

KONAN UNIVERSITY

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ピネス指標

著者	中山 雄司
雑誌名	甲南経営研究
巻	57
号	2
ページ	161-181
発行年	2016-12-10
URL	http://doi.org/10.14990/00002223

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ピネス指標——

中山 雄 司

概要

本ノートでは顧客関係管理研究の最新動向として、Zhang, Bradlow and Small (2015) が提案した来店/購買間隔の不均一性を示すクランピネス (Clumpiness) に注目する。そして、彼らが定義したクランピネス指標を測定し、顧客行動にクランピネスがあるかを検証する手順をまとめ、統計解析環境 R (R Core Team 2016) を使ってそれを実装する方法を提供する。

1 はじめに

顧客関係管理 (Customer Relationship Management 以下、CRM) は、マーケティング・サイエンスの学術研究における重要な分野の一つである。そのことは、マーケティング・サイエンスの歴史を俯瞰したレビュー論文をまとめた Winer and Neslin (2014) において、CRM 研究が10の分野のうちの1つに含まれていることから伺える (Neslin 2014)。Kumar and Reinartz (2012, p. 4) によると、「CRM は企業に対して各顧客がもたらす生涯価値を最大化する実務的手法を策定するために、マーケティング・データベースを分析・利用し、コミュニケーション技術を活用する実践的活動である」と定義される。CRM が重要であるとの認識の背景には、安価で大量のデータ蓄積が可能となり、企業と顧客の双方向のコミュニケーションを容易にする情報通信技術の発展がある。メンバーズ・カードなどを通じて購買履歴を、またダイレクト・メールなどによるプロモーションの実施とそれに対する反応を顧客

データベースに蓄積して、各顧客の購買行動を把握し、特に優良顧客との関係を将来に渡って維持することは、企業活動の長期的な利益基盤を確保する上で重要である。

CRMにおいて最終購入日（Recency）、来店頻度（Frequency）、購入金額（Monetary）⁽¹⁾を顧客ごとに記録したRFM指標は簡潔だが、有用なデータである。RFM指標は1960年代から実務で活用され（Blattberg, Kim and Neslin 2008, p. 323）、近年に至るまで学術研究においても注目されている（Abe 2009, 阿部 2011, Fader, Hardie and Lee 2005）。RFMのそれぞれの指標を組み合わせ、購買行動が類似した顧客ごとにセグメント分けして、顧客を管理することはCRMの第一歩であると言える。

Fader and Hardie (2009)（オリジナルはSchmittlein, Morrison and Colombo 1987）は取引機会が連続的か離散的か、顧客との関係（取引期間）について明示的な契約があるかどうかで、企業と顧客の関係を4つの場合に分類している。RFM指標は顧客との関係に明示的な契約がなく、顧客が将来も購買をする可能性がある確率（「生存」確率）を推定しなければならない場合に不可欠な情報を提供する。

そうした中で、近年RFM指標に加えて、顧客の購買行動を測る新しい尺度がZhang, Bradlow and Small (2015)によって提唱された（以下では、Zhang, Bradlow and Small 2015をZBSと略記する）。それは、クランピネス（Clumpiness）である。クランピネスはZBSでは「(イベントが)均等な間隔に従わない度合い (defined as the degree of nonconformity to equal spacing)」と定義されている（ZBS p. 196）。発生間隔に注目するイベントとしては、来店と購買がある。ZBSは、クランピネスを測定することで、RFMでは捉えられない顧客の購買行動を識別し、特に個別の顧客レベルで将来の購買行

(1) 購入金額は一定期間内での合計購入金額または一来店での平均金額で測られる。

動を予測する力を持つので、RFM 指標にクランピネスを加えた RFMC 指標が CRM において今後重要になると述べている。

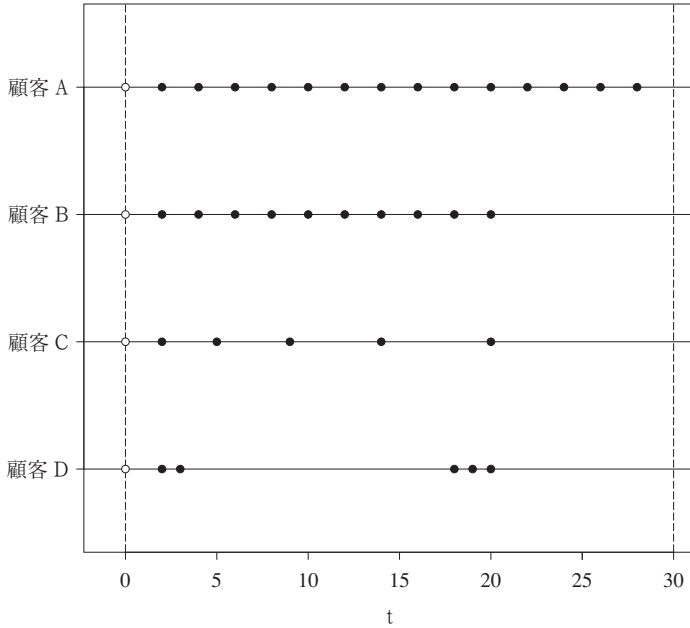
本ノートでは CRM 研究の最新動向として、このクランピネスに注目し、ZBS で定義されたクランピネス指標を測定し、顧客行動にクランピネスがあるかを検証する手順をまとめ、統計解析環境 R (R Core Team 2016) を使って実装する方法を提供する。次節以降の構成は以下の通りである。2 節では、RFM 指標を数値例を使いながらレビューし、RFM 指標では捉えられない顧客の購買行動の特徴があることを指摘する。3 節では、ZBS が提案したクランピネス指標について説明し、購買行動にクランピネスを持つ顧客を識別する手順をまとめる。そして Chen, Sain and Guo (2012) で使われ、Web で公開されている実データにおける顧客の購買間隔にクランピネスがあるかを実際に検証してみる。4 節で本ノートをまとめ、将来の研究課題を挙げる。補論には 3 節における計算を実行するための R コードを掲載する。

2 RFM 指標のレビュー

本節では、RFM 指標を数値例を使いながらレビューし、RFM 指標だけでは捉えられない顧客の購買行動の特徴があることを指摘する。

仮想的な 4 人の顧客 (A, B, C, D) の購買行動を RFM 指標で表してみよう。彼ら全員は初期時点 ($t=0$) で初回の購買を行ったとする。その後、潜在的な継続購買の機会は 30 回あったとしよう ($t=1, 2, \dots, 30$)。図 1 は彼らの購買行動を表している。黒丸●は初回以降に実際に購買が行われた期を示している。顧客 A の購買間隔は安定しており、 $t=2$ から $t=28$ まで 2 期に 1 回のペースで計 14 回購買している。データ期末の $t=30$ から最終購買期 ($t=28$) までの間隔は 2 である。顧客 B も $t=20$ までは安定した購買を行っており、 $t=2$ から $t=20$ まで 2 期に 1 回のペースで計 10 回購買している。ただし、データ期末から最終購買期 ($t=20$) までの間隔は 10 であり、直

図1 仮想的な4人の顧客の購買行動



近の10期では購買がない。顧客 C は初回以降で計 5 回の購買を行っているが、連続する購買時点の間隔は徐々に広がっている。そして、顧客 B と同様にデータ期末から最終購買期 ($t=20$) までの間隔は10であり、直近の10期では購買がない。顧客 D もまた顧客 C と同様に、初回以降で計 5 回の購買を行っており、データ期末から最終購買期 ($t=20$) までの間隔は10であり、直近の10期では購買がない。 $t=4$ 以降で $t=17$ まで購買が途絶えたが、その後は 3 期連続で購買が続いている。

簡単化のために、毎回の購買では 4 人も同額の3000円を支出していたとし、4 人の顧客の購買行動を RFM 指標で表すと表 1 の通りとなる。⁽²⁾ 顧客 A

(2) 初回購買を除いて F と M を測定している。

表1 仮想的な4人の顧客のRFM指標

	R (期)	F (回)	M (万円)
顧客 A	2	14	4.2
顧客 B	10	10	3
顧客 C	10	5	1.5
顧客 D	10	5	1.5

は短い間隔で購買が継続し、データ期末から最終購買期 ($t=28$) までの間隔も短いので、最も優良な顧客と言える。顧客 B はデータ期末から最終購買期 ($t=20$) までの間隔は、顧客 A より長いものの、最終購買期までは顧客 A と同様に短い間隔で購買が継続しており、顧客 A に次いで優良な顧客と言える。顧客 C と顧客 D はデータ期末から最終購買期 ($t=20$) までの間隔は、顧客 B と同様である。また、初回以降の継続購買の回数は顧客 A や顧客 B より少なく 2 人とも 5 回である。顧客 A や顧客 B と比較すると、この 2 人は優良な顧客ではないと判断するのが自然であろう。それでは、顧客 C と顧客 D を比較するとどうだろうか。RFM 指標で判断する限り、両者の購買行動は区別できない。

ここで、連続する購買時点の間隔に注目してみよう。顧客 A と顧客 B は最終購入期まで 2 期に 1 回の安定した間隔で購買している。これに対して、顧客 C の最終購入期までの購買時点の間隔は徐々に広がっている。他方、顧客 D は $t=2, 3$ と購買が続いた後 $t=4$ から $t=17$ まで途絶え、その後は $t=18, 19, 20$ と 3 期連続で購買が続いている。従って、連続する購買時点の間隔に注目すると、顧客 C と顧客 D には大きな違いがあると言える。購買が途絶えた後に一度購買が始まると、連続して塊 (clump) となって購買が続く傾向が ZBS が注目したクランピネスである。

3 クランピネス指標を用いた購買行動の識別

本節では、まず ZBS が提案したクランピネス指標について説明する。次に、来店または購買間隔にクランピネスがある顧客を識別する手順をまとめ、最後に Chen, Sain and Guo (2012) で使われ、Web で公開されている実データ (UCI Machine Learning Repository Online Retail Data Set⁽³⁾) における顧客の購買間隔にクランピネスがあるかを実際に検証してみる。

1 節で述べたように、ZBS はクランピネスを「(イベントが) 均等な間隔に従わない度合い」と定義している。発生間隔に注目するイベントとしては、購買と来店がある。購買時点に関しては、会計時にメンバーズ・カードを提示したり、Web ショップでの購買でも会員としてログインすることが前提となっていることが多くなり、その時点の把握は容易になった。来店時点に関しては、実店舗の場合はそれを完全に把握することは困難であるが、Web ショップの場合には購入しない場合でもページ閲覧時点は記録されていることが多いので、その時点を把握することは容易であろう。以下では注目するイベントは購買であるとして記述する。

ZBS はクランピネスを測る指標が持つべき性質を 4 つ挙げている (ZBS p. 197)。

- 最小性：購買が等間隔で起こる場合、その指標は最小になるべきである。
- 最大性：購買がある時点とその前後に集中して起こる場合、その指標は最大になるべきである。
- 連続性：購買の発生時点が微小に変化した場合、その指標もまた微小に変化すべきである。
- 収束性：購買の発生間隔が狭まる（広がる）につれて、その指標は増加

(3) 以下のサイトを参照のこと。 <https://archive.ics.uci.edu/ml/datasets/Online+Retail>

表2 仮想的な4人の顧客のクランピネス指標の計測

	n	H_p	臨界値
顧客 A	14	0.002	0.112
顧客 B	10	0.109	0.163
顧客 C	5	0.081	0.289
顧客 D	5	0.315	0.289

(減少) するべきである。

そして、ZBS は潜在的な継続購買の機会を N 、実際の購買回数を n として上記の4つの性質を持つ指標

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

を提案している。ただし、 x_i は以下の通り計算する。⁽⁴⁾

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

ここで、 t_i ($i=1, 2, \dots, n$) は i 番目の購買が起こった期を表す。従って、 x_1 はデータの初期時点から1番目の購買までの間隔である。 x_i ($i=2, \dots, n$) は i 番目の購買と $i-1$ 番目の購買の間隔であり、 x_{n+1} はデータの期末から最後の購買までの間隔 (に1を加えたもの) である。ただし、データの初期時点から期末までの間隔が1となるように基準化されていることに注意する ($\sum_{i=1}^{n+1} x_i = 1$)。

2節の仮想的な4人の顧客に対して、クランピネス指標(1)を計算すると

(4) 後述の表3の形式のデータに対して、RFM指標とクランピネス(C)指標を計算するRのコードについては、補論を参照のこと。

表2の通りとなる。顧客Aの数値が最も小さく、顧客Dの数値が最も大きい。イベントが均等な間隔で発生しない度合いを測るのがクランピネス指標であるから、妥当な結果と思える。しかし、顧客Bの数値は顧客Cのそれよりも高い。クランピネスの定義からすると、この結果は整合性に欠くと思えるかもしれないが、実はクランピネス指標(1)は N と n を所与として4つの性質を満たすのである。 $N=30$ は4人に共通であるが、 n は同じではないため、顧客Bの数値が顧客Cのそれよりも高くなることは問題ではない。つまり、 n が異なる顧客Aと顧客Dのクランピネスを比較することはできない。 $n=5$ と購買回数が共通な顧客Cと顧客Dのクランピネスを比較することには意味がある。顧客Dの方が顧客Cよりも購買間隔にクランピネスがあるというのはもっともらしい。

それでは、様々な N と n の組み合わせに対して、顧客の購買間隔にクランピネスがあるかどうかを判断するにはどうすれば良いのだろうか。ZBSはモンテカルロ・シミュレーションを使って統計的検定を行い、それを判断する手順を提案している（ZBS Appendix B）。以下に、説明を若干補足したものを示す。

1. シミュレーション回数 M と有意水準 $\alpha\%$ を定める。
2. n と N を所与として、以下の(a)~(c)を M 回繰り返す。
 - (a) 1から N の数字が書かれた N 枚のカードから、 n 枚のカードを無作為に非復元抽出する。
 - (b) 取り出した n 枚のカードを昇順に並べ替える。
 - (c) カードに書かれた数字をイベントの発生時点とみなして、クランピネス指標(1)を計算して保存する。
3. 保存した M 個のクランピネス指標の経験分布関数から上側 $\alpha\%$ の分位点を計算する。

手順2で計算されるクランピネス指標は、 N 期のうちで n 回のイベント

がランダムに発生するという帰無仮説の下で計算されている。もし、手元のデータから求められたクランピネス指標の値が、上側 $\alpha\%$ の分位点よりも大きいならば、その帰無仮説は棄却されて手元のデータのイベントの発生にはクランピネスが存在すると判断される。⁽⁵⁾ 表2には $(N, M, \alpha) = (30, 20000, 0.05)$ として計算した臨界値も記載している。顧客 D のクランピネス指標だけが、対応する臨界値よりも大きいため、この顧客の購買間隔にのみクランピネスが存在すると判断される。

それでは、Chen, Sain and Guo (2012) で使われたデータを使って、このデータに購買履歴が記録されている顧客の購買行動の RFM 指標とクランピネス指標を計算してみよう。このデータはギフト商品を販売するイギリスのオンライン店舗の取引データである。取引相手の顧客には卸売業者が多く含まれる。データ期間は2010年12月1日から2011年12月9日である。元々のデータは個々の取引ベースで記録されており、複数の商品が一度に購入された場合には、一つの取引が複数行に渡る全54万1909行のデータである。顧客 ID が明記された行に限定し、返品と判断できる取引数量がゼロ以下の行を除いて、顧客ごとに日別に集計すると、1万6766行に集約された。顧客数は4339である。表3は集約されたデータの冒頭部分である。Mはその日の購買金額（単位はポンド）を表す。顧客 ID が12346の顧客は日別で1回のみの購買が記録されているのに対して、顧客 ID が12347の顧客は日別で7回の購買が記録されていることが分かる。

それでは、これらの各顧客について RFMC 指標を計算してみよう。⁽⁶⁾ 表4

(5) (N, M, α) を所与として、 $n=1, 2, \dots, N$ の場合の上側 $\alpha\%$ の分位点を計算する R のコード (関数 `GenerateZtable`) については、補論を参照のこと。

(6) ただし、データの開始日2010年12月1日までに各顧客は会員登録していると仮定し、データ期間内に初めて登録した顧客はいないとする。また、R 指標や購買間隔の計算は日を単位として行うが、R 指標や購買間隔の基本単位はデータに含まれる商品の性質や購買の周期性に応じて、週や月あるいは四半期を使うこともあり得る (ZBS p. 198)。最後に、M 指標はデータ期間中の合計購買金額で測る。

表3 Chen, Sain and Guo (2012) で使用されたデータを整理した冒頭部分

顧客 ID	日付	購買金額 (GBP)
12346	2011/1/18	77183.6
12347	2010/12/7	711.79
12347	2011/1/26	475.39
12347	2011/4/7	636.25
12347	2011/6/9	382.52
12347	2011/8/2	584.91
12347	2011/10/31	1294.32
12347	2011/12/7	224.82

表4 Chen, Sain and Guo (2012) で使用されたデータの RFMC 指標の分布

	最小値	第1四分位点	中央値	平均値	第3四分位点	最大値
R	0	17	50	92.04	141.50	373
F	1	1	2	3.86	4	132
M	0	307	674	2054	1662	280200
C	5.1×10^{-6}	0.092	0.201	0.272	0.399	0.973

は RFMC 指標の分布をまとめたものである。R 指標は 0 から 373 まで分布している。つまり、データ期末に最終購買を行った顧客からデータ期首に購買して以後は購買をしていない顧客までいる。中央値が 50 であるため、顧客の半分はデータ期末から直近 50 日の間に最終購買を行っていることが分かる。F 指標は 1 から 132 まで分布しているが、中央値が 2 であるため、顧客の半分は多くて 2 回しか購入を行っていないことが分かる。M 指標は 0 から約 28 万まで分布している⁽⁷⁾。中央値は 674 である。表 4 にはクラмпネス指標の分布も掲載したが、様々な取引回数の下で計算したクラмпネス指標の分布

(7) 合計購買金額が 0 であるのは不自然であるが、元データを確認すると取引数量は正だが単価が 0 の名前が “ASSTD DESIGN 3D PAPER STICKERS” という商品を購入していた顧客が 1 人だけいた。

をそのまま検討することはできない。前述の統計的検定の手順に従い、購買間隔にクランピネスがある判断されたのは4339人中466人で約1割だった。

図2はRFM指標の行列散布図である。1つの点が1人の顧客を表す。R指標が100以上の顧客はFとMの指標が比較的小さな値に集中している。R指標が100未満で最終購買日がデータ期末に近い顧客についてはFとMの指標のばらつきが大きいことが分かる。FとMに注目すると、人数は少ないがFが大きくMが小さい顧客セグメントとFが小さくMが大きい顧客セグメントの存在が確認できる。図3上はFとMの散布図に有意水準5%でクランピネスと判定された顧客を■で区別して表示したものである。上右側は表示範囲をFの上限は40、Mの上限は5万に狭めている。図3左下はRとMの散布図である。クランピネスと判定された顧客を■で区別して表示したが、多くの点が重なっているため、その区別は難しい。図3右下はクランピネスを持つ場合をTRUE、そうでない場合をFALSEとして顧客を分けてR指標の箱ひげ図を描いたものである。クランピネスを持つと判断された顧客のR指標の値の第3四分位点はそうでない顧客のR指標の中央値より小さいため、前者のR指標は比較的小さい値に集中してはいるが、それでもR指標が大きな外れ値を持つ顧客も中には存在することが分かる。

図3から言えることはRFM指標のみでは同一の購買行動を行う顧客と判断されるが、中には短い期間に購買を集中するとしばらくは購買を行わないクランピネスという特徴を持つ顧客を見逃すということである。購買間隔にクランピネスを持つ顧客は、購買がない期間が続いた後に突然に大量の購買を始める可能性がある。RFM指標のみで優良顧客を識別していると、将来大きな利益をもたらす潜在的可能性を持つ一部の顧客との関係が維持できないかもしれない。そうならないためにも、RFM指標に加えてクランピネス指標に注目することには意味がある。

図2 RFM指標の行列散布図

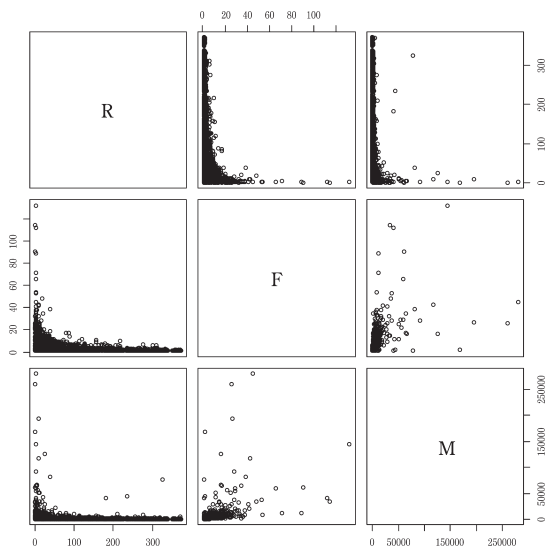
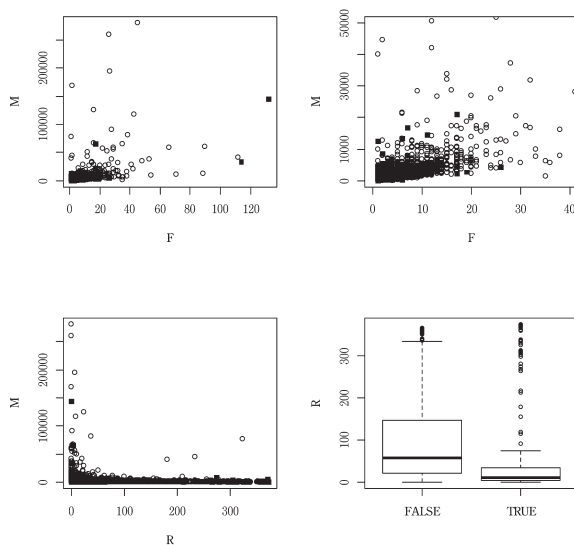


図3 クランピネスを識別したRFM指標



4 お わ り に

本ノートでは顧客関係管理研究の最新動向として、ZBSが提案した来店/購買間隔の不均一性を示すクランピネスに注目した。そして、彼らが定義したクランピネス指標（3節の数式(1)）を測定し、顧客行動にクランピネスがあるかを検証する手順をまとめ、統計解析環境 R（R Core Team 2016）を使ってそれを実装する方法を提供した。また、Chen, Sain and Guo (2012) で使われ、Webで公開されている実データにおける顧客の購買間隔にクランピネスがあるかを実際に検証した。その結果、約1割の顧客の購買間隔にクランピネスがあることが確認された。ZBSは3つの企業（匿名業者、CD Now, Mecoxlane）の他に、Webでプラットフォーム・ビジネスを展開する新しい企業（Hulu, YouTube, Amazon, eBay）の顧客データを使って、クランピネス指標の有用性を明らかにした⁽⁸⁾。特に、後者のWeb時代の新しいビジネスを行う企業の顧客の中で、来店（閲覧）/購買（視聴）間隔にクランピネスがある顧客の割合が4割以上と高いこと（ZBS, Figure 7）、クランピネスがないと判断された顧客よりもクランピネス指標計測後の予測期間でのF指標の値が大きいこと（ZBS, Figure 8）は注目に値する。これは、Web時代の新しいビジネスを行う企業のCRMにおいては、RFM指標に加えてクランピネス指標が重要であることを示唆する。

最後に、今後の研究課題を2つ述べる。第1に、本ノートで用いたクランピネス指標の検証に用いたデータの期間は約1年間と短かったため、データ期間を推定期間と予測期間に分割して、購買間隔にクランピネスがあることとクランピネス指標計測後の購買行動との関係を分析することはできなかった。顧客の長期間の購買行動を把握できる別のデータでそれを試みたい。第

(8) マルチサイド・プラットフォーム（マーケット）における企業の競争戦略については、中田（2013）や中田・西村（2016）の第7章・第8章を参照のこと。

顧客関係管理研究の新動向（中山雄司）

2に、ZBSはイベント発生間隔にクランピネスを生じさせる顧客の行動を確率モデルとして定式化はしなかった。しかし、最近になってPlatzer and Reutterer (2016)がその方向での新しいモデルを提案している。彼らのモデルの検証と更なる拡張は重要な研究課題である。

謝辞

本ノートの執筆に際して、JSPS 科研費基盤研究 (C)「オンライン・オフライン小売店舗の顧客関係管理と消費者の購買意思決定に関する研究」(JP 15K03729)の助成を受けたことを感謝する。また、中田善啓先生からのこれまでの励ましと助言に感謝する。

補論 Rによる実装

1 クランピネス指標の計算

まず、表3の形式の購買履歴データを整理してクランピネス指標を計算する。

```
# データセット（ファイル名「CRM.csv」のcsv形式）をRに読み込み、  
Datasetとして保存する。
```

```
Dataset = read.csv("CRM.csv")
```

```
#CustomerIDをファクタ（カテゴリ変数）に変換。
```

```
Dataset$CustomerID = factor(Dataset$CustomerID)
```

```
#Dateを日付として保存。
```

```
Dataset$Date = as.Date(Dataset$Date, format="%Y/%m/%d")
```

```
# データ構造の確認。
```

```
str(Dataset)
```

```
# データ冒頭8行を表示。
```

```
head(Dataset, n=8)
```

```

#CustomerID の水準を別オブジェクト Customer に保存。
Customer=levels (Dataset$CustomerID)
#Dataset 内の顧客数を ncust に保存し、その値を確認。
ncust=length (Customer)
ncust

# データ期首 (2010年12月1日) を日付データとして Start と名付ける。
Start = as.Date ("2010/12/1", format="%Y/%m/%d")
# データ期末 (2011年12月9日) を日付データとして End と名付ける。
End = as.Date ("2011/12/9", format="%Y/%m/%d")
# データ期間を数値データとして N と名付け、その値を確認。
N = End - Start
N = as.numeric (End - Start +1)
N

# 顧客 ID 別にリスト形式でデータを整理。
CustomerData=NULL
for (i in 1:ncust) {
  # 該当する顧客のデータのみ抽出し保存。
  X = Dataset [Dataset$CustomerID==Customer [i],]
  n = nrow (X) # 購買回数の保存。
  CustomerData [[i]]=list (Customer=Customer [i], X=X, n=n)
}
# 購買日間隔を計算し、CustomerData に追加保存。
for (i in 1:ncust) {
  DiffDate = diff (CustomerData [[i]]$X$date)

```


顧客関係管理研究の新動向 (中山雄司)

```
CustomerData[[i]]$DiffDate = as.numeric(DiffDate)
}

#Clumpiness 指標を計算に必要なデータの準備。3節の数式(2)の計算。
for (i in 1:ncust) {
  n = CustomerData[[i]]$n
  x=rep(NA,n+1)
  if(n>1){# 購買機会が2回以上の場合。
    x[1] = as.numeric(CustomerData[[i]]$X$Date[1] -
      Start + 1)
    x[2:n] = CustomerData[[i]]$DiffDate
    x[n+1]=N+1-as.numeric(CustomerData[[i]]$X$Date
      [n]-Start+1)
  }else{# 購買機会が1回の場合。
    x[1] = as.numeric(CustomerData[[i]]$X$Date[1] -
      Start + 1)
    x[n+1]=N+1-as.numeric(CustomerData[[i]]$X$Date
      [n]-Start+1)
  }
  x = x/ (N+1) # データ期間の長さを1に基準化。
  CustomerData[[i]]$x = x
}

#Clumpiness 指標の計算。
for (i in 1:ncust) {
  n = CustomerData[[i]]$n # 購買機会数の抽出。
  x = CustomerData[[i]]$x #3節の数式(2)の抽出。
```

```

Hp = 1 + sum(log(x) * x) / log(n+1) #3 節の数式 (1) の計算。
CustomerData[[i]]$Hp = Hp
}
# 出力の例 顧客 i (i=1,2,...,ncust)
#i に 1 から ncust の整数を代入すると, 対応する顧客の結果を出力される。
i=1
CustomerData[[i]]

```

2 クランピネスの検証

次に, 購買間隔がランダムであるという帰無仮説の下で, 棄却域の臨界値 (クランピネス指標の上側 $\alpha\%$ の分位点) を出力する関数を作成し, 上記の CustomerData に適用する。

```

# 臨界値を出力する関数の作成。
GenerateZtable = function(N,M,alpha) {
#N 潜在的な購買機会の数
#M シミュレーション回数
#alpha 検定の有意水準
Ztable = rep(NA,N) #n=1,2,...,N の場合の棄却域の臨界値を保存
するオブジェクトの生成。
for(n in 1:N) {
#M 回分のシミュレーションにおけるクランピネス指標を保存する
オブジェクトの生成。
Hp= rep(NA,M)
for(m in 1:M) {#3 節で提示した手順 2 の実行。

```

顧客関係管理研究の新動向 (中山雄司)

```
x = rep(NA, n+1) #3 節の数式(2)を保存するオブジェクトの生成。
tt=sort(sample(1:N, n))
#3 節で提示した手順 2(a) (b) の実行。
if (n>1) {
  x[1]=tt[1]
  x[2:n]=diff(tt)
  x[n+1]=N+1-tt[n]
}else{
  x[1]=tt[1]
  x[n+1]=N+1-tt[n]
}
x=x/(N+1) #3 節の数式(2)を保存。
#3 節で提示した手順 2(c) の実行。
Hp[m]=1 + sum(log(x)*x)/log(n+1)
}
Ztable[n] = quantile(Hp, probs=1-alpha)
#3 節で提示した手順 3 の実行。
}
return(Ztable) #n=1, 2, ..., N の場合の棄却域の臨界値を出力する。
}

# 関数 GenerateZtable を (N=374, M=20000, alpha=0.05) の場合に適用し, Ztable374 に保存。
Ztable374 = GenerateZtable(N=374, M=20000, alpha=0.05)
```

```
#CustomerData 内の顧客の購買間隔にクラumpinessがあるかを検証。
for (i in 1:ncust) {
  n = CustomerData[[i]]$n# 購買機会数の抽出。
  # 保存したクラumpiness指標が臨界値より大きい場合には TRUE を、
  # そうでない場合には FALSE を保存。
  CustomerData[[i]]$Clumpiness = CustomerData[[i]]$Hp >
  Ztable374[n]
}
```

3 RFMC 指標の整理

最後に、CustomerData に保存した結果を使って、RFMC 指標を取り出して整理する。

```
#RFMC 指標を保存するオブジェクトの生成。
RFMC = matrix(NA,nrow=ncust,ncol=6)
colnames(RFMC)=c("ID", "R", "F", "M", "C", "Clumpiness")
RFMC = data.frame(RFMC)
RFMC$ID = factor(RFMC$ID)# 顧客 ID をファクタに変換。

# 顧客別の RFMC 指標をオブジェクト RFMC に保存。
for (i in 1:ncust) {
  n = CustomerData[[i]]$n
  M = sum(CustomerData[[i]]$X$M)
  RFMC[i,1] = CustomerData[[i]]$Customer
  RFMC[i,2:4] = c(CustomerData[[i]]$x[n+1]-1,n,M)
  #RFM 指標の保存。
```

```
RFMC[i,5] = CustomerData[[i]]$Hp # クランピネス指標の保存。  
RFMC[i,6] = CustomerData[[i]]$Clumpiness  
# クランピネスの検証結果の保存。  
}
```

参考文献

- [1] 阿部誠 (2011) 「RFM 指標と顧客生涯価値：階層ベイズモデルを使った非契約型顧客関係管理における消費者行動の分析」日本統計学会誌 41(1): 51-81.
- [2] 中田善啓 (2013) 『プラットフォーム時代のイノベーション：クローズドからオープンビジネスへの進化』同文館出版.
- [3] 中田善啓・西村順二編著 (2016) 『先を読むマーケティング：新しいビジネスモデルの構築に向けて』同文館出版.
- [4] Abe M (2009) Counting your customers one by one: A hierarchical Bayes extension to the Pareto/NBD model. *Marketing Science* 28(3): 541-553.
- [5] Blattberg RC, Kim BD, Neslin SA (2008) *Database Marketing: Analyzing and Managing Customers* (International Series in Quantitative Marketing). Springer.
- [6] Chen D, Sain SL, Guo K (2012) Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining. *Journal of Database Marketing and Customer Strategy Management* 19(3): 197-208.
- [7] Fader PS, Hardie BGS (2009) Probability models for customer-base analysis. *Journal of Interactive Marketing* 23(1): 61-69.
- [8] Fader PS, Hardie BGS, Lee KL (2005) RFM and CLV: Using ISO-value curves for customer base analysis. *Journal of Marketing Research* 42(4): 415-430.
- [9] Kumar V, Reinartz W (2012) *Customer Relationship Management: Concept, Strategy, and Tools, Second Edition* (Springer Texts in Business and Economics). Springer.
- [10] Neslin SA (2014) Customer Relationship Management (CRM) in Winer RS, Neslin SA (2014) *The History of Marketing Science*, Chapter 11.
- [11] Platzer M, Reutterer T (2016) Ticking Away the Moments: Timing Regularity Helps to Better Predict Customer Activity. *Marketing Science* Articles in Advance
- [12] R Core Team (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- [13] Schmittlein DC, Morrison DG, Colombo R (1987) Counting your customers: Who are they and what will they do next? *Management Science* 33(1): 1-24.
- [14] Winer RS, Neslin SA (2014) *The History of Marketing Science* (World Scientific-Now Publishers Series in Business: Vol. 3). World Scientific Publishing.

- [15] Zhang Y, Bradlow ET, Small DS (2015) Predicting customer value using clumpiness: From RFM to RFMC. *Marketing Science* 34(2): 195-208.