経路検索履歴データを用いた事前検索時間の分析

鳥取大学　学生会員　○古川ゆり
鳥取大学　正会員　桑野将司
鳥取大学　非会員　秋元美穂奈
鳥取大学　非会員　菅原一孔

1. 背景と目的

2016 年に内閣府が発表した世論調査[1)]によると，公共交通の経路を調べる際，経路検索システムが最も多く利用されている．このことから検索の際に表示される検索結果が検索者の公共交通の利用に影響を与えていると考えられる．この経路検索システムには，検索者が入力した出発地や目的地，指定した日時等の情報がログデータとして蓄積されている．このログデータは移動に関する希望を含んだデータであり，ニーズに応じた交通計画の立案に寄与すると考えられる．しかし，ログデータを用いた研究の蓄積は少ない[2) 3) 4)]．また既往研究のほとんどが都市圏を対象とした分析であり，地方都市を対象とした研究は極めて少ない．

本研究では鳥取県と島根県に導入されている経路検索システム「バスネット」[5)]の鳥取県内の検索に関するログデータを対象とし，検索者が公共交通を利用する何分前に検索を行っているか，すなわち指定時刻と検索実施時刻の差を分析し，状況別の検索行動の特性を明らかにすることを目的とする．

2. 使用データの概要

有効データは，2016 年 9 月の 1 カ月分のデータ 24,533 件のうちログデータ項目に空白がある検索，2017 年を指定した検索を欠損値として除いた 22,678 件とする．バスネットのログデータ項目を表 1 に示す．

バスネットの起終点の指定方法には地図上でカーソルを動かし起終点を設定する緯度経度検索と特定の場所名を入力するランドマーク検索の 2 種類が存在する．出発地と目的地の双方をランドマークで指定した検索は全体の 91%にあたる 20,690 件である．検索者のほとんどがランドマーク検索を行っていることから，本研究ではランドマーク検索されたデータのみを分析する．

バスネットの時刻指定方法は表 1 の項目「時間指定方法」に示す通り，4 種類存在する．本研究は特定の時刻指定をする検索を対象に，指定時刻と検索実施時刻の差を「事前検索時間」と定義し，事前検索時間に影響を及ぼす要因を分析する．なお，バスネット固定端末である多機能バス停を使用した検索は，システム上，事前検索時間がすべて 0 分となるため分析対象からは除外する．起終点の指定方法，対象地域，時刻指定方法，アクセス端末によってデータをスクリーニングした結果，図 1 に示す通り， 10,402 件の検索ログデータが抽出された．

表 1　バスネットログデータ項目

項目	備考
日時	検索実施時刻，指定時刻
起終点	バス停名，駅名，施設名，緯度経度
アクセス端末	携帯電話，パソコン， スマートフォン，タブレット端末， 多機能バス停，アンドロイドアプリ
発着指定	出発指定，到着指定
時間指定方法	特定の時刻指定， 検索時刻から何分後に出発， 始発時刻，終発時刻
指定交通機関	鉄道を含む，バスのみ
指定バス会社	日ノ丸自動車，日本交通， コミュニティバス，一畑バス
経路検索結果	出発地，目的地，経由地， 利用公共交通の発着時間，所要時間， 待ち時間，利用公共交通の路線

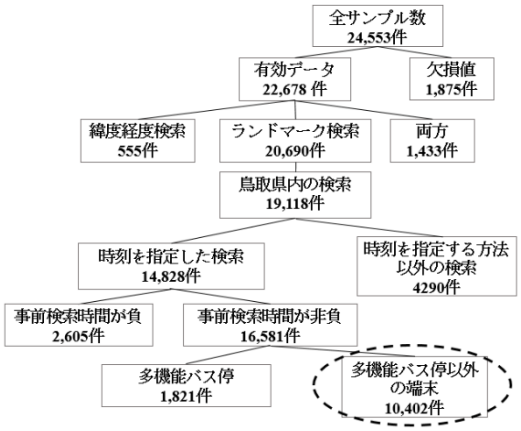

図 1　分析対象データ選定方法

キーワード　指定時間，検索実施時間，生存時間分析
連絡先　〒680-8550　鳥取県鳥取市湖山町南 4 丁目 101
TEL0857-31-531

3. 基礎集計分析

a）検索件数に関する基礎集計分析

1 日あたりの平均検索件数は 346 件であった．指定曜日別の検索件数に着目すると，木曜日を指定した検索が 236 件で最も多く，日曜日を指定した検索が 117 件で最も少ないことがわかり，曜日によって経路検索システムの利用状況が異なることが明らかとなった．

図 2 に乗車希望時間帯別と検索実施時間帯別の 1 日あたりの検索件数を示す．ここで，出発時刻を指定した検索は検索者が指定した時間帯，到着時刻を指定した検索は経路検索結果に表示される検索者が最初に利用する公共交通機関の時間帯を乗車希望時間帯とする．乗車希望時間帯別の検索件数は，8 時台と 12 時台が多く，この時間帯の公共交通機関を検索していることが多いことがわかる．21 時台から 6 時台の公共交通機関を検索しているのは全体の約 4%にあたる 13 件であった．一方，検索実施時間帯に着目すると 10 時台，11 時台に検索を実施しているものが多い．また，21 時台から 6 時台にかけて検索を実施しているのは全体の 21%にあたる 73 件であった．以上より，乗車希望時間帯と検索実施時間帯は必ずしも一致していないことが明らかとなり，検索者は予め公共交通機関のスケジュールを確認していることがわかった．

バスネット利用者が検索に使用したアクセス端末に着目すると，携帯電話・スマートフォンが 57%，パソコンが 37%その他の端末が 6%であり，携帯電話・スマートフォンからのアクセスが最も多いことがわかった．

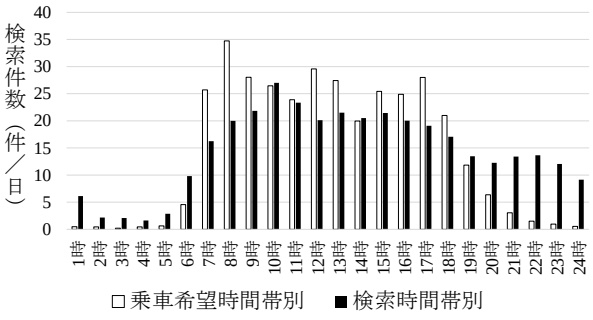

図 2　乗車希望，検索時間帯別検索件数

b）起終点に関する基礎集計分析

図 3 に指定された目的地と出発地の検索件数を示す．なお，ここでは全検索件数 10,402 件に対して 1%にあたる 104 件以上目的地および出発地のどちらかが，検索された地点のみを表示している．図 3 より，出発地，目的地ともに鳥取駅（バス停）を指定した検索が最も多い．また，鳥取大学は出発地に指定されるよりも目的地として指定される件数が多い．このように，出発地または目的地として検索された件数には地点ごとに差が存在し，出発地と目的地では検索時に指定される傾向に違いがあることが明らかとなった．

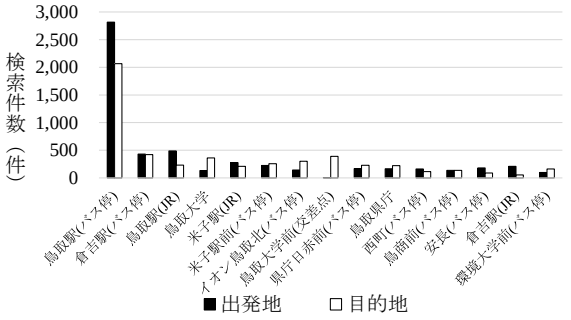

図 3　出発地または目的地として指定された件数

d）事前検索時間に関する基礎集計分析

図 4 に 1440 分（24 時間）まで 5 分単位での事前検索時間の分布を示す．事前検索時間が 0～4 分となる検索が最も多く全体の 20%を占めている．平均事前検索時間は約 4236 分（70 時間 36 分）であった．つまり，公共交通を利用する直前に検索する人もいれば，数日前に検索する人もおり，分布としては右に裾が長い形状となることがわかった．

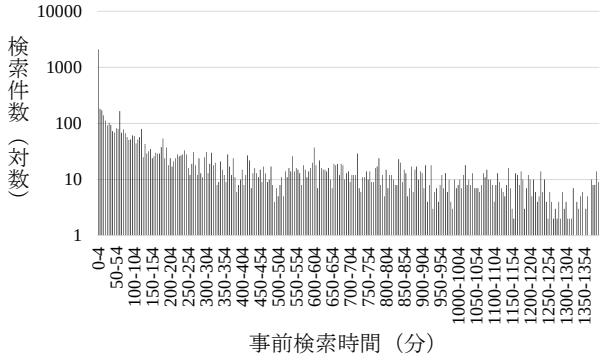

図 4　事前検索時間の分布

4. 生存時間分析の概要

本研究では，対象とする事象が生起するまでの時間（生存時間）を解析の対象とする生存時間分析（パラメトリックアプローチ）[6] を用いる．

生存時間モデルでは，対象とする事象が生起するまでの時間の分布を生存関数，およびハザード関数で表す．本研究では，対象とする事象を事前検索とし，生存時間分析による事前検索時間 (T) のモデル化を試みる．生存関数 $S(t)$ とハザード関数 $h(t)$ は，累積分布関数 $F(t)$ と確率密度関数 $f(t)$ により次のように表される．

$$S(t) = \Pr(T > t) = 1 - \Pr(T \le t) = 1 - F(t) \quad (1)$$

$$h(t) = f(t) / S(t) \quad (2)$$

ここで，T は事前検索の発生時間を表す確率変数，t は事前検索時間を示す．説明変数は次の形で導入される．

$$h(t \mid \mathbf{X}_i) = h_0(t)\exp(-\boldsymbol{\beta}\mathbf{X}_i) \qquad (3)$$

$\mathbf{X}_i$ はサンプル i の説明変数ベクトル，$h(t \mid \mathbf{X}_i)$ は事前検索時間 t におけるサンプル i のハザード関数，h_0 は基準ハザード関数，$\boldsymbol{\beta}$ は未知パラメータベクトルをそれぞれ表す．

モデルの推定に際しては，式（4）に示す尤度関数を最大化することによって，未知パラメータを推定する．

$$L = \prod_i \{h(t_i \mid \mathbf{X}_i) \cdot S(t_i \mid \mathbf{X}_i)\} \qquad (4)$$

5. 推定結果

事前検索時間に影響を及ぼす要因として，本研究では表 2 に示すような 8 項目 56 個の変数を考慮し，これらをモデルの説明変数の候補とする．モデル推定時の説明変数の選定では，パラメータの統計的有意性が 5%以下のものをモデルに残すこととした．

表 2　説明変数候補

項目	備考	個数
天候	雨・曇り	1
指定交通機関	バスのみ	1
発着時刻指定	到着時刻指定	1
使用端末	パソコン，携帯電話	2
指定曜日	土・日	1
指定出発地	鳥取駅，鳥取県庁など	17
指定目的地	鳥取駅，鳥取県庁など	17
乗車希望時間帯	6 時台～21 時台	16
計		56

推定結果を表 3 に示す．ここで，パラメータ値は正で絶対値が大きくなるほど事前検索時間が長くなり，負で絶対値が大きくなるほど事前検索時間が短くなる傾向を意味する．以下パラメータ推定結果の考察を行う．

天候に関して，雨・曇りのダミー変数に関するパラメータ推定値は，負で有意となった．これは，雨の日や曇りの日は，事前検索時間が短くなることを意味する．天候の良い日は公共交通を利用しない人が，天候の悪い日になって手段変更を行い，公共交通の利用直前に経路の検索することが影響していると考えられる．

発着時刻指定に関しては，到着時刻指定に関するダミー変数が正で有意となり，到着時刻を指定した検索は事前検索時間が長くなることがわかった．到着時刻を指定した検索者は，「何時までに到着したい」という移動に対する制約が強く，予め公共交通の時刻表を確認する傾向が高いことを示していると考えられる．

使用端末に関するパラメータ推定値に着目すると，携帯電話からの検索を示すダミー変数は負で有意に，

表 3　推定結果

明変数	推定値		t 値
天候			
雨・曇り	-0.107	**	-5.05
指定交通機関			
（バスのみ＝1，鉄道を含む=0）	0.209	**	9.82
発着時刻指定			
（到着時刻＝1，出発時刻＝0）	0.256	**	10.10
使用端末			
携帯電話	-0.113	**	-2.69
パソコン	0.479	**	10.96
指定曜日			
（土日=1，平日=0）	0.066	**	2.87
指定出発地			
鳥取駅（バス停）	-0.154	**	-5.81
鳥取駅（JR）	0.194	**	2.96
倉吉駅（バス停）	0.449	**	8.56
倉吉駅（JR）	0.545	**	3.93
米子駅（JR）	0.381	**	4.52
県庁日赤前（バス停）	-0.465	**	-5.83
西町（バス停）	0.250	**	3.01
鳥商前（バス停）	-0.341	**	-3.75
安長（バス停）	-0.349	**	-4.48
環境大学前（バス停）	-0.572	**	-5.52
鳥取砂丘コナン空港	1.240	**	5.09
白兎神社・海岸	0.832	**	7.59
鳥取砂丘	0.612	**	6.94
指定目的地			
鳥取駅（バス停）	-0.101	**	-3.53
鳥取駅（JR）	0.243	**	3.07
鳥取大学	-1.114	**	-24.53
米子駅（バス停）	0.167	**	2.30
米子駅（JR）	-0.249	*	-2.83
イオン鳥取北（バス停）	-0.153	*	-2.24
鳥取県庁	0.322	**	3.89
西町（バス停）	0.550	**	5.55
鳥商前（バス停）	-0.594	**	-6.46
安長（バス停）	-0.704	**	-6.22
環境大学前（バス停）	-0.253	**	-3.07
鳥取砂丘コナン空港	0.651	**	0.65
白兎神社・海岸	0.729	**	0.73
鳥取砂丘	0.476	**	0.48
乗車希望時間帯			
6 時台	0.641	**	0.64
7 時台	0.338	**	0.33
8 時台	0.366	**	0.37
9 時台	0.288	**	0.29
10 時台	0.155	**	0.15
11 時台	0.294	**	0.29
12 時台	0.308	**	0.31
13 時台	0.286	**	0.29
14 時台	0.216	**	0.22
16 時台	0.095	*	0.10
18 時台	0.119	**	0.12
尺度パラメータ	0.096	**	18.67
形状パラメータ	0.399	**	134.25
サンプル数	10,402		
対数尤度	-162,061.0		

**:1%有意 *:5%有意

パソコンからの検索を示すダミー変数は正で有意となった．すなわち，携帯電話を使用する検索者はパソコンを使用する検索者に比べ，事前検索時間が短くなる傾向を示しており，直感に即した妥当な結果であるといえる．

指定曜日に関しては，土日を指定した検索を示すダミー変数が正で有意となり，土日を指定した検索は事前検索時間が長くなることがわかった．これは，検索者は土日の予定を計画する段階で経路検索システムを利用することを示していると考えられる．

指定出発地，指定目的地に関するパラメータ推定結果に着目すると，観光客の利用が多いと考えられる鳥取砂丘コナン空港や白兎神社・海岸，鳥取砂丘を指定し検索したことを示すダミー変数は正で有意となった．これらの地点を指定したパラメータ推定値は，他のパラメータ推定値より値が大きく，他の地点を指定した検索者に比べ事前検索時間が長くなる傾向を示している．このことから，観光を目的とする検索者は計画段階で経路を検索していると考えられる．また，鳥取県庁を目的地に指定したことを示すダミー変数は正で有意となっており，事前検索時間が長い傾向にあることを示している．これは手続きや会議を目的とする人が多く，時間を遵守するために，あらかじめ余裕を持って行動する必要を示している．さらに，周辺に商業施設が多く立地している安長（バス停）を出発地，目的地に指定した検索やイオン鳥取北を目的地に指定した検索を示すダミー変数は負で有意となっており，事前検索時間が短い傾向にある．これは，遊びや買い物を目的とする人が多く時間に強くとらわれないためだと考えられる．

乗車希望時間帯に関しては全ての変数が正で有意となった．特に 6，7，8 時台の早い時間帯を指定した検索を示すダミー変数が大きい値で有意となり，この時間帯を指定した検索は事前検索時間が長くなることがわかった．これは朝に移動を希望した検索者は前日の夜の時点で経路を検索していると考えられる．

6. モデルの現況再現性と事前検索時間の予測

推定した事前検索時間の生存関数の現況再現性を検証するために，モデル推定結果による事前検索率と実測値の比較を図 5 に示す．図 5 より推定値と実測値の大きな乖離は見られない．モデル推定結果を用いて，事前検索時間の中央値の算出を行った結果，実測値は 324 分に対して，推定結果を用いた予測値は 320 分であった．実測値と予測値の差は 4 分，誤差は 1.25%と高い精度で予測できていることがわかる．

次に，出発地と目的地の組み合わせ別による事前検索時間中央値を算出した．鳥取駅（JR）を出発地に，目的地に観光地である鳥取砂丘を指定した場合の事前検索時間中央値は 34 時間 42 分であることがわかった．また，日常生活においての検索だと考えられる，鳥取駅（JR）を出発地にイオン鳥取北（バス停）を目的地に指定した検索の場合の事前検索時間中央値は 7 時間 10 分であった．以上のように，モデル推計結果を用いることによって，様々な状況に応じた事前検索時間の算出が可能になったといえる．

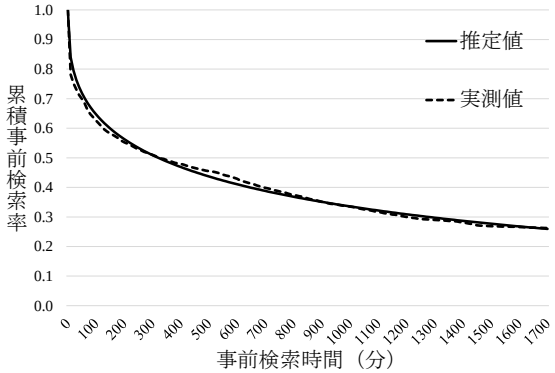

図 5　生存関数の比較

7. 結論

本研究は経路検索システム「バスネット」に蓄積されているログデータに生存時間分析を適用し，事前検索時間の分析を行った．モデルの推計結果を用いて期待事前検索時間の中央値を算出したところ，予測値と実測値の差は小さく，提案した分析方法によって高い精度で事前検索時間の予測が可能であることを示した．また，中央値を用いることで多様な状況に応じた具体的な事前検索時間の算出が可能になった．検索行動である事前検索時間を把握することで，事前に検索する段階で経路案内に乗客のニーズにより合致した情報，例えば周辺の観光地の状況や他の交通手段に関する情報などを付加することができる．これにより，観光誘発や他の交通手段の利用促進に繋げることが可能であると考える．

参考文献

1）公共交通に関する世論調査，内閣府，2016. https://survey.gov-online.go.jp/h28/h28-kotsu/gairyaku.pdf .

2）石村怜美，梶原康至，太田恒平：乗換検索サービスの経路選択経路検索データを用いた公共交通の経路選択行動分析，土木計画学会・講演集，Vol.49，CD-ROM，2014.

3）香川喬之，桑野将司，福山敬，谷本圭志，川村尚生：バス経路検索履歴データを用いた移動特性の分析，交通工学論文集，第 2 巻，第 2 号（特集号 A），pp.A_115-A_124, 2016.

4）石村怜美，太田恒平，富井規雄，経路検索サービスの実績データに基づく近未来の突発的移動需要の検出，土木計画学会・講演集，Vol.49，CD-ROM，2013.

5）バスネット，バスネットについて，（2018 年 4 月 2 日確認）https://www.ikisaki.jp/home/about.

6）J.P.クライン，M.L.メシュベルガー，生存時間分析，シュプリンガー・ジャパン株式会社，2009.