

ETC データを用いた個人の利用傾向の分類とその需要予測への適用可能性の検討

山梨大学土木環境工学科

学生員 ○中沢航太

山梨大学大学院医学工学総合研究部 正会員 佐々木邦明

1. はじめに

日本の高速道路はこれまで日本の国土形成と経済や産業の成長を支え、現在も総距離の延長と共に成長し続けている。平成 17 年道路関係四公団民営化により日本道路公団も NEXCO 三社(東日本高速道路, 中日本高速道路, 西日本高速道路)に分割民営化された。これによりさらに高速道路や高速道路周辺も整備され、より快適なものになりつつある。特に SA, PA は高速道路上にある単なる休憩所ではなく地域や先進的な企業と連携・提携した多様なサービスを提供する一種のコミュニティーの場となっている。また ETC 搭載の助成金や休日特別割引の導入により ETC 利用率、高速道路利用台数は年々増加傾向にある。平成 22 年 1 月 7 日現在で ETC 利用率は 83.2%, 利用台数は一日約 593 万台となっている¹⁾。

平成 13 年の ETC 導入より全国の高速道路で ETC を利用した高速道路の利用が継続的に蓄積されている。ETC データは入出 IC, 利用開始, 利用終了時間, 車種等が記録されている。IC 間 OD が正確かつ継続的に得ることができる。しかし一般的な交通に関する分析を行う上で必要とされる「行動目的」, 「個人属性」, 「OD」が含まれない。そのため道路交通センサスデータなどと同様の需要予測への活用ができない。そこで本研究では個人をベースにデータの縮約を行い、どのような利用パターンがあるのか、またそのパターンの需要予測への適用可能性について検討する。

ETC データを集計量として取り扱うことで、交通量の解析を行うことも可能であるが、今回はミクロ的な利用特性を把握するため個人別のアプローチを取った。ETC データから得られる個人の利用指標としては以下の様なものが挙げられる

- ・年間利用回数
- ・休平日別利用回数
- ・曜日別利用回数
- ・時間帯別利用回数
- ・往復, 片道利用回数
- ・車種

これらは個人の利用によって生成されるものであり、これらを統合して用いることで年間を通した利用パターンの指標として活用が可能であると考えられる。また利用パターンは利用目的によって生成されると考えられるため、これらのパターンは利用目的パターンと深く関係している。このように無限にある利用パターンをベースに高速道路の需要変動を分析することを本研究の目的とする。

2. 高速道路の利用特性

(1) データの概要

利用したデータは、平成 17.4 月~平成 18.3 月に対象区間内で基準 IC を出入りした ETC 搭載の乗用車, 軽乗用車のものであり、記録されている利用の総数は 2852996 である。また識別される ID による対象区間利用台数は 538464 台で、内 443356 台(全体の 82%)が普通車, 軽自動車となっている。

(2) データの基礎的特性

ID 別データの年間利用回数を見てみると 1, 2 回の利用者が全体の 70%を占めている。これらの利用者は各利用回数によって分類するには利用回数が少なすぎ、一括で分類するのは困難である。よって年間 1, 2 回と年間 3 回以上の利用者に分けて分類することとした。

キーワード: ETC, クラスタリング, 重回帰分析, 数値化理論 I 類

連絡先: 〒400-8511 山梨県甲府市武田 4-3-11

山梨大学土木環境工学科交通工学講座

TEL:055-220-8671

E-mail:t06c058@yamanashi.ac.jp

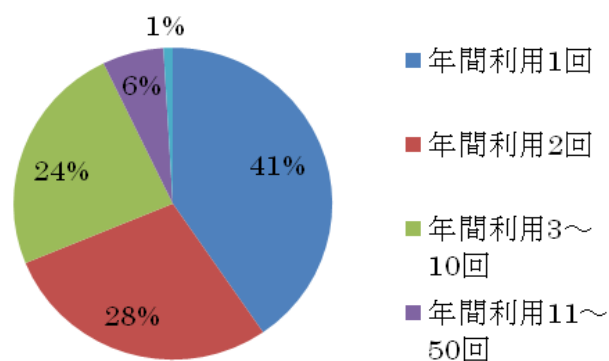


図-1 年間利用回数の比率

3. 事例研究

(1) 個人の利用特性分類の考え方

高速道路利用目的にはそれぞれ利用の仕方に特徴があると考えられる。例えば平日の出勤帰宅時間帯に往復での利用が数多く観測されていれば利用目的は通勤であると推測できる。このように「平日」「出勤時間帯」「帰宅時間帯」「往復」のようにいくつかの指標を組み合わせ分類し、その分類結果の傾向に基づいて、各クラスターの命名を行う。また年間利用回数のように変数の分散が大きい指標 (min.1 max.660) については、分類がその影響だけを受けないよう変数の標準化を行った。

年間利用回数 1, 2 回のクラスタリング結果を図-2 に示す(各クラスターをクラスター 1~4 とする)。クラスター数を 4 に設定し休平日、時間帯の利用回数を変数とした。図の縦軸は最終クラスター中心の値、横軸は分類に使用した各変数である。元々の利用回数が 1, 2 回であり、半数以上が 1 回であるため、利用回数に関する指標は 1 を超えない。図-2 をみると、片道利用回数と往復利用回数については、四つのクラスターすべて同じような傾向にあり、1, 2 回の利用においては往復がほとんど無いことがわかる。クラスターによって違いが出たのは、平休日別の利用回数と利用時間帯であり、それらに基づいて各クラスターの特徴が把握できる。具体的には昼間の利用の多いクラスター 1, 4 は、休平日によって識別されている。クラスター 1 は平日の昼間に 1, 2 回の利用を行ったクラスターであり、クラスター 4 は休日に 1, 2 回の利用を行ったクラスターとなる。それに対して時間帯に

区別無く存在するクラスター 2 は休日利用に特化したものである。また、クラスター 3 は出勤時間帯に片道を平休日関係なく利用したクラスターと考えられる。

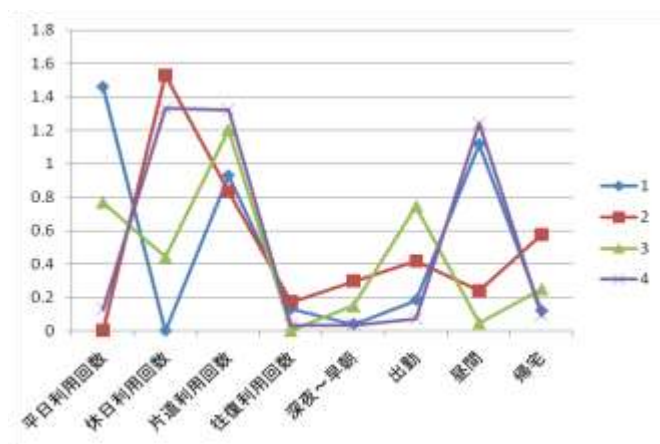


図-2 年間利用回数 1, 2 回クラスタリング結果

年間利用回数 3 回以上のクラスタリング結果を図-3 に示す(各クラスターをクラスター 5~8 とする)。こちらもクラスター数を 4 に設定し、曜日、時間帯、往復・片道の利用率を変数とした。図の縦軸は最終クラスターの中心の値、横軸は各変数である。また図の Z 得点とは変数の値を平均 0, 標準偏差 1 に標準化したことを意味している。平日往復、休日往復、連休往復、片道を中心に各クラスターの特徴がよく表れている。各クラスターの特徴からクラスター 5 は日帰り観光クラスター、クラスター 6 は通勤クラスター、クラスター 7 は宿泊観光クラスターと推定できる。クラスター 8 は片道利用傾向であること以外顕著な特徴がないため明確な利用目的の推定が困難なクラスターである。

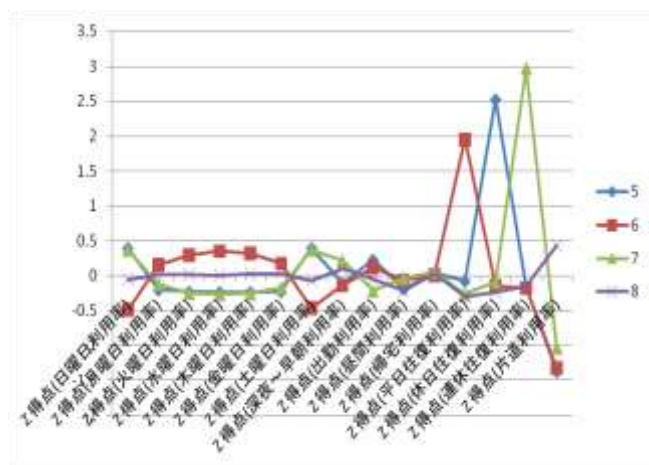


図-3 年間利用回数 3 回以上クラスタリング結果

以上の二つのクラスタリング結果から各クラスタの特徴と割合をまとめて表-1に示す。

表-1 各クラスタの特徴と割合

クラスタ	ケース数	%	年間利用回数	休平日	利用時間帯	往復片道
1	127541	6.6	1,2回	平日	昼間	片道
2	71710	3.7		休日	不定	片道
3	83623	4.3		両方	出勤	片道
4	144687	7.5		休日	昼間	片道
5	169661	8.8	3回以上	休日	出勤帰宅	休日往復
6	441183	22.8		平日	出勤帰宅	平日往復
7	76542	4.0		休日	帰宅時間～早朝	連休往復
8	817422	42.3		両方	帰宅時間～早朝	片道

(2) 利用特性分類に基づく回帰予測

利用データに対応させたクラスタ番号から各日のクラスタの構成比を算出する。被説明変数を各日の利用台数、説明変数を各日のクラスタ構成比とし重回帰分析を行った。つまり各日別の交通量を、その日の交通量を規定する日の特性の代理変数としてクラスタの構成比を用いて、それらの回帰式による予測を行ったものである。その推計値と実績値を図-4に示す。

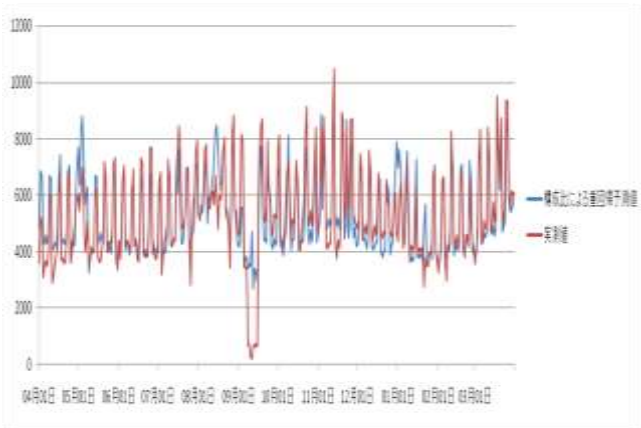


図-4 重回帰分析による予測

分散分析の結果から有意水準 5%で分析による回帰式は有意であると考えられる。重相関係数は 0.846 でやや強い相関が見られる。決定係数は 0.716 となり回帰式の精度は良く、誤差はあるがある程度利用傾向は把握することができている。また補正決定係数は 0.707 となった。これにより各日の利用目的を推定したクラスタ構成比が得られればある程

度の予測が可能であると考えられる。しかし特性を構成比によって表現するという事は、日別の特性をのぞいて 1 年間を同じような日として想定して、利用実績に基づいて特性を決定し、日別の交通量を推定したことになり、利用実績がある場合には予測可能であるが、この構成比が将来にわたっても安定的であるということが必要となってくる。

(3) 数量化理論 I 類を用いた需要予測

(2)から各日のクラスタ構成比がわかればある程度利用台数の予測が可能であることがわかった。これまでの基礎的な集計やクラスタリングの段階で各日のクラスタ構成比は曜日、天候、祝日等の休平日に相関していることが分かっている。そこで各日の特性をこれらの変数を用いて説明することを考える。しかし 5 月の GW、年末年始、長期休業などの季節的な要因は曜日ダミーでは説明することは難しく、その上季節としての線引きが困難である。そこで一年を日曜から土曜の一週間で区切り、52 の週に分割し各週のクラスタ構成の傾向ごとに分類することとした。なおデータが 1 年分であり 4 月の第一週と 3 月の第四週は週として成り立っていない。そこで一年でほぼ同じ変動をするものと仮定し、欠損している 3 月第四週の土曜データに 4/2(土)のデータを充て、4 月第一週と 3 月第四週を同一週とした。この分類にも非階層クラスタリングを用い、クラスタ数を 6、説明変数を各クラスタの変動とし分析した。分類した結果を図-5に示す。

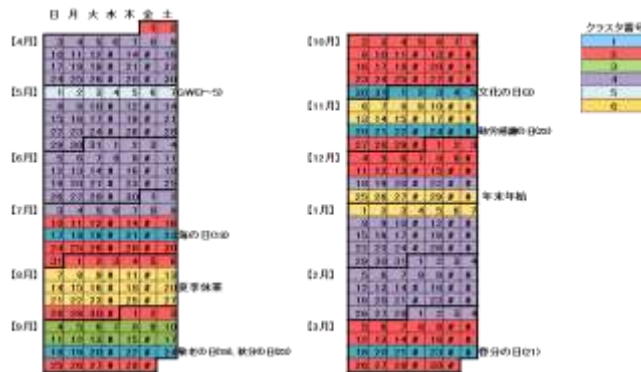


図-5 各週の分類

図-5 からクラスタ 1 は祝日のある週。クラスタ 2 は祝日週の前後にあり、その影響を受ける週。クラスタ 3 は急激に利用データが減少している 9 月の第一、二週。クラスタ 4 は変動の少ない普通の週。

クラスタ 5 は 5 月の GW の週. クラスタ 6 は夏季休業, 年末年始と 11 月の第二, 三週となった. 分析により月を超えた季節的な特徴を明確に表した週ごとの特徴が示された. そこで各日のクラスタ構成比を被説明変数, 週の特徴を表すデータである曜日, 天候, 休平日に加え, 上記の週の分類を説明変数とし数量化理論 I 類によって分析を行った. 分析によって得られた各日のクラスタ構成比の推計値と重回帰分析で求めた回帰式より利用台数の再現を行った. つまり, 週を利用状況によって 6 種類に分類し, これが 1 年間では普遍であると想定した上で, 週ごとの実際の天候などの要素を加味して一日の特性を示すクラスタ構成比に変換し, 推計を行ったものである. その結果を図-6 に示す.

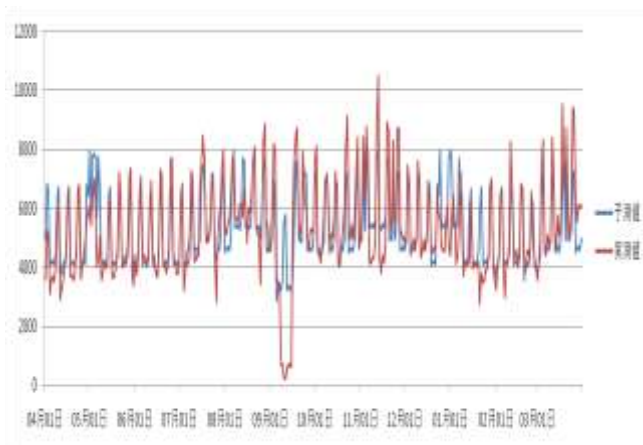


図-6 数量化理論 I 類による予測

こちらもある程度の予測は出来ているものの細部に亘る再現はできておらず, 各日の利用台数を予測出来ているとは言い難い. 数量化理論 I 類による各クラスタ構成比の予測ではクラスタ 1~2, 4~7 の構成比は決定係数より 85%以上でそれぞれの回帰式によって再現できているが, クラスタ 8 は 57%, クラスタ 3 に至っては 50%しか再現できていない. この 2 つのクラスタ構成比の再現の低さが全体の予測精度の低下に影響している. 特にクラスタ 8 は全ケースの 40%を占める大きなクラスタであり, 複数回利用しているが, その利用方法に目立った特徴のないクラスタである. またクラスタ 3 は, 年に 1, 2 回の利用で出勤時間帯に利用するクラスタであり, 突発的な利用であると考えられる. これらの予測精度を上げることが重要になってくる. これらの改善のためには, 今回考慮しなかった利用時間帯別のク

ラスタ構成比の特徴を元に更なる変動要因を導入する必要があると考える. それにより時間帯ごとの予測を行い, より精度の高い予測が行えるのではないかと考察する.

4. おわりに

蓄積された一年分の高速道路利用データを ID ベースのデータに変換し, 非階層のクラスタリングを用い利用目的, 利用傾向別に分類した. クラスタ構成比を説明変数とした重回帰分析からある程度の予測が行える. しかし現状では各日の利用台数を正確には再現できていない.

次に数量化理論 I 類を用い, 質的データからクラスタ構成比の推定を検討した. 曜日, 天候, 休平日に加え季節的な要因を考慮するために週ごとのクラスタリング結果を説明変数として導入した. 推定したクラスタ構成比と重回帰分析によって求めた回帰式によって各日の利用台数の再現を行った. 結果は真のクラスタ構成比による重回帰分析の予測と同等の結果が得られた. しかしこちらも細部に亘る予測はできておらず残差の大きい日の分析から更なる変動要因の導入が必要である.

謝辞

本研究は, NEXCO 三社から高速道路の利用特性分析の検討依頼を受け実施したものである. NEXCO 及び高速道路調査会には, 所要時間に関する貴重なデータ提供と共に有意義なコメントを頂いた. この場を借りてお礼の言葉を申し上げたい.

参考文献

- 1)国土交通省道路局ホームページ
<http://www.mlit.go.jp/road/yuryo/riyou.pdf>
- 2)ETC 総合情報ポータルサイト
<http://www.mlit.go.jp/road/yuryo/riyou.pdf>
- 3)NEXCO 中日本ホームページ
<http://www.c-nexco.co.jp/>