RFMC 分析における Clumpiness 指標の拡張と自社顧客の 行動予測への応用

Clumpiness を活用した離脱時期と競合利用の予測手法の提案

新 美 潤一郎*, 星 野 崇 宏**

Prediction of the Customer Activity with Using the Improved Measure of Clumpiness
Capturing Competitive Usage and Early Churn under Limited Information

Junichiro NIIMI* and Takahiro HOSHINO**

In recent, the concept of clumpiness has received increasing attention as a novel method called RFMC in Customer Relationship Management (CRM). However, it has been not so long since clumpiness is proposed, there still remains difficulties. In this paper we firstly identify the problems of clumpiness in both statistical and practical fields, propose the improved indication of clumpiness and apply it to some of the most preferential problems for firms such as the detection of early churn in smartphone game market and the prediction of consumer's share-of-wallet (SOW) in grocery market to validate its usefulness in marketing analysis.

Key words: Marketing, E-Commerce, purchase prediction, CRM, RFM, Clumpiness, ID-POS, location data

キーワード: マーケティング, 電子商取引, CRM, RFM, Clumpiness, 購買履歴, 位置情報

1. は じ め に

電子商取引の発達やスマートフォンの普及といった環境の変化により, 小売業界ではチャネルの多様化や価格競争などを通じて競争が激化している. このような現状において, 企業では顧客関係管理 (CRM) として自社で保有する顧客のロイヤルティの高さを把握

図 1. RFM 分析では区別できない消費者の異質性

する手法が様々試行されてきた. 特に, 顧客の自社利用に関する利用時期の直近さ (Recency), 利用頻度の多さ (Frequency), 支出額の多さ (Monetary) の 3 つの指標を用いた RFM 分析は, 自社顧客のセグメンテーション等を目的に実務においても広く活用されてきた (Hughes, 1995; Bult & Wansbeek, 1995 等).

しかし図 1 の架空の 2 人の顧客の購買状況が示すように, 同一の RFM 情報でも一方は定期利用, 他方は不定期にまとまった利用を行っており, RFM だけでは識別できない消費者の購買パターンの異質性が存在している. このようなイベント発生不均一性を捉

*名城大学経営学部

(Department of Business Management, Meijo University)

連絡先: 〒468-0073 愛知県名古屋市天白区塩釜口 1 丁目 501 番地 名城大学天白キャンパス 経営学部

E-mail: jniimi@meijo-u.ac.jp

**慶應義塾大学大学院経済学研究科・経済学部, 理化学研究所

(Department of Economics, Keio University & RIKEN AIP)

連絡先: 〒108-8345 東京都港区三田 2-15-45

E-mail: bayesian@jasmine.ocn.ne.jp

えることを目的に, Zhang, Bradlow & Small (2015) で RFM に Clumpiness を加えた RFMC 分析が提案されている (本研究ではこれ以降, 購買等のイベントがまとまって発生する現象それ自体を「Clumpiness」, Clumpiness を伴う振る舞いやそのような行動をする顧客を「clumpy な振る舞い／顧客」, RFMC 分析における Clumpiness 指標を「C 指標」と表記する)。

この Clumpiness は CRM において大きく着目されている手法だが, その一方で提案されて間もない概念であることから未だ不明瞭な点が多く, 活用に課題が残されている。そこで本研究ではこの Clumpiness が抱える問題について, マーケティング的な観点, 統計学的な観点のそれぞれから議論し, 改善した新指標とそれを用いた解析例を提案する。したがって本研究の構成は, 第2節で RFMC を含む CRM 手法全般についての既存研究をレビューし先行研究から Clumpiness の現状を整理するとともに, マーケティング的な観点からの問題意識として, 先行研究では着目されていない Clumpiness 活用の意義について述べる。次に第3節では統計学的な観点からの問題意識として, 離散変数である C 指標ではイベント発生のまとまりの度合いが識別できないことなど実務応用にあたっての統計的な問題点を明らかにするとともに, より適した形での C 指標の拡張について提案する。第4, 5節ではそれぞれ提案指標を用いた顧客行動の予測に関する解析例を示すが, 第4節ではスマートフォンのゲームアプリの利用に関して, 第5節では実店舗での購買行動に関してのモデリングを行うことで, Clumpiness の概念や本研究の提案指標を用いた CRM が, 一般的な購買行動から契約関係を持たないサービス利用にまで幅広く応用可能であることを実データを用いて示す。最後に第6節で提案指標についての考察と今後の課題について述べる。

2. 既存研究

本節では本研究に関係する既存研究について, i) 代表的な CRM 手法, ii) RFM 分析とその問題, iii) Clumpiness と RFMC 分析, の3点にわけて整理する。

2.1. 代表的な CRM 手法

まず CRM の各手法を用いた顧客管理にあたっては, 顧客の自社への忠誠心の高さとしてのロイヤル

図2. Clumpiness を含む解析の構造

ティの推定が不可欠である。ロイヤルティには実際に自社を訪問／利用する頻度／金額等によって表すことのできる行動的ロイヤルティと, 消費者自身の認知・態度に基づき口コミなどに繋がる態度的ロイヤルティが存在する (Jacoby & Chestnut, 1978)。CRM 手法はその多くが行動的ロイヤルティの推定を目標としており, 代表的なものとして単一企業内での顧客の将来期間にわたっての支出の合計金額を割引現在価値として表した顧客生涯価値 (CLV: Pfeifer, Haskins & Conroy, 2005) や, 市場全体における自社への支出額／利用頻度の割合をもって定義されるシェア・オブ・ウォレット (SOW: Du, Kamakura & Mela, 2007), 本研究でも扱う RFM 等が存在する。

ここで本研究で扱う消費者行動に関する領域を俯瞰すると, 図2に示すように本研究で扱う Clumpiness を含めた消費者の行動との関係性を表すことが可能である。それぞれの関係とそれを用いた既存研究について以下に示す。

まず① (自社・競争の利用状況を用いた離脱時期のモデリング) は, CLV を用いた顧客の将来購買や離反時期の予測 (Berger & Nasr, 1998; Blattberg, Kim & Neslin, 2008), RFM を用いたものには CLV の推定やセグメンテーション (Colombo & Jiang, 1999; Sohrabi & Khanlari, 2007; Khajvand, Zolfaghar, Ashoori & Alizadeh, 2011 等多数) がある。次に② (自社で得られない競争他社の利用状況のモデリング) は SOW の文脈で実施されることが多い。代表的なものとしてクレジットカードの支出, 複数銀行間での預金額, 小売企業の売上等に関する SOW の予測 (Du et al., 2007; Gladly & Croux, 2009; Chen & Steckel, 2012; Mägi, 2003, 新美・星野, 2015/2017 など多数) 等がある。

このように既存研究として①, ②のような解析は広

く行われているが、実際の企業活動を考えると自社で保持する顧客といえども一般的には競合利用に関する情報は得られないことから、①の関係性についても不完全情報に基づいて実施することになる。例外として近年では企業が本来得ることのできない自社顧客の競合店舗への訪問履歴をスマートフォンの位置情報から把握することで、実店舗での SOW を考慮した購買予測（新美・星野, 2018）なども行われているが、この手法は実店舗の小売企業において有効なものであり、広く一般の競合関係に対して応用できる手法ではない。また③（消費者の心理的要因を用いた離脱時期のモデリング）、より広くいえば消費者の異質性を用いた離脱予測については、たとえば新美・星野（2015）では消費者の異質性を多様性変数として指標化し離脱時期や競合利用の予測に取り入れているが、多様性指標は行動の幅広さを指標化しているに過ぎず、本研究のように心理的要因にアプローチするものではない。さらに心理的要因を含めた観測できない内的状態を取得するためには追加的な調査が必要になる点からも、こういった心理的要因との関係についても行動ログの解析を通じて分析できる手法が必要である。

そこで本研究では、後述する通り Clumpiness が持つと考えられる 2 つの性質「自社サービスの利用にあたって将来的なブランド・スイッチングに影響を与える心理的要因を内包している可能性」「企業が自社では本来得ることのできない競合他社の利用状況に関する情報を内包している可能性」（図ではそれぞれ④、⑤）の構造から、Clumpiness を本来得られない心理的要因や競合利用の情報の代理変数として活用し顧客の離脱や競合予測に利用すること（図では⑥）で顧客行動の予測精度が高まる可能性について検討する。

2.2. RFM 分析とその問題

本研究では CRM 手法のうち特に RFM 分析に着目している。このような RFM 分析による顧客のスコアリングは企業と消費者の間に明示的な契約関係を持たないような商品に対して特に有効とされている（Mulhern, 1999; Bell, Deighton, Reinartz, Rust & Swartz, 2002）が、そもそも契約関係を必要としない商品では顧客の正確な離反のタイミングを観測することができないという問題点があり、R, F, M の各指標を用いて消費者の購買行動を簡潔に要約し把握できることから広く活用されている（阿部, 2014）。だがその一方で、解析に用いる期間の区切り方によって RFM

表 1. Clumpiness を扱った先行研究での対象と影響の比較

研究	解析対象	Clumpiness の影響
Zhang et al. (2013)	サービス（動画視聴）	CLV の低下
Zhang et al. (2015)	購買（小売業）	離脱時期の二極化
中山 (2017)	購買（小売業）	将来的な大量購買の可能性

の各指標の値が異なってしまう問題（Fader, Hardie & Lee, 2005）や、離反までの期間の長さを統一的に設定してしまうと消費者の異質性を考慮できない問題（阿部, 2014）といった解析の恣意性が指摘されている。さらに 1 節の図にも示したように、消費者の異質性の考慮という点では、R, F, M の各指標上では同一の購買行動にも関わらず実際の利用目的が明らかに異なっている場合がある。

2.3. Clumpiness と RFMC 分析

Clumpiness の概念は Zhang, Bradlow & Small (2013) で最初に提案された。当初は動画ストリーミングサービスのようなメディア利用におけるまとまった視聴 (binge-watching) を捉えることを目的として、繰り返しを伴う行動（視聴、訪問、購買等）に広く適用可能な概念とされたが、その後 Zhang et al. (2015) で顧客のロイヤルティ推定にあたり既存の RFM 分析に加えて考慮すべき 1 指標として改めて提案された。同研究では数多くのデータを用いて Clumpiness と主要なマーケティング変数との関係について広く検討している。具体的には i) 訪問行動の Clumpiness (visit-C) と購買行動の Clumpiness (purchase-C) の関係、ii) Clumpiness とデモグラフィック変数の関係、iii) Clumpiness と離脱の関係、iv) Clumpiness と利用頻度の関係、v) Clumpiness と支出額の関係、vi) Clumpiness と RFM 指標の関係である。

ここで代表的な既存研究から Clumpiness が顧客のロイヤルティに及ぼす影響について表 1 にまとめると、対象とする製品・サービスによって Clumpiness が与える影響は大きく異なっていることがわかる。このように様々な検討結果が示される理由として考えられるのは、Clumpiness があくまで行動の結果を記録した変数でしかないという点である。星野（2009）において購買履歴をはじめとするログデータに収録された結果変数からその行動がどのような嗜好・属性に起因したものかを判別することは困難であることが指摘されていることから、対象とする製品・サービスご

とに Clumpiness の生起要因や影響はそもそも同質ではない可能性がある。例として Hulu のような契約型の動画視聴サービスではあらかじめ決まっていた観たい番組のみをまとめて視聴した上ですぐ解約してしまうような行動が考えられる一方で、中山 (2017) で用いられている家電購買のように日常的に頻繁に利用するとは考えにくい業種ではライフスタイルの変化に伴い定期的な大量購買が必要になる場合がある。これらはいずれも clumpy な行動だが、ロイヤルティという観点で考えればその結果は正反対である。

これら購買タイミングの不均一性という意味での Clumpiness の発生要因としてはいくつか考えられるが、本研究では Clumpiness の生起要因として、i) 製品・サービスの利用に関しての心理的要因による発生、ii) 競合利用に関する不完全情報としての発生の 2 点に着目している。

また Clumpiness 以外に購買タイミングの不均一性を捉える方法として、Platzer & Reutterer (2016) では Pareto/GGG モデルを提案している。Platzer らも述べている通り、Clumpiness がこれら不均一性の指標化を目的としている一方で、Pareto/GGG はこの不均一性を観測できない潜在変数として考慮した行動モデリングの手法である。

2.3.1. 製品・サービスの利用に関しての心理的要因による発生

Clumpiness の生起要因としてまず考えられるのが製品・サービスの利用に関しての心理的要因による発生である。Zhang et al. (2015) ではデモグラフィック変数など様々な変数との関係を幅広く明らかにしているが、繰り返しを伴うあらゆる利用において発生する Clumpiness が、一概に消費者の年齢、性別等の属性のみと関係しているとは考えにくい。ここでは繰り返しを伴うサービス利用の例として、解析 1 でも扱うスマートフォンゲームの利用形態から Clumpiness 発生の心理的要因について述べる。

まずスマートフォンで提供されるゲームアプリは一般的には基本使用料が無料であり、発展的な機能の利用に対して課金する形態を取っていることが多い。このようなゲームアプリの利用パターンには様々な考えられるが、特に Clumpiness が発生しやすいものとして、調査期間の冒頭にゲームを利用し始め、短期間で熱中して利用したのちに比較的早期に離脱してしまうようなパターンがある。これはログデータ上で見ると

期首に clumpy な行動をとったのち平均より早く離脱しているような場合である。このような「熱しやすく冷めやすいプレイヤー」を含め、心理的要因に起因した Clumpiness の発生は十分に考えられる。

そもそもマーケティング分野における消費者の「飽き」はバラエティ・シーキングの文脈で扱われることが多い。バラエティ・シーキングについて詳細にレビューした西原 (2012) では、バラエティ・シーキングに基づくブランド・スイッチングの発生要因として、最適刺激水準 (OSL, Raju, 1980; Raju, 1981) に基づく飽きがあることに触れている。clumpy な行動による短期の離脱に関していえば、プレイに熱中したことから短期間で刺激が通減し OSL を下回ったことで飽きを感じ、結果として早期に離脱するという構造が考えられる。西原 (2012) では高製品関与者における飽き解消のためのバラエティ・シーキングが存在する可能性が示されている。これは Clumpiness を用いて捉えられる可能性がある。

ただし、Zhang らも認めているようにこれら心理的要因などの潜在変数の状態に基づいて Clumpiness が発生すること自体は隠れマルコフモデル (HMM: Netzer et al., 2008; Schweidel et al., 2009) を用いることでも表現可能である。ここで HMM と比較した C 指標の利点は、RFM の各指標と同様に購買の不均一性を数値化できる点である (Zhang et al., 2015)。しかし、それら C 指標を顧客間で比較するには第 3 節で述べるような問題が残されており、大規模な活用は困難なのが現状である。

2.3.2. 競合利用に関する不完全情報としての発生

別の生起要因としては競合利用に関する不完全情報としての発生が考えられる。Clumpiness は RFM の 3 指標とともに用いるべき指標として提案されているが、RFM の 3 指標はそもそも消費者ごとの行動特性に由来して発生した購買行動全体のうち、1 企業について部分的に観測されただけの不完全情報に過ぎない (Fader et al., 2005)。つまり顧客に関して収集した RFM はあくまでも自社の購買履歴データのみに基づく一面的な情報であり、これは Clumpiness に関しても同様であると言える。

例としてスーパーマーケットでの食料品の購買を考えると、食品という性質上定期的な購買が必要となることから、競合利用まで含めた購買パターンの例として図 3 のような 3 つのパターンが考えられる。この 3

図 3. 継続利用が必要な市場での競合を含む利用パターンの例

パターンではいずれも競合購買まで含めると定期的な購買が行われているが、自社の購買履歴データ上では競合購買が観測できないことから、顧客 B、C では利用タイミングに偏りが発生し clumpy な行動だと判定される。

競合利用の情報は重要であるものの、実際の購買履歴データからは自社、競合の利用状況に由来した Clumpiness のみが得られる。消費者が同一市場において 1 日に何店舗もの異なる企業を利用することがないと仮定すれば、継続利用が必要な市場においては自社利用のばらつきが大きい（自社利用に大きな空白期間が存在する）ほど競合店舗を利用している可能性は高まる。この生起要因が正しいとすれば、Clumpiness の情報は企業が本来得ることのできない競合他社の利用状況に関する情報を内包している可能性があり、つまり Clumpiness のマーケティング的な意義として、i) 競合利用の予測にあたって Clumpiness を活用できる可能性、ii) 通常は得ることのできない競合利用の情報の代理変数として Clumpiness を活用して他の予測を行える可能性がある。

3. Clumpiness のモデル上の問題とその拡張

前述の通り Clumpiness の考慮は企業のマーケティング活動に有用な複数の分析に活用できる可能性があるが、その一方で活用にあたっての問題も抱えている。そこで本節では提案された C 指標の実務的な活用にあたっての統計学的な観点から 問題 I. 識別性の問題、問題 II. 同時比較が困難な問題の 2 点について述べるとともに、その解決策としての C 指標の拡張について提案する。

3.1. C 指標の抱える問題

C 指標の抱える問題の 1 点目として、C 指標における識別性の低さがある。そもそも C 指標の算出にあたっては、消費者 p の購買タイミングの不均一性

の指標としての H 指標 H_p と、Zhang et al. (2015) Appendix B のシミュレーションに基づく境界値 H_0 の比較が必要である。

まず H 指標については、離散の単位時間に基づく購買機会 N 回のうち n 回の購買が行われた場合に、 i 回目 ($i \in n$) に購買が行われたタイミングを t_i 、さらに t_i を基準化した x_i を

$$x_i = \begin{cases} t_1 & (i = 1) \\ t_i - t_{i-1} & (i = 2, 3, \dots, n) \\ N + 1 - t_n & (i = n + 1) \end{cases} \quad (1)$$

と定義した場合に

$$H_p = 1 + \frac{\sum_{i=1}^{n+1} \log(x_i)x_i}{\log(n+1)} \quad (2)$$

の形で表すことができる。

次に N 、 n の条件下でのランダムな購買行動のシミュレーションを M 回行い、その経験分布の上側 $\alpha\%$ の分位点を境界値 H_0 とすると、C 指標はこれらと比較して

$$C_p = \begin{cases} 1 & (H_p \geq H_0) \\ 0 & (\text{otherwise}) \end{cases} \quad (3)$$

と計算できる¹⁾。本研究ではすべての境界値の算出にあたり、 N は各解析での購買機会の回数に準じ、実データに C 指標を適用した中山 (2016) に倣って $M = 20000$ 、 $\alpha = 0.05$ でシミュレーションを実施している。ここでも記している通り境界値 H_0 はシミュレーションによって算出される値であることから算出ごとに値が変化するが、十分な試行回数の元では安定的な値をとる。

この過程からも分かるように、C 指標とは購買タイミングの不均一性を有意水準 $\alpha\%$ で検定した結果である。すなわち R、F、M の各指標とは異なり C 指標は 2 値の離散変数として定義されており、イベント発生のみと有無は考慮するがその度合いまでは考慮できないことを意味している。図 4 に示す複数の購買行動を比較した場合にも、H 指標の値から分かるように各消費者での clumpy さの度合いは様々であり、2 値変数に要約することで識別性が大きく失われてしま

¹⁾ Zhang et al. (2015) Appendix B. あるいはそのシミュレーションを実データに適用してレビューしている中山 (2016) を参照

論されているサブスクリプション型の動画視聴サービスのように明示的な契約を行うものが存在する一方で、基本的な使用が無料のスマートフォン用のゲームなども存在している。これら利用形態の全く異なる市場において同一の理由で Clumpiness が発生しているとは考えにくいことから、先行研究では解析の対象とされていない明示的な契約関係を持たないサービス市場の解析の事例として、スマートフォンのゲームアプリからの離脱予測に Clumpiness を応用する。

解析では日本での配信開始とともに大きな流行を見せたスマートフォンのゲームアプリ「Pokémon GO」を対象とする。選定にあたっては、市場に広く普及して十分な利用数が取得でき、なおかつ他のゲームアプリとは競合関係になりにくいような高い独自性を持ったアプリを想定した。参考までに 2016 年 9 月時点のアクティブユーザは 761 万人と推定されており、解析対象として十分な普及段階にあるといえる²⁾。

解析には株式会社インテージの調査パネルとして、スマートフォン利用の人口統計に基づいて大規模に収集されたスマートフォンアプリの利用履歴データである i-SSP モバイル調査を用いる。本データでは匿名化されたユーザごとにアプリを立ち上げた日時や閉じるまでの継続時間を取得しているが、支出情報（アプリ内での有料の機能やサービスの購入）については記録していないため、本分析では RFMC 分析の M 指標は考慮していない。

なお Pokémon GO は Niantic 社による位置情報ベースのゲームである Ingress のゲームシステムを流用していることから厳密には競合になりうるといえるが、得られたログデータからこれら 2 つのアプリを比較したところ、ユーザ数は Pokémon GO の 3% 程度であったため競合関係にないと判断した。

調査期間を 2016 年 9 月から 12 月までの 4 ヶ月として、9 月に 3 回以上 Pokémon GO を起動した消費者を調査対象すると対象者は 937 人であった。表 2 に対象者の継続的な利用状況を示すと、10 月時点で全体の 3 割以上、12 月には 6 割以上が離脱している。Pokémon GO はその話題性の高さから普段ゲームをしていないような消費者にまで広まったことが考えられ、それら極めてライトなユーザや、あるいは周りに合わせて始めただけのユーザが早期に離脱しているこ

表 2. 9 月アクティブユーザの使用状況の追跡（基準化）

	ユーザ数	月間平均起動回数	月間平均利用時間
9 月	100.0%	100.0%	100.0%
10 月	64.2%	116.7%	137.3%
11 月	47.4%	116.0%	149.2%
12 月	39.9%	110.8%	132.4%

とがわかる。その証拠にユーザ数の減少に反して一人当たりの月間平均利用の回数や時間はむしろ高まっている。これらの点からも一旦は顧客として獲得したものの離脱しやすいライトユーザを早期に発見することは重要な課題である。

このようなゲームアプリの離脱要因の解析における C 指標の活用例として、「熱しやすく冷めやすいユーザの発見」が考えられる。具体的には、世間の流行に従い Pokémon GO をプレイし始めるも、初期の短期間に高頻度でプレイしたのちに一定期間すると離脱するような飽きやすいユーザである。これら早期の離脱が予想されるユーザを事前に発見できることは、顧客獲得に比べ顧客維持コストの方が低く（Bhattacharya, 1998; Colgate, & Danaher, 2000 等）、なおかつ顧客維持率の 1% の改善で企業価値が平均 5% 改善する（Gupta, Lehmann, & Stuart, 2004）といった観点からマーケティング・コストの効率化につながるといえる。

そこで本解析では、Clumpiness を用いて初月 9 月のうちまとまった短期間に高頻度でゲームを利用している clumpy なユーザを識別し、この振る舞いが離脱までの時間に影響するかどうかを生存時間分析を用いて検証する。C 指標を扱うためには時間を離散化する必要があるが、例えば 1 時間単位で集計することを考えると仕事・学校や睡眠などの生活リズムに起因した習慣性が発生してしまうことから、熱中しているかどうかに関わらず Clumpiness が生まれやすいという問題がある。したがって本研究では 1 日を単位時間として離散化とした。

4.2. 解析方法

解析には顧客の購買、訪問、離脱等のイベント発生タイミングの予測やそれらに影響する要因の分析に広く用いられている Cox 比例ハザードモデル（Cox, 1972）を用いる。このモデルを用いた研究として、本研究と同様に調査期間の終了を右側打ち切りとしたハ

²⁾ 株式会社ヴァリューズ『「ポケモン GO」の今と、この 1 年の振り返り調査』（<https://www.valuesccg.com/knowledge/report/marketing/034/>）より引用

表 3. 解析 1: アプリ利用状況

変数名	平均値	標準偏差	最大値	最小値
解析 1 (サンプルサイズ: 937 人)				
自社 R 指標 (日)	7.56	6.33	27.00	1.00
自社 F 指標 (回)	248.58	342.95	1815.00	1.00
自社 C 指標	0.48	0.50	1.00	0.00
自社 C 比率	0.92	0.48	2.70	0.00
競合 F 指標 (回)	18.17	10.53	30.00	1.00

ザードモデルで顧客の満足度と離脱までの期間の関係をモデリングした Bolton (1998) や、ヨーロッパの金融サービスにおける銀行口座からの顧客の離脱をモデリングした Van den Poel & Lariviere (2004)、購買発生タイミングをモデリングした Jain & Vilcassim (1991)、小売のロイヤルティプログラムが顧客の離脱時期と支出額に及ぼす影響を調べた Meyer-Waarden (2007) 等がある。

離脱の判定にはある 1 回の使用から次の使用までの期間が 30 日間以上ある場合を基本として、30 日経過せずに調査期間末尾まで到達した場合には離脱の観測ができなかったものとして右側打ち切りとしている。この基準で離脱を観測した場合に、937 人の調査対象者のうち 890 人 (95%) が期間中に 1 回離脱する結果となった。この離脱よりあとの利用については離脱後の復帰と判定しているが、この 890 人のうち期間中に復帰したのは 44 人 (離脱者のうち 4.9%) とごく少数であったことから本研究では解析から除外した。結果として全体の 95% が期間中に一度は離脱、平均的な生存時間は 94.91 日となった。表 3 に要約統計量を示すと、第一に F 指標について自社/競合で大きな差異が認められる。これは一つには Pokémon GO の話題性やロイヤルティの高さ故に自社 F 指標が向上していることがあるだろうが、その一方で位置情報を用いたゲームという性質上、色々な場所で何度もアプリを起動することからも値の増加につながっていることが考えられる。心理的要因のモデリングのためには競合が少なく独自性の高いゲームアプリの選定が必要だが、これらの理由から Pokémon GO は対象として最適である。また C 指標/比率を比較すると、Clumpiness の発生割合は 48% だが、同じ clumpy と判定された顧客でも実際には境界値の 2.7 倍の強い不均一性も確認できるなど、C 比率により消費者の異質性を捉えられている。

離脱時期のモデリングにあたり、共変量として 9 月

中のアプリ使用履歴から集計した各ユーザについての R 指標 (9 月中の最終使用日から月末までの日数)、F 指標 (9 月中の総起動回数)、C 指標 (9 月中の利用における Clumpiness の有無)、C 比率の 4 変数、さらにそれらを同時期の人気ゲームアプリ (同時期に調査対象者の利用頻度が最も高かった 10 点のスマホゲーム) についても R、F、C の 3 指標を計測した競合利用の計 7 変数を用意し、これら因子が離脱までの期間の長さには及ぼす影響について検証した。また共変量の組み合わせによるモデルの適合度を AIC と AUC³⁾ を用いて比較、各変数の影響とともに最適なモデルについて評価する。ただし R、F 指標は正規分布に近づけることを目的に対数化した上で投入している。

4.3. 解析結果

表 4 にパラメータ推定値、AIC、AUC のモデル間比較の結果について示すと、まず全てのモデルに共通する傾向として、自社 R/F 指標は常に有意水準 0.1% でそれぞれ正/負の符号となった。ハザードモデルであることから、この結果は調査期間中に実際に利用している期間が短くなるほど離脱が早まり、一方で調査期間中の利用頻度が増えるほど離脱が遅まると解釈でき、これは感覚に即した結果であるといえる。Model 4, 5, 6 においては競合するゲームアプリの利用頻度との関係についても検証しているが、いずれも離脱時間に対して有意に影響を与えるものではなかった。

次に C 指標と C 比率での変化について確認すると、Model 1 から 3、そして Model 4 から 6 ままで比較した場合に、いずれの比較でも Clumpiness を考慮しない場合、C 指標を用いた場合、C 比率を用いた場合の順に AIC、AUC とともに改善する傾向が確認できる。また C 比率を用いた場合には有意水準 0.1% で離脱に対して負の影響を与えており、ゲーム市場においては Clumpiness の度合いの高さが離脱を早めるという当初の仮説を支持するものとなっている。また Model 2 の結果から Clumpiness の有無による生存率の変化について図 5 に示す。この結果から、Clumpiness を有するプレイヤーではそうでないプレイヤーと比べて期首時点での離脱率が大きく高まっており、初月で半数以上が離脱していることがわかる。

³⁾ AUC (Area Under the Curve) は ROC 曲線下の面積の割合であり、0 から 1 までの値域で分類の精度を示す指標。0.5 より高い場合にランダムより高い分類精度を表す。

表 4. 解析 1 パラメータ推定値と予測精度の比較

	Model 1 (R, F)		Model 2 (R, F, C)		Model 3 (R, F, C 比)		Model 4 (R, F, 競)		Model 5 (R, F, C, 競)		Model 6 (R, F, C 比, 競)	
自社 R	0.2555	***	0.2605	***	0.2894	***	0.2461	***	0.2512	***	0.2837	***
自社 F	-0.3532	***	-0.4123	***	-0.4682	***	-0.3561	***	-0.4159	***	-0.4711	***
自社 C	-		0.4377	**	-		-		0.4415	**	-	
自社 C 比率	-		-		0.7349	***	-		-		0.7362	***
競合 R	-		-		-		0.0354		0.0373		0.0251	
競合 F	-		-		-		0.0123		0.0174		0.0236	
競合 C	-		-		-		0.0319		0.023		-0.0199	
AIC	2971.512		2966.721		2952.388		2971.507		2968.588		2954.349	
AUC	0.754		0.759		0.770		0.753		0.759		0.769	

* 有意水準 5% ** 有意水準 1% *** 有意水準 0.1%

図 5. Clumpiness の有無によるユーザーの離脱率の比較

これらの結果は、長期的な優良ユーザーの育成のために必ずしも単なる高頻度の利用が有利に働くわけではないことを示しており、同一の利用頻度でもその利用パターンに顧客の離脱までの期間が影響を受けることを意味している。このように、初期の利用履歴データから離脱する確率の高いユーザーのみを早期に発見可能であるという点で、従来の RFM 分析に加えて Clumpiness を考慮することはマーケティング戦略上重要であるといえる。

またこの結果は、離脱時間のモデリングにあたり必ずしも競合利用の情報が重要ではないことを示しており、それでもなお Clumpiness の考慮によりモデルが改善していることから、Clumpiness が競合利用以上の情報（この解析例で言えば熱しやすく冷めやすいという意味でのプレイヤーの心理的要因）を含有してい

る可能性がある。そして従来の C 指標でもその性質自体を捉えることは可能だが、本研究で提案する C 比率を用いることで、その程度の大小をより詳細に捉えることが可能であるといえる。

5. 解析例 2：実店舗での購買予測への RFMC 分析の応用

次に解析 2 として、本研究で提案する C 比率を購買行動のモデリングに応用し、競合利用の代理変数となる可能性について実データを用いて検証する。

5.1. 解析の目的と概要

解析 1 ではサービス利用を対象に解析したが、解析 2 では競合関係にある複数の小売店での購買行動を対象に、解析 2-1 として自社での将来期間の購買回数、解析 2-2 として競合他社での購買回数のモデリングを行う。まず解析 2-1 では、一般的に入手が困難な競合利用の情報よりも自社データから得ることのできる Clumpiness を用いた方が高い精度で顧客行動の予測が可能なることを示すとともに、競合利用の代理変数としての Clumpiness の有用性を示すことを目的としている。次に解析 2-2 として観測できない競合他社の利用状況の予測を行うが、これは先にも述べたとおり Clumpiness が自社と競合の利用状況に由来する不完全情報となりうることから、Clumpiness の有無やその度合いを考慮することが競合予測においても高い有用性を示すと考えられるためである。解析には実店舗

表 5. 解析 2 : 8 月の T 社/競合他社の利用状況

変数名	平均値	標準偏差	最大値	最小値
解析 2 (サンプルサイズ: 1059 人)				
T 社 R 指標 (日)	14.82	2.89	26.00	12.00
T 社 F 指標 (回)	4.62	2.32	15.00	3.00
T 社 M 指標 (円)	12083.50	20297.91	348224.00	0.00
T 社 C 指標	0.10	0.30	1.00	0.00
T 社 C 比率	0.58	0.32	2.03	0.07
競合 F 指標 (回)	14.56	17.42	86.00	0.00

での競合利用を含めた購買のモデリングとして、スーパーマーケットの利用を対象とした解析を行う。

5.2. 解析方法

実際の解析では、スーパーマーケットにおける購買履歴データとその周辺地域の競合他社への訪問履歴のデータとして九州を中心に展開する大手スーパーマーケットチェーン T 社の購買履歴データと、スマートフォンの位置情報データから競合店舗への訪問頻度を取得した。これら 2 つは別個のデータでありそのまま紐づけることはできないが、加藤・星野・松村 (2017) では実店舗の ID-POS データと位置情報データで購買と訪問のマッチングを実施しており、新美・星野 (2018) でも同様の手法で顧客の買い回りのモデリングを行なっている。これらに倣って T 社店舗での決済から 30 分以内に位置情報データ上で店舗に該当するエリアから退出しているような消費者の組み合わせを算出し、期間中に少なくとも 3 回以上重複していることを条件として購買と訪問の重複回数の多い消費者から順にマッチングしたところ、1059 人のサンプルを得た。対象期間を 2017 年 8 月から 10 月までの 3 ヶ月間、買い忘れを考慮して 1 日に複数回の購買は 1 購買とした。ここで 8 月の利用状況を表 5 に示す。C 指標と C 比率を比較すると、離散値の C 指標としては全体の 1 割程度となっている clumpy な顧客だが、C 比率として連続値で見るとその内部にもばらつきがあることがわかる。さらに Clumpiness の有無による購買回数の差の比較を図 6 に示すと、Clumpiness のみられる消費者では購買回数の分散が大きいものの、Clumpiness の見られない群と比べて平均的に購買回数が増加していることがわかる。この Clumpiness あり群の分散の大きさからも、顧客の行動理解にあたっての購買タイミングの不均一性を連続値で比較することの重要性が示されていると言える。

図 6. Clumpiness の有無による自社購買の比較

図 7. H 指標の分布

図 8. C 比率の分布

また参考までに本データとして得られている期間全体 (8 月~10 月より $N = 92$ 日間) での、H 指標とそれを元に算出した C 比率でのデータの分布について図 7, 8 に示す。いずれの散布図でも点それぞれが 1 顧客の期間中の購買行動を示している。2 つの図を比較すると、C 比率では境界値となる $C = 1$ を基準に値が標準化されていることがわかる。図 8 において C 比率の最大値が揺らいでいるのは、シミュレーションを元に算出した境界値の揺らぎが拡大したものである。

解析 2-1 では 8 月 1 ヶ月間の T 社/競合それぞれの利用状況から 9 月の T 社の購買回数を、解析 2-2 では 8 月の T 社の利用状況を用いて同月中の競合の購買回数を予測する。解析では一般化線形モデル (GLM) として、購買回数のモデリングに広く用いられているポアソン重回帰分析を用いる。予測にあたっては説明変数の組み合わせを変化させた複数の予測モデルを作成し、モデルの適合度を情報量基準 AIC, BIC を用いて比較する。モデルごとに使用した変数を表 6 に示す。

Model I として一般的な RFM の各指標を用いたモデル, Model II として Zhang et al. (2015) による RFMC の各指標を用いたモデル, Model III として

表 6. 解析 2 で作成したモデルと使用した変数

モデル名	使用変数
解析 2-1：自社での将来購買の予測	
Model I	自社 RFM
Model II	自社 RFMC
Model III	自社 RFMC 比
Model IV	自社 RFM+ 競合 F
Model V	自社 RFMC+ 競合 F
Model VI	自社 RFMC 比 + 競合 F
解析 2-2：競合での累積購買の予測	
Model I	自社 RFM
Model II	自社 RFMC
Model III	自社 RFMC 比

Model I に C 比率を加えたモデル, Model IV として Model I に競合利用の情報を用いたモデル, Model V として Model II に競合利用の情報を用いたモデル, Model VI として Model III に競合利用の情報を用いたモデルの 6 つについて, 情報量規準 AIC と BIC を用いた比較を行った.

5.3. 解析結果

解析 2-1, 2-2 の結果について, それぞれ表 7, 8 に示す. まず全体の傾向について俯瞰すると, いずれの解析でも自社 RFM に提案指標を加えた Model III で情報量規準が最も改善する結果となっている. またいずれの解析でも全てのモデルにおいて, 自社 R, F 指標については有意水準 0.1% で購買行動に影響を与えている. clumpy な行動について比較すると, 解析 2-1, 2-2 とともに, C 指標, C 比率のどちらを用いた場合にも有意に影響を与えているものの, 情報量規準で比較すると競合変数の投入の有無にかかわらず C 比率を用いた場合にモデルがより改善していることがわかる.

また解析 2-2 の結果から, 競合予測という観点でも当初の仮説通り, 異なる N , n の値を持つ顧客間での Clumpiness の度合いについて連続値で比較することにより, 従来手法に比べて高い精度で予測が可能であることが示されたと言える.

これらの結果から, 日常的な継続利用が必要な市場においては自社の将来購買や競合購買の予測にあたり本研究の提案指標を用いることが有効であるといえる.

6. ま と め

最後に実データ解析の結果から本研究で得られた知見についてまとめるとともに, 今後の課題について述べる.

6.1. 本研究のまとめ

本研究では Zhang et al. (2013) で提案された Clumpiness 指標の実務応用にあたっての課題について, マーケティング的な観点, 統計学的な観点の双方から整理した. まずマーケティング的な観点として先行研究で着目されていない RFMC 分析のマーケティング上の意義について提案し, 次に統計学的な観点として実際の活用における課題から指標の拡張を行った. そして提案したマーケティング上の課題を本指標を用いて捉えることを目的に, 性質の異なる複数の実データに提案指標を適用して解析を行い, その有用性を示した.

具体的には, 本来であれば競合他社にしか持ち得ない競合利用やそもそも入手することが困難な心理的要因の情報を自社の行動データから算出可能な Clumpiness を用いて代替することで, これら入手不可能な情報を代用した場合よりも高い精度で顧客行動を予測することができた. これは図 1 における④, ⑤の構造であり, 不完全情報しか持ち得ない状況下でも Clumpiness の活用により従来より高い精度で顧客行動を予測可能であることを示している. よって本研究の貢献は次の 3 点に集約できる. i) 先行研究の少ない Clumpiness について新たなデータを用いて解析しその有用性を示した点, ii) 顧客の離脱や競合購買の予測に Clumpiness を応用するマーケティング上の意義について提示した点, iii) Clumpiness の実務利用上の課題について検討し, 指標の改善を通じて高い精度での予測を実現した点である.

6.2. 今後の課題

本研究では商品購買から契約関係を持たないサービス利用まで, Clumpiness の概念が広く応用可能であることを H 指標に基づく新指標の提案とそれを用いた実データ解析を通じて検証したが, そもそも Clumpiness の計算過程としてのこの H 指標自体に未だ多くの課題が残されている. 一例として H 指標はまとまった利用でも調査期間の期首または期末に発生した場合に値が大きく増加してしまうなど, まとも

表 7. 解析 2-1 パラメータ推定値とモデルの比較

	Model I (RFM)			Model II (RFMC)			Model III (RFMC 比)			Model IV (RFM+ 競合)			Model V (RFMC+ 競合)			Model VI (RFMC 比 + 競合)		
自社 R	-0.0153	***		-0.0145	***		-0.0137	***		-0.015	***		-0.0142	***		-0.0136	***	
自社 F	0.0588	***		0.0573	***		0.0536	***		0.0588	***		0.0575	***		0.0541	***	
自社 M	0.0183			0.0368			0.0677	**		0.0173			0.0339			0.0615	*	
自社 C	-			0.2531	**		-			-			0.2279	**		-		
自社 C 比率	-			-			0.2803	***		-			-			0.2512	***	
競合 F	-			-			-			0.0002	***		0.0002	**		0.0002	*	
AIC	2907.929			2902.070			2890.673			2901.034			2896.696			<u>2887.637</u>		
BIC	-4021.403			-4022.675			<u>-4034.072</u>			-4023.711			-4023.463			-4032.522		

* 有意水準 5% ** 有意水準 1% *** 有意水準 0.1%

表 8. 解析 2-2 パラメータ推定値とモデルの比較

	Model I (R, F, M)			Model II (R, F, M, C)			Model III (R, F, M, C 比)		
自社 R	-0.0053	***		-0.0039	***		-0.0029	***	
自社 F	0.0054	***		0.0023	**		-0.0112	***	
自社 M	0.0699	***		0.0996	***		0.1903	***	
自社 C	-			0.5765	***		-		
自社 C 比率	-			-			0.8183	***	
AIC	139451.103			138523.443			<u>135488.945</u>		
BIC	131336.773			130413.698			<u>127379.200</u>		

* 有意水準 5% ** 有意水準 1% *** 有意水準 0.1%

り自体の発生時期によりその大小が変化してしまう。RFM 分析に関して先行研究でも指摘されているように、この性質では期間の区切り方で結果が変化するため解析結果に恣意性が発生してしまう。Clumpiness は利用時期に関係なくまとまった利用を検出するべきであり、今後も継続的な改善が必要である。また先行研究でも示したように、未だ Clumpiness の応用例は稀有でありその有用性は完全には明らかにされていない。今後も様々な製品、サービスやその競合関係に幅広く指標を応用してその性質について精査していく必要がある。

本研究では解析例としてゲームアプリを扱ったが、一般的なゲームアプリではいわゆる「スタミナ」「ログインボーナス」と呼ばれるような機能を用いて、一度にプレイできるゲーム量を制限したり利用の習慣化を狙ったりといった試みが広く試行されている。本研究で提案する C 比率ではゲーム利用の不均一性を比

較可能な形で捉えられることから、これら機能を用いてプレイを習慣化させることが長期的なロイヤルティにつながるかと言った点についても今後検証することが可能だろう。

また解析では Clumpiness が心理的要因と競合利用の 2 つの情報を内包している可能性について検討した。これは企業が収集するログデータのみでこのような解析が実現できるという点で今後の活用の可能性も大きいと言えるが、スマートフォンゲームのうち比較的競合が少ないと考えられる Pokémon GO、あるいは心理的要因の発生しにくいスーパーマーケットでの日用品購買を対象にしたとはいえ、現状ではこれら 2 つの要因を分離して解析することはできていない。将来的には調査データ等と紐付けた解析を行うなどして消費者の認知・態度等との関連を見ることでこれら要因を個別に検討する必要がある。

謝 辞

本研究の実施にあたり、株式会社インテージの中野 暁様には、i-SSP モバイル調査のデータをご提供頂きました。また編集担当委員および 2 名の匿名の査読者の先生より有益なコメントをいただきました。この場を借りて御礼申し上げます。

本研究は科学研究費 (18H03209/17K18568) および博士課程教育リーディングプログラム PhD プロフェッショナル登龍門 (名古屋大学) の助成を受けております。

参 考 文 献

- 阿部誠 (2014). RFM データを用いた顧客生涯価値の算出：既存顧客の維持介入と新規顧客の獲得（特集ソーシャルビジネスとマーケティング）. マーケティングジャーナル, **34**, 1, 73–90.
- Bhattacharya, C. B. (1998). When customers are members: Customer retention in paid membership contexts. *Journal of the academy of marketing science*, **26**, 1, 31–44.
- Bell, D., Deighton, J., Reinartz, W. J., Rust, R. T., & Swartz, G. (2002). Seven barriers to customer equity management. *Journal of service Research*, **5**, 1, 77–85.
- Berger, P. D., & Nasr, N. I. (1998). Customer lifetime value: Marketing models and applications. *Journal of interactive marketing*, **12**, 1, 17–30.
- Blattberg, R. C., Kim, B., & Neslin, S. A. (2008). Database Marketing: Analyzing and Managing Customers. *International Series in Quantitative Marketing*.
- Bolton, R. N. (1998). A dynamic model of the duration of the customer's relationship with a continuous service provider: The role of satisfaction. *Marketing science*, **17**, 1, 45–65.
- Bult, J. R., & Wansbeek, T. (1995). Optimal selection for direct mail. *Marketing Science*, **14**, 4, 378–394.
- Chen, Y., & Steckel, J. H. (2012). Modeling credit card share of wallet: Solving the incomplete information problem. *Journal of Marketing Research*, **49**, 5, 655–669.
- Colgate, M. R., & Danaher, P. J. (2000). Implementing a customer relationship strategy: The asymmetric impact of poor versus excellent execution. *Journal of the Academy of Marketing Science*, **28**, 3, 375–387.
- Colombo, R., & Jiang, W. (1999). A stochastic RFM model. *Journal of Interactive Marketing*, **13**, 3, 2–12.
- Cox, D. R. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society, Series B.*, **34**, 2, 187–220.
- Du, R. Y., Kamakura, W. A., & Mela, C. F. (2007). Size and share of customer wallet. *Journal of Marketing*, **71**, 2, 94–113.
- Fader, P. S., Hardie, B. G., & Lee, K. L. (2005). RFM and CLV: Using iso-value curves for customer base analysis. *Journal of marketing research*, **42**, 4, 415–430.
- Ganesh, J., Arnold, M. J., & Reynolds, K. E. (2000). Understanding the customer base of service providers: an examination of the differences between switchers and stayers. *Journal of marketing*, **64**, 3, 65–87.
- Gladly, N., & Croux, C. (2009). Predicting customer wallet without survey data. *Journal of Service Research*, **11**, 3, 219–231.
- Gupta, S., Lehmann, D. R., & Stuart, J. A. (2004). Valuing customers. *Journal of marketing research*, **41**, 1, 7–18.
- 星野崇宏 (2009). 調査観察データの統計科学：因果推論・選択バイアス・データ融合. 岩波書店.
- Hughes, A. M. (1995). Making a database pay off using recency, frequency, and monetary analysis. *Journal of Database Marketing*, **3**, 1, 77–89.
- Jacoby, J., & Chestnut, R. W. (1978). Brand loyalty: Measurement and management. John Wiley & Sons Incorporated.
- Jain, D., & Vilcassim, N.J. (1991). Investigating Household Purchase Timing Decisions: A Conditional Hazard Function Approach. *Marketing Science*, 1–23.
- 加藤諒・星野崇宏・松村優哉 (2017). 位置情報と購買履歴のマッチングによる競合他社での購買行動の理解, 日本マーケティング・サイエンス学会第 102 回研究大会.
- Khajvand, M., Zolfaghar, K., Ashoori, S., & Alizadeh, S. (2011). Estimating customer lifetime value based on RFM analysis of customer purchase behavior: Case study. *Procedia Computer Science*, **3**, 57–63.
- Mägi, A. W. (2003). Share of wallet in retailing: the effects of customer satisfaction, loyalty cards and shopper characteristics. *Journal of retailing*, **79**, 2, 97–106.
- Meyer-Waarden, L. (2007). The effects of loyalty programs on customer lifetime duration and share of wallet. *Journal of Retailing*, **83**, 2, 223–236.
- Mitrović, S., & Singh, G. (2016). Predicting branch visits and credit card up-selling using temporal banking data. *arXiv preprint*, arXiv:1607.06123.
- Morrison, D. G., & Silva-Risso, J. (1995). A latent look at empirical generalizations. *Marketing Science*, **14**, 3, G61–G70.
- Mulhern, F. J. (1999). Customer profitability analysis: Measurement, concentration, and research directions. *Journal of interactive marketing*, **13**, 1, 25–40.
- 中山雄司 (2016). 顧客関係管理研究の新動向：来店/購買間隔の不均一性を測るクランビネス指標. 甲南経営研究, **57**, 2, 161–181.
- 中山雄司 (2017). RFMC 分析：ISMS Durable Goods Dataset 1 を用いた事例. 日本マーケティング・サイエンス学会第 101 回研究大会.
- Netzer, O., Lattin, J. M., & Srinivasan, V. (2008). A hidden Markov model of customer relationship dynamics. *Marketing science*, **27**, 2, 185–204.
- 新美潤一郎・星野崇宏 (2015). ユーザ別アクセス・パターン情報の多様性を用いた顧客行動の予測とモデリング. 応用統計学, **44**, 3, 121–143.
- 新美潤一郎・星野崇宏 (2017). 顧客行動の多様性変数を利用した購買行動の予測, 人工知能学会論文誌, **32**,

- 2, B-G63_1.
- 新美潤一郎・星野崇宏 (2018). 位置情報データによる競合店舗の利用状況の多様性を用いた購買予測手法の提案, 2018 年度人工知能学会全国大会 (第 32 回).
- 西原彰宏 (2012). バラエティ・シーキング要因の探索的研究: 製品関与概念を手がかりに, 産研論集, **39**, 79-89.
- Pfeifer, P. E., Haskins, M. E., & Conroy, R. M. (2005). Customer lifetime value, customer profitability, and the treatment of acquisition spending. *Journal of managerial issues*, 11-25.
- Platzer, M., & Reutterer, T. (2016). Ticking Away the Moments: Timing Regularity Helps to Better Predict Customer Activity. *Marketing Science*, **35**, 5, 779-799.
- Raju, P. S. (1980). Optimum Stimulation Level: Its Relationship to Personality, Demographics, and Exploratory Behavior. *Journal of Consumer Research*, **7**, December, 272-282.
- Raju, P. S. (1981). Theories of Exploratory Behavior: Review and Consumer Research Implications. *Research in Marketing*, **4**, 223-249.
- Schweidel, D. A., Bradlow, E. T., & Fader, P. S. (2011). Portfolio dynamics for customers of a multiservice provider. *Management Science*, **57**, 3, 471-486.
- Sohrabi, B., & Khanlari, A. (2007). Customer lifetime value (CLV) measurement based on RFM model. *Iranian Accounting & Auditing Review*, **14**, 47, 7-20.
- Van den Poel, D., & Lariviere, B. (2004). Customer attrition analysis for financial services using proportional hazard models. *European journal of operational research*, **157**, 1, 196-217.
- Zhang, Y., Bradlow, E. T., & Small, D. S. (2013). New measures of clumpiness for incidence data. *Journal of Applied Statistics*, **40**, 11, 2533-2548.
- Zhang, Y., Bradlow, E. T., & Small, D. S. (2015). Predicting Customer Value Using Clumpiness: From RFM to RFMC. *Marketing Science*, **34**, 2, 195-208.

(2019 年 9 月 24 日受理, 2020 年 2 月 3 日採択)
(この間審査 3 回・審査期間合計 77 日)