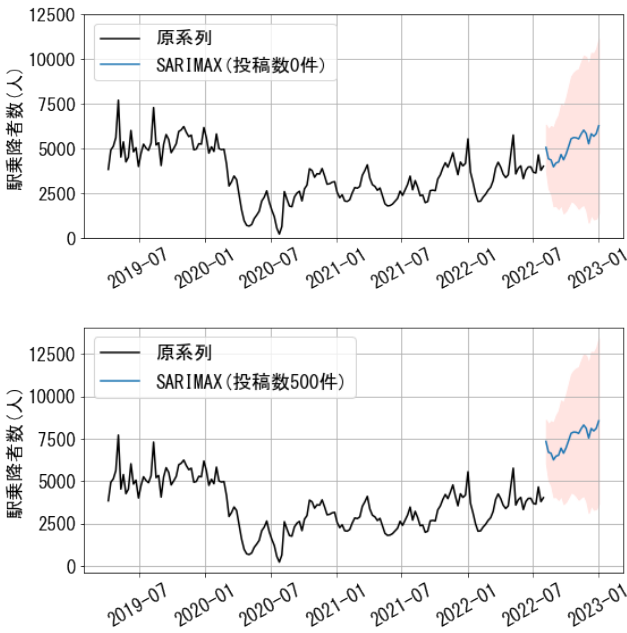


図-3 丸亀駅感度分析結果

より乗降者数にプラスの影響を与えない可能性がある。したがって、投稿内容の感情分析を行い、投稿内容がポジティブな投稿数、ネガティブな投稿数を分離して評価する必要がある。

謝辞

データ提供にご協力いただいた四国旅客鉄道株式会社の方々に、心より感謝の意を表します。