

ダイレクトメールの配信効果と顧客行動に関する研究

A Study on the Effectiveness of Direct Mail Delivery and Customer Behavior

経営システム工学専攻 北嶋 弓月

1 研究背景・目的

1990年代以降、インターネット普及は広告業界にも影響を及ぼしている。従来主流であった新聞、チラシといった広告媒体は変化し、現代ではインターネット広告が主流となっている [1][2]。インターネット広告には大きく分けて3つの特徴（購買の即応性、双方向コミュニケーション、広告効果の可視化）があるとされている [3]。とりわけ広告効果の可視化は従来の広告と比較して最も異なる特徴と言える。インターネット広告には大きくわけて「Web 広告」と「ダイレクトメール（DM）広告」の2種類が存在する。Web 広告は新規顧客獲得や企業の認知率向上を目的に広告出稿を行っている。一方、DM 広告は自社サービスを利用する既存顧客を対象としており、リピーター、離反防止を目的に広告出稿を行っており、配信の目的が異なる [3]。しかし、一般的にこれらの広告効果測定指標にはインプレッション数やクリック数、コンバージョン数等同じ指標が用いられる。これらの評価指標を用いた Web 広告に関する研究は多くなされている一方、同じ評価指標を持つ DM に関する研究はそう多くはされていない。Web 広告と DM 広告は目的が異なることから、DM に関する研究を行うことは重要であると考えられる。

本研究では、DM 広告に着目し、過去に配信した DM、また ID-POS（購買・予約データ）から、どのような対象やコンテンツにすることがクリックや購買の向上につながるのかを特定することを目的とする。本研究では、DM 配信手法の異なる企業2社（A 社、B 社）のデータを用い、異なる分析手法を用いて2社それぞれの DM 効果推定を行う。

2 A 社の事例

本節では、A 社の分析事例について説明する。A 社は家電量販店の企業である。

2.1 A 社のデータセット

本研究では A 社のデータのうち、ID-POS データ、DM 配信データ、ユーザー属性データを用いた（表1~3）。

表 1: ID 付き POS データ概要

データ	ID-POS データ
期間	2017/01/01~2021/08/31
カラム	会員番号、取引時間、ストアコード、商品名、金額、etc...

表 2: DM 配信データ

データ	DM 配信データ
期間	2021/06/01~2021/06/30
カラム	会員番号、送信開封フラグ、送信日時、配信タイトル、etc...

表 3: ユーザー属性データ

データ	ユーザー属性データ
カラム	会員番号、性別、年齢、住所（都道府県）、住所（市区町村）、メールマガジン送信可否フラグ、初回利用日、最終利用日、カード利用回数
ユーザー数	998,600 ユーザー

本研究では、A 社の DM 配信データのうち、配信数や内容の関係からキャンペーンと新商品情報に関する7通の DM を用いた。DM の情報と送信ユーザー数は表4の通りである。

表 4: A 社の使用 DM 一覧

No.	DM 名	送信ユーザー数	フラグ
1	ゲーム発売に関するメール 1	30,103	g1
2	ゲーム発売に関するメール 2	29,793	g2
3	映像/音楽発売に関するメール 1	9,153	m1
4	映像/音楽発売に関するメール 2	9,723	m2
5	映像/音楽発売に関するメール 3	9,768	m3
6	パソコン販売に関するメール 1	125,667	s1
7	パソコン販売に関するメール 2	125,148	s2

2.2 A 社の分析

A 社のデータでは、クリックに影響を及ぼす変数に関する分析を行った。分析は次に示す流れで行った。まず、対象顧客を3セグメントに分類した。次に、クリックに影響を及ぼす変数の選択を行った。この結果を用いて、A 社の DM 効果推定モデルを構築した。

2.2.1 顧客の選定

まず、RFM 分析 [4] を用いてユーザーを3セグメントに分類した。3セグメントの分類方法としては、初めに RFM の各指標を5段階に分類し、これらの合計値を3セグメントに分類した。分類結果は表5の通りである。また、優良度は数値が高くなるほど優良度が高いものとした。

表 5: A 社の使用 DM 一覧

優良度	R (Recency)	F (Frequency)	M (Monetary)	SUM	セグメント
1	327 日以上	1~3 回	17,064 円以下	3~7	セグメント 1
2	326 日以下	4~5 回	17,065 円以上 54,970 円以下		
3	187 日以下	6~7 回	54,971 円以上 92,305 円以下	8~11	セグメント 2
4	108 日以下	8~19 回	92,306 円以上 238,490 円以下		
5	41 日以下	20 回以上	238,491 円以上	12~15	セグメント 3

2.2.2 モデルに用いる変数の選択

次に、モデル構築の際に用いる変数を決定するために、回帰分析による変数選択を行った。具体的にはロジスティック回帰分析 [4] を行い、ステップワイズ法による変数選択を行った。データはユーザーごとになっており、目的変数にはクリックの有無、説明変数には、A 社が取り扱うカテゴリ商品に対する過去の購買有無、RFM 指標、ユーザー属性を用いた。分析の結果、最も AIC が低くなった結果を表 6 に示す。

表 6: ステップワイズ法による変数選択結果

変数	Estimate	Std. Error	z value	$Pr(> z)$
(Intercept)	-0.92	0.14	-6.79	1.15×10^{-10}
last_days	-1.50×10^{-3}	2.32×10^{-4}	-6.46	1.16×10^{-10}
frequent	0.0021	0.0019	1.09	0.27
amount	-1.67×10^{-7}	1.36×10^{-7}	-1.23	0.22
SERVICE	0.26	0.15	1.69	0.09
HOUSE.EQ	-0.58	1.05	-0.55	0.58
sex_label	-0.020	0.02	-0.92	0.36
age_label	0.19	0.02	9.70	$< 2 \times 10^{-16}$
PC	0.10	0.06	1.68	0.093
ELECTRONICS	0.15	0.23	0.64	0.521
MEMORIES	0.39	0.18	2.14	0.033

2.2.3 階層ベイズモデルを用いた DM 効果測定モデルの構築

さらに続いて、表 6 の結果を用い、個人差を考慮するために階層ベイズモデルによる DM 効果測定モデルの構築を行った。パラメータの推定方法として、マルコフ連鎖モンテカルロ法 (MCMC) [5] を用いた。各変数の事前分布のは表 7 の通りに設定した。モデルの目的変数には DM に記載されている URL のクリックの有無を用いている。

表 7: モデルの事前分布の設定

変数	変数名	変数	データ形式	事前分布
目的変数	CLICK	y_i	クリックの有無 (有:1, 無:0)	ベルヌーイ分布
説明変数	SERVICE	β_j	過去の購買有無 (有:1, 無:0)	コーシー分布
	HOUSE.EQ			
	ELECTRONICS			
	PC			
	MEMORIES			
	Recency	β_j	数値	一様分布
	Frequency			
	Monetary			
	sex_label			
	age_label			
Intercept	β_0	-	-	t 分布

表 7 に記載した変数を用いて、DM 効果推定モデルの構築を行った。本研究では、DM のフラグをランダム効果として加味したランダム切片モデルを用いた。モデルは式 (1)~(3) の通りである。

$$r_{1k} \sim t(3, 0, 2.5) \quad (1)$$

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \dots + \beta_j x_{ij} + r_k \quad (2)$$

$$y_i \sim \text{Bernoulli}(\theta) \quad (3)$$

ここで、 x_{ij} はユーザー i 、変数 j の値、 β_j はパラメータ (ただし β_0 は切片)、 r_k はランダム切片を表し、 y_i はク

リックの有無を表す (詳細は表 7 記載)。本研究ではこのモデルを RFM 分析の各セグメントごとに用いた。また、本研究では R の brms を用いて実装を行った。iter 数は 15000、warm up 期間は 10000、間引き数を 5 とした。

2.3 階層ベイズモデルによる推定結果

本抄録では、セグメント 1 の推定結果についてのみ記載する。セグメント 1 のモデルパラメータ推定結果を表 8 に示す。表 8 から、全ての変数においてパラメータの収束を示す \hat{R} が 1.1 以下であることがわかる。また、商品カテゴリのうち、メモリや住宅設備のパラメータ推定値が高く、反対にプリンタなどの情報家電が含まれるエレクトロニクス、PC は低くなっていることがわかる。

表 8: セグメント 1 のパラメータ推定結果

Variable Name	Estimate	\hat{R}
Intercept	-1.19	1.00
last_days	0.00	1.00
frequent	0.00	1.00
amount	0.00	1.00
PC	-0.12	1.00
HOUSE.EQ	0.25	1.01
MEMORIES	0.47	1.00
ELECTRONICS	-0.26	1.00
SERVICE	0.02	1.00
sex_label	0.03	1.00
age_label	0.20	1.00
ml_flag	0.18	1.00

表 9: 精度検証の結果

	セグメント 1	セグメント 2	セグメント 3
Accuracy	0.58	0.62	0.61

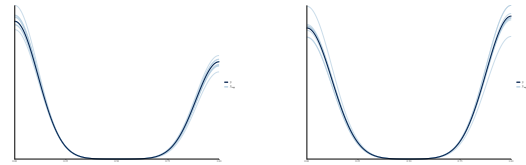


図 1: クリックの事後分布 (セグメント 1, 2)

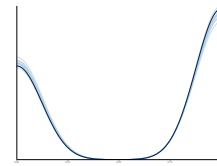


図 2: クリック率の事後分布 (セグメント 3)

本研究ではモデルの精度検証に Accuracy を用いた。表 9 はセグメント 1~3 の予測精度を検証した結果である。平均して Accuracy は 0.6 程度であることがわかる。また、事後分布の結果を図 1, 2 に示す。メールフラグとクリックの有無のデータ数は合わせているものの、セグメントごとにモデルを実行していることから、特にセグメント 1 と 3 においては不均衡であることが読み取れる。

3 B社の事例

次に、B社の事例について説明する。A社ではユーザー個別の行動について分析を行ったが、実際には必ずしも全ての企業がセグメント配信を行っているわけではない。そのため、B社では全体の行動に着目する。B社は生活関連サービスを取り扱う企業である。

3.1 B社のデータセット

本研究では、B社のデータセットから予約に関するデータとDM配信データを用いた(表10, 11)。

表10: 予約に関するデータ

データ	予約データ
期間	2018/01/01~2021/09/30
カラム	会員番号, 取引時間, ストアコード, サービス名, 金額, etc...

表11: DM配信データ

データ	DM配信データ
期間	2018/01/01~2021/09/30
カラム	配信時間, 配信タイトル, 本文, 開封率(日別), クリック数(日別), etc...

本研究では、B社の予約カテゴリのうち、過去にDM配信したことのある14カテゴリを対象とした。カテゴリにはインハウスサービスやアウトハウスサービス、レジャーサービスなどが含まれる。また、本研究では、DMの対象期間を2019年11月9日~2021年6月30日までとし、この期間に配信されたDM127通を対象とした。さらにDMを目視によって6種類に分類し、分析に用いた。種類ごとのDM数を表12に示す。

表12: B社のDM種類と配信数

DM種類	配信数
カテゴリ紹介	38
生活役立ち記事特集	28
キャンペーン	25
季節	24
人気カテゴリ特集	6
その他	6

3.2 B社の分析

B社の分析について説明する。B社の分析では、まず対象顧客の選定と予約数の定義を行った。その後、DM予約販促効果モデルの構築を行った。

3.2.1 対象顧客の選定

本研究では、DMの対象期間が2019年11月9日からであることから、対象顧客を2019年の11月9日以前に会員登録をしているユーザーのみとした。また、4年間での累計予約数が3回以上のユーザーを対象とした。したがって、本研究で対象としたユーザー数は約30万ユーザーであった。

3.2.2 予約数の定義

B社では、基本的に一斉送信を行っておりユーザーのIDは紐付け不可である。本研究内ではB社のDMは送

信してすぐに予約が行われるという仮定の下、次のように各DMの予約数を定義した(式(4))。

$$Res_m = res_t + res_{t+1} \quad (4)$$

ここで、 Res_m は m 通目のDM予約数を表しており、 t は m 通目のDMを送信した日にちを表す。この定義に従って、DM予約数を時系列グラフに表したものが図3である。

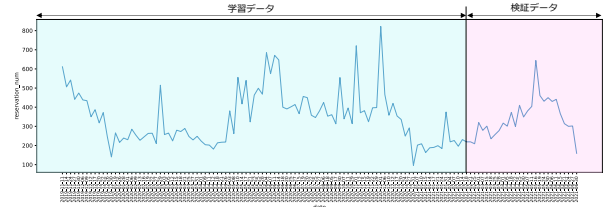


図3: B社のDMごとの予約数分布

3.2.3 予約販促効果推定モデルの構築

前述した対象ユーザーと予約数の定義の下、本研究では自己回帰モデル[6]を用いた予約販促効果推定モデルの構築を行った。本研究では、山根ら[7]の先行研究を参考にモデルの構築を行っている。まず、 t 地点におけるDM予約数(目的変数)を y_t 、DM種類 n の有無を $z_t^{(n)}$ 、DM単体の効果(潜在変数)を x_t とする。DMの効果 x_t は2期の自己回帰性を持つものと仮定する(式(5))。また、DMの種類を加味する(式(6))。 a_k は k 期前のDM影響パラメータ、 b_n はDM種類 $z_t^{(n)}$ の販促効果を表すパラメータである。数パターンのモデルを試したのち、本研究におけるDM予約販促効果推定モデルは式(7)とした。なお、 c は切片、 e_t は残差を表す。本研究のモデルの決定には、精度指標MAE, RMSEを用いた。

$$x_t = \sum_{k=1}^2 a_k x_{t-k} + e_t \quad (5)$$

$$y_t = \{x_t\} \sum_{n=1}^6 b_n z_t^{(n)} \quad (6)$$

$$y_t = \left\{ c + \sum_{k=1}^2 a_k y_{t-k} \right\} \sum_{n=1}^6 b_n z_t^{(n)} \quad (7)$$

本研究では、式(7)を最小二乗法を用いて推定した。また、パラメータ推定では、残差平方和 $\sum e_t^2$ を最小にするようにパラメータの推定を行った。パラメータ推定の実行には、Rのoptim関数を用いた。自己回帰モデルは初期値依存があることから、初期値にはNelder-Mead法による推定値を用い、パラメータ推定にはBFGS公式による準ニュートン法を用いた。

3.3 DM予約販促効果推定モデルによる分析結果

パラメータを推定した結果を表13に示す。DMの影響を表すパラメータ a_k は、1つ前のDM影響を表す a_1

よりも、2つ前のDM影響を表す a_2 の方が高かった。また、DMの種類効果を表すパラメータ b_n では、あまり大きな変化は見られないものの、「季節」、「人気カテゴリ速報」のパラメータ値が少し高いことがわかる。

また、学習データにモデルをフィッティングさせた分布と、検証データに当てはめた予測分布を図4、5に示す。

表 13: パラメータ推定結果

Parameters	Name	Predict
Intercept	—	194.06
a_1	1つ前のDM影響	1.07
a_2	2つ前のDM影響	1.11
b_1	カテゴリ紹介	0.85
b_2	生活役立ち記事特集	0.84
b_3	キャンペーン	0.85
b_4	季節	0.86
b_5	人気カテゴリ速報	0.88
b_6	その他	0.84

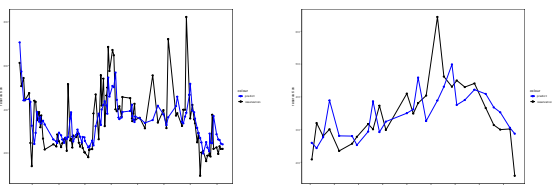


図 4: 学習データに当てはめた分布比較
図 5: 検証データに当てはめた分布比較

時系列モデルであることから、多少の周期ずれは見られるが、全体の傾向は捉えられていることがわかる。一方、傾向は捉えられているものの、急増、急降の挙動全ては捉えきれていないこともわかる。

4 考察

4.1 A社の分析結果に対する考察

2.3節の分析結果と、ランダム効果別事後平均の結果(表14)から考察を行う。s1, s2のメールのうち、セグメント1ではs2の事後平均の方が低いのにに対して、セグメント2, 3ではs1の事後平均の方が高い値を取っている。s1, s2のDM内容は同じ内容であり、s2はs1の後追いとして送信されたメールである。このことを踏まえると、優良顧客には同内容の再送付も効果的であるが、非優良顧客には1通目のDMが重要な役割を果たすと推察できる。

また、A社のセグメント配信には現状として商品購買の有無、ECサイト閲覧履歴の有無などを指標としている。しかし、表14から、セグメント3では週ごとのDMでも反応があるのに対してセグメント1, 2ではパソコンのメールでのみ大きな反応があることがわかる。したがって本研究で用いたRFM指標によるセグメントもDMの配信指標として取り入れることが有効であると推察できる。

表 14: ランダム効果別の事後平均結果

カテゴリ	メールフラグ	セグメント1	セグメント2	セグメント3
ゲーム	g1	-0.061	-0.072	0.032
ゲーム	g2	-0.011	-0.044	0.047
映像	m1	-0.034	-0.199	0.039
映像	m2	-0.006	0.022	-0.022
映像	m3	-0.089	0.151	-0.015
パソコン	s1	0.140	0.055	-0.093
パソコン	s2	0.057	0.111	-0.044

4.2 B社の分析結果に対する考察

B社の分析結果からは、1つ前のDM影響パラメータよりも2つ前のDM影響パラメータの方が相対的に高い値であった。仮説とは異なり、全員がすぐに予約を行うのではなく、半数以上の顧客は日数をおいて予約する傾向にあるのではないかと推察した。昨今の新型コロナウイルス感染症によりサービス内容だけではなく、衛生面にも考慮する顧客が増加している可能性があり、このことが影響しているとも考えられる。

5 結論

本研究では、DM配信方法が異なる2つの企業を対象に、それぞれに対してDM配信の効果測定を行った。A社では階層ベイズモデルを用いたDM効果推定モデルを構築し、B社ではDM全体に対する予約販促効果推定モデルを構築した。全体の結果から、DMの配信方法やデータの取得方法によって分析手法を臨機応変に変える必要性があると考えられる。

参考文献

- [1] 総務省, 「令和3年情報通信白書」, <https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r03/html/nd242120.html>, 2021, 最終閲覧日: 2022/2/7
- [2] 電通, 「2020年日本の広告費」, <https://www.dentsu.co.jp/news/release/pdf-cms/2021012-0225.pdf>, 2020, 最終閲覧日: 2022/2/7
- [3] Frost, R., Strauss, J., 麻田孝治(翻訳), 「インターネット・マーケティング概論—ネット時代の新たなマーケティング戦略と手法」, ピアソンエデュケーション, 2000
- [4] 川端一光, 岩間徳兼, 鈴木雅之, 「Rによる多変量解析入門 データ分析の実践と理論」, オーム社, 2018
- [5] 花田政範, 松浦壮, 「ゼロからできるMCMC マルコフ連鎖モンテカルロ法の実践的入門」, 講談社, 2020
- [6] 沖本竜義, 「経済・ファイナンスデータの計量時系列分析」, 朝倉書店, 2010
- [7] 山根智之, 菅原光太郎, 西村直樹, 小林健, 吉田佑輔, 高野祐一, 中田和秀, 「時系列モデルによる商品販促効果の分析」, オペレーションズ・リサーチ, Vol. 61, pp. 65-70, 2016