

## 事例研究 [論文]

# 検索語の確率分布に基づく 時期別の母親の関心事に合わせた情報推薦

岩永 二郎, 吉田 晏大, 高橋 彩音

## 1. はじめに

母親は妊娠、出産、育児にかけて不安を抱えることが多く、心配事や困難な場面が多い。そのような母親の課題を解決するために、さまざまな調査や研究が行われてきた。ベネッセ教育総合研究所の調査 [1] では、妊娠中から出産後までの母親の生活とサポート状況について調査結果を報告している。また、川井らの研究 [2] では、母親の育児不安を明確化することを目的に調査を行い、育児不安には不安・抑うつ感や育児困難感があることを示した。

産前産後の母親の課題解決を目的とした研究には、子どもの年齢や月齢に注目した研究がある。唐田らの研究 [3] では、0 歳から 6 歳までの乳幼児をもつ母親を対象に調査を行い、子どもの成長・発達に合わせた悩みが年齢によって推移することを述べている。また、藤田らの研究 [4] では、母親の育児ストレスの原因として核家族化の進行に伴うソーシャルサポートの低下に注目し、乳幼児をもつ核家族の母親を対象に 4 ヶ月、1 歳 6 ヶ月、3 歳 6 ヶ月の乳幼児健康診査を利用して調査を行っている。このように産後の母親の課題を研究対象とする場合、子どもの年齢や月齢に注目することが重要である。

出産後の子どもの月齢だけでなく出産前の母親の妊娠週数にも注目し、母親の関心を幅広く分析する研究

がある。岩永らの研究 [5] では、母親の不安や心配を含む関心事が検索語に現れることに注目し、出産日付きの検索履歴を数理的な方法で解析した。具体的には、各検索語に対して出産日を起点とする確率分布を単峰性、二峰性をもつ確率分布として精緻に推定できることを示した。すなわち、産前産後の時期別に母親の関心事をより正確に把握できることを示唆している。

本研究の目的は、産前産後の時期別に母親の関心事に合わせて適切な情報を推薦する仕組みを構築することで母親の課題を解決することである。本論文では、先行研究 [5] で推定した単峰性、二峰性の確率分布の適用事例として推薦カレンダーの作成を提案する。推薦カレンダーとは、出産前から出産後の各月に対して、推薦する複数のキーワードが対応するテーブルである。具体例として、表 1 に出産 1 ヶ月前 (-1 ヶ月)、出産月 (0 ヶ月)、および出産 1 ヶ月後 (1 ヶ月) の 3 期間に対して五つのキーワードを割当てた推薦カレンダーを示す。このとき、出産前後で母親自身への関心から子どもへの関心に移りかわるような推薦カレンダーの作成を目標とする。各月に推薦されるキーワードは、キーワード候補となる検索語の確率分布を入力として、検索量、確率値、分布の傾き、累積分布関数を考慮した割当て問題を解くことで決定される。

時期別の母親の関心事に合わせた情報推薦ができれば、母親の不安や心配事を先回りして支援することができる。また、企業は母親が困る時期に合わせて商品やサービスを推薦する効率の良いマーケティングが可能になる。たとえば、子育て支援アプリケーションで推薦カレンダーを利用することができる。子どもの誕生日が既知である利用者に対して 1 日一つのキーワードに関する情報をアプリケーション上で発信することで利用者は有益な情報を取得できる。一方、アプリケーション運営側は利用者が毎日アプリケーションを起動する動機付けができ、離脱防止に繋げることができる。

本論文の貢献は次の 2 点にまとめられる。一つ目に、先行研究 [5] で推定した単峰性、二峰性をもつ確率分

いわたが じろう  
(株) エルデシュ  
〒 222-0033 神奈川県横浜市港北区新横浜 2-3-12 新横浜スクエアビル 14F  
iwana@erdos-the-book.com  
よしだ やすひろ  
東京工業大学工学院経営工学系  
〒 152-8550 東京都目黒区大岡山 2-12-1  
yoshida.y.bi@m.titech.ac.jp  
たかはし あやね  
筑波大学理工学群社会工学類  
〒 305-8573 茨城県つくば市天王台 1-1-1  
takahashi.ayane.sw@alumni.tsukuba.ac.jp  
受付 22.7.15 採択 22.10.20

表 1 推薦カレンダーの例

出産 1 ヶ月前 (-1 ヶ月)	出産月 (0 ヶ月)	出産 1 ヶ月後 (1 ヶ月)
陣痛待ち	おしゃぶり	予防接種
内祝い	ミルク	うつ伏せ練習
出産報告	授乳間隔	首すわり
予定日	睡眠時間	よだれ
母子同室	おむつ	ベビー服

布の具体的な適用事例として推薦カレンダーを提案し、実用上の有効性を示した。二つ目に、推薦システムの研究において多期間に複数商品を同時に推薦する問題を扱い、推薦カレンダー作成問題を確率分布を入力とする割当問題に帰着させた点に新規性がある。

本論文の構成を説明する。2 節では関連研究について述べる。3 節では分析データの説明と前処理である確率分布の推定方法について述べる。4 節では数理最適化を用いた推薦カレンダーの作成方法について提案する。5 節で推薦カレンダーの考察を行い、6 節でまとめと今後の研究課題について述べる。

## 2. 関連研究

### 2.1 時期別の母親の関心事の推定

本論文が提案する推薦カレンダーの作成には、先行研究 [5] で提案されている単峰性最適化モデルと二峰性最適化モデルが前処理に利用される。先行研究 [5] では、出産日付きの検索履歴から集計して得られる確率分布に対して、混合整数凸二次計画問題を解くことで単峰性、二峰性をもつ確率分布を推定する方法が提案されている。推定した確率分布はノイズが除去されるため、精緻な確率値や累積分布関数が得られる。さらに、極値点の前後の区間で分布が単調性をもつため、上昇下降傾向の解釈性が向上する。しかし、解釈性の向上を利用した具体的な応用事例については言及していない。本論文では、推定した単峰性、二峰性の確率分布が実用上、有効であることを示す。

先行研究 [5] のように、経験分布に対して領域知識に基づく制約を課した数理最適化問題を解くことでノイズを除去する研究がある。情報推薦の研究である文献 [6-8] では、商品閲覧履歴から作成した最新度 (Recency) と頻度 (Frequency) の二つの特徴量に対して購買確率を対応付ける 2 次元確率表を経験分布として求める。そして、特徴量に単調性制約を課した数理最適化問題を解くことでノイズを除去した。精緻な 2 次元確率表を用いて商品を推薦することで推薦精度が向上することが示されている。また、岩永ら [9] ではノイズを除去した精緻な 2 次元確率表を協調フィルタリ

ングの評価値行列に適用することで推薦精度を改善できることを示した。上記のように経験分布を入力とし、数理最適化問題を解くことでノイズを除去する方法は、さまざまなタスクの精度改善が期待できる。

### 2.2 推薦システム

推薦システムには内容ベースフィルタリングや協調フィルタリングなどの代表的な推薦手法 [10] があり、電子商取引サイトにおける商品推薦や、ストリーミングサービスにおける音楽や映画の推薦 [11] などに利用されている。これらの研究は特定の時点で複数の情報を推薦する際に利用される。また、時系列を考慮した推薦システムには、マイクロブログにおいて適切な投稿タイミングを推薦する研究がある [12, 13]。これらの研究は一つの情報に注目して多期間の中で最適な時点で推薦する際に利用される。

しかし、本研究で作成する推薦カレンダーのように多期間で複数の情報を同時に推薦する応用研究はない。推薦カレンダーを利用することで適切な時期に適切な情報を適切な量だけ推薦することができる。

### 2.3 割当問題

割当問題とは、ある集合の要素を別の集合の要素に割当てて組合せを決定する問題であり、一般に 0-1 整数計画問題に定式化される。割当問題の応用事例には、作業員への仕事割当問題 [14]、ライドシェアのルーティング問題 [15]、時間割作成問題 [16] など幅広く研究されている。

割当問題の中には確率を扱った問題として、飛行機ゲートの割当問題 [17]、看護師の在宅看護スケジューリング問題 [18] などがある。これらの研究では、飛行機の到着遅れや患者のケアに必要な時間の不確実性を確率として扱い、割当問題に帰着している。しかし、本研究のように複数の確率分布を入力として確率値、分布の傾き、累積分布関数を同時に扱う割当問題の研究はない。

## 3. 分析データと確率分布の推定

### 3.1 分析データ

本研究では、経営科学系研究部会連合協議会主催令

和3年度データ解析コンペティションで提供されたコネヒト株式会社<sup>1</sup>が運営する新生児をもつ親向け情報サイト「ママリ」の検索履歴データを使用する。本データを加工することで出産日付きの検索履歴を作成できる。検索語には母親の関心事が現れるため、出産日を起点として時期別に母親の関心が解析できる。本研究で用いる出産日付き検索履歴は、1人の子どもをもつ母親が2019年1月1日から2020年12月31日までの2年間で行った1語の検索1,895,283件を対象とする。

### 3.2 確率分布の推定

出産日付きの検索履歴を加工することで、各検索語に対して出産日を起点とする検索月の経験確率分布を作成できる。これらの経験確率分布は、概形が単峰型または二峰型でノイズを多く含む分布である。ノイズが含まれると、確率値や累積分布関数の値を精緻に扱えないだけでなく、分布の傾きについて議論することができない。そこで、先行研究 [5] で提案されている単峰性最適化モデルと二峰性最適化モデルを利用して経験確率分布からノイズを除去し、極値点の前後で単調性をもつ性質の良い確率分布を取得する。

まず、各検索語に対して、学習期間の出産日付き検索履歴から作成される経験確率分布に単峰性最適化モデルと二峰性最適化モデルを適用して確率分布を推定する。次に検証期間の出産日付き検索履歴から作成される経験確率分布と比較して二乗平均平方根誤差が小さいモデルを選択する。最後に、学習期間と検証期間のすべての出産日付き検索履歴から作成される経験確率分布に、選択されたモデルを適用して、推薦カレンダー作成に inputs する確率分布を取得する。

実際、学習期間は2019年1月1日から2019年12月31日までの1年間とし、検証期間は2020年1月1日から2020年12月31日までの1年間とした。対象とする検索語は実務での利用を想定して検索量が上位の固有名詞と名詞の中からマーケティングに関連する1,307件を抽出した。ただし、固有名詞は主に母親の関心事である商品名や医療系の語を抽出した。各検索語に対してモデル選択を行ったところ単峰性の確率分布1,024件、二峰性の確率分布283件が得られた。なお、確率分布の推定では、混合整数凸二次計画問題を最適化ソフトウェア Gurobi Optimizer<sup>2</sup>(ver. 9.5.1)を用いて解いた。

### 3.3 推定した確率分布の特徴

単峰性最適化モデル、および二峰性最適化モデルを用いて推定した確率分布の特徴について確認する。推定した確率分布は期間集合を  $T$ 、確率分布を  $(p_t | t \in T)$  とすると次の性質をもつ。

単峰性最適化モデルから得られる確率分布

$$\begin{aligned} 0 &\leq p_t \leq 1 & (\forall t \in T) \\ \sum_{t \in T} p_t &= 1 \\ p_{\min(T)} &= 0 \\ p_{\max(T)} &= 0 \\ p_t &\leq p_{t+1} & (\forall t \in T, t+1 \leq t_0) \\ p_t &\geq p_{t+1} & (\forall t \in T, t_0 \leq t, t < \max(T)) \end{aligned} \quad (1)$$

ただし、極大値を取る期を  $t_0$  とする。

二峰性最適化モデルから得られる確率分布

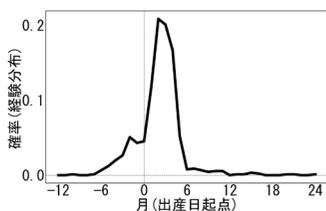
$$\begin{aligned} 0 &\leq p_t \leq 1 & (\forall t \in T) \\ \sum_{t \in T} p_t &= 1 \\ p_{\min(T)} &= 0 \\ p_{\max(T)} &= 0 \\ p_t &\leq p_{t+1} & (\forall t \in T, t+1 \leq t_0) \\ p_t &\geq p_{t+1} & (\forall t \in T, t_0 \leq t, t+1 \leq t_1) \\ p_t &\leq p_{t+1} & (\forall t \in T, t_1 \leq t, t+1 \leq t_2) \\ p_t &\geq p_{t+1} & (\forall t \in T, t_2 \leq t, t < \max(T)) \end{aligned} \quad (2)$$

ただし、一つ目の極大値を取る期を  $t_0$ 、極小値を取る期を  $t_1$ 、二つ目の極大値を取る期を  $t_2$  とする。なお、単峰性最適化モデル、および二峰性最適化モデルから得られる確率分布について、 $p_{\min(T)} = 0$  は期間  $T$  の開始時期では検索が始まっておらず、 $p_{\max(T)} = 0$  は期間  $T$  の終了時期では検索が終わっていることを表す。

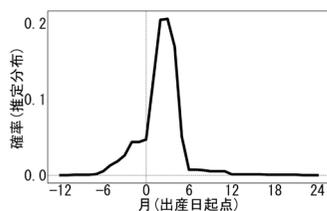
図1は、検索語「ベビージム」の確率分布であり、(a)は経験確率分布を(b)は推定確率分布を表している。経験確率分布から2ヶ月目にピークとなる単峰型の概形をもつことが確認でき、「ベビージム」が乳幼児向けの玩具であることと合致している。また、経験確率分布と推定確率分布を比較すると経験確率分布の裾のノイズが推定確率分布では除去されており、式(1)を満たすことが確認できる。図2は、検索語「チーズ」の確率分布であり、(a)は経験確率分布を(b)は推定確率分布を表している。経験確率分布から-6ヶ月目と12ヶ月目にピークとなる二峰型の概形をもつことが確認できる。-6ヶ月目のピークは母親が妊娠中にチーズの摂取にリスクがあるか検索することで生じる。一

<sup>1</sup> <https://connehito.com/>

<sup>2</sup> <https://www.gurobi.com/products/gurobi-optimizer/>

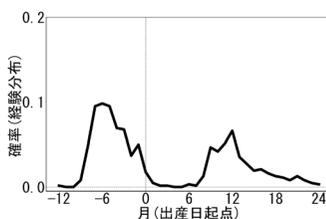


(a) 「ベビージム」の経験確率分布

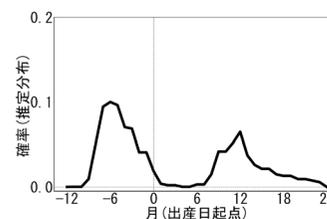


(b) 「ベビージム」の推定確率分布

図1 単峰性最適化モデルによる推定の例



(a) 「チーズ」の経験確率分布



(b) 「チーズ」の推定確率分布

図2 二峰性最適化モデルによる推定の例

方、12ヶ月目のピークは乳幼児がチーズを食べはじめる時期を検索することで生じる。経験確率分布のピーク以外のノイズが推定確率分布では除去されており、式(2)を満たすことが確認できる。

#### 4. 提案手法

本節では、3節で得られた確率分布を用いて、推薦カレンダーを作成する方法を提案する。

##### 4.1 推薦カレンダー作成問題

推薦カレンダーとは、出産前から出産後の各月に対して、推薦する複数のキーワードが対応するテーブルである。本論文では出産月を起点(0ヶ月)として出産月の6ヶ月前から6ヶ月後にわたる13ヶ月間を対象に、キーワード候補の中から推薦するキーワードを決定する。ただし、キーワード候補は3節で作成した確率分布に対応する1,307件の検索語である。

各月に推薦するキーワードは、キーワード候補となる検索語の確率分布を入力として、検索量、確率値、分布の傾き、累積分布関数を考慮して決定する。推薦カレンダーの入力となる検索月の確率分布は、検索利用者の時期別の関心の推移を表すと考え、次のように推薦カレンダーの要件を定める。

##### (A) キーワードの重複禁止

推薦カレンダーの中でキーワードは重複して推薦しない。

##### (B) 各月の推薦キーワード数

1日一つのキーワードを推薦することを想定して、

各月に推薦するキーワードは30件とする。

##### (C) 関心の減少月への推薦禁止

各キーワードは関心が減少傾向にある月には推薦しない。

##### (D) 関心量が少ない早い期間と遅い期間への推薦禁止

各キーワードは関心量が少ない期間、すなわち、早すぎる期間と遅すぎる期間には推薦しない。

##### (E) 一般的なキーワード優先

できるだけ検索量の多いキーワードを推薦する。

##### (F) キーワードの推薦タイミング優先

各キーワードはできるだけ関心が急増加する月に推薦する。

##### (G) 適切な関心量の月への推薦を優先

各キーワードは適切な関心量をもつ月からできるだけ外れないように推薦する。

##### 4.2 推薦カレンダーの数理最適化モデル

推薦カレンダー作成問題の数理最適化モデルを提案する。4.1節で述べた推薦カレンダー作成問題を分割問題に帰着させ、多目的最適化問題として定式化する。

##### 集合と定数の定義

まず、集合の定義を行う。期間の集合  $T$  は、出産月を0とした月の集合で

$$T := \{-6, -5, \dots, -1, 0, 1, \dots, 5, 6\}$$

と定める。また、キーワード候補の集合を

$W$

とする。

次に定数の定義を行う。すべてのキーワード  $w$  の検索量を昇順に並べて分位数を取った値を

$$v_w \quad (w \in W)$$

とする。  $v_w$  は  $0 \leq v_w \leq 1$  を満たす正規化された値である。また、キーワード  $w$  の推定した確率分布について、すべての月  $t$  の傾きを

$$d_{w,t} (w \in W, t \in T)$$

とし、さらに  $d_{w,t}$  を昇順に並べて分位数を取った値を

$$s_{w,t} \quad (w \in W, t \in T)$$

とする。  $s_{w,t}$  は  $0 \leq s_{w,t} \leq 1$  を満たす正規化された値である。次に、キーワード  $w$  に関する確率分布について月  $t$  に関する累積分布関数の値を

$$c_{w,t} \quad (w \in W, t \in T)$$

とする。  $c_{w,t}$  は  $0 \leq c_{w,t} \leq 1$ ,  $c_{w,-6} = 0$ ,  $c_{w,6} = 1$  を満たし、  $t$  について単調増加である。

#### 変数と制約の定義

まず、決定変数を定義する。キーワード  $w$  を月  $t$  に推薦する場合に 1、推薦しない場合に 0 を取る 0-1 整数変数を

$$x_{w,t} \in \{0, 1\} \quad (w \in W, t \in T)$$

とする。推薦カレンダーの形式上の条件として要件 (A) から

$$\sum_{t \in T} x_{w,t} \leq 1 \quad (\forall w \in W) \quad (3)$$

を、要件 (B) から

$$\sum_{w \in W} x_{w,t} = 30 \quad (\forall t \in T) \quad (4)$$

を制約に追加する。

次に推薦カレンダー作成における禁止制約を追加する。要件 (C) から各キーワードは検索確率の傾きが負となる月への推薦を禁止するので、

$$x_{w,t} = 0 \quad (\forall w \in W, \forall t \in T, d_{w,t} < 0) \quad (5)$$

を制約に追加する。また、要件 (D) は累積分布関数を用いて関心量が少ない期間を決める。本問題設定では、累積分布関数の値が 0.05 より小さければ早すぎるとし、0.5 より大きければ遅すぎると仮定して

$$x_{w,t} = 0 \quad (\forall w \in W, \forall t \in T, c_{w,t} < 0.05) \quad (6)$$

$$x_{w,t} = 0 \quad (\forall w \in W, \forall t \in T, 0.5 < c_{w,t}) \quad (7)$$

を制約に追加する。

ここで中間変数を導入する。要件 (G) については累積分布関数を用いて適切な関心量をもつ月を決める。本問題設定では、累積分布関数の値が 0.25 となる時期が適切であると仮定して、できるだけ累積分布関数の値が 0.25 に近くなる月に推薦を行う。そこで累積分布関数の値が 0.25 より大きくなる月に推薦する場合の違反量を連続変数

$$y_w \in \mathbb{R}_{\geq 0} \quad (w \in W)$$

とし、累積分布関数の値が 0.25 より小さくなる月に推薦する場合の違反量を連続変数

$$z_w \in \mathbb{R}_{\geq 0} \quad (w \in W)$$

とする。このとき基準値 0.25 からの差異を表す制約式として

$$\sum_{t \in T} x_{w,t} c_{w,t} - 0.25 \sum_{t \in T} x_{w,t} \leq y_w \quad (\forall w \in W) \quad (8)$$

$$0.25 \sum_{t \in T} x_{w,t} - \sum_{t \in T} x_{w,t} c_{w,t} \leq z_w \quad (\forall w \in W) \quad (9)$$

を追加する。

#### 目的関数の定義

まず、多目的最適化問題として定式化するための指標の準備をする。要件 (E) から最大化する指標

$$\sum_{w \in W} v_w \left( \sum_{t \in T} x_{w,t} \right)$$

を与える。次に要件 (F) からキーワードを推薦する月はできるだけ分布の傾きが大きい月に推薦するので、最大化する指標

$$\sum_{w \in W, t \in T} s_{w,t} x_{w,t}$$

を与える。最後に要件 (G) から式 (8), (9) で定まる  $y_w$  と  $z_w$  を用いて最小化する指標

$$\sum_{w \in W} (y_w + z_w)$$

を与える。

最後にそれぞれの指標に対する重みを  $\alpha, \beta, \gamma$  として多目的最適化問題となるように最大化問題として、

$$\begin{aligned} & \alpha \sum_{w \in W} v_w \left( \sum_{t \in T} x_{w,t} \right) \\ & + \beta \sum_{w \in W, t \in T} s_{w,t} x_{w,t} \\ & - \gamma \sum_{w \in W} (y_w + z_w) \end{aligned} \quad (10)$$

を目的関数に定める。

式(3)–(10)をまとめて、推薦カレンダー作成問題は混合整数計画問題として次のように定式化できる。

$$\begin{aligned} \min. & \quad \alpha \sum_{w \in W} v_w \left( \sum_{t \in T} x_{w,t} \right) \\ & + \beta \sum_{w \in W, t \in T} s_{w,t} x_{w,t} \\ & - \gamma \sum_{w \in W} (y_w + z_w) \\ \text{s.t.} & \quad \sum_{t \in T} x_{w,t} \leq 1 \quad (\forall w \in W) \\ & \quad \sum_{w \in W} x_{w,t} = 30 \quad (\forall t \in T) \quad (11) \\ & \quad \sum_{t \in T} x_{w,t} c_{w,t} - 0.25 \sum_{t \in T} x_{w,t} \leq y_w \quad (\forall w \in W) \\ & \quad 0.25 \sum_{t \in T} x_{w,t} - \sum_{t \in T} x_{w,t} c_{w,t} \leq z_w \quad (\forall w \in W) \\ & \quad x_{w,t} = 0 \quad (\forall w \in W, \forall t \in T, d_{w,t} < 0) \\ & \quad x_{w,t} = 0 \quad (\forall w \in W, \forall t \in T, c_{w,t} < 0.05) \\ & \quad x_{w,t} = 0 \quad (\forall w \in W, \forall t \in T, 0.5 < c_{w,t}) \end{aligned}$$

なお、混合整数計画問題は最適化ソフトウェア **Gurobi Optimizer** を用いて解くことができる。

## 5. 推薦カレンダーの考察

本節では混合整数計画問題 (11) を解くことで得られる推薦カレンダーの考察をする。

目的関数 (10) の重みを  $\alpha = 1, \beta = 1, \gamma = 1$  として作成した基準となる推薦カレンダーを表 2 に示す。一方、 $\alpha = 1, \beta = 10, \gamma = 1$  として作成したタイミング重視の推薦カレンダーを表 3 に示す。タイミング重視の推薦カレンダーは、目的関数 (10) における分布の傾きの項  $s_{w,t}$  にかかる重みを大きくすることでより傾きの大きな月、すなわち関心が急増加する月にキーワードを推薦させる意図をもつ。なお、表 2、表 3 は -6 ヶ月、-2 ヶ月、0 ヶ月、2 ヶ月、6 ヶ月の 5 ヶ月間を抜粋している。各月にはそれぞれ 30 件のキーワードが推薦されており、検索量が多いほど上位に掲載されている。また、表 2、表 3 に掲載されていない月も含めて、どちらかの推薦カレンダーにのみ現れるキーワードを太字で記した。

### 5.1 基準となる推薦カレンダーの考察

表 2 の基準となる推薦カレンダーの考察を行う。-6 ヶ月に注目すると「胎動」や「むくみ」、「母親学級」など、母親の体調や活動に関連するキーワードが多く

推薦されている。次に -2 ヶ月では、「陣痛アプリ」や「授乳服」、「出生届」のように母親の体の変化や出産準備に関するキーワードが推薦されている。出産月である 0 ヶ月では、「産後うつ」や「産後ダイエット」など母親の話題も含まれるが、出産直後の乳幼児に関する話題である「乳児湿疹」や「おしゃぶり」、「おむつ」といったキーワードも推薦されている。その後、2 ヶ月では「お食い初め」や「セルフねんね」、6 ヶ月では「つかまり立ち」や「後追い」のように、子どもの成長に関連するキーワードが多く推薦される。

このように推薦カレンダーは産前産後にかけて母親の話題から子どもの話題へ推移し、時期特有の関心に合致した適切なキーワードが推薦されていることが確認できる。

### 5.2 タイミング重視の推薦カレンダーとの比較

表 2 の基準となる推薦カレンダーと表 3 のタイミング重視の推薦カレンダーの比較を行う。基準となる推薦カレンダーはタイミング重視の推薦カレンダーに比べて一般的なキーワードが多い。これは基準となる推薦カレンダーのほうが相対的に検索量の多いキーワードが選択されているためである。実際、基準となる推薦カレンダーのキーワードの検索量平均は 3,677.3 件 (中央値 2,056.5 件)、タイミング重視の推薦カレンダーのキーワードの検索量平均は 3,230.6 件 (中央値 1,661.0 件) となっている。一方、タイミング重視の推薦カレンダーは分布の傾きが大きな月にキーワードが推薦されている。実際、タイミング重視の推薦カレンダーのキーワードの傾き平均は 0.1172、基準となる推薦カレンダーのキーワードの傾き平均は 0.0975 となっている。

より詳細に比較するために互いの推薦カレンダーに共通して現れるキーワード 313 件のうち推薦月が変化したキーワード 87 件に注目した。87 件のキーワードについて、基準となる推薦カレンダーの傾き平均は 0.0496、タイミング重視の推薦カレンダーの傾き平均は 0.1010 となっている。

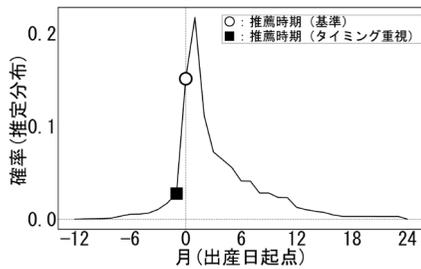
具体的な推薦月の変化を調べるために推薦月が変化した「ミルク」と「お宮参り」に注目した。図 3 の (a) は「ミルク」に対する推薦月を付与した確率分布、(b) は「お宮参り」に対する推薦月を付与した確率分布である。「ミルク」の推薦月は基準となる推薦カレンダーでは 0 ヶ月であるが、タイミング重視の推薦カレンダーでは -1 ヶ月に前倒しされている。これは -1 ヶ月から 0 ヶ月にかけての分布の傾きが 0.1236 である一方、0 ヶ月から 1 ヶ月にかけての分布の傾きが 0.0659 となっており、推薦月を前倒しする方が関心が急増化する

表2 基準となる推薦カレンダー ( $\alpha = 1, \beta = 1, \gamma = 1$ )

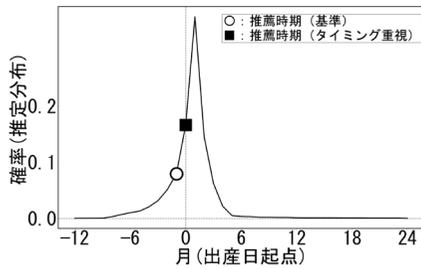
-6ヶ月	-2ヶ月	0ヶ月	2ヶ月	6ヶ月
胎動	前駆陣痛	乳児湿疹	お食い初め	断乳
性別	帝王切開	おしゃぶり	ヒップシート	フォローアップミルク
抱っこ紐	便秘	産後うつ	股関節脱臼	ファーストシューズ
口唇口蓋裂	新生児	完ミ	スリーパー	後追い
骨盤ベルト	出産報告	完母	ジョイントマット	ベビーフード
前置胎盤	学資保険	乳腺炎	絶壁	夜間断乳
腹帯	臨月	吐き戻し	バンボ	卵アレルギー
羊水検査	出産手当金	ミルク	点頭てんかん	歯磨き
エコー	おっぱいマッサージ	授乳間隔	こどもちゃれんじベビー	つかまり立ち
トコちゃんベルト	高位破水	髄鞘炎	セルフねんね	ベビーダノン
性別判定	破水	骨盤矯正	寝言泣き	パン粥
酸化マグネシウム	予定帝王切開	母乳量	抜け毛	ヨーグルト
むくみ	ラズベリーリーフティー	産後クライシス	爪切り	マミーボコパンツ
胎児スクリーニング	正産期	スワドルアップ	産後太り	はちみつ
スクリーニング検査	陣痛アプリ	湿疹	ウエスト症候群	チュアベルト
母親学級	授乳ケープ	混合	ロコイド	ハイハイ
抱き枕	予定日	背中スイッチ	バスチェア	ハイハイン
胎教	へその緒	エルゴオムニ 360 クールエア	プロベト	納豆
オーガズム	おしりふき	睡眠時間	ディズニーデビュー	食パン
姓名判断	スクワット	混合育児	ベビーマッサージ	コロコロうんち
前期破水	内診グリグリ	鼻吸い器	哺乳瓶拒否	たまごボーロ
妊娠中期	巨大児	乳頭混乱	エアリコ	卵
マタニティヨガ	育児日記	授乳時間	先天性股関節脱臼	パンがゆ
たばこ	授乳服	モロー反射	蒙古斑	ビングル
両親学級	体温計	頻回授乳	成長曲線	すくすくチェア
ラクソベロン	乳頭マッサージ	液体ミルク	ヘパリン類似物質	育休明け
切迫	出生届	おむつ	掛け布団	歯ブラシ
リステリア菌	孫フィーバー	母乳外来	キンダベート	赤ちゃんせんべい
ベビーグッズ	内診	産後ダイエット	レッグウォーマー	卵黄
ジェンダーリベール	出産予定日	寝不足	オーボール	ハミコ

表3 タイミング重視の推薦カレンダー ( $\alpha = 1, \beta = 10, \gamma = 1$ )

-6ヶ月	-2ヶ月	0ヶ月	2ヶ月	6ヶ月
胎動	前駆陣痛	乳児湿疹	お食い初め	後追い
性別	学資保険	おしゃぶり	ヒップシート	ベビーフード
ベビーカー	臨月	お宮参り	バンボ	夜間断乳
抱っこ紐	出産	完ミ	パパ見知り	卵アレルギー
口唇口蓋裂	おっぱいマッサージ	吐き戻し	こどもちゃれんじベビー	歯磨き
骨盤ベルト	会陰マッサージ	乳幼児突然死症候群	セルフねんね	つかまり立ち
前置胎盤	予定帝王切開	メルシーボット	抜け毛	ベビーダノン
腹帯	妊娠高血圧症候群	髄鞘炎	テアトルアカデミー	パン粥
むくみ	ラズベリーリーフティー	骨盤矯正	爪切り	ヨーグルト
胎児スクリーニング	正産期	産後クライシス	ウエスト症候群	五月人形
尿漏れ	陣痛アプリ	コニー	ロコイド	ハイハイ
張り	計画無痛分娩	スワドルアップ	ディズニーデビュー	ハイハイン
スクリーニング検査	胎児発育不全	湿疹	水分補給	マルツエキス
母親学級	初産	内祝い	哺乳瓶拒否	納豆
抱き枕	スクワット	背中スイッチ	マカロンバス	食パン
胎教	妊娠高血圧	ギャン泣き	歯固め	コロコロうんち
オーガズム	内診グリグリ	エルゴオムニ 360 クールエア	エアリコ	たまごボーロ
姓名判断	巨大児	睡眠時間	先天性股関節脱臼	卵
体重増加	育児日記	ジーナ式	ヘパリン類似物質	カッテージチーズ
妊娠中期	授乳服	鼻吸い器	掛け布団	ビングル
マタニティヨガ	乳頭マッサージ	スリング	背ばい	すくすくチェア
たばこ	出生届	ベビーモニター	キンダベート	歯ブラシ
両親学級	ニューボーンフォト	マザーズリック	レッグウォーマー	赤ちゃんせんべい
切迫	内診	液体ミルク	オーボール	アッフルチェア
ベビーグッズ	限度額適用認定証	おむつ	四種混合	卵黄
仰向け	外回転術	K2 シロップ	ミアミリー	ハミコ
転倒	ベビーリング	桶谷式	コアラシート	コンセントカバー
ジェンダーリベール	ベビー返し	寝不足	タミータブ	BF
羊水	股関節	でべそ	リーメンビューゲル	ヘッドガード
マタニティコーデ	骨盤	ナッパナッパ	ドケルバン病	ツナ缶



(a) 「ミルク」の確率分布



(b) 「お宮参り」の確率分布

図3 確率分布における推薦時期

るタイミングでキーワードを推薦できるためである。「お宮参り」の推薦月は基準となる推薦カレンダーでは-1ヶ月であるが、タイミング重視の推薦カレンダーでは0ヶ月と先送りされている。これは-1ヶ月から0ヶ月にかけての分布の傾きが0.0869である一方、0ヶ月から1ヶ月にかけての分布の傾きが0.1927となっており、推薦月を先送りする方が関心が急増するタイミングでキーワードを推薦できるためである。

このようにタイミング重視の推薦カレンダーはより分布の傾きが大きくなる月、すなわち、関心が急増する月にキーワードを推薦していることがわかる。

### 5.3 推薦カレンダーの柔軟性

目的関数(10)の重み $\alpha, \beta, \gamma$ の重みを変化させると推薦キーワードがどの程度入れ替わるか考察を与える。

図4は、基準となる推薦カレンダーに対して $\alpha, \beta, \gamma$ を増加させた場合に推薦キーワードが入れ替わる数を示している。 $\alpha, \beta, \gamma$ の重みを増加させると推薦キーワードが入れ替わる数がおおよそ単調に増加することが確認できる。特にそれぞれの重みを十分に大きな値に設定したとき、390件のキーワードのうち $\alpha = 100$ では53件、 $\beta = 100$ では122件、 $\gamma = 100$ では230件のキーワードが入れ替わる。

このように目的関数の重みを調整することで実務の用途に応じた推薦カレンダーを作成することができる。

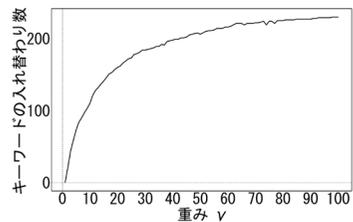
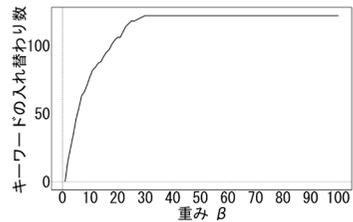
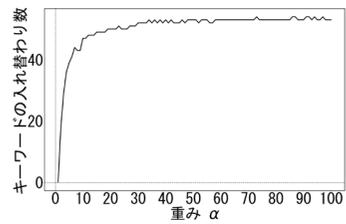


図4 キーワードの入れ替わり数の推移

たとえば、出産経験がない初産の母親にはできるだけ一般的なキーワードを含む人気重視の推薦カレンダーを利用して情報推薦を行い、第二子以降を出産する知識豊富な母親にはより細かい知識を補うことができるタイミング重視の推薦カレンダーを利用して情報推薦を行うことができる。また、重み調整が有効に機能するのは確率分布からノイズを除去することで確率値や累積分布関数、特に分布の傾きを精緻に利用できるようになったためと考えられる。

## 6. まとめ

本論文は時期別の母親の課題を解決することを目的に推薦カレンダーを作成した。まず、各検索語について出産日を起点とする確率分布を作成し、単峰性最適化モデルと二峰性最適化モデルを用いてノイズを除去する。そして推定した確率分布を入力とする割当問題を解くことで推薦カレンダーを作成した。推定した確率分布は、確率値、分布の傾き、累積分布関数を精緻に扱える。多目的最適化問題を解いて得られる推薦カレンダーは、目的関数の重みを調整することで柔軟に変更できる。具体的には、一般的なキーワードを優先する推薦カレンダーやキーワードの推薦タイミングを優先する推薦カレンダーを柔軟に作成できる。推薦カ

レンダーは多期間に複数のキーワードを同時に推薦できる。適切な時期に適切な情報を適切な量だけ推薦することができれば、母親の課題を先回りして支援できる。一方、企業は母親の課題が発生する時期に合わせて適切な商品やサービスを推薦する効率の良いマーケティングができる。実務では以前より母親の状態や時期に応じてさまざまな施策が検討されてきた。その多くは出産経験者や産科の医療従事者の知見を元に人手で施策立案されている。本論文で提案する推薦カレンダーの作成方法は、個人の経験や知識を大規模データから得られる集合知に置き換え、人手による施策立案を自動化できる可能性がある。自動化が実現できれば施策立案の高速化だけでなく、その時代の流行や特徴的な課題への即時対応も期待できる。

次に本研究で取り上げなかった課題について触れる。月粒度の推薦カレンダーの作成を行ったが、産前産後の母親の関心事は移り変わりが早いいため、週粒度の推薦カレンダーも有効であると考えられる。また、本研究ではキーワード候補に対する類義語処理を行っていない。そのため、「パン粥」と「パンがゆ」、「ベビーフード」と「BF」などが異なるキーワードとして扱われている。実用の際には必要に応じて名寄せすることが好ましい。

最後に今後の研究課題について述べる。本研究ではすべての母親に共通して適用できる推薦カレンダーの作成を目指したが、母親にパーソナライズした推薦カレンダーの作成も検討したい。たとえば、第1子出産のときの関心と第2子出産のときの母親の関心事は異なると考えられる。また、居住地域や出産した季節によっても母親の関心が異なると考えられるため、状況に応じて最適な推薦カレンダーを検討することができる。

**謝辞** 本研究では、経営科学系研究部会連合協議会主催令和3年度データ解析コンペティションで提供されたコネヒト株式会社の検索履歴データを使用した。関係各位に謝意を表する。

## 参考文献

- [1] ベネッセ教育総合研究所, 「産前産後の生活とサポートについての調査」, [https://berd.benesse.jp/up\\_images/research/sanzensango.pdf](https://berd.benesse.jp/up_images/research/sanzensango.pdf) (2022年6月22日閲覧)
- [2] 川井尚, 庄司順一, 千賀悠子, 加藤博仁, 中野恵美子, 恒次欽也, “育児不安に関する基礎的検討,” 日本総合愛育研究所紀要, **30**, pp. 27–39, 1994.
- [3] 唐田順子, 森田明美, “乳幼児をもつ母親の子育てに関する困りごとや悩みごとに関する研究—児の年齢別, 初経産別による検討—,” 東洋大学人間科学総合研究所紀要, **7**, pp. 249–263, 2007.
- [4] 藤田大輔, 金岡緑, “乳幼児を持つ母親の精神的健康度及ばすソーシャルサポートの影響,” 日本公衆衛生雑誌, **49**, pp. 305–313, 2002.
- [5] 岩永二郎, 西村直樹, 鮎川矩義, 高野祐一, “出産前後の情報検索の分析—数理最適化による検索日の確率分布推定—,” 人工知能, **37**, D-L74.1–11, 2022.
- [6] 岩永二郎, 鍋谷昂一, 梶原悠, 五十嵐健太, “関心度と忘却度に基づくレコメンド手法—単調性制約付きレコメンドモデルの構築—,” オペレーションズ・リサーチ: 経営の科学, **59**, pp. 72–80, 2014.
- [7] J. Iwanaga, N. Nishimura, N. Sukegawa and Y. Takano, “Estimating product-choice probabilities from recency and frequency of page views,” *Knowledge-Based Systems*, **99**, pp. 157–167, 2016.
- [8] N. Nishimura, N. Sukegawa, Y. Takano and J. Iwanaga, “A latent-class model for estimating product-choice probabilities from clickstream data,” *Information Sciences*, **429**, pp. 406–420, 2018.
- [9] J. Iwanaga, N. Nishimura, N. Sukegawa and Y. Takano, “Improving collaborative filtering recommendations by estimating user preferences from clickstream data,” *Electronic Commerce Research and Applications*, **37**, 100877, 2019.
- [10] C. C. Aggarwal, *Recommender Systems: The Text-Book*, Springer, 2016.
- [11] J. Bennett and S. Lanning, “The netflix prize,” In *Proceedings of KDD Cup and Workshop*, **2007**, pp. 3–6, 2007.
- [12] 下山雄大, 小林亜樹, “フォロワーの目に留まるツイート投稿タイミング推薦手法,” 情報処理学会研究報告. データベース・システム研究会報告, **17**, pp. 1–8, 2013.
- [13] 向井友宏, 黒澤義明, 目良和也, 竹澤寿幸, “マイクロログの分析に基づくユーザの嗜好とタイミングを考慮した情報推薦手法の提案,” 言語処理学会 第17回年次大会発表論文集, **17**, pp. 452–455, 2011.
- [14] 三輪冠奈, “コンビニエンスストアの業務と人員配置に関する研究,” 名古屋学院大学論文集社会科学篇, **51**, pp. 143–158, 2015.
- [15] 愛甲聡美, T. Phathinan, 瀬尾亨, 朝倉康夫, “アクティビティパターンを与件としたシェアリング車両の最適割り当て問題,” 土木学会論文集 D3 (土木計画学), **73**, pp. I\_1233–I\_1242, 2017.
- [16] S. Daskalaki, T. Birbas and E. Housos, “An integer programming formulation for a case study in university timetabling,” *European Journal of Operational Research*, **153**, pp. 117–135, 2004.
- [17] M. Seker and N. Noyan, “Stochastic optimization models for the airport gate assignment problem,” *Transportation Research Part E: Logistics and Transportation Review*, **48**, pp. 438–459, 2012.
- [18] A. Errarhout, S. Kharraja and C. Corbier, “Two-stage stochastic assignment problem in the home health care,” *IFAC-PapersOnLine*, **49**, pp. 1152–1157, 2016.