

# 傾向スコアを用いた広告効果の シミュレーション ～マクドナルドを事例として～

---

株式会社ミスミ  
服部達哉

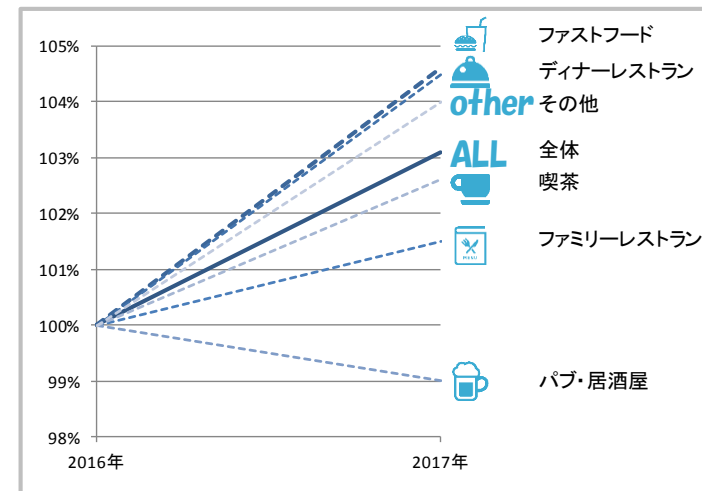
# Contents

---

- 外食業界の動向
- マクドナルドのここ最近でのテレビCM出稿量
- CM視聴とマクドナルド利用頻度に対する共変量の存在
- 傾向スコア逆確率重み付け法による共変量調整
- ATE/ATT/ATUの定義と分析への展開
- ATE/ATT/ATUによる推定の具体的手順
- リサーチクエスチョンとアプローチ
- 分析に用いたデータ
- 傾向スコアの推定
- IPW推定結果
- CMの特徴も踏まえた考察
- 今後の課題
- 引用・参考資料

# 外食業界の動向

- 一般社団法人日本フードサービス協会「外食産業市場動向調査」によると、2017年の外食需要は対前年比で堅調に推移し、とりわけファストフードは2016年比で104.6%と各業態の中で最も高い伸び率だった。



※一般社団法人日本フードサービス協会会員社による外食産業市場動向調査 2017年年間結果報告

- そのファストフード業界では、タコベルやシェイクシャックの日本進出など、競争環境は厳しさを増している。
- そんな中でファストフードの雄である日本マクドナルドは、一時期の苦境を乗り越え、業績は回復してきており、「夜マック」などの新商品を投入して厳しい競争に打ち勝とうとしており、当然ながら新商品他のテレビCMを投下して来店を促している。

# マクドナルドのここ最近でのテレビCM出稿量

- 日本マクドナルドのCM出稿状況を、野村総研インサイトシグナルマーケティング分析コンテスト用のCM出稿データで2018年1月から3月分を確認すると下記のとおりである。尚、ここでの「タイプ」は、各々のCMが特定の商品のCMか、企業としての取り組みなどを示したものを表している。

	CM	タイプ	開始時期	終了時期	出稿回数
1	おてごろマック	商品	2018/2/13	2018/2/27	117
2	カケテミーヨ チーズボロネーゼ	商品	2018/2/27	2018/3/12	138
3	カフェラテ	商品	2018/1/25	2018/3/31	196
4	クルー体験会	企業・CSR	2018/3/7	2018/3/15	184
5	チキンタツタ/チキンタルタ	商品	2018/2/6	2018/2/26	144
6	ハッピーセット	商品	2018/1/25	2018/3/31	235
7	ハッピーりぼーんプロジェクト	企業・CSR	2018/2/22	2018/3/31	250
8	はみだすハムてりたま/チーズてりたま/てりたま	商品	2018/3/6	2018/3/26	117
9	プレミアムローストコーヒー	商品	2018/2/23	2018/3/2	133
10	プレミアム三角チョコパイ ヘーゼルナッツ	商品	2018/1/25	2018/2/5	95
11	マックシェイク×カフェオーレ	商品	2018/3/28	2018/3/31	64
12	マックフルーリー 超オレオ	商品	2018/2/20	2018/3/5	144
13	マック謎とき探検隊 おいしさのヒミツ	企業・CSR	2018/1/30	2018/3/19	48
14	夜マック	商品	2018/3/19	2018/3/31	135

# CM視聴とマクドナルド利用頻度に対する共変量の存在

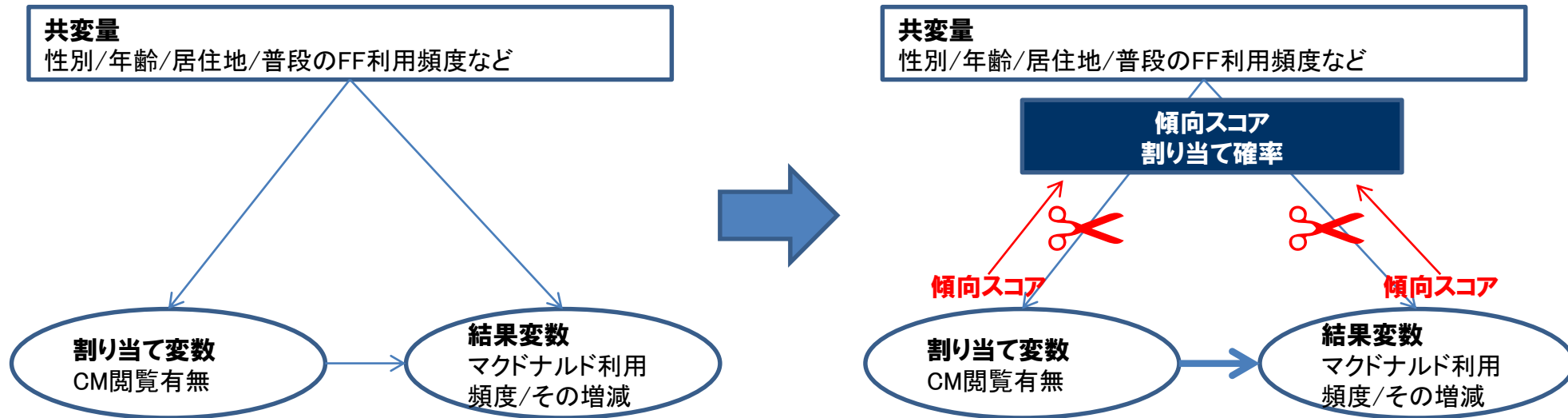
- 前頁におけるマクドナルドCMを見たか見ないか、加えてマクドナルドへの来店頻度を人口動態特性や行動特性別に集計すると以下のとおりである。
- 普段のファストフード利用頻度や性別、職業によって、マクドナルドのCM閲覧有無やマクドナルドの利用頻度に違いがあることが分かる。性別・職業・普段のファストフード利用頻度が、マクドナルドのCMを見るか見ないか、マクドナルドの利用頻度が高いか低いかの双方に影響を与える、「共変量」となっていることが分かる。共変量の影響があると、CMと利用頻度の関係を正しく推定できないとされている。

		普段のファストフード利用頻度				性別				職業			
		週1回以上		週1回未満		男性		女性		専業主婦・主夫		それ以外	
		件数	縦%	件数	縦%	件数	縦%	件数	縦%	件数	縦%	件数	縦%
3/8～3/25間のマクドナルドCM閲覧	計	485	100%	2330	100%	1546	100%	1454	100%	456	100%	2544	100%
	テレビCMを見た	93	19%	513	22%	415	27%	270	19%	59	13%	626	25%
	テレビCMを見ていない	392	81%	1817	78%	1131	73%	1184	81%	397	87%	1918	75%
3/7時点での直近1か月内でのマクドナルド利用頻度	計	441	100%	2188	100%	1449	100%	1306	100%	415	100%	2340	100%
	週に1回以上	61	14%	26	1%	62	4%	30	2%	6	1%	86	4%
	月に2～3回程度	127	29%	141	6%	155	11%	128	10%	48	12%	235	10%
	月に1回程度	136	31%	578	26%	394	27%	365	28%	124	30%	635	27%
	ここ1ヶ月で利用していない	116	26%	1382	63%	795	55%	761	58%	235	57%	1321	56%
	利用したことがない	1	0%	61	3%	43	3%	22	2%	2	0%	63	3%

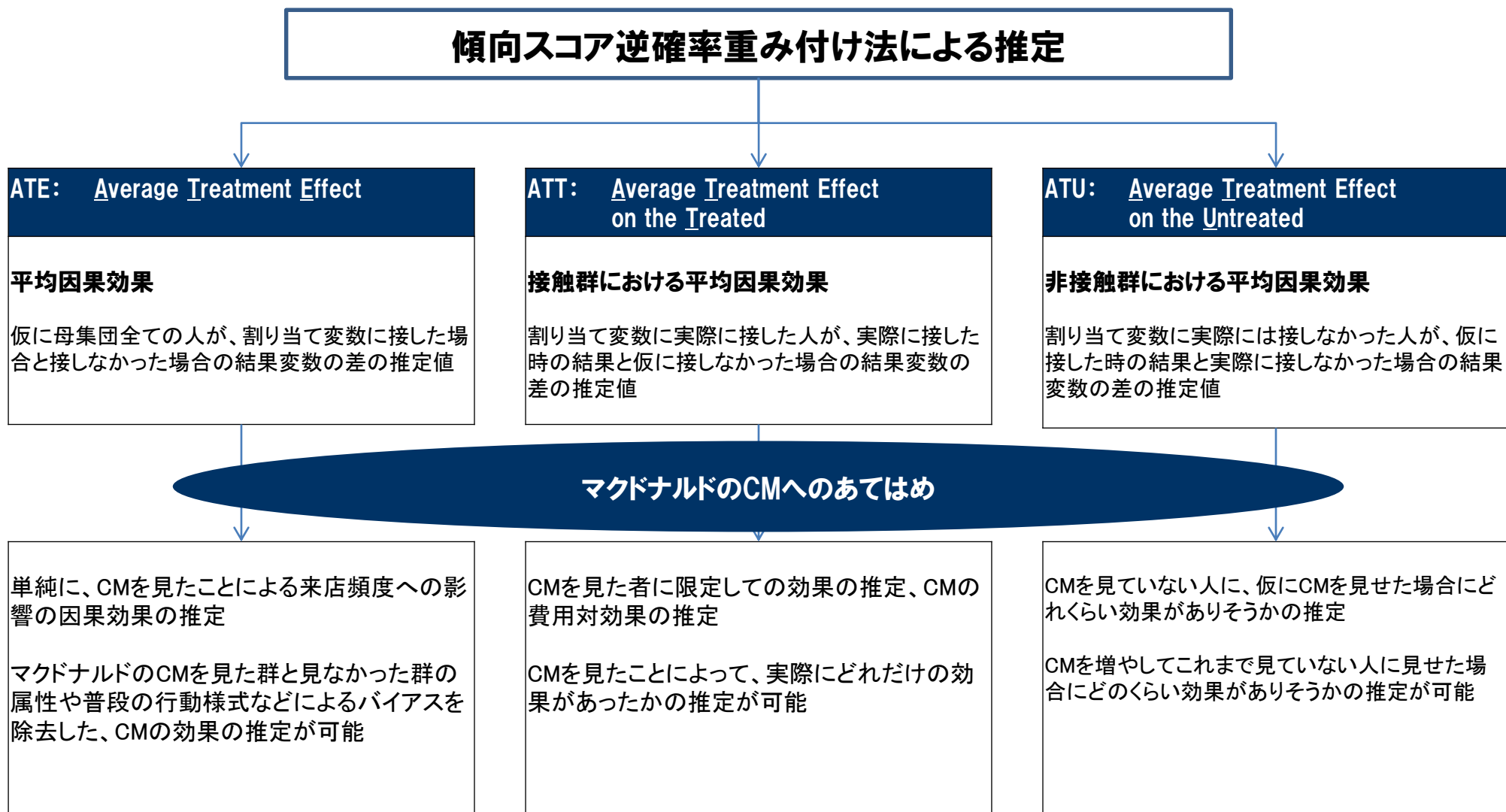
※「マクドナルドのCM閲覧」は、当該期間中にマクドナルドのテレビCMが投稿されているテレビ番組を視聴したか否かにより定義。ただし、「クルー体験会」のCMは趣旨が異なるため除く。

# 傾向スコア逆確率重み付け法による共変量調整

- 様々な共変量が割り当て変数（CM閲覧有無）と結果変数（マクドナルド来店頻度）に与える影響（バイアス）を取り除き、割り当て変数と結果変数の因果関係を推定する方法として、傾向スコア逆確率重み付け法がある。
- 傾向スコアとは、様々な共変量をもとに「ある個体（回答者）が割り当て変数に割り当てられる確率」のことを示している。その確率で割り戻して（即ち確率の逆数を乗じて）集計分析することで、割り当て変数に全対象者が等しい確率で割り当てられたとみなし、純粋な割り当て変数による因果効果を推定することを可能とするのがIPW（Inverse Probability Weighting: 逆確率重み付け法）である。



# ATE/ATT/ATUの定義と分析への展開



# ATE/ATT/ATUによる推定の具体的手順

## 傾向スコアの推定

共変量を説明変数に、割り当て変数を目的変数に、ロジスティック回帰分析で「CM接触群となる確率」を推定



## IPW推定

- ATEは接触群の個体(個別の回答者)を傾向スコアで、非接触群の回答者を1-傾向スコアで、それぞれ「当該群となる確率」とし、その逆数で重み付け。
- ATTでの接触群の値とATUでの非接触群の値は真の値であるため、ATEの式を下記のように展開して各個体に対するウェイト値を算出。
- ATE/ATT/ATUそれぞれ、下記のウェイト値を個別の回答者に重み付けして推定

	接触群	非接触群
ATE	$1/\text{傾向スコア}$	$1/(1-\text{傾向スコア})$
ATT	1	$\text{傾向スコア}/(1-\text{傾向スコア})$
ATU	$(1-\text{傾向スコア})/\text{傾向スコア}$	1

接触群を「1」にするため、ATEの係数の両辺に傾向スコアを乗じる

非接触群を「1」にするため、ATEの係数の両辺に1-傾向スコアを乗じる



# リサーチクエストとアプローチ

---

本研究においては、マクドナルドのCMについて以下の点を考察する。

- ① バイアスを除去した、純粹なCMの因果効果はどの程度と推定されるか？
  - ATEを用いて、単純なクロス集計と比較するとどのような傾向がみられるか？
- ② どのクリエイティブが、実際にどの程度効果があったと推定されるか？
  - ATTを用いて、個別のクリエイティブの効果はどの程度だったか？
- ③ どのクリエイティブの出稿量を増やせば、さらなる効果が見込めるか？
  - ATUを用いて、ATTで効果の見られたクリエイティブの出稿量増加シミュレーションはどの程度になるか？

# 分析に用いたデータ

## データセット

「野村総研インサイトシグナルマーケティング分析コンテスト2018」  
分析用アンケートデータおよびTVCM出稿データ



## 有効サンプルと 変数の加工

- 全3000サンプル中、2018/3/7時点での直近1か月でのマクドナルド利用頻度と、2018/3/26時点での直近1か月でのマクドナルド利用頻度、並びに基本属性などで欠測値がない、2479サンプルを有効サンプルと定義。
- 結果変数としての利用頻度は名義尺度になっているため、増減や差分の把握のしやすさを考慮して、「月当たり利用回数」に換算したデータを用いた。
- この2479サンプルに、TVCM出稿データから、2018年3月8日～2018年3月25日にマクドナルドのCMを見たか否かを定義。
- マクドナルドのテレビCMは、商品の宣伝・訴求ではない「クルー体験会」のようなCMを除き、以下のCMを指定期間内に見たか否かをダミー変数化。

CMクリエイティブ	開始日	終了日	放送回数
カケテミーヨ チーズボロネーゼ	3月8日	3月12日	11
カフェラテ	3月20日	3月25日	89
ハッピーセット	3月8日	3月25日	50
ハッピーリぼーンプロジェクト	3月8日	3月11日	12
はみだすハムてりたま/チーズてりたま/てりたま	3月8日	3月24日	103
マック謎とき探検隊 おいしさのヒミツ	3月13日	3月19日	16
夜マック	3月19日	3月25日	121

# 傾向スコアの推定

- 年齢や性別、世帯構成、世帯年収により、マクドナルドの選好度は違うはず。また、居住地により店舗数の多寡も異なるはず。そもそも、ファストフードによく行くか否かによってもマクドナルドの利用頻度は異なるはず。この仮説に基づき、左表の変数を傾向スコア推定のロジスティック回帰説明変数とした。
- 前頁のCMクリエイティブのいずれかを見た場合に、「マクドナルドのCMを見た」として、また商品の個別CMについては放送回数が50回以上(前頁参照)と多いCMについて、それぞれ視聴有無を割り当て変数として傾向スコアを推定した。

ロジスティック回帰説明変数
年齢
男性ダミー
東京都ダミー
神奈川県ダミー
専業主婦(主夫)ダミー
単身世帯ダミー
夫婦と未婚の子世帯ダミー
年収400万円以上ダミー
ファストフード「ほとんど毎日」ダミー
ファストフード「週2~3回程度」ダミー
ファストフード「週に1回程度」ダミー
ファストフード「月に1~2回程度」ダミー

※上記の「ダミー」は、on/offのダミー変数の意。



割り当て変数
マクドナルドのCM いずれかを視聴
カフェラテCMを視聴
ハッピーセットCMを視聴
はみだすハムてりたま/チーズてりたま/てりたまCMを視聴
夜マックCMを視聴

# IPW推定結果～ATEによる全体比較

- 3/7時点と3/26時点の結果変数「直近1か月以内のマクドナルド平均利用回数」を、マクドナルドCM閲覧別に単純平均並びにIPW推定結果で見たものが以下である。
- CM閲覧有無で分けた両群の3/7時点のマクドナルド利用頻度は、単純平均に比べてIPW推定で差が小さくなっていることから、バイアスが補正されていることが分かる。IPW推定においては、3/7時点と3/26時点でのCMを見なかった群における利用頻度にほとんど差がみられないことから、IPW推定でバイアスが補正されていると推察される。
- IPW推定でのCMを見た群におけるマクドナルド利用頻度は、3/7時点から3/26時点で多くなっており、CMによる来店頻度増加の効果があったと言える。

	直近1か月マクドナルド平均利用頻度(単位:回)					
	単純平均			IPW推定		
	3/7時点	3/26時点	増減	3/7時点	3/26時点	増減
CMいずれかを見た群	0.6951	0.7705	0.0755	0.6859	0.7623	0.0764
CMいずれかも見なかった群	0.6187	0.6525	0.0338	0.6430	0.6497	0.0068
差	0.0763	0.1180	0.0417	0.0429	0.1125	0.0696
p値	0.148	0.027		0.458	0.038	

※マクドナルド月利用頻度の平均値は、選択肢「週に1回以上」を4.5回、「月に2～3回」を2.5回、「月に1回程度」を1回、「ここ1か月で利用していない」「これまで利用したことがない」をそれぞれ0回として算出。また、平均値の差の検定は、 $p < 0.05$ で判断し、IPW推定においてはRのsurveyパッケージを利用。次頁以降も同様。

# IPW推定結果～ATTによる個別クリエイティブの効果

- カフェラテCMとハッピーセットCMでは、3/7時点も3/26時点も、見た群の実測値と彼らが仮に見なかった場合での利用回数において有意な差は見られない。一方で、てりたまCMでは、3/7時点では両群で有意な差はないものの、3/26時点では有意な差がみられる。特に見た群では利用回数が増加しており、特にこのてりたまCMで、CMを見た者にとって実際の来店頻度増加の効果があったと言える。
- 夜マックCMでは、CM放映開始(3/19)前の3/7時点で既に双方に有意な差が見られていることから、傾向スコア推定の説明変数では除去しきれなかったバイアス(例えば夕食での外食頻度、など)があることが示唆される。

	直近1か月マクドナルド平均利用頻度(単位:回) IPW推定											
	カフェラテCM			夜マックCM			てりたまCM			ハッピーセットCM		
	3 / 7 時点	3 / 26 時点	増減	3 / 7 時点	3 / 26 時点	増減	3 / 7 時点	3 / 26 時点	増減	3 / 7 時点	3 / 26 時点	増減
CMを実際に見た場合	0.6896	0.7674	0.0778	0.6976	0.7638	0.0663	0.6923	0.7745	0.0822	0.7370	0.7864	0.0495
CMを実際に見た者が仮に見なかった場合	0.6449	0.7018	0.0569	0.5970	0.6703	0.0733	0.6425	0.6560	0.0135	0.6637	0.7327	0.0690
差	0.0447	0.0655	0.0209	0.1006	0.0935	-0.0071	0.0498	0.1185	0.0687	0.0732	0.0538	-0.0195
p値	0.301	0.138		0.029	0.043		0.333	0.019		0.184	0.321	

# IPW推定結果～ATUによるシミュレーション

- 実際にCMを見た群で、仮に見なかった場合に比べて(ATT)の来店頻度向上が認められた「てりたまCM」だが、実際にはCMを見なかった者が仮に見た場合にどの程度の来店頻度増加効果が見られそうかをシミュレーション(ATU)すると、以下のとおりである。参考として、同時期のCM放映だった「ハッピーセット」でもATUを推定した。
- てりたまCMでは、3/7時点では、実際に見なかった者の来店頻度と彼らが仮に見た場合の頻度の差はわずかだが、3/26時点では差は拡大、差も有意になっている。仮に見た場合の推定値自体も、3/7から3/26で大きく向上している。
- ハッピーセットCMにおいては、実際に見なかった場合と仮に見た場合の差は、3/7時点と3/26時点ともに、有意なものとはなっていない。
- てりたまCMでは、実際に見なかった者にとって見る機会が増えるような出稿量の増加は、さらなる来店頻度向上が期待できると考えられる。

	直近1か月マクドナルド平均利用頻度(単位:回) IPW推定					
	てりたまCM			ハッピーセットCM		
	3/7時点	3/26時点	増減	3/7時点	3/26時点	増減
CMを実際に見なかった者が仮に見た場合	0.6652	0.7579	0.0927	0.7381	0.7938	0.0557
CMを実際に見なかった場合	0.6173	0.6540	0.0367	0.6685	0.7342	0.0657
差	0.0479	0.1039	0.0561	0.0696	0.0596	-0.0100
p値	0.332	0.038		0.176	0.254	

# CMの特徴も踏まえた考察

- ATTを検証したCMの特徴を右表のように整理した。
- ATTにより因果効果が認められたてりたまCMと、認められなかったハッピーセットでは、CMの放送開始日は同じだったため、事後調査までの放送日数も同様に18日間である。両者の違いは商品の特徴であり、ハッピーセットは定番商品である一方、てりたまは季節限定商品である。このことから、定番商品のCMは新たに来店を促す効果は期待できないが、季節限定商品(今回の研究では、夜マックではほかにもバイアスが存在する可能性が示唆されたが、既存の定番商品ではないという意味で場合によっては新商品も)では、来店を促す効果が期待できる、と言える。
- ATUにおいても、てりたまCMは非接触者に仮に接触させた場合に来店頻度向上の効果が期待できる。CMの機能は単に来店を促すだけにとどまらないが、CM投下の予算が限られている中で、仮に広告効果の最大化を図るのであれば、定番商品よりも、季節限定商品や新商品の出稿量を一定の費用内で最大化することが有効と考えられる。

CM	商品タイプ	事後調査 (3/26時点) までの放送日数
カフェラテ	定番商品	6
夜マック	新商品	7
てりたま	季節限定	18
ハッピーセット	定番商品	18

# 今後の課題

---

- 本研究では、共変量によるバイアスを除去することで、CMによる因果効果を正しく推定できることが示された。一方で、下記の点については本研究では考慮されておらず、今後の課題と考えられる。
- CMを見たか見ないかを、そのTV番組を視聴しているか否かで定義しており、そのCMを本当に見たか見なかったかという実態レベルでの因果効果を検証できていない。すなわち、例えばTV番組を視聴していたが、CM中だけ他のチャンネルに合わせた場合や所用を済ませていた場合など、CMそのものは実際には見ていなかった可能性を考慮できていない。これらをデータ化し同様の分析を行えば、異なる結果が得られる可能性がある。
- 個別のCMでのATTの検証を行ったが、事前調査日時3/6と事後調査日時3/26の間に、複数のCMを視聴している場合もありうる。組み合わせでの来店喚起効果も想定されるが、その考察はできていない。組み合わせの効果を検討すると、異なる結果が得られる可能性がある。



# 引用・参考資料

---

- 一般社団法人日本フードサービス協会 <http://www.jfnet.or.jp/data/online.html>
- 『岩波データサイエンス Vol.3』岩波データサイエンス刊行委員会編、岩波書店、2016  
同、サポートサイト <https://sites.google.com/site/iwanamidatascience/vol-3/vol3-ingasuiron>  
<https://github.com/iwanami-datascience/vol3>
- 『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（シリーズ確率と情報の科学）』  
星野崇宏、岩波書店、2009
- WEBサイト“読書日記：統計ソフトの「ウェイト」は調査の「ウェイト」ではない”  
<http://elsur.jpn.org/mt/2013/05/001729.html>
- 『原因と結果』の経済学』中室牧子・津川友介、ダイヤモンド社、2017
- 『統計的因果推論（統計解析スタンダード）』岩崎学、朝倉書店、2015
- 『実証分析入門 データから「因果関係」を読み解く作法』森田果、日本評論社、2014