

シングルソースデータが可能にする 広告効果測定・広告戦略策定について

ベイジアンネットワークを活用したグラフィカルモデリングと確率推論

慶應義塾大学大学院政策・メディア研究科
田中 渚子

1 はじめに

- L (1) 問題意識と研究目的
- L (2) 対象商品と市場環境

2 分析の準備

- L (1) 分析の枠組みと流れ
- L (2) 変数の作成と整理
- L (3) モデルの概要と利点

3 データ分析

- L (1) データの探索
- L (2) 仮説の導出
- L (3) 広告効果分析
- L (4) 広告戦略への展開

4 まとめ

- L (1) 分析からの示唆
- L (2) モデルの有用性と課題
- L (3) 参考文献など

1-1. 問題意識

1 消費者のメディア接触行動の複雑化

スマートフォンやソーシャルメディア、動画サイトなどの普及により、若年層を中心に消費者のメディア接触行動は複雑化している。このような状況下で企業は様々な形態の広告戦略を実施しているが、**広告効果指標は広告チャネルによって分断**されているため、**チャネルを横断した消費者起点の広告効果測定ができず**、チャネル間での適切な広告予算配分が難しくなっている。

2 成熟市場における消費者ニーズの多様化

成熟市場においては、消費者のニーズはより洗練され、多様化する傾向にある。そのような状況下では、商品の実用性だけでなく、ライフスタイルに合うか、他人にどう見られるかなど、**自身のアイデンティティへの欲求が購買行動に大きく影響している**と考えられる。さらに、同じ消費者であっても、時と場合、また、商品ジャンルによって、統一された意思決定基準で購買するとは限らない。こうした消費者の性質を理解するためには、購買に対する価値観や、日常生活行動、特定のブランドへの関心度合いなど、メディア接触や購買行動などの**行動データとして顕在化しにくい内面的な要素について、多面的に考慮する**必要があり、**ターゲットセグメンテーションが難しくなっている**。

3 分析業務の煩雑化

上記のような課題に対し、デジタル技術の発展や、シングルソースデータの活用により、広告効果測定に有用なデータは飛躍的に増加している。しかし、**互いに性質の異なる膨大なデータ**について、目的に応じて見るべき変数を選別し、適切に表示する作業が煩雑になっているのも事実である。膨大なデータから有用な知見を引き出し、それを**広告戦略策定上の意思決定につなげることが難しくなっている**。

1-1. 研究目的

前項の問題意識を考慮し、本研究の目的を以下の3点とした。

1 広告チャネルを横断した消費者起点のモデル化

テレビCMと雑誌広告のように異なる広告チャネル間での相乗効果を表現する広告効果測定モデルを構築する。また、マスメディアやソーシャルメディアなどのメディア利用状況を一括してモデルに取り入れることで、マスメディアだけでなく、インターネットメディアにおける広告戦略策定の示唆となる分析を行う。

2 購買行動の背後にある内面的要素のモデル化

たとえ同程度の広告に接触したとしても、購買に対する価値観やその商品への関心度合いなどによって、広告効果は異なると考えられる。価値観や購買意向などの内面的なデータと、メディア接触などの行動データを結び付けた広告効果測定モデルを構築することで、ニーズが顕在化していない潜在ターゲット像を発見し、広告効果の最大化に寄与する分析を行う。

3 広告戦略策定上の意思決定プロセスへの貢献

互いに異質、かつ多数の変数の関係性をわかりやすく視覚化し、モデルから得られた知見を広告戦略策定における意思決定につなげられるよう、ビジネス上の実務に貢献するプロセスを提案する。

1-2. 対象の商品ジャンル：シャンプー

本研究における問題意識と共通する課題を抱えている点に着目し、分析対象の商品ジャンルを「シャンプー・コンディショナー」とした。

消費者ニーズの多様化の問題に直面

近年シャンプー利用は、家族単位から個人単位に移行しており、頭皮ケア、ボリュームアップなど、個人のニーズにあったシャンプー選びが行われるようになった。

→低価格で万人向けの製品が、高付加価値を持つ多様な製品に入れ替わっている。

メディア接触行動の複雑化の問題に直面

・ソーシャルメディアなどの口コミ力が商品の選択肢を広げ、地肌にやさしい点を訴求した中小メーカーのノンシリコンシャンプーが売れ筋上位を占めるようになった。

→宣伝力のある大手の独壇場だったシャンプー市場において、宣伝力のない中小メーカーの製品が、ソーシャルメディアを通して台頭している。



大手メーカーは苦戦し、2013年に相次いで主力商品を刷新している。
2013年3月：資生堂（TSUBAKI）、4月：P&G（パンテーン）、8月：ユニリーバ（ラックス）

1-2. 対象商品：ラックス

調査期間の直前に大規模なブランドイメージのリニューアルを実施している点に着目し、分析対象の商品をラックスのシャンプー・コンディショナーとした。

ユニリーバは、主力商品であるラックスにおけるユーザーの高年齢化傾向を懸念し、ブランドの若返りを図るべく20~30代の女性に向けたメッセージングを大きく変更している。リニューアルは調査データの取得期間の直前である2013年の秋から行われているため、広告効果が特徴的に表れていることが想定される。

リニューアル前の
テレビCM

レッドカーペット
ドレス
憧れ



リニューアル後の
テレビCM

オフィス街
ハイヒール
親近感

リニューアル後のテレビCM



<http://socialgirlslab.com/2014/04/15/lux/>

<http://www.isssc.com/interview/cg/044/>

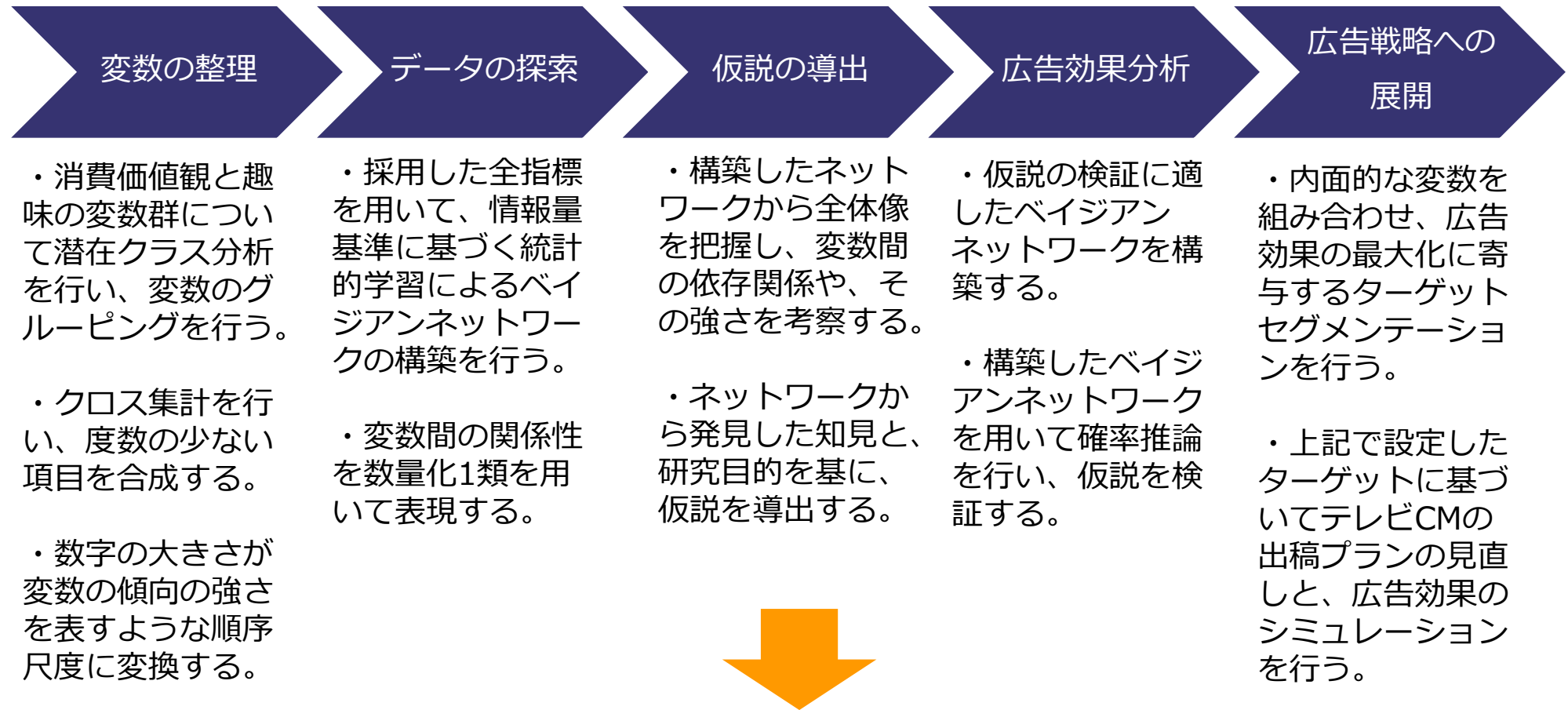
http://www.youtube.com/watch?v=bMz6UUAkdFU&list=UUtay9nkHuUEjJ__Qx6xwv-Q

<http://www.unilever.co.jp/aboutus/mediacentre/pressreleases/2014/onnewoasobe.aspx>

http://www.youtube.com/watch?v=vfWobTKg7YE&list=UUtay9nkHuUEjJ__Qx6xwv-Q

2-1. 分析の枠組みと流れ

本項から、データ分析の準備に入る。本研究では、以下の流れで分析を実施する。



一連のプロセスを通じ、複雑化したメディア接触行動および、内面的要素を用いた多面的なターゲットセグメンテーションを考慮した広告効果測定モデルを構築し、広告戦略策定における意思決定プロセスへの貢献を目指す。

2-2. 変数の作成：趣味

潜在クラス分析を用いて、消費価値観と趣味の変数群について変数をグルーピングした。

- ・情報量基準としてBICを用い、クラス数を2～9に設定して5回ずつBICの変化を調べ、最適なクラス数を判断した。
- ・各変数の、クラスへの反応確率を基に、各クラスの特徴を解釈し、ラベルを付与した。
- ・ベイジアンネットの構築用に、各サンプルについて、各クラスへの所属確率を5段階で離散化した。

ラベル→	多趣味	インテリ系	アウトドア系	ヤング系	無趣味
クラスのサイズ→	9.7%	17.4%	16.9%	30.0%	26.1%
釣り	0.09	0.01	0.12	0	0.01
アウトドア、キャンプ	0.15	0.06	0.12	0.01	0.01
日曜大工、機械・模型いじり	0.17	0.01	0.14	0.03	0
園芸、庭いじり	0.22	0.08	0.18	0.09	0.03
競馬	0.12	0.01	0.12	0.05	0.02
パチンコ	0.09	0.04	0.16	0.04	0.03
宝くじ	0.26	0.03	0.11	0.07	0
マッサージ、エステ	0.27	0.11	0.06	0.06	0.02
外食・グルメ・食べ歩き	0.74	0.5	0.23	0.24	0.05
国内旅行	0.88	0.93	0.3	0.05	0.04
海外旅行	0.34	0.43	0.03	0.01	0.02
映画・演劇・美術鑑賞	0.8	0.4	0.05	0.44	0.04
読書	0.48	0.37	0.2	0.41	0.1
編み物、料理	0.22	0.11	0.06	0.15	0.03
スポーツ・フィットネス	0.27	0.18	0.15	0.1	0.09
ゴルフ	0.16	0.08	0.1	0.01	0.04
スキー	0.18	0.07	0.11	0.03	0.01
ドライブ	0.58	0.2	0.3	0.07	0.01
遊園地、テーマパーク	0.52	0.17	0.09	0.05	0.01
カラオケ	0.41	0.09	0.03	0.15	0.03
ビデオ、DVD鑑賞	0.7	0.29	0.26	0.55	0.03
パソコン	0.93	0.71	0.76	0.81	0.26
音楽鑑賞（コンサートも含む）	0.72	0.35	0.05	0.47	0.05
テレビゲーム	0.32	0.05	0.15	0.19	0.02
スポーツ観戦	0.31	0.19	0.14	0.13	0.04
写真・ビデオ撮影	0.39	0.08	0.07	0.1	0.01
楽器演奏、合唱	0.17	0.05	0.02	0.09	0.02
ボランティア活動	0.09	0.03	0.02	0.04	0
囲碁・将棋・麻雀	0.05	0.01	0.03	0.01	0
書道・茶道・華道・絵画・俳句	0.06	0.04	0.01	0.02	0

クラス数	BICスコア				
2	53398.63	53398.63	53398.63	53398.63	53398.63
3	52936.81	52910.4	52910.4	52910.4	53185.05
4	52629.93	52629.93	52629.93	52629.93	52629.93
5	52534.55	52534.76	52534.76	52534.76	52534.76
6	52596.95	52579.23	52591.16	52596.95	52579.22
7	52666.1	52666.1	52641.7	52636.18	52636.18
8	52749.15	52779.02	52759.9	52707.8	52766.26
9	52821.96	52850.72	52825.6	52976.48	52828.85

- ・多趣味
ほぼすべての趣味を持つ非常にアクティブな層。
- ・インテリ系
エステや食べ歩き、旅行や映画鑑賞を好む層。
- ・アウトドア系
釣りやキャンプ、競馬やパチンコを好む層。
- ・ヤング系
カラオケやDVD鑑賞、テレビゲームを好む層。
- ・無趣味
特に趣味を持たない層。

2-2. 変数の作成：消費価値観

ラベル→	合理的先進層	情緒的先進層	価格重視層	安心重視層	自分重視層	こだわりなし
クラスのサイズ→	5.9%	7.1%	23.1%	10.4%	21.6%	32.0%
商品を買う前にいろいろ情報を集めてから買う	0.87	0.71	0.56	0.43	0.34	0.1
よい情報を得るためにはお金を払うのが当然である	0.26	0.26	0.04	0.08	0.07	0.02
商品や店舗に関する情報をよく人に教える方である	0.2	0.4	0.07	0.04	0.05	0.01
探している商品が見つからない場合は、すぐに店員に聞く方である	0.55	0.51	0.25	0.36	0.15	0.06
とにかく安くて経済的なものを買う	0.23	0.24	0.64	0.12	0.03	0.24
価格が品質に見合っているかどうかをよく検討してから買う	0.93	0.67	0.69	0.56	0.4	0.16
多少値段が高くて、品質のよいものを買う	0.88	0.71	0.07	0.77	0.43	0.03
多少値段が高くて、利便性の高いものを買う	0.77	0.64	0.08	0.5	0.27	0.03
同じ機能・値段であるならば、外国製品よりも日本製品を買う	0.76	0.6	0.61	0.83	0.25	0.13
できるだけ長く使えるものを買う	0.95	0.65	0.52	0.63	0.43	0.18
安全性に配慮して商品を買う	0.69	0.4	0.23	0.38	0.2	0.05
多少値段が高くて、アフターサービスが充実している方がよい	0.54	0.41	0.05	0.36	0.13	0.02
有名な人がよいと言っているものを選ぶことが多い	0.01	0.2	0.03	0	0.02	0.01
周りの人がよいと言っているものを選ぶことが多い	0.07	0.36	0.15	0.14	0.02	0.03
使っている人の評判が気になる	0.26	0.46	0.26	0.24	0.07	0.02
流行にはこだわるほうである	0.09	0.36	0.04	0	0.06	0.02
周りの人が持っているものを持っていないと気になる	0.01	0.18	0.02	0.01	0	0.02
使いやすい（着やすい）かどうかよりも、色やデザインを重視して商品を買う	0.07	0.46	0.12	0.04	0.1	0.04
テレビやパソコンなどの商品でも、色やデザインを重視して商品を買う	0.32	0.61	0.18	0.17	0.18	0.05
名の通ったブランドやメーカーの商品であれば、そのぶん多少値段が高くてよい	0.24	0.66	0.03	0.34	0.09	0.03
いつも買うと決めているブランドがある	0.44	0.56	0.16	0.29	0.18	0.06
無名なメーカーの商品よりは、有名なメーカーの商品を買う	0.35	0.66	0.2	0.67	0.13	0.06
周りの人と違う个性的なものを選ぶ	0.36	0.4	0.12	0.06	0.14	0.03
自分のライフスタイルにこだわって商品を選ぶ	0.87	0.67	0.31	0.22	0.4	0.07
自分のためにオーダーメイドされた商品をよく買う	0.08	0.1	0.01	0	0.03	0.02
環境保護に配慮して商品を買う	0.36	0.15	0.08	0.12	0.07	0.01
使い捨て商品をよく買う	0.06	0.14	0.07	0.03	0	0.04
レンタルやリースをよく利用する	0.09	0.11	0.05	0.02	0.02	0.02
中古製品やリサイクル品をよく買う	0.19	0.15	0.2	0.09	0.06	0.07
プライベートブランド（小売店が独自に販売しているブランド）をよく買う	0.34	0.36	0.38	0.15	0.06	0.1
すぐに使える現金や預貯金がないときに、クレジットカードで高額買い物をすることがよくある	0.24	0.48	0.12	0.21	0.11	0.06
自分の好きなものは、たとえ高価でもお金を貯めて買う	0.73	0.73	0.16	0.29	0.27	0.04

2-2. 変数の作成：消費価値観

クラス数	BICスコア				
2	67947.22	67947.22	67947.22	67947.22	67947.22
3	67238.39	67238.39	67238.39	67238.39	67238.39
4	66885.79	66885.79	66885.79	66885.79	66885.79
5	66892.73	66876.48	66876.48	66883.47	66887.54
6	66868.88	66868.88	66904.14	66868.89	66868.88
7	66914.69	66936.43	66914.69	66914.21	66960
8	67002.35	66978.37	66978.37	67161.03	66980.07
9	67070.61	67115.07	67149.11	67110.25	67070.61

- ・合理的先進層：情報収集に積極的。品質・耐久性・安全性を重視する。
- ・情緒的先進層：情報収集に積極的。流行・評判・デザイン・ブランドを重視する。
- ・価格重視層：情報収集に積極的ではない。価格を最重視する。
- ・安心重視層：情報収集に積極的ではない。日本製・アフターケアを重視する。
- ・自分重視層：情報収集に積極的ではない。ライフスタイルに合うか、周りとは違う個性、などを重視する。
- ・こだわりなし：全体的にどの消費価値観についてもこだわらない。

次項に、本研究において使用した変数の一覧を示す。なお、提供されたコンテストデータのうち、データ分析に使用したデータは下記のとおりである。

アンケートデータ：815サンプル（関東1都6県、女性のみ、欠損値を除く）

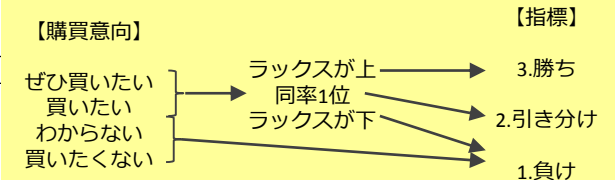
調査期間：2014年2月8日～2014年4月5日

また、広告効果については、1回目の調査時と2回目の調査を比較した際の購買実態及び購買意向におけるポジティブな変化を「態度変容」と定義し、態度変容をもって広告効果とみなすことにした。

2-2. 使用した変数一覧

変数カテゴリ	変数名	説明
シャンプー 購買	購買実態	1回目の調査時のデータ (2/14)。 1.非認知(知らない) 2.認知(お店ではみかけてないが名前は知っている) 3.店舗接触(ここ1ヶ月以内にお店でみかけた) 4.検討(ここ1ヶ月以内にお店で手にとって、見た) 5.購買(ここ1ヶ月以内に買った)
	購買意向	※(1)参照。1回目の調査時のデータ (2/14)。
	態度変容(実態)	1回目と2回目の調査を比較し、上で定義した購買実態または購買意向において、ポジティブな変化が見られれば1.あり、変化がない、またはネガティブな変化が見られる場合は0.なしとする。
	態度変容(意向)	
	ノンシリコン(実態)	0.使っていない 1.使っている
	ノンシリコン(意向)	1.使いたくない 2.どちらともいえない 3.使いたい 4.ぜひ使いたい
広告接触	テレビCM	1回目の調査日(2/14)の翌日から、2回目の調査日(3/14)の前日までの、ラックスのシャンプー(ボディソープ、トリートメントやオイルは除く)に関連する広告への接触。
	雑誌広告	テレビ広告: 0.非接触 0.1.1~5回 0.6.6~10回 1.1.11~15回 1.6.16~20回 2.1.21~25回 2.6.26~30回 3.1.31~35回 3.6.36~40回 4.1.41回以上 雑誌広告: 0.非接触 1.接触
メディア 利用状況	テレビ視聴時間	※(2)参照。
	女性誌閲読数	調査期間全体において、全サンプルが閲読している雑誌のうち、女性の割合が75%以上であり、週刊誌や住宅関連の雑誌を除く約100種類の雑誌の閲読数。 0.閲読なし 1.1冊 2.2冊 3.3冊 4.4冊 5.5冊以上
基本属性	年齢	20.20-24歳 25.25-29歳 30.30-34歳 35.35-39歳 40.40-44歳 45.45-49歳 50.50-54歳 55.55-59歳 60.60-64歳 65.65-69歳
	職業	1.無職・フリーター・その他 2.学生 3.パート・アルバイト 4.派遣 5.主婦 6.自営業 7.会社員 8.公務員・教職員 9.医者・弁護士・経営者
	世帯年収	1.300万円未満 2.300-500万円 3.500-700万円 4.700-1000万円 5.1000万円以上
	未婚既婚	0.独身(未婚・離婚・死別) 1.既婚
	子供有無	0.いない 1.いる
	居住地	0.北関東(茨城・栃木・群馬) 1.千葉 2.埼玉 3.神奈川 4.東京
	居住形態	1.社宅・公営住宅他 2.賃貸 3.マンション(住宅ローン無し) 4.一戸建て(住宅ローン無し) 5.マンション(住宅ローン返済中) 6.一戸建て(住宅ローン返済中)
消費価値観	合理的先進層	
	情緒的先進層	
	価格重視層	潜在クラス分析における各クラスへの所属確率を1~5までの5段階で離散化。 数値が大きいほどクラスへの所属確率が高い。
	安心重視層	
	自分重視層	
	消費先進度	新しい商品やサービスに対する態度。1.あまり関心がない 2.一般に普及してから 3.少し様子をみてから 4.人よりも先に
趣味	多趣味	
	インテリ系	潜在クラス分析における各クラスへの所属確率を1~5までの5段階で離散化。 数値が大きいほどクラスへの所属確率が高い。
	アウトドア系	
	ヤング系	
CGM利用頻度	mixi	
	Twitter	
	Facebook	
	モバゲー	1.利用しない
	GREE	2.ときどき利用している(週1回程度)
	YouTube	3.日常的に利用している(ほぼ毎日)
	ニコニコ動画	
シャンプー関連 サイト利用状況	LINE	
	アットコスメ	0.訪問なし 1.訪問あり(調査期間全体において少なくとも1回以上アクセスしたか否か)
	ドラッグストア(ネット)	0.訪問なし 1.訪問あり(ケンコーコムまたは爽快ドラッグに対し、調査期間全体において少なくとも1回以上アクセスしたか否か)
購買チャネル 利用頻度	ラックスサイト	0.訪問なし 1.訪問あり(ラックスのブランドサイトに対し、調査期間全体において少なくとも1回以上アクセスしたか否か)
	ドラッグストア	1.年に1回以下 2.半年に1~2回程度 3.月に1~2回程度 4.週に1回程度 5.週に2~3回以上
	小スーパー	1.半年に1~2回以下 2.月に1~2回程度 3.週に1回程度 4.週に2~3回程度 5.ほとんど毎日(主に食料品・日用品を販売しているスーパーマーケット)
	大スーパー	1.年に1回以下 2.半年に1~2回程度 3.月に1~2回程度 4.週に1回程度 5.週に2~3回以上(衣料品・家電等も販売している総合的なスーパーマーケット(イトーヨーカドー、イオン、ダイエー等))
	ショッピングセンター	1.ほとんど利用していない 2.年に1回程度 3.半年に1~2回程度 4.月に1~2回程度 5.週に1回以上(衣料品店、飲食店、映画館などからなる総合的なショッピングセンター、ショッピングモール)

※(1) TSUBAKI・バンテーンの購買意向と比較した指標とする。

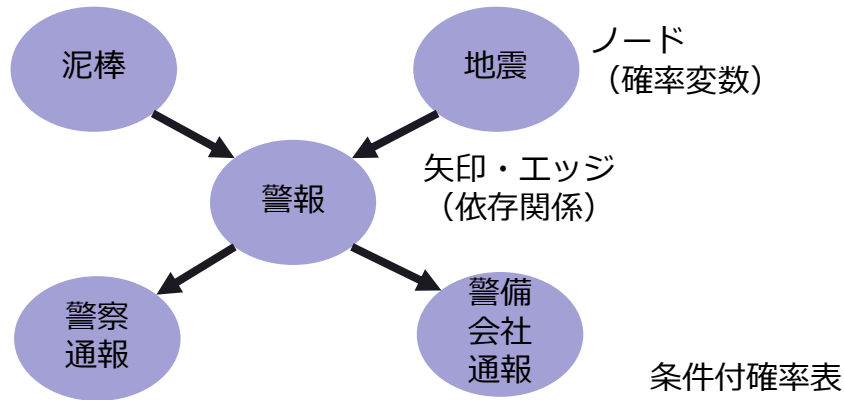


※(2)平日の19時-23時台(プライムタイム)に放送が開始する番組のうち、少なくとも一つの番組の視聴があった場合、その時間は1時間すべて視聴したとみなす。
たとえば19:00-20:54などの2時間番組の場合は、19時台と20時台すべて、つまり2時間視聴したとみなす。
この時間数の調査期間全体における総和を、調査期間全体における平日数で割った時間数をテレビ視聴時間の指標とする。

2-3. モデルの概要と利点

本研究におけるデータ分析の手法として、ベイジアンネットワークを用いた。
ベイジアンネットワークの概要と広告効果測定における優位性は以下のとおりである。

【概要】



警報	P(警備会社通報 警報)	
	しない	する
ならない	0.9	0.1
なる	0.3	0.7

- ・ 確率変数をノードで、依存関係を矢印（エッジ）で表した有向グラフで変数間の因果関係を説明するモデル。
例：原因：泥棒・地震→結果：警報作動
- ・ 各ノードは観測データによる条件付確率表で表される。
- ・ あるノードに確定データを入力し、ほかのノードの確率を推論する。
例：泥棒がいるとき、警報のなる確率は？

【優位性】

1 多彩な変数間の関係性を視覚化できる

- └ 条件付確率表による確率的な依存関係を用いるため、非線形な関係性も柔軟に表現できる。
- └ 多くの変数をグラフィカルモデルとして視覚化できる。

2 説明変数と目的変数が固定されていない

- └ 多数の指標を含む複雑なモデルについても、様々な切り口から分析が可能。
- └ 複数の広告チャネル間の相乗効果の分析や、複数の変数を組み合わせた多面的なセグメンテーションが可能。

3 統計的学習・人間の知識の双方から構築可能

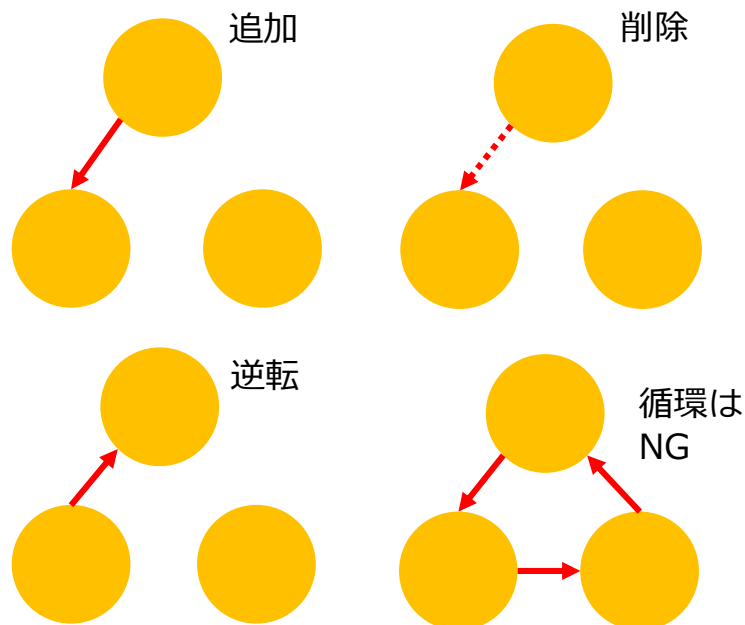
- └ 情報量基準を用いた統計的学習による構築ができるため、探索的な分析が可能。
- └ ノード・エッジの指定・禁止を設定できるため、分析者の事前知識を加味するなど、目的に応じたアドホックな分析も可能。

4 確率推論によるシミュレーションが可能

- └ あるノードに確定データや確率分布を入力することで、ほかのノードの事後分布を推論し、事前分布と比較することができる。
- └ 構築したモデルに対し、設計した広告施策の確率分布を入力することで、広告効果を推論することができる。

3-1. データの探索：モデルの構築方法

【山登り法】



1. 初期状態として（基本は）エッジの引かれていない空のグラフを定義する。

2. （エッジが循環しない範囲で）エッジを追加・削除・逆転をし、情報量基準が改善したものを採用する。

3. 情報量基準が最も改善するまで繰り返す。

※情報量基準として、BDeスコアを採用した。
(Bayesian Dirichlet equivalence)

参考文献[3]参照

本項から、実際のデータ分析に入る。
ベ이지アンネットの統計的学習による構築に使用したアルゴリズムと、構築の際の禁止条件について、下記に示す。

【禁止条件】

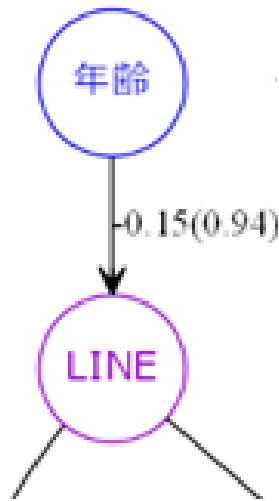
構築の際にブラックリストとして指定した禁止条件は下記のとおりである。

エッジの方向に対し、原因→結果とすると

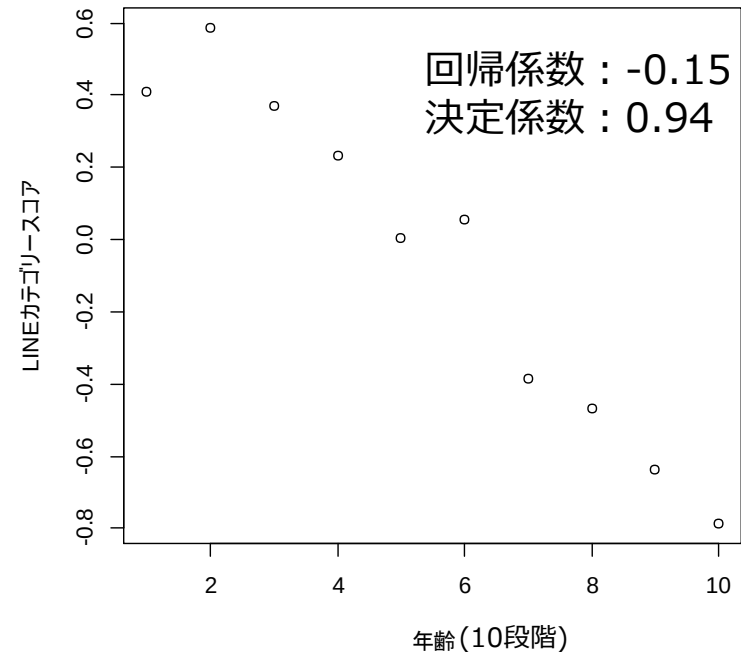
- ①年齢は、全指標に対して結果になってはいけない。
- ②基本属性の変数は、他カテゴリの変数に対して結果になってはいけない。
- ③潜在クラス分析をした消費価値観、趣味の変数については、同一カテゴリ間で因果関係になってはいけない。
- ④態度変容に関する変数は、全指標に対して原因になってはいけない。
- ⑤広告接触とメディア利用状況に関する変数は、（時系列的な問題から）態度変容を除く他のカテゴリの変数に対して原因になってはいけない。
- ⑥テレビCMはテレビ視聴時間の、雑誌広告は女性誌閲読数の原因になってはいけない。

3-1. データの探索：ノード間の関係性の表現

ノード間の関係性や、関係性の強さを表現するために、各変数を順序尺度として捉え、数量化1類を用いた。その上で、カテゴリースコアの回帰係数と決定係数の算出を行い、エッジ部分に記載した。



カテゴリースコア	
年齢.20-24	0.409846
年齢.25-29	0.585692
年齢.30-34	0.370687
年齢.35-39	0.233772
年齢.40-44	0.005422
年齢.45-49	0.055845
年齢.50-54	-0.38528
年齢.55-59	-0.46763
年齢.60-64	-0.63893
年齢.65-69	-0.78528
定数項	1.785276



上で示しているのは、年齢とLINEの利用頻度に関するノードである。この例では、年齢を説明変数、LINEの利用頻度を目的変数として数量化1類を用いて分析を行った。その上で、年齢を10段階で離散化した1～10の段階を説明変数、カテゴリースコアを目的変数とした場合の回帰係数と決定係数を算出し、エッジ部分に表記している。

この例では、年齢が20-24歳の段階→25-29歳の段階といったように一段階あがるごとに、LINEの利用頻度（1～3の順序尺度で、大きいほど利用頻度が高い）は0.15ポイント下がることを示している。

回帰係数がノード間の関連性を表し、決定係数が関連性の強さを表す。

※ただし、順序尺度として捉えることが適切でない変数については当てはまらない。

次項で、p.10の変数一覧のすべてを使用して構築したベイジアンネットと、そこから得られる知見について示す。

3-1. データの探索：モデルの構築と考察

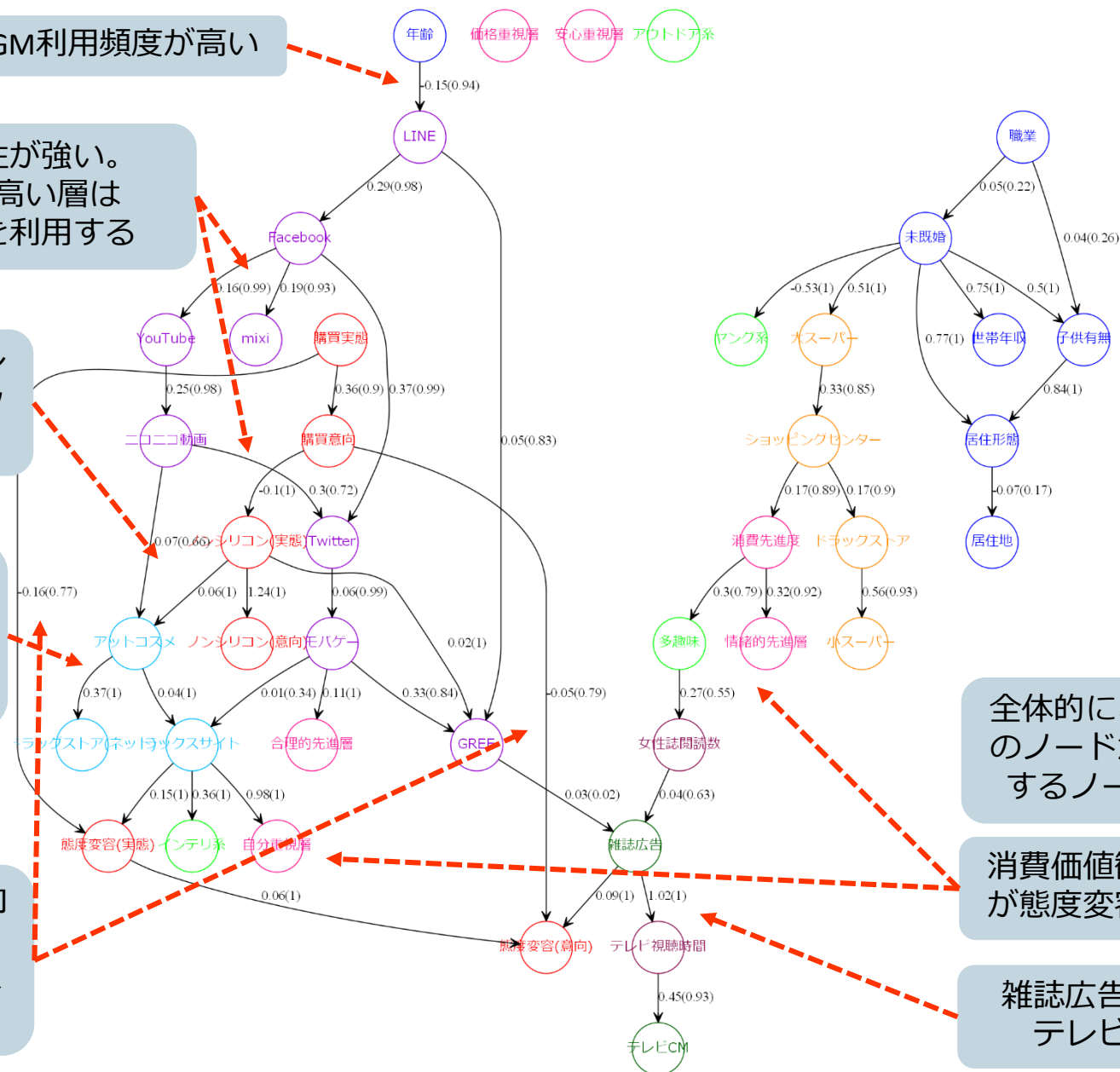
若年層ほどCGM利用頻度が高い

CGM同士の関連性が強い。
CGM利用頻度の高い層は
複数のメディアを利用する

ノンシリコンチャン
プを使う層はアッ
トコスメを使う

アットコスメを使う
層はネット通販のド
ラッグストアを利用
する

購買実態や購買意向
が上がるほど、
態度変容は起きにく
くなる



全体的に、左側にデジタル関連
のノードが、右側に実生活に関
するノードが繋がっている

消費価値観や趣味に関する変数
が態度変容ノードの近くにある

雑誌広告に接触している層は
テレビの視聴時間も長い

3-2. 仮説の導出

前項に記載した通り、統計的学習によって得られたネットワークから、複数の知見を得ることができた。これらの知見と、研究目的で挙げている「広告チャネルを横断した消費者起点のモデル化」、「購買行動の背後にある内面的要素のモデル化」という視点を基に、以下の仮説を設定した。

CGM同士の関連性が強い。
CGM利用頻度の高い人は複数の
メディアを平行して利用する

雑誌広告に接触している人は
テレビの視聴時間も長い

購買実態や購買意向が上がるほ
ど、態度変容は起きにくくなる

消費価値観や趣味に関する変数
が態度変容ノードの近くにある

仮説①

ある特定のメディアの利用頻度が高いユーザーは、他のメディアに関しても利用頻度が高い傾向にある。異なる広告チャネルで広告に接触することで、相乗効果が生まれているのではないか？

仮説②

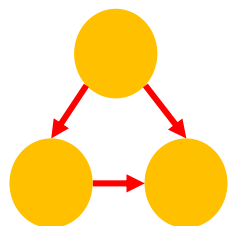
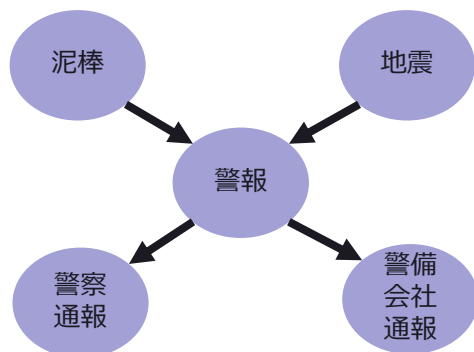
同じ量の広告に接触したとしても、購買実態と購買意向が違えば、広告効果は異なるのではないか？

仮説③

消費価値観や趣味に関する内面的な要素に関する変数を基にターゲットセグメンテーションを行い、広告施策の見直しを行うことで、広告効果をより大きくすることはできるのではないか？

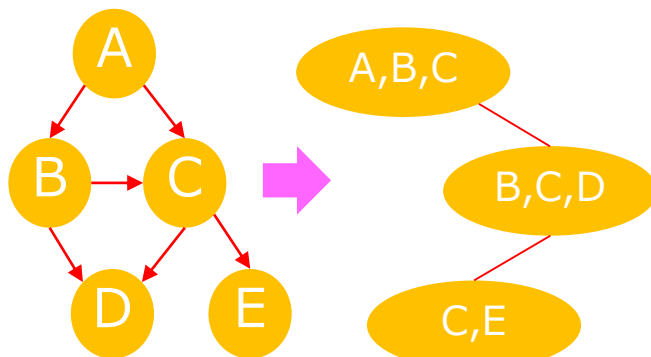
3-3. 確率推論について

次項以降の仮説の検証に際し、ベイジアンネットワークの確率推論を行った。
確率推論の概要と、本研究で使用したアルゴリズムを下記に示す。



multiply
connected

Junction tree



クリーク

ベイジアンネットワークにおける確率推論は、観測されたデータからの確率伝搬によって、各変数の確率分布を更新していく方法が取られる。

■ 確率伝播の流れ

- ① 観測された変数の値 e (エビデンス) をノードにセットする。
- ② 親ノードもエビデンスも持たないノードには事前確率分布を与える。
- ③ 知りたい対象の変数 X の事後確率 $P(X|e)$ を得る。

例：泥棒が「いる」というエビデンスをセットし、警報になる事後確率を得る。さらに、警報の事後確率の分布をもとに警察への通報や警備会社の通報についても確率分布を更新することができる。

しかし、この方法は、multiply connectedなネットワーク構造の場合、計算の収束性が保証できなくなる。

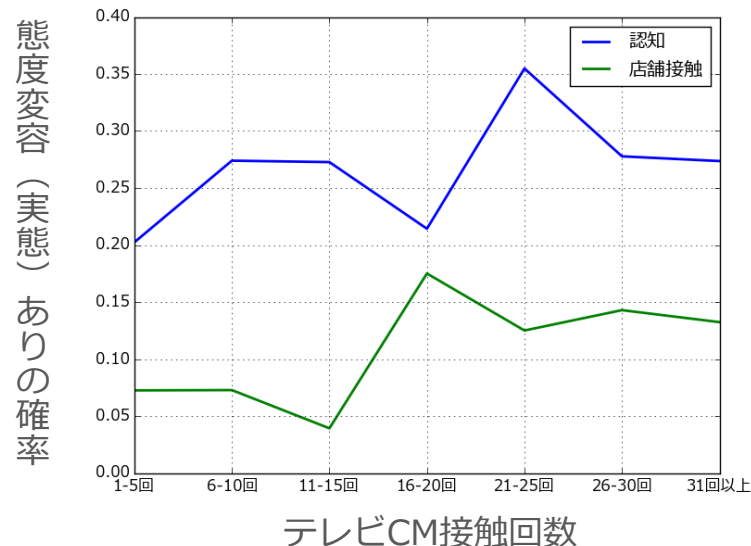
※multiply connected：リンクの向きを考慮しないでネットワークを見たときに、どこか一つでもパスがループしている部分がある時、このベイジアンネットワークはmultiply connected と呼ばれる。

こういった場合の確率推論のアルゴリズムの代表的なものとしてJunction treeアルゴリズムが挙げられる。

Junction treeアルゴリズムは、複数のノードをクリークとして結合し、singly connected な木構造からなる無向グラフへ変換する。この木構造にしたがってクリークごとに確率伝播を行うことで、確率推論を行う。

3-3. 仮説①②：広告効果測定

仮説①②の検証について、下記に示す。



×：非接触、○：接触

購買実態	テレビCM	雑誌広告	目的変数
店舗接触	×	×	態度変容（実態）
	○	×	
	○	○	
	×	×	態度変容（意向）
	○	×	
	○	○	
検討	×	×	態度変容（実態）
	○	×	
	○	○	
	×	×	態度変容（意向）
	○	×	
	○	○	

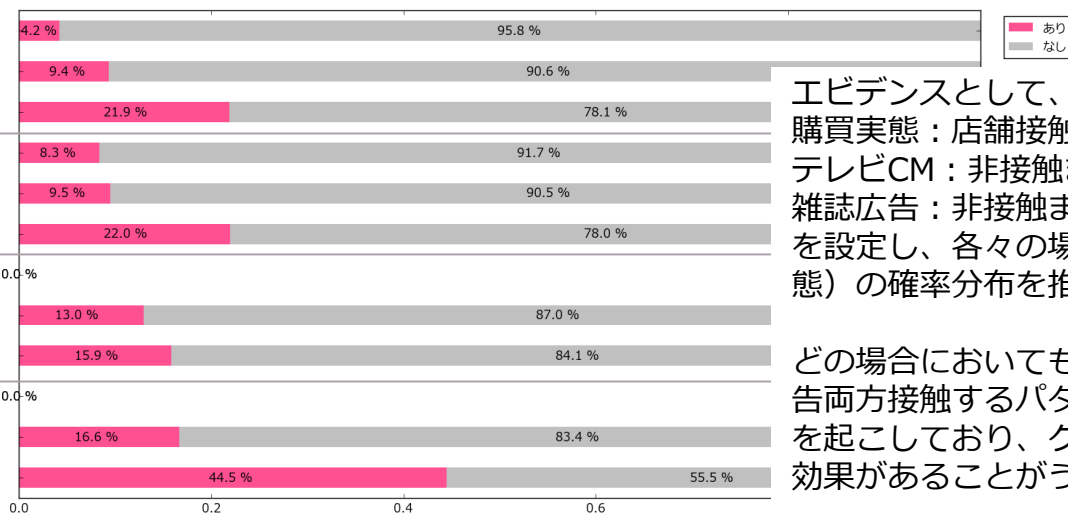
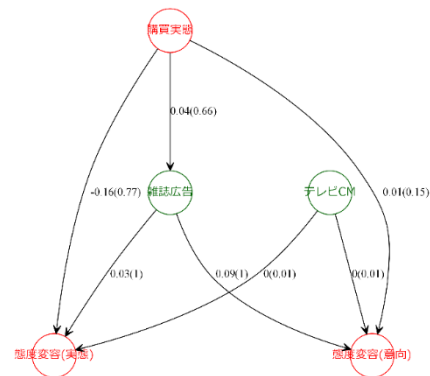
0%の部分は該当するエビデンスが成立する確率が極めて低く、推論の安定性を欠くと判断し、削除した。

エビデンスとして、
購買実態：認知・店舗接触
テレビCM接触回数 を設定し、各々の場合の態度変容（実態）ありの確率を推論した。

認知層は店舗接触層よりも態度変容をする確率が高く、全体ネットワークにおける知見と矛盾しない結果となった。（仮説②）

また、態度変容をする確率は、認知ユーザーは21回、店舗接触ユーザーは16回を境に増加しており、CM接触回数を増やすことで、広告効果を上げられることがうかがえる。

購買実態・テレビCM・雑誌広告の各々から、態度変容（実態・意向）に矢印が引かれるよう指定し、ネットワークを構築した。次項以降で構築したモデルも含めて、アルゴリズムや禁止条件はp.12と同様である。



態度変容の確率分布※グラフは確率分布を百分率で表示

エビデンスとして、
購買実態：店舗接触または検討
テレビCM：非接触または1回以上の接触
雑誌広告：非接触または接触
を設定し、各々の場合の態度変容（実態）の確率分布を推論した。

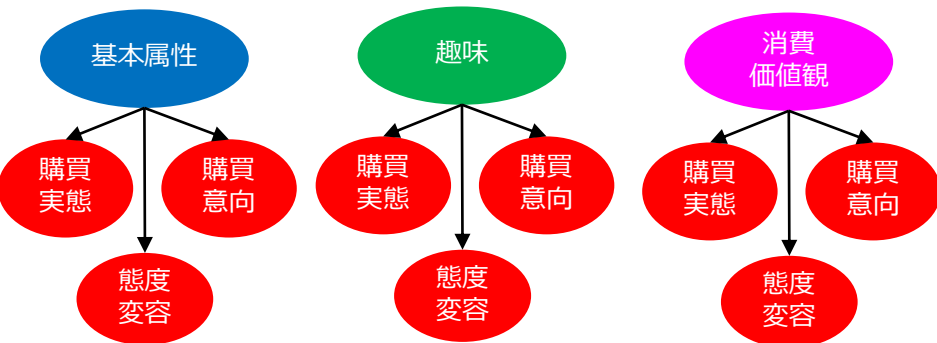
どの場合においてもテレビCMと雑誌広告両方接触するパターンが最も態度変容を起こしており、クロスメディアの相乗効果があることがうかがえる。（仮説①）

3-4. 仮説③：広告戦略への展開：ターゲットセグメンテーション

本項より、仮説③の検証に入る。

広告効果をより高めるために、態度変容を起こしやすく、それが購買につながりやすいようなターゲットセグメンテーションを目指す。まず、属性、趣味、シャンプー関連などの各変数カテゴリから複数の変数を抜粋し、それらの変数を説明変数とし、購買実態・意向・態度変容（実態）を目的変数とするようなネットワークを、変数カテゴリごとに構築した。

なお、構築したモデルの説明力を検証するため、説明変数に全サンプル分の観測データをエビデンスとして設定し、目的変数を推論した場合の正解率を算出した。



【購買意向】

予測値

観測値	予測値		
	負け	引き分け	勝ち
	負け	553	12
	引き分け	113	57

正解率
78%

【購買実態】

予測値

観測値	予測値				
	非認知	認知	店舗接触	検討	購買
	非認知	14	9	4	0
	認知	2	296	46	0
	店舗接触	5	121	204	0
	検討	0	28	12	9

66%

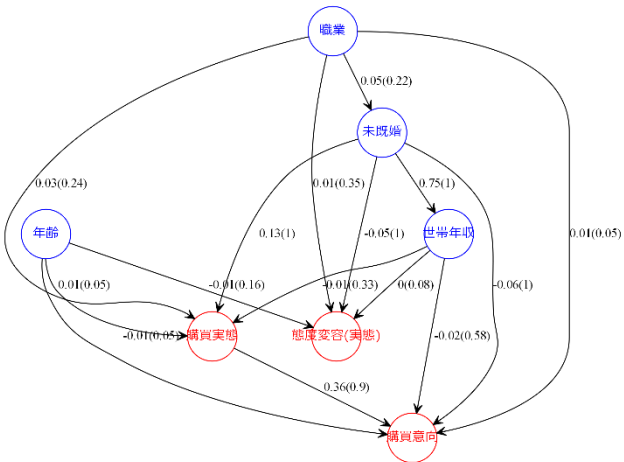
【態度変容（実態）】

予測値

観測値	予測値	
	0	1
	0	661

86%

基本属性とメディアの利用状況に関する変数の正解率が高い。これらの変数が、購買実態や意向、態度変容に影響していることを示唆している。



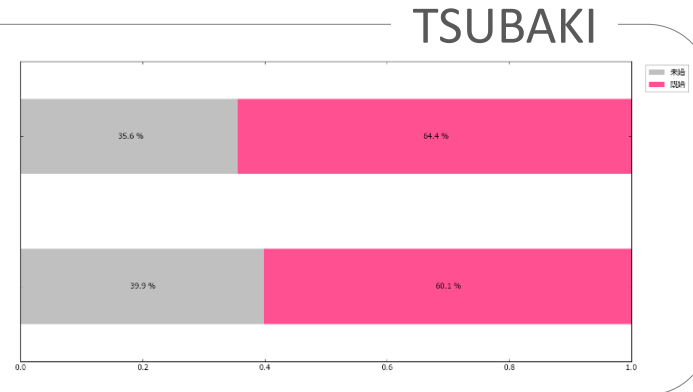
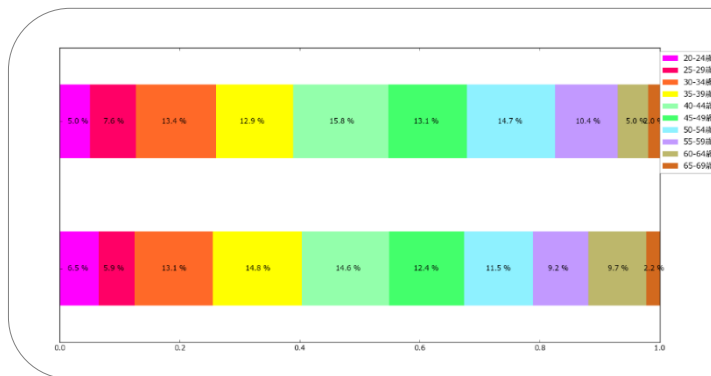
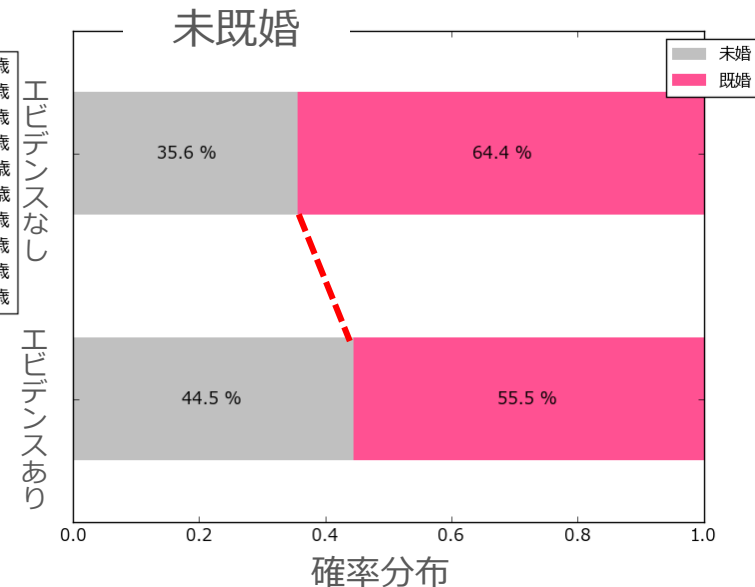
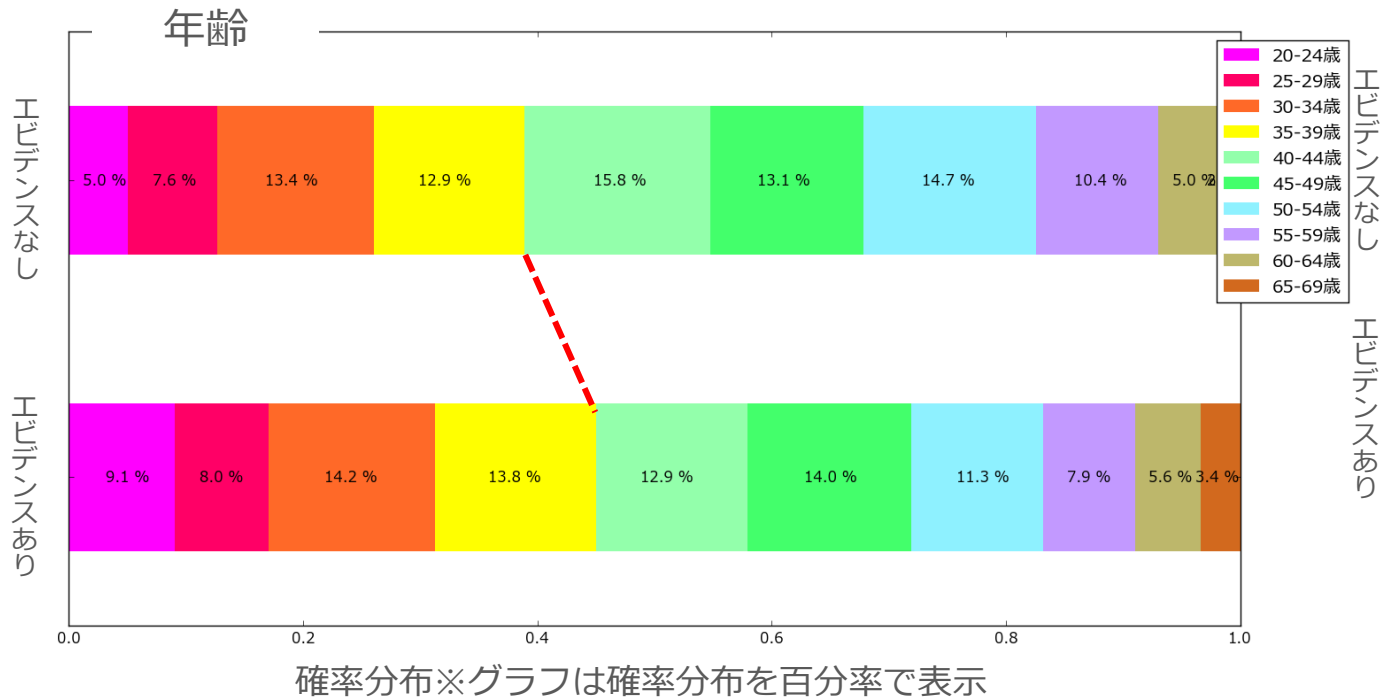
正解率一覧

	基本属性	趣味	消費価値観	ノンシリコン シャンプー利用	CGM利用頻度 テレビ視聴時間
購買意向	78%	71%	72%	70%	78%
購買実態	66%	47%	54%	44%	65%
態度変容(実態)	86%	82%	84%	82%	87%

3-4. 仮説③：ターゲットセグメンテーション：基本属性

購買実態・意向が高く、さらに態度変容を起こしている顧客層に着目した際に、影響の大きい変数を抽出したい。
エビデンスを下記のように設定し、基本属性の各変数について事前分布と事後分布を比較した。

購買実態：店舗接触または検討、購買意向：引き分けまたは勝ち、態度変容(実態)：あり
基本属性カテゴリで使用した変数：年齢・未既婚・世帯年収・職業



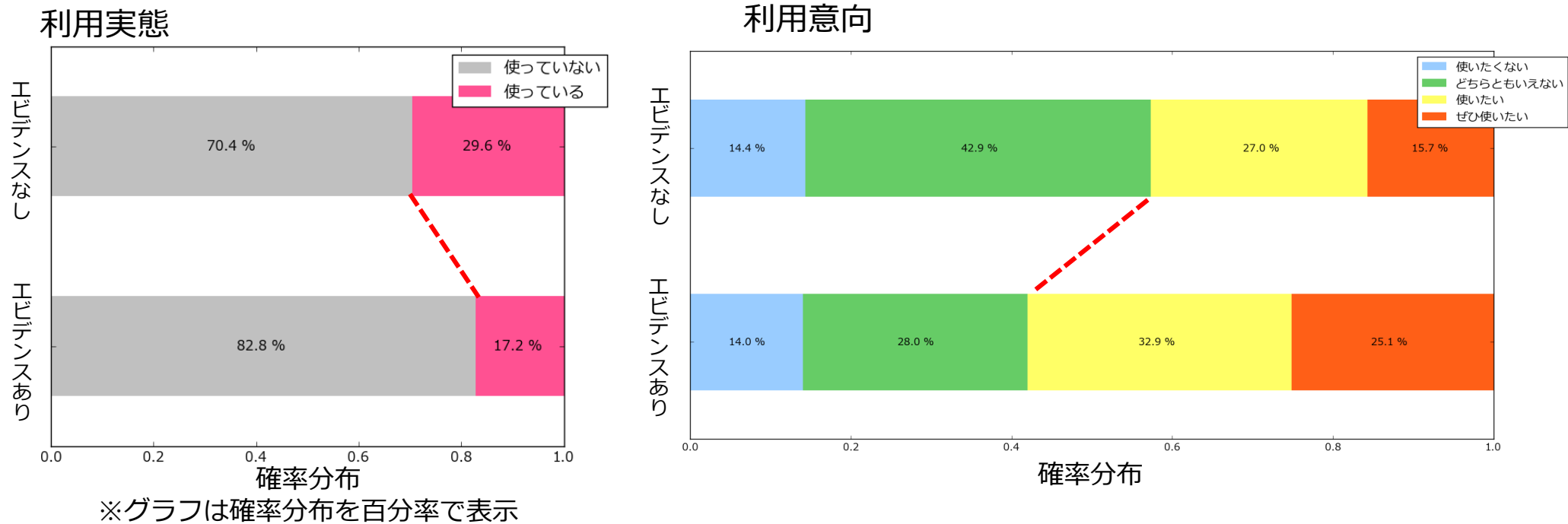
購買実態・意向共に高く、態度変容を起こしている顧客層は20~30代の独身女性の分布が高い。

同様の推論を行ったTSUBAKIと比較して、特に若年層の分布が高くなっている。

ラックスのブランドイメージの若返り戦略は成功していることを支持する結果が得られた。

3-4. 仮説③：ターゲットセグメンテーション：ノンシリコンシャンプー

前項と同様のエビデンスを設定し、ノンシリコンシャンプーの利用実態と利用意向について事前分布と事後分布を比較した。



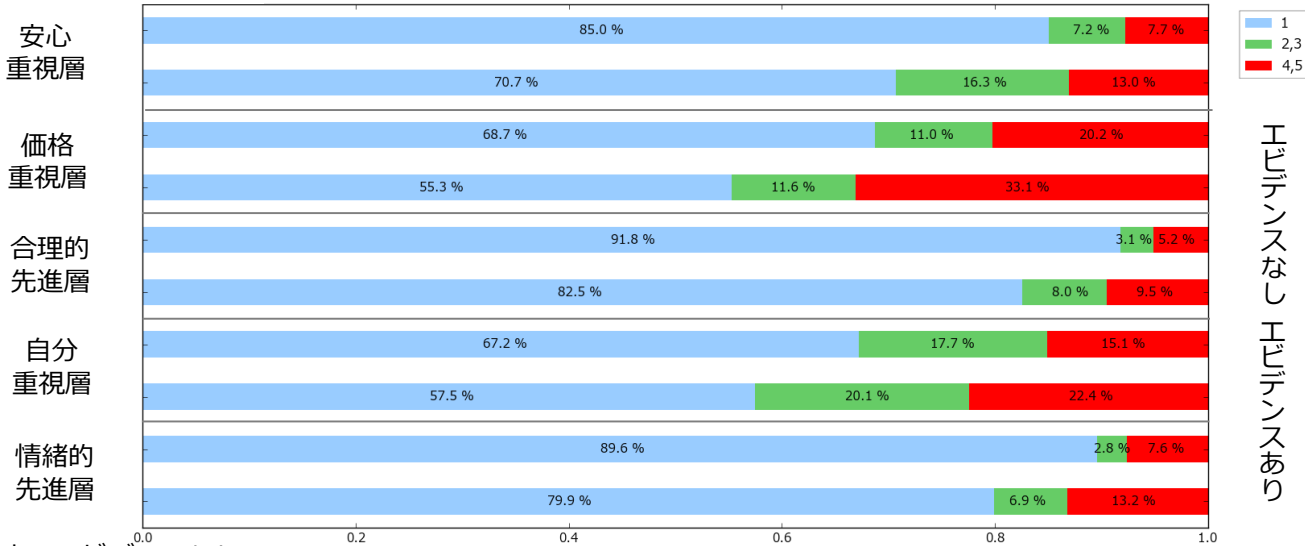
ラックスへの購買実態・意向が高く、態度変容を起こしている顧客層は、ノンシリコンシャンプーを使っていないが、利用意向は高いという推論結果となった。この結果から、ラックスへの関心度合いの高い層は、ノンシリコンシャンプーにも興味を示していると考えられる。

視点を変えれば、ノンシリコンシャンプーへの関心が強い層であっても、シリコンの入ったシャンプーに対する購買意向は決して低いわけではないという見方もできる。

3-4. 仮説③：ターゲットセグメンテーション：消費価値観と趣味

前項と同様のエビデンスを設定し、消費価値観と消費先進度、趣味について事前分布と事後分布を比較した。

消費価値観

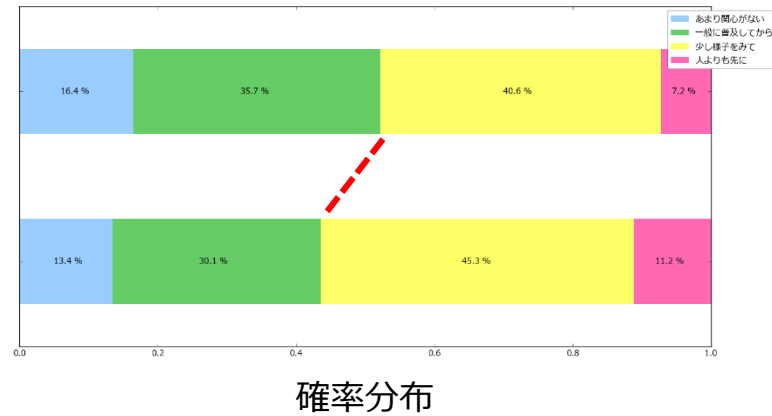


上：エビデンスなし
下：エビデンスあり

確率分布※グラフは確率分布を百分率で表示

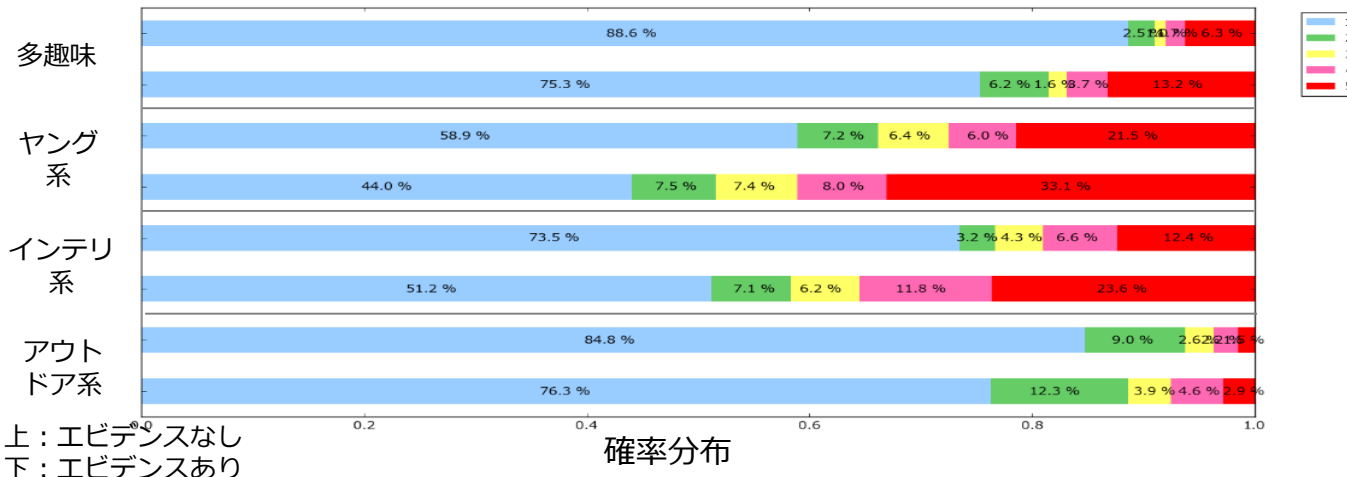
エビデンスなし
エビデンスあり

消費先進度



確率分布

趣味



上：エビデンスなし
下：エビデンスあり

確率分布

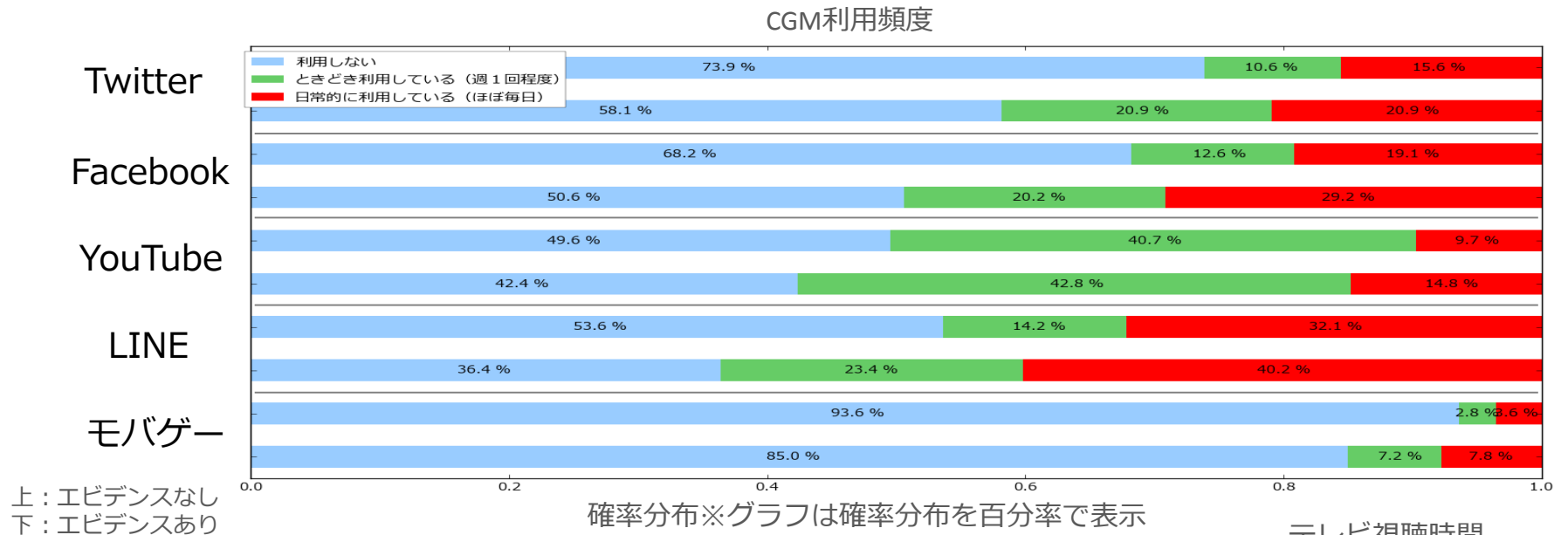
事前分布と比較すると、消費先進度は人よりも先に使用したり、少し様子を見てから使用する層に増加が見られ、新しい商品の消費に関して積極的な層が態度変容を起こしていることを示唆している。

消費価値観と趣味については、すべての変数について、傾向が強い層の伸びが見られた。

伸び率が大きい変数は、合理的先進層、情緒的先進層、多趣味、インテリ系の変数だった。

3-4. 仮説③：ターゲットセグメンテーション：メディア利用状況

前項と同様のエビデンスを設定し、テレビ視聴時間とCGM利用頻度について事前分布と事後分布を比較した。
CGM利用頻度カテゴリで使用した変数：Twitter、Facebook、YouTube、LINE、モバゲー

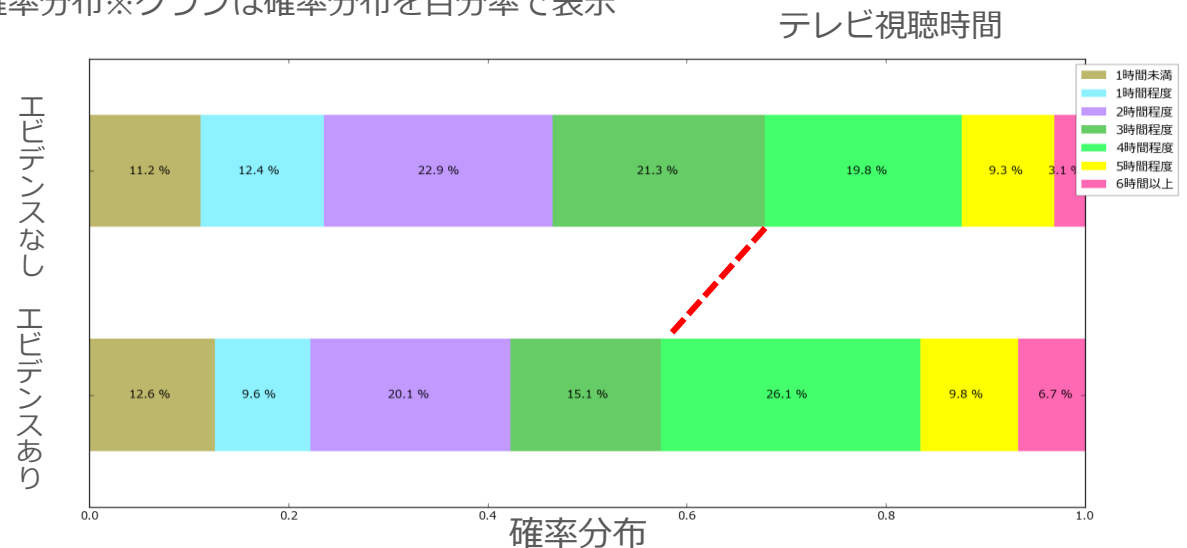


ラックスへの購買実態・意向が高く、態度変容を起こしている顧客層は、事前分布と比べるとCGMの利用頻度が高い。

特に、TwitterとFacebookについては、差が顕著に表れ、エビデンスありの顧客層は40～50%程度が週に1回程度以上は利用しているという推論結果となっている。

また、テレビの視聴時間に関しては、4時間以上視聴する確率が増加している。

エビデンスありの顧客層については、CGM、テレビともに利用傾向が高いことを示唆している。



3-4. 仮説③：ターゲットセグメンテーション：まとめ

ラックスについて、すでに店舗接触や検討段階にあり、さらに、購買意向が高く、かつ、態度変容（実態）を起こしたという条件で各変数カテゴリについて推論を行った。

このように、事前分布と事後分布を比較した際に特徴的な違いが見られた変数に着目することで、購買の手前におり、さらに、態度変容を起こしやすい層、つまり、より広告の効果が購買につながりやすい層のターゲットセグメンテーションが可能になると考えられる。

【考察】

- ・20代~30代の独身女性で、ノンシリコンシャンプーについては現在は使っていないが、使ってみたいと思っている。このような層は、シャンプーへの関心が高く、現在のブランドからの買い替えを検討している可能性が高い。
- ・消費先進度は高く、一般に普及するのを待たずに試してみたいという層である。
- ・消費価値観としては、情報収集に積極的な合理的先進層または情緒的先進層の傾向を持つ。
合理的先進層：情報収集に積極的。品質・耐久性・安全性を重視する。
情緒的先進層：情報収集に積極的。流行・評判・デザイン・ブランドを重視する。
- ・趣味は、多趣味、インテリ系の傾向を持つ。
多趣味：ほぼすべての趣味を持ち、非常にアクティブな層
インテリ系：エステや食べ歩き、旅行や映画鑑賞を好む層。

すべての要素を兼ね備えるセグメントは非常にサンプル数が少なくなることが想定されるため、今回はライフスタイルや価値観などの内面的な要素に着目し、ターゲットセグメントの条件を

【消費先進度：少し様子を見て以上】かつ【消費価値観：情緒的先進層 または 趣味：インテリ系】とした。

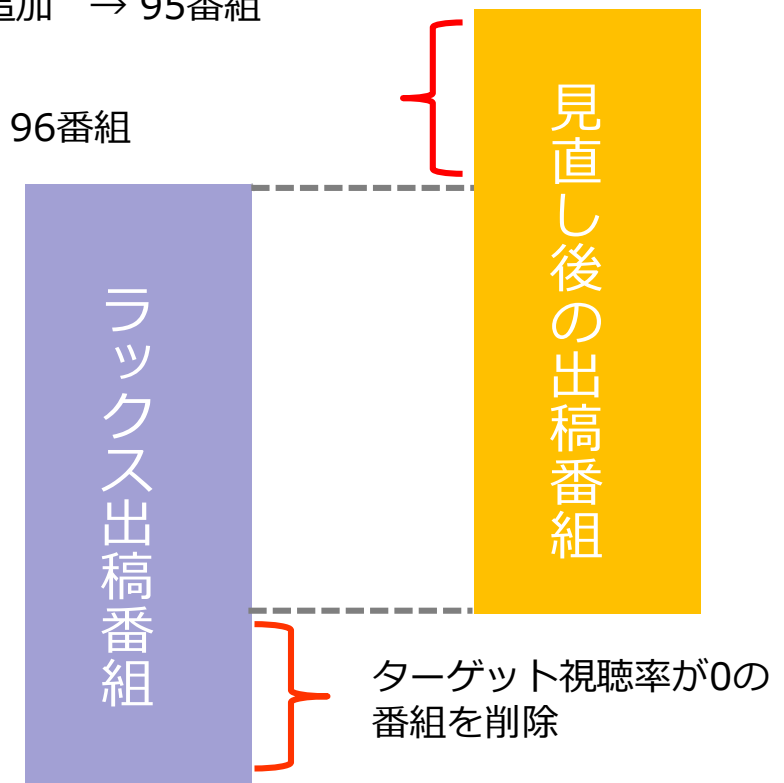
3-4. 仮説③：広告戦略への展開：テレビCM出稿番組の見直し

前項で設定したターゲットセグメントへのリーチが高まるよう、下記の通りテレビCMの出稿番組の見直しを行った。

1. 設定したターゲットセグメントに合致するサンプルID群（以下、ターゲット）を抽出 → 114サンプル（全体の約14%）
2. ターゲットの視聴したテレビ番組（2/15-3/13）を抽出 → 3297番組
3. さらに、ターゲット内視聴率が15%以上の番組を抽出 → 109番組
4. ラックスが出稿しているテレビ番組を抽出 → 379番組
5. 3. で抽出した109番組のうち、ラックスが出稿していない番組を追加 → 95番組
※ソチオリンピック関連番組、NHKによる番組を除く
6. ラックス出稿番組のうちターゲット視聴率が0%の番組を削除 → 96番組
7. 見直し後の出稿番組は、378番組

ターゲット視聴率が15%以上の番組を追加

ターゲット内視聴率	番組名	放送局
26%	相棒 season 12 #17	テレビ朝日
25%	サザエさん	フジテレビ
25%	踊る踊る踊る！さんま御殿！！SP	日本テレビ
25%	相棒 season 12 #16	テレビ朝日
25%	王様のブランチ	TBS
25%	マツコ&有吉の怒り新党	テレビ朝日
24%	マツコ&有吉の怒り新党	テレビ朝日
24%	失恋ショコラティエ #08	フジテレビ
24%	マツコ&有吉の怒り新党	テレビ朝日
23%	めちゃ2感謝してるッ！SP	フジテレビ
23%	相棒 season 12 #18	テレビ朝日
22%	日曜劇場「S-最後の警官-」 第6話	TBS



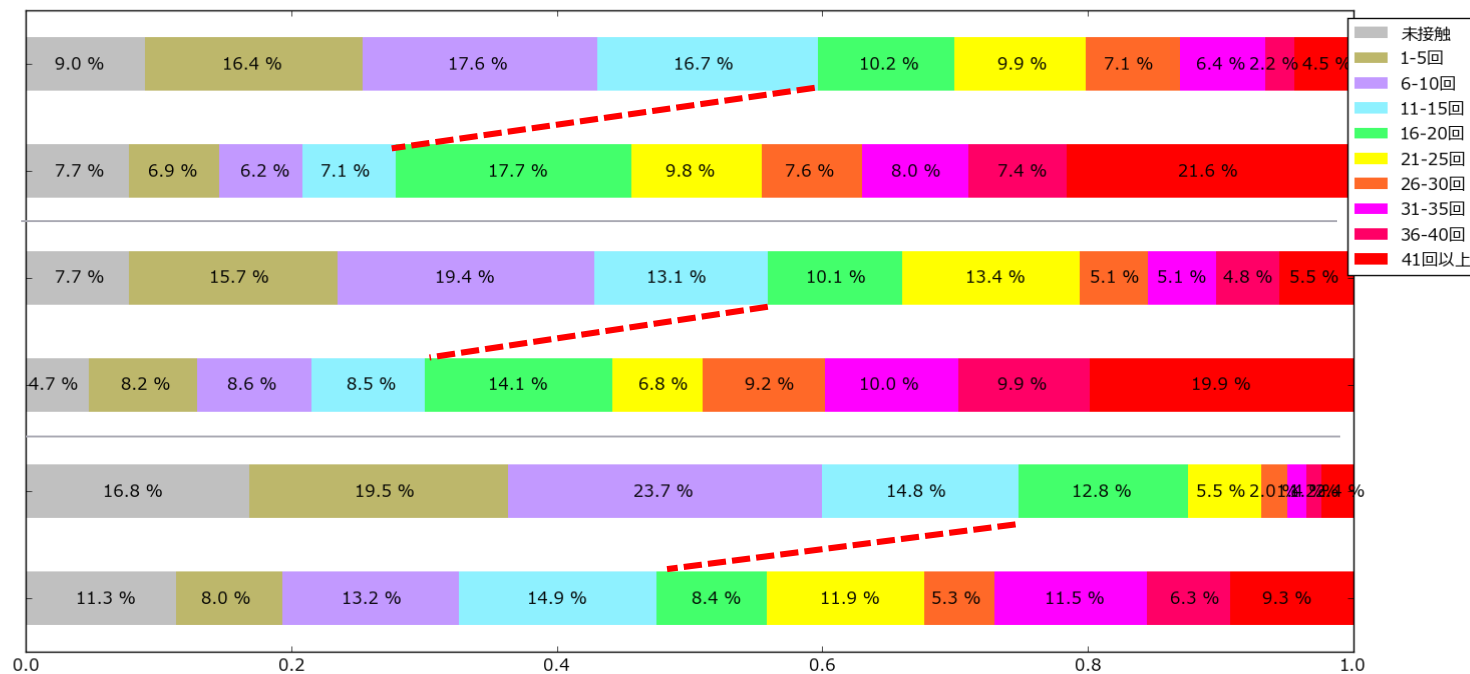
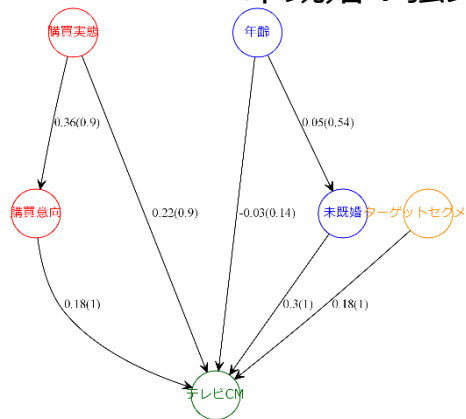
3-4. 仮説③：テレビCM出稿番組の見直し後の視聴回数分布

ラックスの実際の出稿番組と、見直しを行った場合の出稿番組について、それぞれネットワークを構築し、購買実態・意向と年齢未既婚、ターゲットをそれぞれエビデンスとして設定した場合のテレビCM接触回数の分布を推論し、比較した。なお、ネットワークの構築の際は、購買実態・意向と年齢未既婚、ターゲットの各変数からテレビCMにエッジが引かれるように指定した。

購買実態：店舗接触または検討
購買意向：引き分けまたは勝ち

ターゲットと合致

年齢：20~30代
未既婚：独身



上：ラックスの出稿番組
下：見直し後の出稿番組

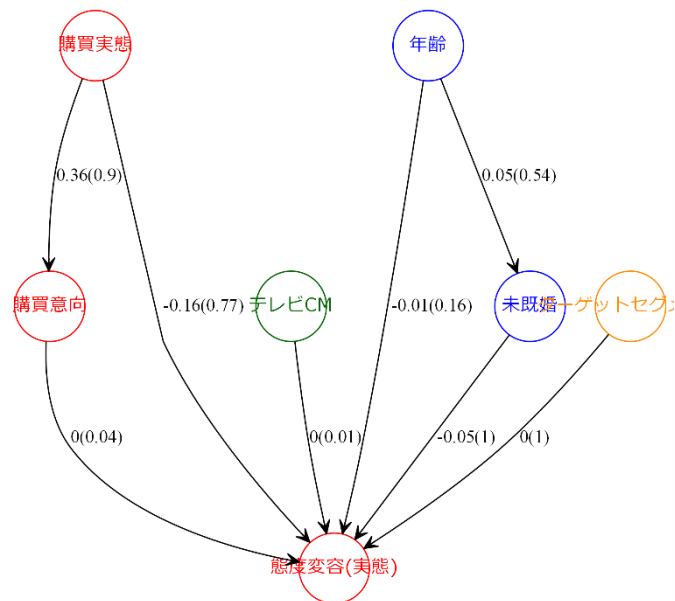
確率分布※グラフは確率分布を百分率で表示

出稿番組を見直すことで、各エビデンスの条件についてテレビCMの接触回数16回以上の分布が大きく増加していることがわかる。

3-4. 仮説③：テレビCM出稿番組の見直し後の態度変容確率

購買実態・意向・年齢・未既婚・ターゲットなどの変数と、テレビCMの接触回数から態度変容の確率分布を推論するネットワークを構築した。

その上で、上記のネットワークに対し、前項と同じ条件で購買実態・意向、年齢・未既婚、ターゲットについてエビデンスを設定し、その各々の場合について、前項の、実際の出稿番組の視聴回数分布と、出稿番組見直し後の視聴回数分布をエビデンスとして追加して推論した場合で態度変容確率を比較した。



購買実態：店舗接触または検討
購買意向：引き分けまたは勝ち

38.1% → 40.4% に改善
(実際の出稿) (見直し後)

ターゲットと合致

41.6% → 43.3% に改善
(実際の出稿) (見直し後)

年齢：20~30代
未既婚：独身

31.7% → 34.8% に改善
(実際の出稿) (見直し後)

見直し後のテレビCM接触回数の分布をエビデンスとして推論すると、右記の条件において、態度変容の起こる確率が増加している。内面的要素を考慮したターゲットセグメンテーションに基づく出稿番組の見直しが一定の効果を上げていることを示唆している。（仮説③）

4-1. まとめ：分析からの示唆

本研究で行ったデータ分析から得られた知見を下記にまとめる。

【広告チャネルを横断した消費者起点のモデル化】

- 「CGM同士」や、「アットコスメとインターネット上のドラッグストア」、「女性誌閲読数とテレビ視聴時間」など、メディア間での依存関係の高さが示唆されていた。(p.14)
- CGMの利用頻度とテレビの視聴時間を用いて、購買実態や意向、態度変容を推論した際は高い正解率を保っており、メディアの利用状況が購買行動に影響することを示唆する結果が出ている。(p.18)
- 雑誌広告とテレビCMは両方接触すると態度変容の確率が高まるという結果が出ており、異なる広告チャネル間の相乗効果を示唆している。(p.17)
- 態度変容の有無からメディア利用状況を推論したところ、購買に近い状態で態度変容をした顧客層はFacebookやTwitterの利用傾向が強く、かつ、テレビの視聴時間も長いという結果となった。このような顧客層には、テレビを見ながらソーシャルメディアを利用するといったようなマルチスクリーン視聴の傾向があることを示唆している。したがって、広告チャネルの予算配分については、テレビとソーシャルメディアの双方で相乗効果を生むようなキャンペーン施策が効果的と思われる。(p.22)

【購買行動の背後にある内面的な要素のモデル化】

- 購買実態によって、同程度のテレビCMの接触回数であっても態度変容の確率が異なるという結果となった。また、検討段階の顧客層に関しては雑誌広告の態度変容確率が高いなど、広告チャネルによって、より効果的な購買実態の段階があることも示唆されている。どのような購買実態の顧客層に態度変容を促したいのかという広告戦略に応じて広告チャネルを選択し、予算を配分する必要性があると思われる。(p.17)
- 趣味や消費価値観などの内面的な変数を基に設定したターゲットセグメントを用いてテレビCM出稿番組の見直しを行うことで、態度変容の確率を増加させられるという結果が得られた。内面的な要素にまで踏み込んだ顧客のセグメンテーションと、メディア接触行動や購買行動などの行動データを組み合わせることで、より深い広告戦略の展開が可能になると考えられる。(p.26)

4-2. モデルの有用性と課題

本研究で用いたベイジアンネットワークについて、広告効果測定における手法の有用性と課題をまとめた。

- 本研究では約40個の変数を用いてベイジアンネットを構築したが、ノードやエッジを用いて視覚化を行うことで、多数の変数の関係性の全体像をつかむことができた。そして、そこから複数の知見や仮説を導出している。このような探索的なアプローチは、見るべき変数が飛躍的に増加している現在の広告効果測定環境において、有効な方法であると考えられる。
- ベイジアンネットは条件付確率表を用いて依存関係を表現する比較的自由度の高いモデルのため、アンケートデータや基本属性データ、広告接触回数など、性質の異なる多様なデータが混在するシングルソースデータの分析に関して有効な方法であると考えられる。
- ベイジアンネットの特徴として、多段の依存関係を表現した上で、様々な切り口で確率推論ができる点が挙げられる。今回の分析のように、態度変容の有無から消費価値観を推論しターゲットセグメンテーションを行ったり、テレビCMの接触回数分布から態度変容確率を推論したりと、目的に応じて様々な視点からモデルを構築したり、推論したりすることができる。このような柔軟性についても、広告効果測定における優位性の一つといえるだろう。
- ターゲットセグメンテーションやテレビCM出稿番組の見直しのプロセスに関しては、マーケティング戦略やテレビCMの出稿金額、態度変容あたりの利益などを考慮した上で推論を行うことで、費用対効果の算出が可能になると思われる。そういった具体的な指標を取り入れることができれば、広告戦略策定における意思決定の実務的なプロセスにより貢献できるようになるだろう。
- 課題としては、推論の際にエビデンスの条件を多数設定していくと、エビデンス条件の成立する確率が少なくなり、確率推論の安定性が著しく低下してしまう点にある。データを離散化する際は、極端に度数の少ない値をなくすような配慮が必要である。

4-3. 参考文献など

【ベイジアンネットワーク関連】

[1] 本村陽一：“ベイジアンネットワーク：入門からヒューマンモデリングへの応用まで”，日本行動計量学会第7回春の合宿セミナー，2004.

<https://staff.aist.go.jp/y.motomura/paper/BSJ0403.pdf>

[2] 植野真臣：“ベイジアンネットワーク”，コロナ社，2013.

[3] Radhakrishnan Nagarajan, Marco Scutari, Sophie Lèbre：“Bayesian Networks in R with Applications in Systems Biology”，Springer，2013.

【Rパッケージなど】

モデル構築：bnlearn：<http://www.bnlearn.com/>

確率推論：gRain：<http://cran.r-project.org/web/packages/gRain/>

潜在クラス分析：poLCA：<http://cran.r-project.org/web/packages/poLCA/>

数量化1類：<http://aoki2.si.gunma-u.ac.jp/R/qt1.html>

【シャンプー市場について】

2012/9/21 日経MJ（流通新聞）016ページ

2013/2/15 日経MJ（流通新聞）002ページ

2013/3/19 日本経済新聞朝刊 039ページ

2013/7/19 日経MJ（流通新聞）006ページ

【ラックスについて】

<http://www.issc.com/interview/cg/044/>

<http://socialgirlslab.com/2014/04/15/lux/>

http://www.youtube.com/watch?v=bMz6UUAkdFU&list=UUtay9nkHuUEjJ__Qx6xwv-Q

<http://www.unilever.co.jp/aboutus/mediacentre/pressreleases/2014/onnawoasobe.aspx>

http://www.youtube.com/watch?v=vfWobTKg7YE&list=UUtay9nkHuUEjJ__Qx6xwv-Q