

< 修 士 論 文 >

アンケートに基づく顧客満足度モデルの
構築とその担当者割当への応用

滋 賀 大 学 大 学 院

デ ー タ サ イ エ ン ス 研 究 科

デ ー タ サ イ エ ン ス 専 攻

修了年度:2021年度

学籍番号:6020102

氏 名:池田 拓郎

指導教員:岩山 幸治

提出年月日:2022年1月 12 日

1	序論	1
1.1	研究の背景	1
1.2	研究の目的	2
1.3	本論文の構成	2
2	先行研究のレビュー	3
2.1	顧客満足度の期待-不一致モデル	3
2.2	ゲームのスキル推定アルゴリズム	4
3	顧客満足度のモデル化	8
3.1	ポアソン回帰モデル	8
3.2	順序付きプロビットモデル	9
3.3	マルコフ連鎖モンテカルロ法	10
4	階層ベイズモデルの逐次近似推定	12
4.1	提案モデル	12
4.2	担当者の成果の平均化	19
4.3	実データによる閾値の最尤推定	19
5	実証分析	21
5.1	データの説明と推定の概略	21
5.2	ポアソン回帰モデル	22
5.3	順序付きプロビットモデル	23
5.4	提案モデル	24
5.5	予測精度の比較	28
5.6	案件関与回数と予測精度の関係	29
5.7	項目別推定値による担当者と顧客の分類	32
5.8	考察	34
6	結論	36
6.1	まとめ	36
6.2	今後の課題と展望	36
	謝辞	38
	参考文献	39

1 序論

1.1 研究の背景

従来のマーケティングにおいて、企業の商品やサービスに対する顧客の満足度 (Customer satisfaction, CS) は、顧客の態度変容や商品・サービスの購買意欲の基本的な決定要因として重要視されてきた [1]。顧客満足度が高ければ高いほど顧客の定着率は高まり [2]、ポジティブな口コミ (word of mouth, WOM) も増える [3]。そして最終的には、顧客満足度が企業の将来の財務指標の成長性を高めるといった研究もなされている [4]。したがって今後、社会のグローバル化・デジタル化・自由化が進み、様々な市場で商品やサービスの競争が増えるにつれて、顧客満足度の重要性はますます高まっていくと予想される。

医療・美容・デザイン・コンサルタント・データ分析等といった、顧客に対して企業側の特定の人員 (もしくはチーム) が担当者となり、性能の定まった商品を持たない「不定形のサービス」を提供する「顧客-担当者型サービス」においては、その一連のビジネス (以降、案件と呼称) の品質水準が、その案件の担当者 (チーム) の能力 (スキル) に少なからず依存することは想像に難くない。一方で仮に同水準のサービスを提供したとしても、顧客が満足するかどうかは顧客の受け止め方次第であり、顧客の期待していた水準を超えているかどうかによって左右されることも、同意を得られやすいと考えられる。過去、顧客満足度の研究においてはこの顧客の期待水準と満足度の関係が注目されてきた [5]。

これに対し「顧客-担当者型サービス」の特徴として、顧客は自身の受けるサービスの水準がサービス購買後・契約後・利用後にしか分からず、予め自分の期待水準にあったサービスや担当者を選べないことがあり、また企業側も顧客の期待水準について事前に知ることができないこともあるという構造が挙げられる。これにより「顧客-担当者型サービス」においては、ビジネス規模の拡大に伴って顧客や担当者が増加すると、以下の事象が発生する。

- 担当者の人員の入れ替わりや教育コストにより、担当者間のスキルに一定の格差が生じる
- 担当者数が多くなると、すべてのスキルを正確には把握できない
- 顧客増加により、期待水準の幅が広がる。また顧客すべての期待水準を把握できない

この状況下では、「高い期待水準の顧客」に「低いスキルの担当者」があたってしまう事象すなわち顧客と担当者のミスマッチが発生する可能性が高まる。その結果、顧客満足度の低下、ロイヤルティの低下・離反・クレーム・悪い口コミを引き起こし、サービスを提供する企業全体のブランドイメージに悪影響を及ぼすこともある。この問題はサービス業界に限らず、市場シェアが高い企業が顧客層の幅広さ故に異質な顧客のニーズに応えきれず、結果的に顧客の不満足を招きかねないジレンマとして提唱されている [6]。以上から、顧客-担当者型サービス企業はそのビジネス拡大の過程に「顧客と担当者のミスマッチによる顧客満足度の低下リスクが増加する」という構造的な課題を抱えていると言える。

1.2 研究の目的

本論文では前項の課題の解決の一助となることを目指し、「顧客-担当者型サービス」における顧客の満足度をアンケート調査したデータを用いて、ベイズ推定により担当者スキルと顧客期待水準を定量的に推定することを目的とする。また推定した担当者スキル・顧客期待水準を用いた案件の顧客満足度の予測から、「期待水準の高い顧客に低いスキルの担当者を任命したことにより顧客満足度が低くなる」といった mismatches を事前察知して回避する可能性について検討する。これにより、構造的に顧客満足度を安定させづらい「顧客-担当者型サービス」についても、満足度を一定水準以上に保つことを目指せるものとする。また副次的に、推定された担当者スキルや顧客期待水準の担当者や顧客の管理への活用についても考察する。

1.3 本論文の構成

第2章では、本論文に関する先行研究・事例を整理、検討する。第3章では、顧客満足度に対する担当者スキル・顧客期待水準の関係をポアソン回帰モデル・順序プロビットモデルによりモデル化し、マルコフ連鎖モンテカルロ法によりパラメータを推定するための手法を示す。第4章では、顧客満足度と担当者スキル・顧客期待水準の関係を期待-不一致モデル [5]・TrueSkill [7]の手法を元にした階層モデルを用いて記述し、逐次的にパラメータを近似推定するための手法を示す。第5章では、提案した手法に基づき、リサーチ会社の株式会社マクロミルの有する案件顧客満足度アンケートデータを用いた実証分析を行い、その結果を考察する。第6章では、本論文のまとめとして第5章から得られた成果に基づき提案手法の課題や今後の展望について述べる。

2 先行研究のレビュー

本章では、顧客満足度のメカニズムについての先行研究を紹介するとともに、「顧客-担当者サービス」における顧客満足度のモデル化と親和性があると考えられる、ゲームにおけるプレイヤースキル推定(レーティング)のアルゴリズムについて説明する。

2.1 顧客満足度の期待-不一致モデル

マーケティングにおける顧客満足度の研究で代表的なものとして、「期待-不一致モデル(expectation disconfirmation model)」[5]が挙げられる。期待-不一致モデルとは、商品購入前の期待(expectation)と購入後に感じられた成果(performance)の一致-不一致によって顧客満足度を評価するモデルであり、具体的には成果が期待通りか期待以上であれば顧客は満足を得、逆に成果が期待以下の場合には、顧客は不満を感じるというモデルである。モデルの概念図を図 1 に示す [8]。

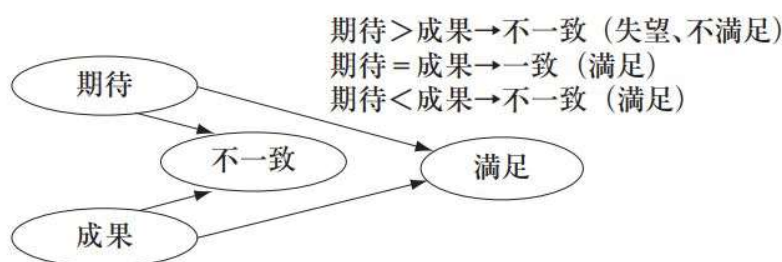


図 1 期待-不一致モデル [8]

日本においては、この期待-不一致モデルを元に 2010 年から日本版顧客満足度指数 JCSI (Japanese Customer Satisfaction Index) の運用が始められた [9]。JCSI は ACSI (American Customer Satisfaction Index) を基礎に、「顧客期待(利用前の予想・期待)」と「知覚品質(利用した際の品質評価)」「知覚価値(価格への納得感)」を基軸に顧客満足度を測定し、また顧客満足度の達成から「ロイヤリティ(継続的な利用意向)」と「クチコミ」の生起を仮定する。図 2 にその概要を示す [8]。

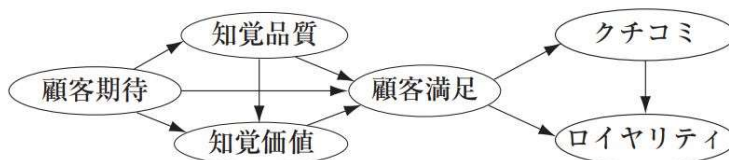


図 2 日本版顧客満足度指数のモデル [8]

また小野は期待-不一致モデルについて、知覚品質と知覚価値(2つを統合して知覚パフォーマンスと表現)を単純に向上させれば顧客満足度が上がるわけではなく、以下の3つの属性が混合して影響すると提唱している [6]。

- 最低条件:一定の水準を下回ったら商品・サービスとしての存在意義すら疑われる、最低許容水準を維持することが求められる属性。一定水準以上は顧客満足度への影響は薄い。
- 満足因子:水準を上げれば上げるほど、比例的に満足度が高くなる関係をもつ属性。競合が同様の企業努力をしている中では差別化が難しい。
- 感動因子:一定水準を超えると急激に顧客満足度を高める効果を持つ属性。商品やサービスのコモディティ化が進んだ市場において重要。

上記の3つの属性について、本論文で扱う「顧客-担当者型サービス」における優先順位を考えると、最低条件を下回った場合には顧客のロイヤルティが低下し離反リスクが高まる、また負の方向のロコミにより市場内でのブランド価値低下が発生する可能性があるため、まずは最低条件以上の価値を獲得することが優先され、それが達成できたのちに満足因子、感動因子の順に取り組むことがビジネス上妥当な戦略であると考えられる。また感動因子は顧客の期待を大きく上回ることで発生する属性であるため、ビジネスとして再現性を担保することは容易ではない。よって本論文では既存の顧客満足度の研究に基づき、「顧客-担当者型サービス」における顧客の満足度について以下を仮定する。まず、顧客の満足度は、期待-不一致モデルから、担当者が提供するサービスの品質と、顧客の期待の一致もしくは不一致により決定され、サービスの品質が顧客の期待する水準、すなわち最低条件を下回った場合に著しく満足度を下げるものとする。この期待水準は顧客によって異なるものとする。また、サービスの品質は担当者のスキルによって決まるものとする。これらの仮定の下で、サービスの品質が顧客の許容する最低条件を下回るリスクの回避を、最優先の課題とする。

この課題を達成するため、顧客の期待水準と担当者のスキルから顧客の満足度が決定される過程のモデルに基づき、顧客の期待水準、担当者のスキルとそこから決まるサービスの品質を推定する。そして、推定された顧客の期待水準と担当者のスキルの情報を用いて、顧客満足度が著しく低くならないように顧客と担当者のマッチングすることを目指す。

2.2 ゲームのスキル推定アルゴリズム

前項で提示した「顧客-担当者型サービスにおける顧客の期待水準と担当者スキルの推定と、そのマッチング」という問題は、顧客と担当者をそれぞれゲームのプレイヤー、案件をゲームの試合、顧客の満足度をゲームの試合結果、すなわち満足度が高ければ担当の勝利、低ければ顧客の勝利と考えると、「人間同士の対戦ゲームにおけるプレイヤースキルの推定とマッチング」問題と類似した問題と考えられる。このとき、顧客の期待水準と担当者のスキルがそれぞれプレイヤースキルに対応する。ゲームにおけるプレイヤースキルの推定(レーティング)とマッチングは、以下の効用をもたらすとされる [7]。

- プレイヤーが同じようなスキルを持つ他のプレイヤーとマッチングすることで、面白くパランスの取れた試合を行うことができる。

- プレイヤーのスキルレーティングをプレイヤー自身や一般の人々に公開することで、ゲームに関する興味やプレイヤー同士の競争心を刺激することができる。

- レーティングはトーナメントの参加資格の基準として使用することもできる。

これを顧客-担当者型ビジネスに置き換えて解釈すると、以下のように解釈できる。

- 顧客に対し、その期待水準に見合ったスキルを有する担当者をマッチングすることで、期待水準と品質の不一致を抑え、最低水準以上の顧客満足度を担保するとともに、サービスの過剰品質も抑えられる。

- 担当者のスキルを担当者自身に公開することで、自身の強み・弱みを把握し学習意欲を刺激することができる。なおゲームと異なり、スキルの組織内への共有は避けたほうが妥当と思われる。

- 担当者スキルを、高度な期待水準を持つ顧客の案件を担当する資格の基準として使用することができる。

この解釈を踏まえると、ゲームにおけるプレイヤースキル推定の問題と、顧客-担当者型サービスにおける期待水準・スキル推定の問題の親和性は高いように考えられる。そこで以下では、ゲームにおけるプレイヤースキルのレーティングについての先行研究を紹介する。

ゲームにおけるスキルレーティングの歴史は、1959年に Arpad Elo がチェスの統計的なレーティングシステム(Elo システム) [10]を開発し、1970年に世界チェス連盟 FIDE で採用されたものが代表的な始まりとされる。Elo システムの主なアイデアは、可能なゲーム結果の確率を2人のプレイヤーのスキル評価 s_1 と s_2 の関数としてモデル化するものであったが、あくまで1対1の対戦という制約があった。一方で昨今隆盛したインターネットを用いた大規模なネットワーク対戦型ゲームにおいては、複数プレイヤーが所属するチーム同士の対戦もあり、これらに対応するため Herbrich らは、チーム対戦の試合結果から各プレイヤーの実力を推定するアルゴリズム「TrueSkill」を提案した [7]。TrueSkill はマイクロソフトにより Xbox Live でのコンピュータゲームのマッチング用に開発されたプレイヤーの強さを評価(レーティング)するためのアルゴリズムであり、Elo システムとは異なり、3人以上でプレイするゲームも想定して設計されている。「顧客-担当者型ビジネス」においても、一人の顧客に対して営業や分析担当など複数の担当者がチームとなってサービスを提供する事例が少なくないため、この TrueSkill について説明する。

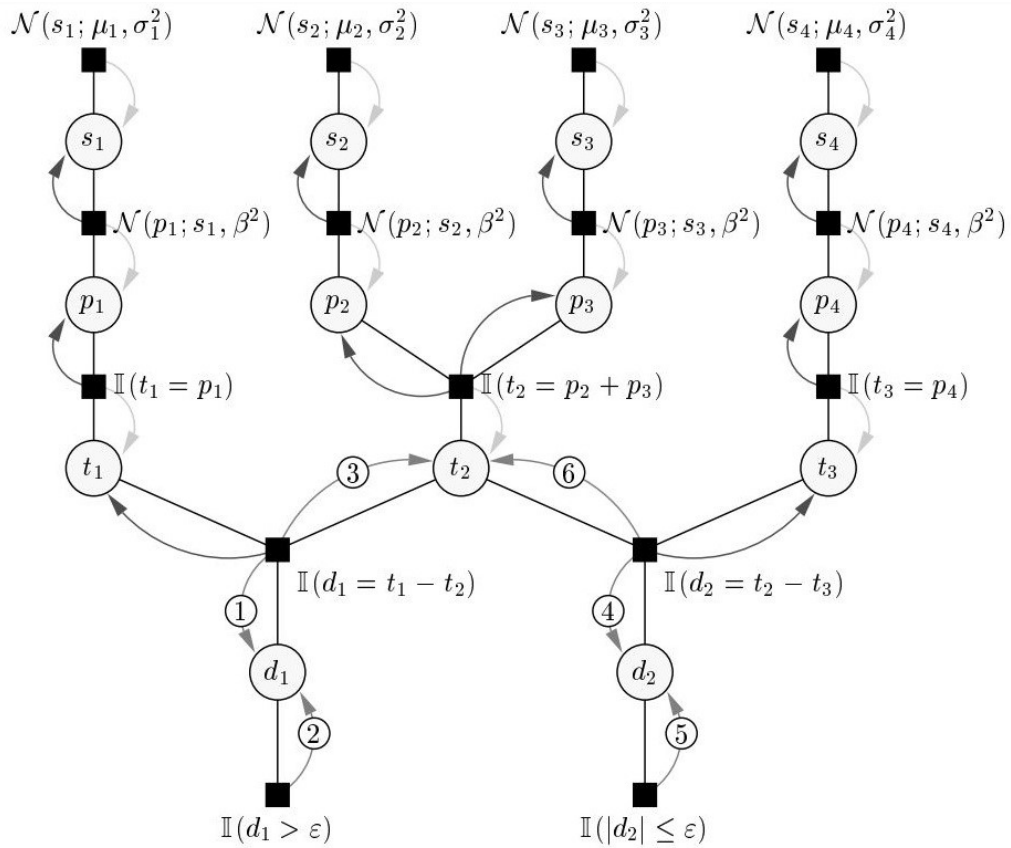


図3 TrueSkill の因果グラフ [7]

TrueSkill では n 人のプレイヤーが k 個のチームに分かれて順位を競うゲームについて、図3の因子グラフで表されるモデルを考える。TrueSkill モデルには4つの変数が存在する。図中の s_i はプレイヤーのゲームスキル、 p_i はプレイヤーのゲームパフォーマンス、 t_i はチームのパフォーマンスの合計、 d_j は対戦するチーム間のパフォーマンスの差を表す。各チームのゲームの結果の順位は $\mathbf{r} := (r_1, \dots, r_k)^T \in \{1, \dots, k\}$ で表し、もし $r_i = r_j$ のように順位が等しければ、その2チームの試合結果は引き分けであることを意味する。各チームのプレイヤー編成を $A := \{A_1, \dots, A_k\}$ とする。このとき、チーム編成 A 、プレイヤースキル \mathbf{s} が与えられた下で、結果の順位が \mathbf{r} となる確率を $P(\mathbf{r}|\mathbf{s}, A)$ とすると、プレイヤーのスキル \mathbf{s} の事後分布はベイズの法則より以下の式で推定される。

$$P(\mathbf{s}|\mathbf{r}, A) = \frac{P(\mathbf{r}|\mathbf{s}, A)P(\mathbf{s})}{P(\mathbf{r}|A)}$$

TrueSkill では、プレイヤー全員のスキルについて、

$$P(\mathbf{s}) = \prod_{i=1}^n \mathcal{N}(s_i; \mu_i, \sigma^2)$$

という正規分布の事前分布を仮定する。ここで \mathcal{N} は正規分布の密度関数を表す。またそれぞれのプレイヤーが発揮するパフォーマンス p_i が、期待値がそのプレイヤーのスキル s_i であり、各プレイヤーに共通の分散 β^2 を持つ正規分布に従うと仮定し、

$$p_i \sim \mathcal{N}(p_i; s_i, \beta^2)$$

と表す。チーム j のパフォーマンスの合計 t_j は、プレイヤーのパフォーマンスの和として

$$t_j := \sum_{i \in A_j} p_i$$

で表す。このとき、チームの順位 r は以下を満たすように決定される:

$$\begin{aligned} t_i - t_j \geq \epsilon &\Rightarrow r_i < r_j, \\ |t_i - t_j| < \epsilon &\Rightarrow r_i = r_j. \end{aligned}$$

つまり、チームのパフォーマンスが高いほどチームの順位が高くなる。また、チーム i とチーム j 間のパフォーマンスの差 $d = |t_i - t_j|$ が一定の値 ϵ 以上でなければ、引き分けとなると考える。なおこの閾値 ϵ は実データから予め設定された「引き分けになる確率」から計算により得られるものとする。

TrueSkillでは1回のゲームごとにプレイヤースキルを推定するために、ガウシアン密度フィルタを参考にしたオンライン(逐次的)学習法 [11]を用い、事後分布を正規分布に近似して求め、それを次のゲームにおける各プレイヤーのスキルの事前分布として用いる。

また、もしプレイヤーのスキルが時間とともに変化する(次のゲームをプレイするまでの間に、スキルが上達or低下する)と予想される場合は、次ゲームのスキルが前ゲームのスキルを期待値とする正規分布 $\mathcal{N}(s_{i,t+1}; s_{i,t}, \gamma^2)$ に従うと仮定し、ゲームごとにスキルの事前分布の分散に一定値 γ^2 を加算するといった工夫もなされている。

本論文ではTrueSkillのプレイヤースキルからチーム同士の勝敗が決するモデルについて、プレイヤーを担当者あるいは顧客、チームを複数担当者のチームあるいは顧客、勝敗を顧客満足度として、担当者のスキルと顧客の期待水準の推定へ応用する。本論文で扱う顧客満足度のモデルでは、対戦は常に担当者と顧客の2チームのみの間で行うため、実際の推定アルゴリズムは簡略化される。

3 顧客満足度のモデル化

本章では、顧客満足度と担当者スキル・顧客期待水準の関係を回帰モデルで表し、マルコフ連鎖モンテカルロ法を用いて顧客満足度の結果情報からスキルと期待水準の事後確率を推定する手法を説明する。

3.1 ポアソン回帰モデル

前章の TrueSkill モデルに基づくモデルの前に、より単純なモデルとして、担当者と顧客の組み合わせをダミー説明変数、顧客満足度をポアソン分布に従う目的変数としたポアソン回帰モデルを検討する。「顧客-担当者型ビジネス」において1回のビジネスを「案件」と呼称する。いま、 L 件の案件について、それぞれの担当者・顧客・顧客満足度アンケート結果のデータを得られたとする。ここで全担当者数は N 人、全顧客数は M 人とする。このデータについて、案件の担当者を L 行 N 列のダミー一行列により、

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1N} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{iN} \\ \vdots & & \vdots & & \vdots \\ a_{L1} & \cdots & a_{Lj} & \cdots & a_{LN} \end{pmatrix}$$

$$a_{ij} \in \{0,1\}$$

と表現する。ここで担当者 j が案件 i に参加している場合は、 $a_{ij} = 1$ となり、不参加の場合は $a_{ij} = 0$ となる。また1つの案件を複数人チームで担当する場合は、複数箇所を1とすることで表現する。次に顧客について担当者と同様に L 行 M 列のダミー一行列により、

$$C = \begin{pmatrix} c_{11} & \cdots & c_{1j} & \cdots & c_{1M} \\ \vdots & & \vdots & & \vdots \\ c_{i1} & \cdots & c_{ij} & \cdots & c_{iM} \\ \vdots & & \vdots & & \vdots \\ c_{L1} & \cdots & c_{Lj} & \cdots & c_{LM} \end{pmatrix}$$

$$c_{ij} \in \{0,-1\}$$

と表す。顧客 j が案件 i の顧客である場合は $c_{ij} = -1$ となり、それ以外は $c_{ij} = 0$ となる。ここで -1 を用いるのは、期待-不一致モデルから顧客期待水準は高いほど満足度が低くなるためである。

いま実データの顧客満足度 $R = \{r_1, r_2, \dots, r_L\}$ がポアソン分布に従うとすると、ある案件 i についての担当、顧客満足度の関係は以下のポアソン回帰モデルで表される。

$$\begin{aligned} \ln(\lambda_i) &= \beta_0 + \mathbf{C}_i \boldsymbol{\beta}_c + \mathbf{A}_i \boldsymbol{\beta}_a \\ \boldsymbol{\beta}_c &= (\beta_{c1}, \dots, \beta_{cM})^T \\ \boldsymbol{\beta}_a &= (\beta_{a1}, \dots, \beta_{aN})^T \\ r_i &\sim \text{Pois}(\lambda_i) \end{aligned}$$

ここで、行列 A と C の i 行目の行ベクトルをそれぞれ A_i 、 C_i と表記した。 β_0 は切片項、パラメータ β_c は顧客の期待水準の高さ、 β_a は各担当者のスキルの高さを表し、事前分布として期待値0の正規分布に従うとして、その事後分布は後述のマルコフ連鎖モンテカルロ法(MCMC)を用いて推定する。

3.2 順序付きプロビットモデル

前項では顧客満足度 r がポアソン分布に従うと仮定したが、実際の顧客満足度の結果情報はアンケートによって取得されることが多いため、 $r \in \{1,2,3,4,5\}$ のような上限下限の定められた整数値として現れる。この場合、ポアソン分布を仮定するよりも、順序付きプロビットモデルを用いるほうが適していると考えられる。

順序付きプロビットモデルは、目的変数が離散値をとる場合のモデルとして用いられるロジットモデル・プロビットモデルの拡張であり、目的変数が「順序のある離散値」を取る場合のモデルである[12]。概念としては図4に示したように、正規分布を仮定した潜在変数 Y_i^* の値と閾値 m_i の比較により、目的変数 Y_i の値が決定づけられるとするモデルである。

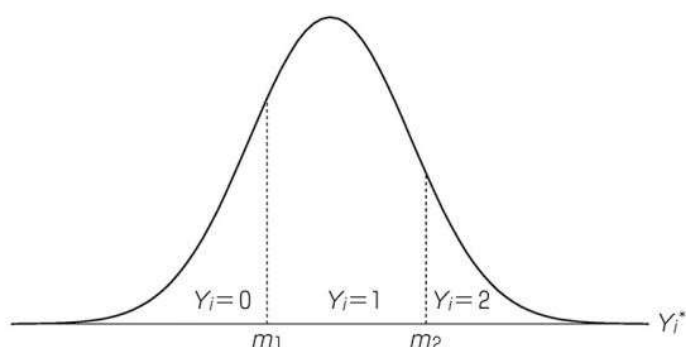


図4 順序付きプロビットモデルの概念図 [13]

ポアソン回帰モデルでは担当と顧客のダミー行列とパラメータからアンケート結果 r の分布が求められるが、順序付きプロビットモデルでは潜在変数 v_i が、

$$v_i = \beta_0 + C_i\beta_c + A_i\beta_a + u_i$$

$$u_i \sim \mathcal{N}(0,1)$$

で得られると仮定する。 u_i は標準正規分布に従う誤差項である。この潜在変数 v_i は顧客と担当の割当てによって決定される量であり、期待-不一致モデルの考えから案件の潜在的な評価を表すと言える。

ここで、潜在変数 v_i とアンケート結果 r_i の間に閾値 $t = \{t_1, t_2, t_3, t_4\}$ をもとにした以下の関係があると仮定する。

$$r_i = \begin{cases} 1 & \text{if } v_i \leq t_1, \\ 2 & \text{if } t_1 < v_i \leq t_2, \\ 3 & \text{if } t_2 < v_i \leq t_3, \\ 4 & \text{if } t_3 < v_i \leq t_4, \\ 5 & \text{if } t_4 < v_i \end{cases}$$

この仮定のもとで、 r_i がそれぞれの値を取る確率は、下記の通りとなる。

$$\begin{aligned} P(r_i = 1) &= P(v_i \leq t_1) = P(\beta_0 + \beta_c C^T + \beta_a A^T + u_i \leq t_1) \\ &= \Phi(t_1 - \beta_0 - \beta_c C^T - \beta_a A^T) \end{aligned}$$

$$\begin{aligned} P(r_i = 2) &= P(t_1 < \beta_0 + \beta_c C^T + \beta_a A^T + u_i \leq t_2) \\ &= \Phi(t_2 - \beta_0 - \beta_c C^T - \beta_a A^T) - \Phi(t_1 - \beta_0 - \beta_c C^T - \beta_a A^T) \end{aligned}$$

...

$$\begin{aligned} P(r_i = 5) &= P(t_4 < \beta_0 + \beta_c C^T + \beta_a A^T + u_i) \\ &= 1 - \Phi(t_4 - \beta_0 - \beta_c C^T - \beta_a A^T) \end{aligned}$$

ここで Φ は標準正規分布の分布関数(累積密度関数)である。本論文では、この順序付きプロビットモデルのパラメータについて、ポアソン回帰モデルと同様に事前分布として期待値 0 の正規分布を設定し、その事後分布をMCMCを用いて推定する。

3.3 マルコフ連鎖モンテカルロ法

マルコフ連鎖モンテカルロ(Markov Chain Monte Carlo, MCMC)法とは、マルコフ連鎖を利用して所望の確率分布に従う乱数(サンプル)を生成する手法である [14]。以下では、MCMC法の基本的な手法の一つであるメトロポリス・ヘイスティングス(Metropolis Hastings, MH)法を用いて、モデルに含まれるパラメータを近似推定する方法について説明する。

あるモデルに含まれるパラメータ θ について、データ D が与えられた元での事後分布の確率密度関数を $f(\theta|D)$ 、事前分布の確率密度関数を $f(\theta)$ 、尤度関数を $f(D|\theta)$ と表記する。ここで、ベイズの公式から、 θ の事後分布は以下で表せる。

$$f(\theta|D) \propto f(D|\theta)f(\theta) = \text{Kernel}(\theta)$$

これに対し、MH法では以下のステップを実施する。

1. パラメータ θ の初期サンプル $\hat{\theta}_1$ をランダムに定める。
2. 次のサンプル候補として、初期サンプルに正規分布に従う誤差を加えたサンプルを発生させる。具体的には $\hat{\theta}'_2 = \hat{\theta}_1 + \mathcal{N}(\epsilon|0, \sigma^2)$ に従うサンプル $\hat{\theta}'_2$ を生成する。 σ は任意の定数。
3. パラメータのサンプル $\hat{\theta}_1$ を使ったモデルの事後分布の確率密度を、ステップ2で生成したサンプル候補 $\hat{\theta}'_2$ を使ったモデルの事後分布の確率密度で割り、比率 $rate$ を計算する。ここでそれぞれの事後分布は「事後分布 = 事前分布 * 尤度 / 周辺尤度」の形であるため、お互いの周辺尤度が打ち消しあい、下記のようにKernelのみを用いて計算できる：

$$rate = \frac{\text{Kernel}(\hat{\theta}'_2)}{\text{Kernel}(\hat{\theta}'_1)}$$

4. ステップ 3 で計算された $rate$ の値によって、次のサンプルを以下のように決定する。
【 $rate > 1$ 】であれば、 $\hat{\theta}'_2$ を用いた事後分布の方が大きく、パラメータとして適していると言えるため、次のサンプルとしては $\hat{\theta}'_2$ を採用する。
【 $rate < 1$ 】であれば、 $\hat{\theta}'_1$ の方がパラメータとして適していると言えるが、この時も確率 $rate$ で $\hat{\theta}'_2$ を次のサンプルとして採用し、確率 $(1 - rate)$ で $\hat{\theta}'_1$ を採用する。なお $\hat{\theta}'_1$ が採用された場合は、サンプルの値は変化しない。このように、過去1時点前の値に応じて次の値の確率分布が決まる確率過程をマルコフ連鎖と呼ぶ。
5. 以降、あらかじめ決められたサイズのサンプルが得られるまでステップ 2 からステップ 4 を繰り返す。ここでステップ 4 のルールにより定義されるマルコフ連鎖はサンプル生成を繰り返すことにより、定常分布と呼ばれる確率分布に収束し、それが事後分布 $f(\theta|D)$ と一致することが知られている。
6. パラメータ $\hat{\theta}$ の決められたサイズのサンプルが得られた後、サンプルの前半の所定数を定常分布への収束までの経過として破棄し、後半の収束後のサンプル $\hat{\theta}$ で事後分布を近似する。例えば、収束後のサンプル $\hat{\theta}$ の平均値がパラメータ θ の近似的な推定値となる。

このようにMH法では、解析的には計算できないパラメータの事後分布について、所望の確率分布が定常分布となるようなマルコフ連鎖によって生成したサンプルで、分布を近似する。

なおMH法では、サンプル生成に使用する乱数の分散によっては収束まで長い時間がかかる場合があるという欠点がある。本論文の実証分析で使用した確率的プログラミング言語「Stan」では、これらの課題を解決するため、マルコフ連鎖における次ステップのサンプル候補をランダムに生成するのではなく、物理学におけるハミルトン系の数値計算を利用してほぼ確実に採択されるサンプルを生成するハミルトニアン・モンテカルロ法Hamiltonian Monte Carlo: HMC)の1実装であるNUTS (No-U-Turn Sampler) [15]が用いられている。

4 階層ベイズモデルの逐次近似推定

本章では、顧客満足度に関する期待-不一致モデルと、ゲームにおけるマッチングアルゴリズム TrueSkill [7]を参考に、担当者スキル・顧客期待水準・顧客満足度を階層ベイズモデルで表し、逐次的に上記パラメータを近似推定する手法について説明する。

4.1 提案モデル

前章で提案したポアソン回帰モデル・順序付きプロビットモデルのパラメータの推定では MCMC を用いていたため、一回の推定に多くの計算時間を要するという問題があった。そのため、1 つの案件が完了するごとに担当者のスキルや顧客の期待水準の推定を更新するといった運用が困難である。加えて、現実の担当者のスキルや顧客の期待水準は時間とともに変動すると考えられるが、これらのモデルではスキルや期待水準を一定と仮定しているため、その変動を捉えるためにはデータを一定期間ごとに分割して推定を繰り返すことが必要となる。

これらの問題を解決するため本論文では、MCMC により全データを一括で用いてパラメータを推定するのではなく、顧客満足度研究の代表的なモデル「期待-不一致モデル」と、ゲームのスキルレーティングアルゴリズム「TrueSkill」を組み合わせた階層ベイズモデルを用いて、1案件ごとにパラメータ(担当者スキル・顧客期待水準)の事後分布を近似推定する逐次的な手法を提案する。図 5 に提案モデルのグラフィカルモデルを示す。

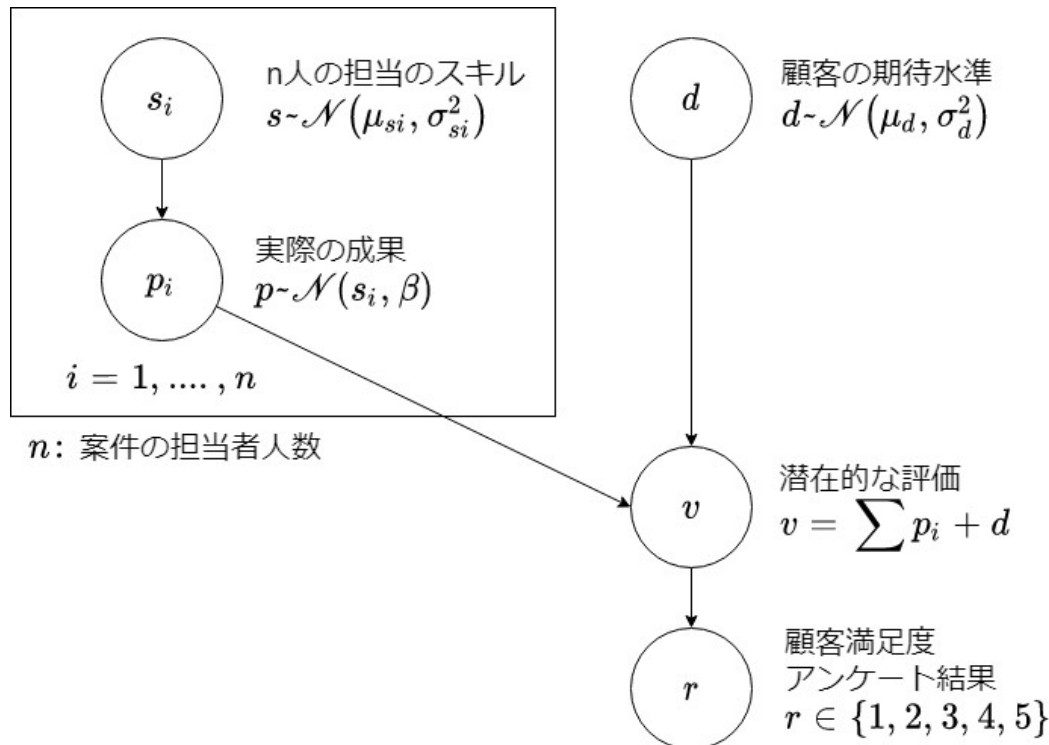


図 5 提案モデル

このモデルでは、ある案件の担当者*i*のスキルを s_i 、成果を p_i 、顧客の期待水準を d 、案件の潜在的な評価を v 、最終的な顧客満足度アンケート結果を r とする。スキルは担当者ごとに個別の平均 μ_{s_i} と分散 $\sigma_{s_i}^2$ をもつ正規分布、顧客期待水準も同様に平均 μ_d 、分散 σ_d^2 を持つ正規分布に従うものとする。成果は担当者スキルを平均として分散 β をもつ正規分布に従うとする。ここで成果の分散について、どの担当者・どの案件でも同様の業務を行うため、スキルから成果へのばらつきは一定値であると仮定した。

以上から、ある1案件についての担当者スキル、成果、顧客の期待水準は下記で表される。

$$\begin{aligned} P(s_i) &= \mathcal{N}(\mu_{s_i}, \sigma_{s_i}^2), \\ P(p_i|s_i) &= \mathcal{N}(s_i, \beta), \\ P(d) &= \mathcal{N}(\mu_d, \sigma_d^2). \end{aligned}$$

また、全担当者のスキルと顧客の期待水準をまとめてベクトル $\tilde{\mathbf{s}} = (s_1, \dots, s_n, d)^T$ で表す。このとき、 $\tilde{\mathbf{s}}$ は以下の多変量正規分布に従う：

$$\begin{aligned} P(\tilde{\mathbf{s}}) &= \mathcal{N}(\tilde{\mathbf{s}}|\tilde{\boldsymbol{\mu}}_s, \Sigma_s) = \mathcal{N}(\tilde{\mathbf{s}}|\tilde{\boldsymbol{\mu}}_s, \Lambda_s^{-1}) \quad (1) \\ \tilde{\boldsymbol{\mu}}_s &= [\mu_{s_1}, \dots, \mu_{s_n}, \mu_d]^T \\ \Sigma_s &= \text{diag}[\sigma_{s_1}^2, \dots, \sigma_{s_n}^2, \sigma_d^2] \\ \Lambda_s^{-1} &= \Sigma_s \end{aligned}$$

また、成果 p_i はスキルを平均とした正規分布に従うので、スキルで条件付けられていない成果の確率は、

$$P(p_i) = \mathcal{N}(p_i|\mu_{s_i}, \sigma_{s_i}^2 + \beta)$$

で表される。ここで期待-不一致モデルから、満足度は顧客の期待水準と担当者の成果の差で表現されるため、案件の潜在的な満足度、すなわち評価 v は成果の和と期待水準により以下で表される。

$$v = \sum_{i=1}^n p_i + d$$

ただし顧客期待水準は高ければ高いほど評価が下がるため、期待水準が高いほど d は負の方向に大きな値を持つこととなる。上記より複数の担当者による案件の潜在的な評価の確率分布 $P(v)$ は以下で表される。

$$\begin{aligned} P(v) &= \mathcal{N}(v|\mu_{s_1} + \dots + \mu_{s_n} + \mu_d, (\sigma_{s_1}^2 + \beta) + (\sigma_{s_2}^2 + \beta) + \dots + (\sigma_{s_n}^2 + \beta) + \sigma_d^2) \\ &= \mathcal{N}(v|\mathbf{1}^T \tilde{\boldsymbol{\mu}}_s, \mathbf{1}^T \Sigma_s \mathbf{1} + n\beta) \\ &= \mathcal{N}(v|\mu_v, \sigma_v^2) \\ &= \mathcal{N}(v|\mu_v, \lambda_v^{-1}) \end{aligned}$$

ただし上式において以下を用いた。

$$\sigma_v^2 = \mathbf{1}^T \Sigma_s \mathbf{1} + n\beta = \lambda_v^{-1}$$

$$\begin{aligned}\mu_v &= \mathbf{1}^T \tilde{\boldsymbol{\mu}}_s \\ \mathbf{1} &= [1, \dots, 1]^T\end{aligned}$$

一方で、 \boldsymbol{s} が与えられたもとでの評価の条件付き確率 $P(v|\boldsymbol{s})$ は、 \boldsymbol{s} を用いて以下で表すことができる。

$$P(v|\boldsymbol{s}) = \mathcal{N}(v|\mathbf{1}^T \boldsymbol{s}, n\beta) \quad (2)$$

この評価 v から、最終的な顧客満足度アンケート結果 r が得られるとする。

アンケート結果を正規分布と仮定した推定

ここで、一旦簡単のため顧客満足度アンケート結果 r が正規分布に従うと仮定し、実際に得られた r から担当者スキルと顧客期待水準を推定する。まず仮定により r は以下のように v を平均、分散を定数 γ とした正規分布で表される。

$$P(r) = \mathcal{N}(r|v, \gamma)$$

また \boldsymbol{s} が与えられたもとでの条件付き確率 $P(r|\boldsymbol{s})$ は以下の通りとなる。

$$\begin{aligned}P(r|\boldsymbol{s}) &= \mathcal{N}(r|\mathbf{1}^T \boldsymbol{s}, n\beta + \gamma) \\ &= \mathcal{N}(r|\mathbf{1}^T \boldsymbol{s}, \lambda_r^{-1}) \\ \lambda_r^{-1} &= n\beta + \gamma\end{aligned}$$

ベイズの公式より、顧客満足度 r が与えられた元でのパラメータ \boldsymbol{s} の事後分布は、以下で与えられる。

$$P(\boldsymbol{s}|r) \propto P(r|\boldsymbol{s})P(\boldsymbol{s}) = \mathcal{N}(r|\mathbf{1}^T \boldsymbol{s}, \lambda_r^{-1})\mathcal{N}(\boldsymbol{s}|\tilde{\boldsymbol{\mu}}_s, \Lambda_s^{-1})$$

上記の両辺の対数を取り、 \boldsymbol{s} について整理する。

$$\begin{aligned}\ln P(\boldsymbol{s}|r) &= \ln(\mathcal{N}(r|\mathbf{1}^T \boldsymbol{s}, \lambda_r^{-1})) + \ln \mathcal{N}(\boldsymbol{s}|\tilde{\boldsymbol{\mu}}_s, \Lambda_s^{-1}) + \text{const.} \\ &= -\frac{1}{2}\{(r - \mathbf{1}^T \boldsymbol{s})^2 \lambda_r\} - \frac{1}{2}\{(\boldsymbol{s} - \tilde{\boldsymbol{\mu}}_s)^T \Lambda_s (\boldsymbol{s} - \tilde{\boldsymbol{\mu}}_s)\} + \text{const.} \\ &= -\frac{1}{2}\{-2r\mathbf{1}^T \boldsymbol{s} \lambda_r + \boldsymbol{s}^T \mathbf{1}\mathbf{1}^T \boldsymbol{s} \lambda_r + \boldsymbol{s}^T \Lambda_s \boldsymbol{s} - 2\boldsymbol{s}^T \Lambda_s \tilde{\boldsymbol{\mu}}_s\} + \text{const.}\end{aligned}$$

上記において const は \boldsymbol{s} に関係の無い定数項をまとめたものを示す。これをさらに整理して、

$$\ln P(\boldsymbol{s}|r) = -\frac{1}{2}\{\boldsymbol{s}^T (\mathbf{1}\mathbf{1}^T \lambda_r + \Lambda_s) \boldsymbol{s} - 2\boldsymbol{s}^T (\lambda_r r \mathbf{1} + \Lambda_s \tilde{\boldsymbol{\mu}}_s)\} + \text{const.} \quad (3)$$

を得る。以上から、アンケート結果 r を正規分布に従うと仮定した場合の担当者スキルと顧客期待水準の事後平均と事後分散を更新する式として、

$$\begin{aligned}P(\boldsymbol{s}|r) &= \mathcal{N}(\boldsymbol{s}|\hat{\boldsymbol{\mu}}_s, \hat{\Lambda}_s^{-1}) \\ \hat{\Lambda}_s &= \mathbf{1}\mathbf{1}^T \lambda_r + \Lambda_s \\ \hat{\boldsymbol{\mu}}_s &= \hat{\Lambda}_s^{-1}(\lambda_r r \mathbf{1} + \Lambda_s \tilde{\boldsymbol{\mu}}_s)\end{aligned}$$

が得られる。

アンケートを順序付き離散値と仮定

前節ではアンケート結果 r について評価 v を平均とした正規分布に従うと仮定したが、実際のアンケートでは r は $\{1,2,3,4,5\}$ といった上限下限のある順序付きの離散値で得られることが多いため、 r は v を潜在変数とした順序付きプロビットモデルに従うと考えた方が適当である。順序付きプロビットモデルにおいては、 v を潜在変数として r が得られる確率は $P(v)$ の累積分布関数によって、

$$P(r = k) = \int_{t_k}^{t_{k+1}} P(v)dv$$

で表される。なお t_k は $r = k$ となる閾値であり、 $t_1 = -\infty, t_6 = +\infty$ である。

ここで推定したいパラメータ \mathfrak{s} と v の事後分布 $P(\mathfrak{s}, v|r)$ は、アンケートデータ r に対する v の事後分布 $P(v|r)$ と、 v に対する \mathfrak{s} の事後分布 $P(\mathfrak{s}|v)$ に分解できる：

$$P(\mathfrak{s}, v|r) = P(\mathfrak{s}|v)P(v|r). \quad (4)$$

ここで式(4)の右辺を構成する二つの事後分布は以下のように求められる。

評価の条件付きスキル事後分布の計算

式(4)の右辺の $P(\mathfrak{s}|v)$ 、すなわち評価 v で条件づけられたスキル \mathfrak{s} の事後分布は、ベイズの定理より、 \mathfrak{s} で条件づけられた v の条件付き分布 $P(v|\mathfrak{s})$ と、 \mathfrak{s} の事前分布 $P(\mathfrak{s})$ によって下記で表される

$$P(\mathfrak{s}|v) \propto P(v|\mathfrak{s})P(\mathfrak{s})$$

ここで式(1)と式2)から、

$$P(\mathfrak{s}|v) \propto \mathcal{N}(v|\mathbf{1}^T \mathfrak{s}, n\beta) \mathcal{N}(\mathfrak{s}|\tilde{\boldsymbol{\mu}}_s, \Lambda_s^{-1})$$

を得る。ここで両辺の対数を取って \mathfrak{s} を含まない項を定数として整理すると、以下が得られる：

$$\begin{aligned} \ln P(\mathfrak{s}|v) &= -\frac{1}{2} \{ (v - \mathbf{1}^T \mathfrak{s})^2 (n\beta)^{-1} + (\mathfrak{s} - \tilde{\boldsymbol{\mu}}_s)^T \Lambda_s (\mathfrak{s} - \tilde{\boldsymbol{\mu}}_s) \} + \text{const.} \\ &= -\frac{1}{2} \{ -2v \tilde{\boldsymbol{\mu}}_s^T \mathbf{1} (n\beta)^{-1} + (n\beta)^{-1} \tilde{\boldsymbol{\mu}}_s^T \mathbf{1} \mathbf{1}^T \mathfrak{s} + \mathfrak{s}^T \Lambda_s \mathfrak{s} - 2 \tilde{\boldsymbol{\mu}}_s^T \Lambda_s \mathfrak{s} \} + \text{const.} \end{aligned}$$

これをさらに整理して、

$$\ln P(\mathfrak{s}|v) = -\frac{1}{2} \{ \mathfrak{s}^T (\mathbf{1} \mathbf{1}^T (n\beta)^{-1} + \Lambda_s) \mathfrak{s} - 2 \tilde{\boldsymbol{\mu}}_s^T (v \mathbf{1} (n\beta)^{-1} + \Lambda_s \tilde{\boldsymbol{\mu}}_s) \} + \text{const.} \quad (5)$$

を得る。したがって、潜在的な評価 v が与えられたもとでのスキルの期待水準 \mathfrak{s} の事後分布 $P(\mathfrak{s}|v)$ として、

$$P(\mathfrak{s}|v) = \mathcal{N}(\mathfrak{s}|\tilde{\boldsymbol{\mu}}'_s, \Lambda'_s{}^{-1})$$

$$\begin{aligned}\Lambda'_s &= \mathbf{1}\mathbf{1}^T(n\beta)^{-1} + \Lambda_s \\ \tilde{\boldsymbol{\mu}}'_s &= \Lambda'_s{}^{-1}(v\mathbf{1}(n\beta)^{-1} + \Lambda_s\tilde{\boldsymbol{\mu}}_s)\end{aligned}\quad (6)$$

が得られる。

顧客満足度の条件付き評価の計算

次に式(4)の右辺 $P(v|r)$ は、データ r が与えられた状態での潜在評価の事後分布 $P(v)$ であり、これは v が $r = k$ となる閾値 $t_k \leq v < t_{k+1}$ の範囲のいずれかの値を確率的にとる、すなわち確率分布 $P(v)$ を閾値 $t_k \leq v < t_{k+1}$ で切断した確率分布(切断正規分布)であるといえる。一般の切断正規分布は、

$$f(x; \mu, \sigma, a, b) = \frac{1}{\sigma} \frac{\phi\left(\frac{x-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)}$$

で表される。ここで ϕ は標準正規分布の密度関数、 Φ は標準正規分布の累積分布関数を表す。これを $P(v)$ に当てはめると $P(v|r)$ は、

$$\begin{aligned}P(v) &= \mathcal{N}(v|\mu_v, \sigma_v^2) \\ P(v|r = k) &= \frac{1}{\sigma_v} \frac{\Phi\left(\frac{v-\mu_v}{\sigma_v}\right)}{\Phi\left(\frac{t_{k+1}-\mu_v}{\sigma_v}\right) - \Phi\left(\frac{t_k-\mu_v}{\sigma_v}\right)}\end{aligned}\quad (7)$$

で表される。

アンケートデータ r が得られたもとのスキル $\tilde{\mathbf{s}}$ の事後分布 $P(\tilde{\mathbf{s}}|r)$ は式(4)を v について積分することで得られる:

$$P(\tilde{\mathbf{s}}|r) = \int P(\tilde{\mathbf{s}}|v)P(v|r)dv.$$

しかし、右辺は上記のとおり正規分布と切断正規分布の積となるため、この積分を解析的に求めることはできない。そこで本論文では確率分布の近似手法であるモーメントマッチング法を用いることとする。

モーメントマッチング法

モーメントマッチング法は、ある複雑な確率分布 $p(z)$ を、より簡単な分布 $q(z)$ を用いて近似する方法である [16]。近似分布 $q(z)$ を、下記のような指数型分布族とする:

$$q(z; \eta) = h(z)\exp(\eta^T t(z) - a(\eta)).$$

さらに近似したい目標分布 $p(z)$ と近似分布 $q(z)$ を近づけるため、これらの間の Kullback-Leibler ダイバージェンス:

$$D_{\text{KL}} [p(z)||q(z; \eta)] = -E_p[\ln q(z; \eta)] + E_p[\ln p(z)]$$

を自然パラメータ η に関して最小化する。上式を展開して η に関して整理すると、

$$D_{KL} [p(z)||q(z; \eta)] = -\eta^T E_p[t(z)] + a(\eta) + \text{const.}$$

となる。 const は η に関係の無い定数項である。上式の η に対する勾配を計算しゼロとおくと、

$$\nabla_{\eta} a(\eta) = E_p[t(z)]$$

を得る。ここで指数型分布族において $a(\eta)$ の勾配は十分統計量の期待値と一致することから、

$$E_q[t(z)] = E_p[t(z)]$$

が求まる。よって指数型分布族を用いて分布 $p(z)$ を近似するためには、 $p(z)$ の十分統計量の期待値 $E_p[t(z)]$ 、すなわち $p(z)$ のモーメントを計算し、その結果を使って $q(z)$ のパラメータを決定すれば良い。

スキル・期待水準の近似事後分布導出

本論文では上記のモーメントマッチング法を用いて、スキル・顧客期待水準の事後分布 $P(\tilde{s}|r)$ を、独立な正規分布 $q(\tilde{s})$ で近似する:

$$P(\tilde{s}|r) \approx q(\tilde{s}) = \prod_{i=1}^{n+1} \mathcal{N}(\tilde{s}_i | \hat{\mu}_i, \hat{\sigma}_i^2)$$

事後分布 $P(\tilde{s}|r)$ の期待値 $E[\tilde{s}|r]$ は $P(\tilde{s}|r)$ を v について周辺化したものの期待値であるから、以下で表される。

$$\begin{aligned} E[\tilde{s}|r] &= \int \int \tilde{s} P(\tilde{s}|v) P(v|r) dv d\tilde{s} \\ &= \int \left(\int \tilde{s} P(\tilde{s}|v) d\tilde{s} \right) P(v|r) dv \end{aligned}$$

ここで、上式括弧内は正規分布 $P(\tilde{s}|v)$ の期待値であるため、式(6)より、

$$\begin{aligned} E[\tilde{s}|r] &= \int \tilde{\mu}'_s P(v|r) dv \\ &= \int (\Lambda'_s{}^{-1} (v \mathbf{1} (n\beta)^{-1} + \Lambda_s \tilde{\mu}_s)) P(v|r) dv \\ &= \Lambda'_s{}^{-1} \Lambda_s \tilde{\mu}_s + \Lambda'_s{}^{-1} \mathbf{1} (n\beta)^{-1} \left(\int v P(v|r) dv \right) \end{aligned}$$

を得る。ここで上式の括弧内の積分は式(7)で表される切断正規分布 $P(v|r)$ の期待値である。一般に期待値 μ 、分散 σ^2 の正規分布を区間 (a, b) で切断した切断正規分布の期待値は、

$$\mu + \sigma \frac{\phi\left(\frac{a-\mu}{\sigma}\right) - \phi\left(\frac{b-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)}$$

で与えられる [17]。ここから $P(v|r = k)$ の期待値は閾値 $t_k \leq v < t_{k+1}$ を用いて、

$$\begin{aligned} \left(\int v P(v|r) dv \right) &= E(v|r = k) \\ &= \mu_v + \sigma_v \frac{\phi\left(\frac{t_k - \mu_v}{\sigma_v}\right) - \phi\left(\frac{t_{k+1} - \mu_v}{\sigma_v}\right)}{\Phi\left(\frac{t_{k+1} - \mu_v}{\sigma_v}\right) - \Phi\left(\frac{t_k - \mu_v}{\sigma_v}\right)} \end{aligned}$$

で得られる。これを元の式に代入すると、担当者スキル・顧客期待水準の事後分布の期待値は、

$$E[\tilde{\mathbf{s}}|r = k] = \Lambda_s'^{-1} \Lambda_s \tilde{\boldsymbol{\mu}}_s + \Lambda_s'^{-1} \mathbf{1} (n\beta)^{-1} \left(\mu_v + \sigma_v \frac{\phi\left(\frac{t_k - \mu_v}{\sigma_v}\right) - \phi\left(\frac{t_{k+1} - \mu_v}{\sigma_v}\right)}{\Phi\left(\frac{t_{k+1} - \mu_v}{\sigma_v}\right) - \Phi\left(\frac{t_k - \mu_v}{\sigma_v}\right)} \right)$$

で得られる。

次に、担当者スキル・顧客期待水準の事後分布の分散共分散行列を導出する。定義により分散共分散行列は、

$$\begin{aligned} \text{Var}[\tilde{\mathbf{s}}|r] &= E[(\tilde{\mathbf{s}} - E[\tilde{\mathbf{s}}|r])(\tilde{\mathbf{s}} - E[\tilde{\mathbf{s}}|r])^T] \\ &= \int \int (\tilde{\mathbf{s}} - E[\tilde{\mathbf{s}}|r])(\tilde{\mathbf{s}} - E[\tilde{\mathbf{s}}|r])^T P(\tilde{\mathbf{s}}|v) P(v|r) d\tilde{\mathbf{s}} dv \\ &= \int \left(\int (\tilde{\mathbf{s}} - E[\tilde{\mathbf{s}}|r])(\tilde{\mathbf{s}} - E[\tilde{\mathbf{s}}|r])^T P(\tilde{\mathbf{s}}|v) d\tilde{\mathbf{s}} \right) P(v|r) dv \end{aligned}$$

で与えられる。ここで上式の括弧内の積分は $P(\tilde{\mathbf{s}}|v)$ の分散共分散行列であるため、式(6)より $\Lambda_s'^{-1}$ で置き換えると、

$$\begin{aligned} \text{Var}[\tilde{\mathbf{s}}|r] &= \int \Lambda_s'^{-1} P(v|r) dv \\ &= \int (\mathbf{1}\mathbf{1}^T (n\beta)^{-1} + \Lambda_s)^{-1} P(v|r) dv \\ &= (\mathbf{1}\mathbf{1}^T (n\beta)^{-1} + \Lambda_s)^{-1} \int P(v|r) dv \\ &= (\mathbf{1}\mathbf{1}^T (n\beta)^{-1} + \Lambda_s)^{-1} \end{aligned}$$

が得られる。以上により、顧客満足度アンケートデータ r が与えられた元での担当者スキル・顧客期待水準の事後分布 $P(\tilde{\mathbf{s}}|r)$ の以下の近似が得られた：

$$P(\tilde{\mathbf{s}}|r = k) \approx \prod_{i=1}^{n+1} \mathcal{N}(\tilde{s}_i | \hat{\mu}_i, \hat{\sigma}_i^2)$$

$$\hat{\mu}_i = \left\{ \Lambda_s'^{-1} \Lambda_s \tilde{\mu}_s + \Lambda_s'^{-1} \mathbf{1} (n\beta)^{-1} \left(\mu_v + \sigma_v \frac{\Phi\left(\frac{t_k - \mu_v}{\sigma_v}\right) - \Phi\left(\frac{t_{k+1} - \mu_v}{\sigma_v}\right)}{\Phi\left(\frac{t_{k+1} - \mu_v}{\sigma_v}\right) - \Phi\left(\frac{t_k - \mu_v}{\sigma_v}\right)} \right) \right\}_i$$

$$\hat{\sigma}_i^2 = \{(\mathbf{1}\mathbf{1}^T (n\beta)^{-1} + \Lambda_s)^{-1}\}_{ii}.$$

また逐次的な推定では案件1件ごとにスキル・期待水準の事後分布を更新し、これを事前分布として次の案件を用いた推定を行うこととする。

4.2 担当者の成果の平均化

前節まで担当者が複数の案件について、チームのゲームパフォーマンスをメンバーの和として扱う TrueSkill に習い、案件の潜在的な評価は担当者の成果の和から得られるとしてきた。この仮定では案件の担当者数を増やせば増やすほど成果が増し評価が向上するということになるが、現実には担当者数が多い案件は作業規模が大きいいため、一人の担当者が案件の質や評価に及ぼす影響力は相対的に小さくなることも多い。これを考慮し、複数人担当案件の成果を各担当者の成果の和とするのではなく、平均と仮定してモデルを修正した。

修正モデルにおいて案件を n 人で担当した場合、評価 v は以下のように表される。

$$v = \frac{1}{n} \sum_{i=1}^n p_i + d$$

$$P(v) = \mathcal{N}\left(v \mid \frac{1}{n}(\mu_{s1} + \dots + \mu_{sn}) + \mu_d, \frac{1}{n^2}((\sigma_{s1}^2 + \beta) + (\sigma_{s2}^2 + \beta) + \dots + (\sigma_{sn}^2 + \beta)) + \sigma_d^2\right) \quad (8)$$

$$= \mathcal{N}(v \mid \mathbf{E}^T \tilde{\mu}_s, \mathbf{E}^T \Sigma_s \mathbf{E} + \beta')$$

$$= \mathcal{N}(v \mid \mu_v, \sigma_v^2) = \mathcal{N}(v \mid \mu_v, \lambda_v^{-1})$$

ただし上式において、

$$\mathbf{E} = \left(\frac{1}{n}, \dots, \frac{1}{n}, 1\right)^T \left[n \text{個の } \frac{1}{n} \text{と、1つの } 1 \text{を要素にもつベクトル} \right]$$

$$\beta' = \frac{\beta}{n}$$

とした。なお \mathbf{E} において顧客期待水準にかかる項は平均を取らないので人数 n で割らず1としている。また前項までに導出した式において $\mathbf{1}$ を \mathbf{E} 、 β を β' とそれぞれ置き換えることで、この修正モデルにおいても全て同様の式が成立する。

4.3 実データによる閾値の最尤推定

提案モデルでは、潜在的な評価 v は閾値 t によって離散値のアンケート回答 r に変換されるとしたが、推定にあたってはこの閾値を何らかの定数に設定する必要がある。本論文ではこの閾値の設定方法として、アンケート回答の実データを用いた最尤推定を行う。

いま担当者スキルと顧客期待水準の事前分布をそれぞれ $s_i \sim \mathcal{N}(s_i|0,1)$; $d \sim \mathcal{N}(d|0,1)$ と仮定すると、ある案件 i の評価の従う確率分布は式(8)より、

$$\begin{aligned} P(v_i) &= \mathcal{N}\left(v_i \mid 0, \frac{1}{n^2}(n + n\beta) + 1\right) \\ &= \mathcal{N}(v_i \mid 0, \sigma_{vi}^2) \end{aligned}$$

で表せる。ただし、分散 σ_{vi} は案件の担当者人数によって変動する。この評価 v_i からアンケートの結果 r_i が得られる尤度は、順序付きプロビットモデルによって以下で表される。

$$\ell_i(\{t_k\}_{k=0}^5) = P(r_i) = \int_{t_{r_i-1}}^{t_{r_i}} \mathcal{N}(v_i \mid 0, \sigma_{vi}^2) dv_i$$

ここで t_0, t_1, \dots, t_5 はアンケート回答の閾値を表しており、 $t_{k-1} \leq v_i < t_k$ のとき $r_i = k$ を満たす。また、 $t_0 = -\infty, t_5 = +\infty$ とする。したがってデータ全体についての閾値の対数尤度は、

$$\begin{aligned} \mathcal{L}(\{t_i\}_{k=0}^5) &= \sum_i \ln \int_{t_{r_i-1}}^{t_{r_i}} \mathcal{N}(v_i \mid 0, \sigma_{vi}^2) dv_i \\ &= \sum_{k=1}^5 \sum_{i \in \{j \mid r_j = k\}} \ln \int_{t_{k-1}}^{t_k} \mathcal{N}(v_i \mid 0, \sigma_{vi}^2) dv_i \end{aligned}$$

で表される。この対数尤度を最大化するように、 t_1, t_2, t_3, t_4 について最適化を行うことで、仮定した事前分布と実データから評価をアンケート結果に変換するための閾値を推定することができる。なお、この最適化の解析解を求めることは困難であるため、実際の推定では数値的に最適化を行った。

5 実証分析

本章ではポアソン回帰モデル・順序付きプロビットモデル・提案モデルの3つのモデルを用いて、実際の顧客満足度アンケート結果から担当者のスキルと顧客の期待水準を推定する。さらに推定したスキルと期待水準を元に、案件の顧客満足度アンケート結果を予測し、実際の結果と比較することで各提案モデルの精度を評価する。

5.1 データの説明と推定の概略

実証分析では、ネットリサーチ会社の株式会社マクロミルの案件顧客満足度アンケートを利用した。具体的には、2018年1月から12月までの間に株式会社マクロミルで受注し顧客に納品したネットリサーチ案件について、リサーチ結果の納品後に顧客に対して依頼した満足度アンケートに回答があったもののうち、一部を抜粋したデータを対象としている。サンプルサイズは1021件であり、データには表1に示したイメージのとおり、案件の顧客満足度、担当者名、顧客IDが含まれている。担当者と顧客の列の値は、各案件に関与する場合は1、それ以外は0である。満足度列は「今回ご利用いただいたサービス全般について、お客様の満足度をお応えください」という質問に対する「5:満足,4:やや満足,3:どちらともいえない,2:やや不満,1:不満」という顧客の回答値が入力されている。

表1 顧客満足度アンケートのイメージ

案件ID	満足度	担当者1	担当者2	担当者3	担当者4	担当者5	担当者6	担当者7	中略	顧客1	顧客2	顧客3	顧客4	後略
PJ-1	5	1	1	1	0	0	0	0	...	1	0	0	0	...
PJ-2	5	0	0	0	1	1	0	0	...	0	1	0	0	...
PJ-3	4	0	0	0	0	0	1	1	...	0	0	1	0	...
...

また、この顧客満足度アンケートの満足度の分布を図6に示す。今回のデータは全体からの抜粋であり、また満足度1と2はごく稀にしか現れないため、全体の60%を満足度5が占めるという偏った分布となっている。なおアンケートの回答は任意であるため、アンケートに回答がある時点で案件に対して好意的なバイアスがかかると仮定すると、その上で満足度3が得られた場合は相対的に「かなり低い満足度」と捉えるのが妥当と考えられる。

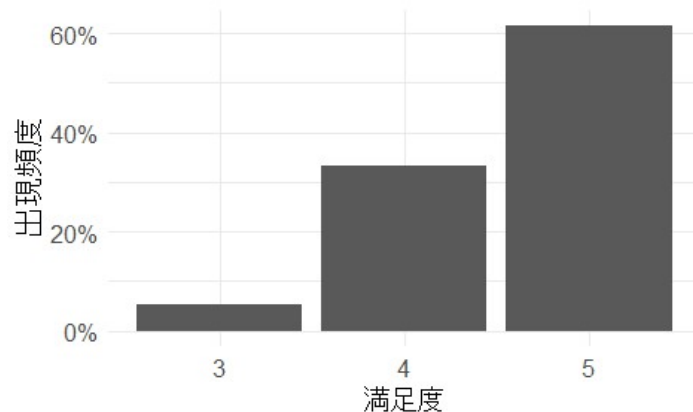


図 6 実データの満足度の分布

このデータについて、前章までに提案した「ポアソン回帰モデル」「順序プロビットモデル」「提案モデル」の3モデルを用いて推定と予測精度の検証を行った。手順としては、全データ 1021 件のうち最初の 500 件を推定用データとし、残りの 521 件を予測精度検証用のテストデータとして用いた。ただし提案モデルについては逐次的な推定となるため、推定用データで閾値の最尤推定と担当者スキルと顧客期待水準の事後分布の推定を行ったのち、予測精度検証においても案件 1 件ごとに事後分布を更新し、1 件前までの推定結果を用いた顧客満足度予測を実際の結果と比較した。またいずれのモデルも担当者スキル・顧客期待水準の事前分布は $\mathcal{N}(0,1)$ に統一した。なお、実際のデータでは案件ごとに受注価格が異なるが、本論文に用いられたデータにおいては事前分析により価格の満足度に及ぼす影響は担当者や顧客に比べて小さいことが確認されている。

5.2 ポアソン回帰モデル

まずポアソン回帰モデルを用いて実際の顧客満足度アンケートデータを元に MCMC によるパラメータの推定を行った。MCMC で生成する系列の長さは 2000 とし、独立な 2 つの系列でサンプルを生成した。各系列について前半の 1000 点は収束前として破棄するため、各パラメータにつき 2000 サンプルを生成した。MCMC で生成されたサンプルの診断に一般的に用いられる有効サンプルサイズ比率と \hat{R} を確認したところ、全てのパラメータについて有効サンプルサイズ比率は 0.5 以上、 \hat{R} は 1.05 以下と、いずれも問題がない水準であることを確認した。なおサンプリングに要した時間は CPU (Interl® Core™ i5-7300U 2.60GHz)、メモリ 8.00GB の PC で約 1 時間であった。

この推定により得られた担当者スキルと顧客期待水準の事後平均の分布を図 7 に示す。

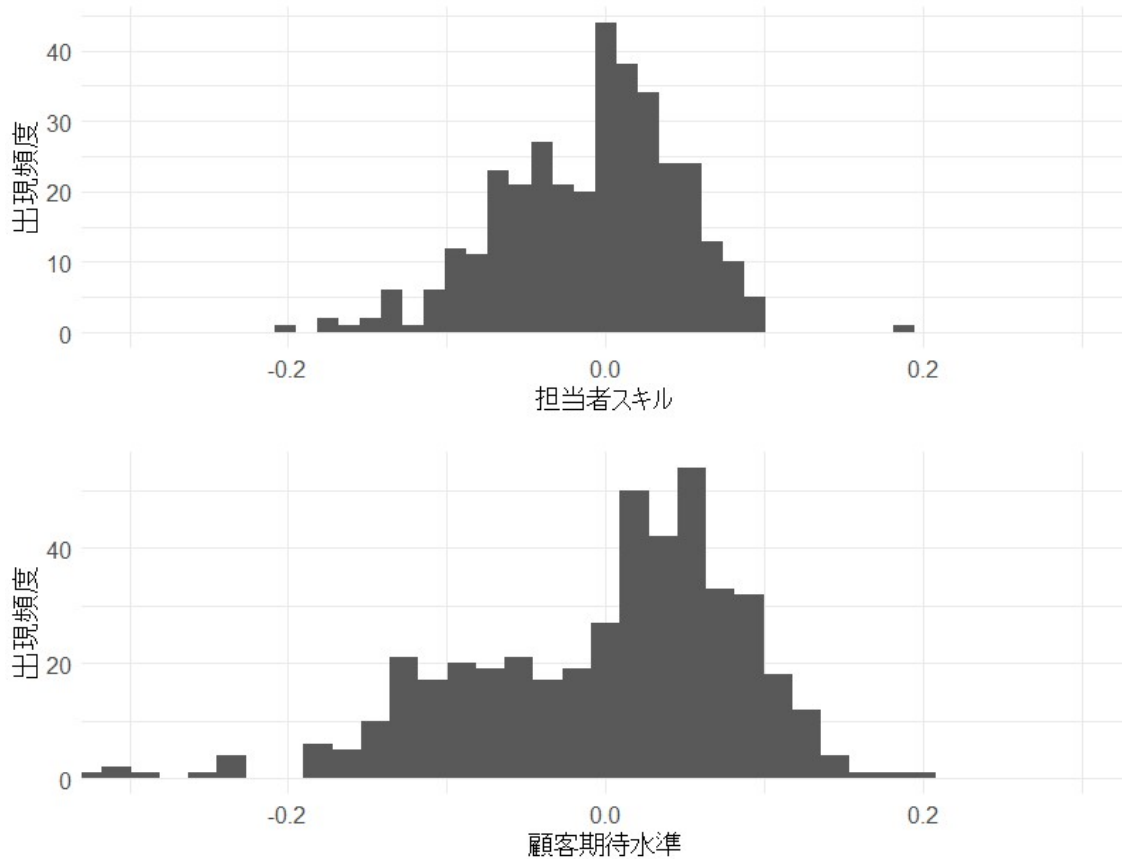


図7 ポアソン回帰モデルの推定結果

5.3 順序付きプロビットモデル

次に順序付きプロビットモデルを用いて、顧客満足度アンケートデータを元にMCMCによるパラメータの推定を行った。ポアソン回帰モデルの推定と同様の設定で、MCMCで各パラメータにつき2000サンプルを生成した。また、MCMCで得られたサンプルは、全てのパラメータについて有効サンプルサイズ比率は0.5以上、 \hat{R} は1.05以下と、いずれも問題がない水準であることを確認した。なおサンプリングに要した時間は、ポアソン回帰モデルで1時間だったところ、順序付きプロビットモデルでは約10時間であった。

この推定により得られた担当者スキルと顧客期待水準の事後平均の分布を図8に示す。

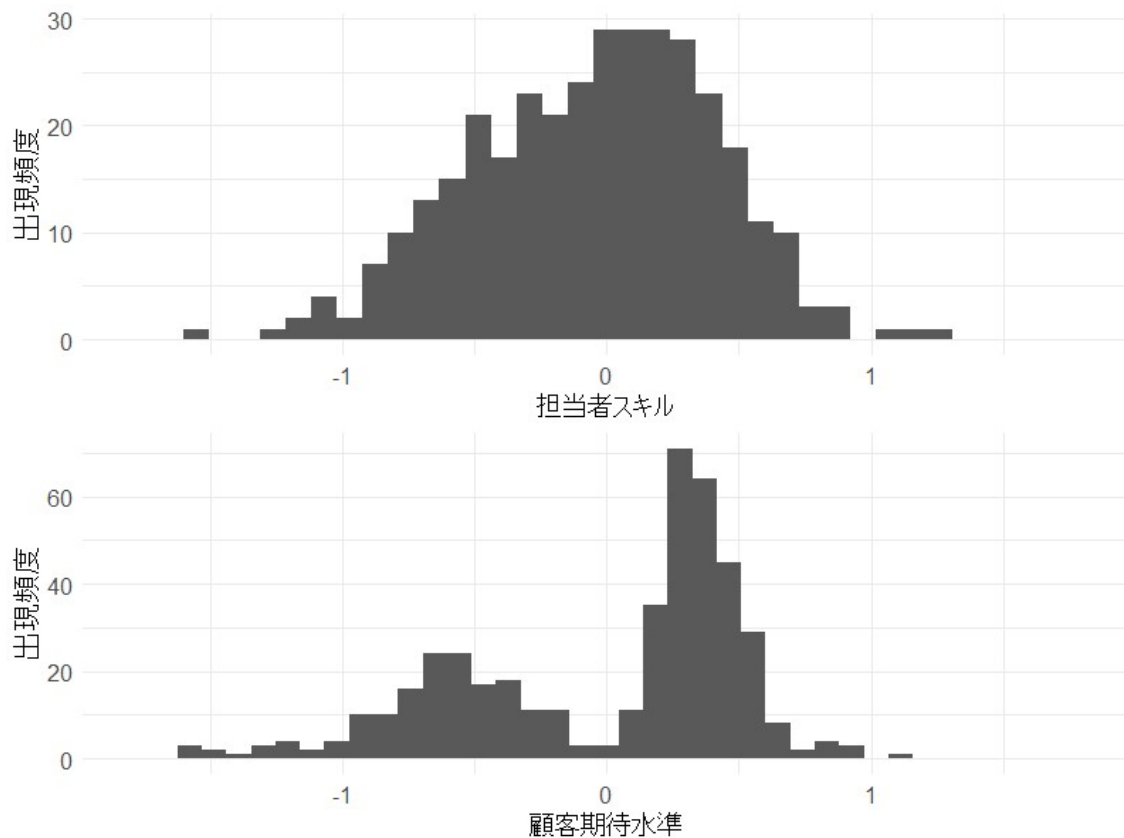


図 8 順序付きプロビットモデルの推定結果

5.4 提案モデル

定数パラメータの最尤推定と最適化

最後に提案モデルを用いて顧客満足度アンケートデータを元にした担当者スキルと顧客の期待水準の推定を行った。スキルと期待水準の推定の前に、モデル内で定数として扱っているパラメータであるアンケート結果 r の閾値 t 、スキルから成果が生成される際の分散 β 、案件ごとに担当者スキル・顧客期待水準の分散を増加させる定数 γ の最適化を行った。手順としては、 β をある値に固定した元で推定用データを用いた最尤推定により閾値 t を推定し、その後、この t と β を用いたアンケート結果 r の予測値と実測値との平均二乗誤差を推定用データについて計算した。表 2 に示した探索の結果、誤差最小となった $\beta = 1$ を採用した。

表 2 成果の分散 β と平均二乗誤差

β の値	0.01	0.1	0.2	0.5	0.8	1	1.2	2	5	10
平均二乗誤差	0.472	0.472	0.472	0.472	0.472	0.47	0.472	0.472	0.472	0.48

γ については β と t を決定した上で同様の誤差最小化による探索を行った。表 3 に示した結果より γ は一定値を下回ると誤差への影響がほぼ一定となるため、誤差が最小になる 0 ではない値として、 $\gamma = 0.0001$ を採用した。

表 3 分散の増加定数 γ と予測誤差

γ の値	0	0.00001	0.0001	0.001	0.005	0.01	0.1	0.2	0.3
平均二乗誤差	0.47	0.47	0.47	0.472	0.486	0.488	0.488	0.488	0.488

逐次推定の経過

上記により得られた定数 β, t, γ を用いて、推定データとテストデータを元にスキル・期待水準の逐次的な推定を行った。この推定に要した時間は推定用データとテストデータの合計 1024 件で 1 分程度であった。

案件を担当した回数の多いある担当者のスキル推定の過程を図 9 に示す。図 9 において、黒い実線は事後平均 $\hat{\mu}$ 、灰色の領域は事後平均から $\pm\hat{\sigma}$ の領域を示す。グラフの開始地点はこの担当者が初めて案件を担当した時点であり、その後、案件を担当するたびにスキルの事後平均 $\hat{\mu}$ が変動するとともに、事後分散 $\hat{\sigma}^2$ が減少していく様子が見て取れる。また、この担当者が案件を担当していない事後平均が一定の期間においても、灰色の領域がわずかに拡大している、すなわち事後分散がわずかに増加しているのは、TrueSkill でも導入されている「時間経過によるスキル変動の可能性」を反映するため、担当の有無に関わらず推定ごとに定数 γ をスキル分散に加算しているためである。

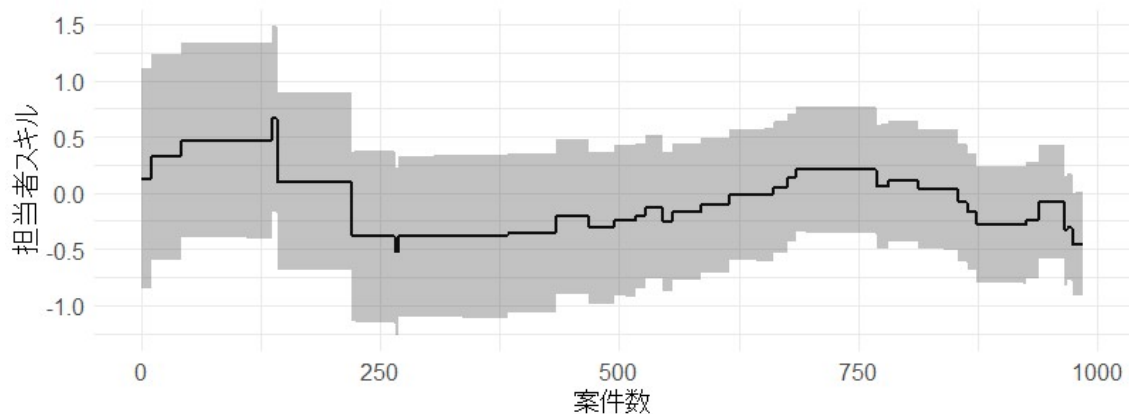


図 9 提案モデルによる担当者スキル推定推移

案件の担当回数とスキル推定の分散の関係を図 10 に示す。案件の担当数が増えるとスキルの事後分布の分散は減少する、つまりスキルの推定精度は上昇する傾向が見られる。

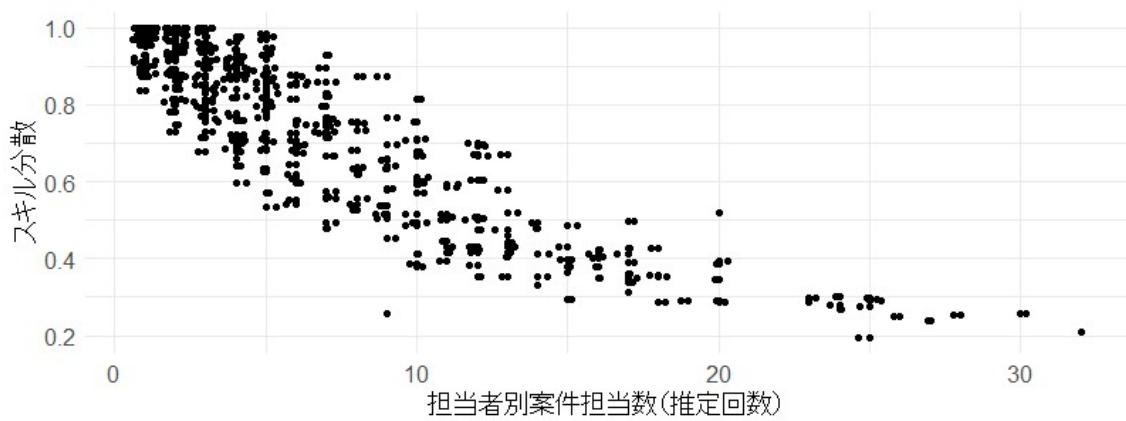


図 10 担当の案件担当回数とスキル分散の Jitter plot

次に顧客について、担当のスキルと同様に案件関与の多いある顧客の期待水準の推定過程を図 11 に示す。担当者スキル同様、顧客要求水準も案件に関与するたびに事後平均が変動するとともに事後分散が減少していく様子が見て取れる。

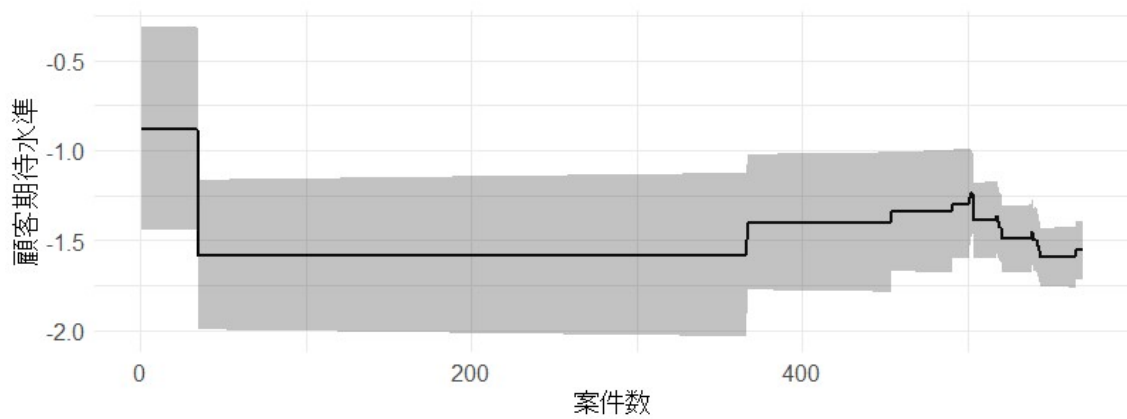


図 11 提案ベイズモデルによる顧客要求水準の平均の推移

顧客の案件関与回数と期待水準の分散の関係は図 12 に示したとおり、案件関与回数が増える
と分散は減少する、つまり期待水準の推定精度は上昇する傾向が見られる。

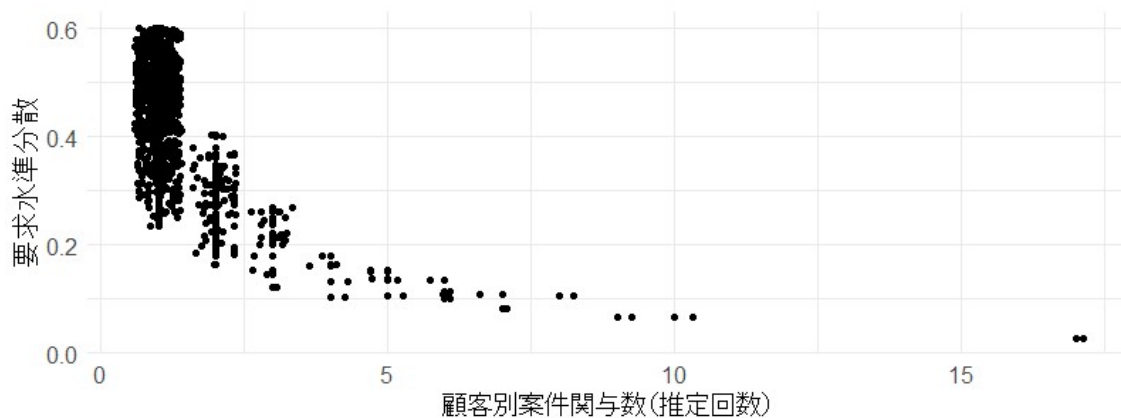


図 12 顧客の案件関与回数と期待水準分散の Jitter plot

推定用データの全案件が完了した時点での担当者スキル・顧客期待水準の事後平均値の分布
を図 13 に示す。

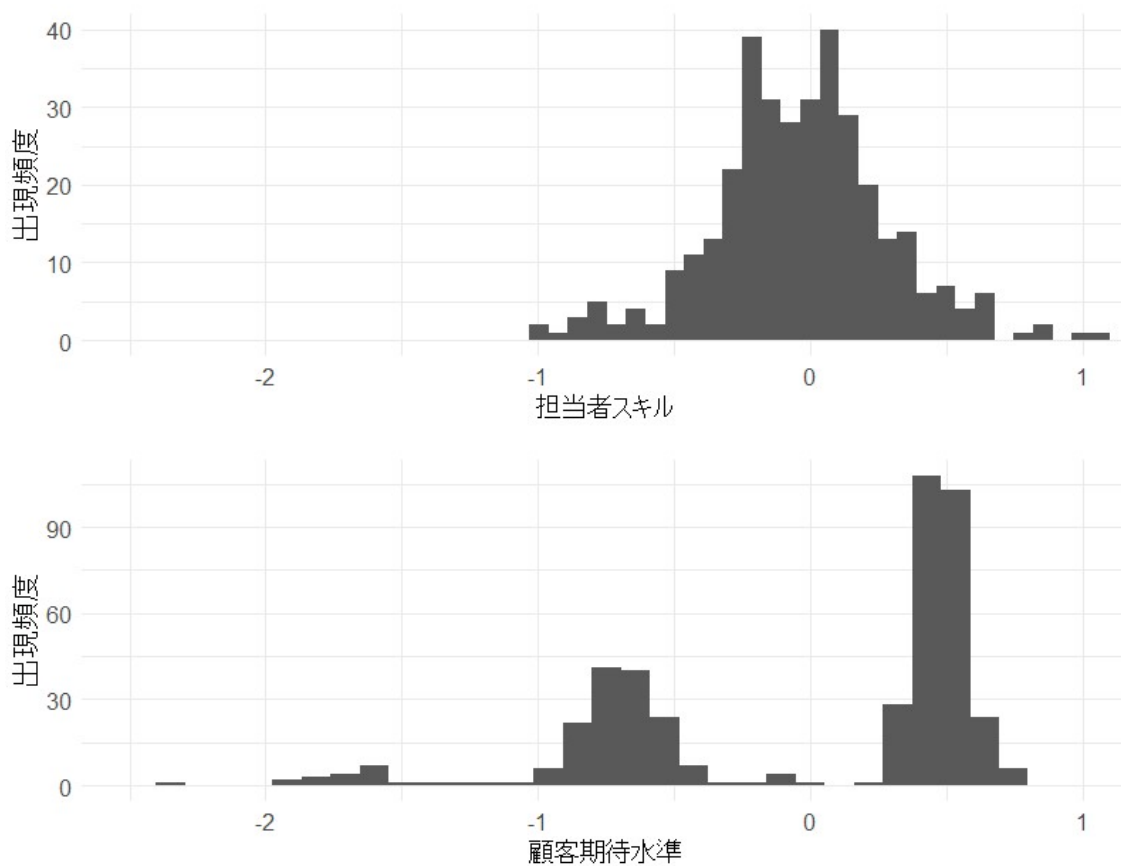


図 13 提案モデルによる担当者スキル・顧客期待水準の推定結果

5.5 予測精度の比較

本論文で扱った3つのモデルについて、実際の顧客満足度アンケートデータによる推定を行ったのち、テストデータを用いてアンケート結果の予測を行い、精度を比較した結果を表 4 に示す。ここでアンケート結果の予測値は $r \in \{1,2,3,4,5\}$ の値の出現確率という形で出力されるが、本論文ではその中でもっとも出現確率(予測確度)の高い r の値を予測値として採用し、これが実際の結果と一致すれば正解、不一致であれば不正解とした。またアンケート結果は順序付きの数値データとも捉えられるので、予測値と実測値の平均二乗誤差も併記した。

表 4 モデルの顧客満足度アンケート結果予測精度比較

モデル	正解率	平均二乗誤差
ポアソン回帰モデル	60.08%	0.68
順序付きプロビットモデル	57.77%	0.51
提案モデル	61.80%	0.49

比較の結果、正解率と平均二乗誤差ともに提案モデルが最も高い結果となった。しかし図 6 のとおり実際の顧客満足度アンケートデータは「満足度5」が 60%を占めるため、全予測を「満足度5」とした場合であっても正解率は 60%となることを考慮すると、この結果自体は実運用において有用な水準の正解率とは言えない。

アンケート結果の予測確度が 80%以上と 80%未満のものに分けて集計した正解率と平均二乗誤差を表 5 に示す。提案モデルの予測確度が 80%以上の場合に正解率は 73.8%と、他2つのモデルよりも顕著に高くなっている。また確度 80%未満の場合、正解率 60.2%と上記の全予測を「満足度5」とした倍の正解率に近い値となっている。なお、ポアソン回帰モデルにおいては、予測確度が 80%を超えることがなかった。

表 5 予測確度 80%以上の予測精度比較

モデル	予測確度	正解率	平均二乗誤差	サンプルサイズ
ポアソン回帰モデル	80%未満	60.08%	0.681	521
順序付きプロビットモデル	80%未満	53.94%	0.566	343
	80%以上	65.17%	0.416	178
提案モデル	80%未満	60.22%	0.509	460
	80%以上	73.77%	0.311	61

より詳細に予測結果について確認するため、各モデルの混同行列を表 6 に示す。括弧内は予測値に対する実測値の割合を表す。ポアソン回帰モデルではほぼ全ての予測値が 5 に偏っているため、予測値 5 に対して実測値は 3 といった大幅な誤りが多い。順序付きプロビットモデルでは予測値 3~5 が満遍なく予測されているが、予測値 4 に対する実測値 5 の割合が最大であ

った。提案モデルではやや予測値 5 に偏っているものの、予測値 4 に対する実測 4 の割合が最大となっているなど、平均二乗誤差の小ささにつながる正確性が現れている。

表 6 各モデルの混同行列

モデル	予測値	実測 3	実測 4	実測 5
ポアソン回帰モデル	1	3 (37.5%)	5 (62.5%)	0 (0.0%)
	2	3 (33.3%)	5 (55.6%)	1 (11.1%)
	3	0 (0.0%)	1 (33.3%)	2 (66.7%)
	5	23 (4.6%)	165 (32.9%)	313 (62.5%)
順序付きプロビットモデル	3	4 (50.0%)	4 (50.0%)	0 (0.0%)
	4	9 (8.7%)	38 (36.5%)	57 (54.8%)
	5	16 (3.9%)	134 (32.8%)	259 (63.3%)
提案モデル	3	0 (0.0%)	1 (100.0%)	0 (0.0%)
	4	11 (19.3%)	26 (45.6%)	20 (35.1%)
	5	18 (3.9%)	149 (32.2%)	296 (63.9%)

5.6 案件関与回数と予測精度の関係

提案モデルは逐次的な推定を行うため、推定に用いるアンケートデータのサンプルサイズが大きくなると、データの期間内に各担当者・顧客が関与する案件数(以下、案件関与回数)が増加する。その結果、期待水準やスキルの推定回数が増え予測の精度は向上すると考えられる。この確認のため、前節まで扱っていたデータに 2019 年と 2020 年のデータを加え、3 カ年(3089 件)として、長期的な逐次パラメータ推定を行い、その予測精度の変化を確認した。なお定数 β, t, γ の推定は前節までと変わらず 2018 年データの前半 500 件を使用して行った。

前節までに案件関与回数(推定回数)が増すことで推定値の分散が減少する傾向は確認したが、案件関与回数と予測精度の関係についても確認するため、案件ごとのアンケート結果予測と正解との平均二乗誤差を目的変数、その案件に関与した担当者ならびに顧客の案件関与回数(複数人担当の場合は平均値)を説明変数とした線形回帰分析を行った。その結果の回帰係数を表 7 に、データの分布を図 14 に示す。ここでは担当者の案件関与回数は推定誤差に有意な影響はなく、顧客の関与回数の影響があることが認められた。

表 7 予測の平均二乗誤差と過去案件関与回数の線形回帰係数

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.530	0.027	19.287	0.000
担当者:案件関与回数	-0.002	0.002	-1.198	0.231
顧客:案件関与回数	-0.012	0.004	-2.732	0.006

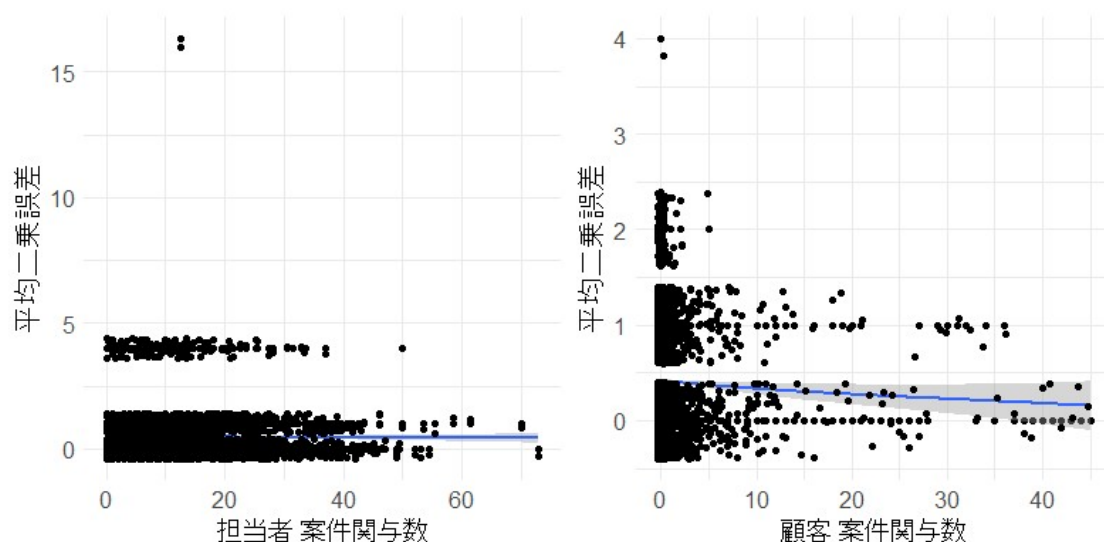


図 14 案件関与回数と精度(平均二乗誤差)の Jitter plot

上記より予測精度に対する顧客の案件関与回数の影響が認められたため、表 8 に顧客案件関与回数ごとの予測精度の変化を示す。ここで、サンプルサイズシェア 70%以上を締める関与回数 0 回顧客については精度が低く、関与回数が 1 回になると平均二乗誤差が改善し、2 回以上では正解率も 66%以上と予測精度が向上することが見て取れる。また予測の確度も関与回数が増えるに伴って増加している。

表 8 顧客案件関与回数別 精度比較

顧客 案件関与回数	正解率	平均 二乗誤差	平均 予測確度	サンプル サイズ	サンプル サイズシェア
0 回	62.49%	0.53	62.2%	2253	73%
1 回	63.41%	0.42	71.3%	410	13%
2-5 回	66.19%	0.39	73.9%	281	4%
6-10 回	71.05%	0.29	76.8%	76	4%
11 回以上	70.59%	0.29	77.3%	68	3%

以上より、案件関与回数が予測精度に与える影響が大きかった。より詳細に、顧客の関与回数別の予測値と実測値の比較結果を表 9 に示す。ここで、関与回数 0 回における 5 と予測したうち実測も 5 であった割合は 63.1%と低水準だが、関与回数 1 回では同 72.8%、11 回以上では

86.7%と、案件関与回数に従って精度が向上する様子が見て取れる。また、5 と予測したうち実際には3であった割合は、関与回数1回以上では0～2.1%と低いため、顧客満足度が満足度3の低水準となるリスクを回避する目的からは、過去一度でも案件に関与があり期待水準が推定できた顧客に関する案件について、予測値が5であればほぼ安全と判断できる。

表9 案件関与回数別の混同行列

案件関与回数	予測値	実測 3	実測 4	実測 5
0回(新規)	4	3 (5.6%)	20 (37.0%)	31 (57.4%)
	5	109 (5.0%)	701 (31.9%)	1388 (63.1%)
1回	3	1 (20.0%)	1 (20.0%)	3 (60.0%)
	4	14 (11.1%)	56 (44.4%)	56 (44.4%)
	5	5 (1.8%)	71 (25.4%)	203 (72.8%)
2-5回	3	1 (33.3%)	1 (33.3%)	1 (33.3%)
	4	8 (9.4%)	45 (52.9%)	32 (37.6%)
	5	4 (2.1%)	49 (25.4%)	140 (72.5%)
6-10回	3	0 (0.0%)	1 (100.0%)	0 (0.0%)
	4	4 (17.4%)	13 (56.5%)	6 (26.1%)
	5	0 (0.0%)	11 (21.2%)	41 (78.8%)
11回以上	3	0 (0.0%)	3 (100.0%)	0 (0.0%)
	4	11 (31.4%)	22 (62.9%)	2 (5.7%)
	5	0 (0.0%)	4 (13.3%)	26 (86.7%)

最後に、3カ年データを年度別に分け期間ごとの予測精度を算出した結果を表10に示す。ここでは時間的に後の期間ほどそれまで蓄積した担当者スキルや顧客期待水準の推定結果を使えるため、高い精度が得られることが期待されたが、結果としては正解率がやや改善したものの平均二乗誤差は大差がなく、単純にアンケートデータが貯まりさえすれば精度が向上するわけではない。

表 10 データ年別の精度比較

データ期間	正解率	平均二乗誤差	サンプルサイズ
2018 年後半	61.80%	0.486	521
2019 年	63.65%	0.506	1095
2020 年	64.34%	0.480	973

上記の背景として、表 11 に示した「データ期間別の案件関与回数」を見ると、2018 年後半でも 2020 年でも、過去関与回数 0 回の新規顧客の割合が変わっていないことがわかる。これは本論文で用いたアンケートデータが一部抜粋である影響を除けば、時間の経過とともに顧客が次々に入れ替わっている、もしくは同一顧客が顧客満足度アンケートに回答してくれる回数には限度があるといった可能性が示唆される。このデータの特性により、アンケート結果を時間的に蓄積しても顧客の案件関与回数が増加せず、結果として予測精度も上がらないといった状況が生じている。

表 11 データ期間別の案件関与回数

データ期間	顧客の案件関与回数				サンプル サイズ
	0回(新規)	1-5 回	6-10 回	11 回以上	
2018 年後半	69.48%	26.30%	3.07%	1.15%	521
2019 年	70.41%	22.74%	2.92%	3.93%	1095
2020 年	70.09%	25.08%	2.88%	1.95%	973

5.7 項目別推定値による担当者と顧客の分類

ここまではサービスの顧客満足度を1個の指標として扱ってきたが、顧客満足度のアンケート設計としてサービスを複数の項目に分割して、それぞれの満足度を調査することも可能である。本論文で用いたアンケートデータにおいても、案件について「コミュニケーションの的確さ」「企画段階での提案」「調査知識」「事業業界理解」などの項目を設け、それぞれの満足度を聴取している。これにより、一面的な総合的顧客満足度という評価だけでなく、各項目について細かく顧客期待水準と担当者スキルを推定することができる。図 15 に実際のアンケート結果を元に項目別に推定した担当者スキルと顧客満足度の分布を示す。ここから特に顧客期待水準については項目によって推定値分布の形が異なることが見て取れる。

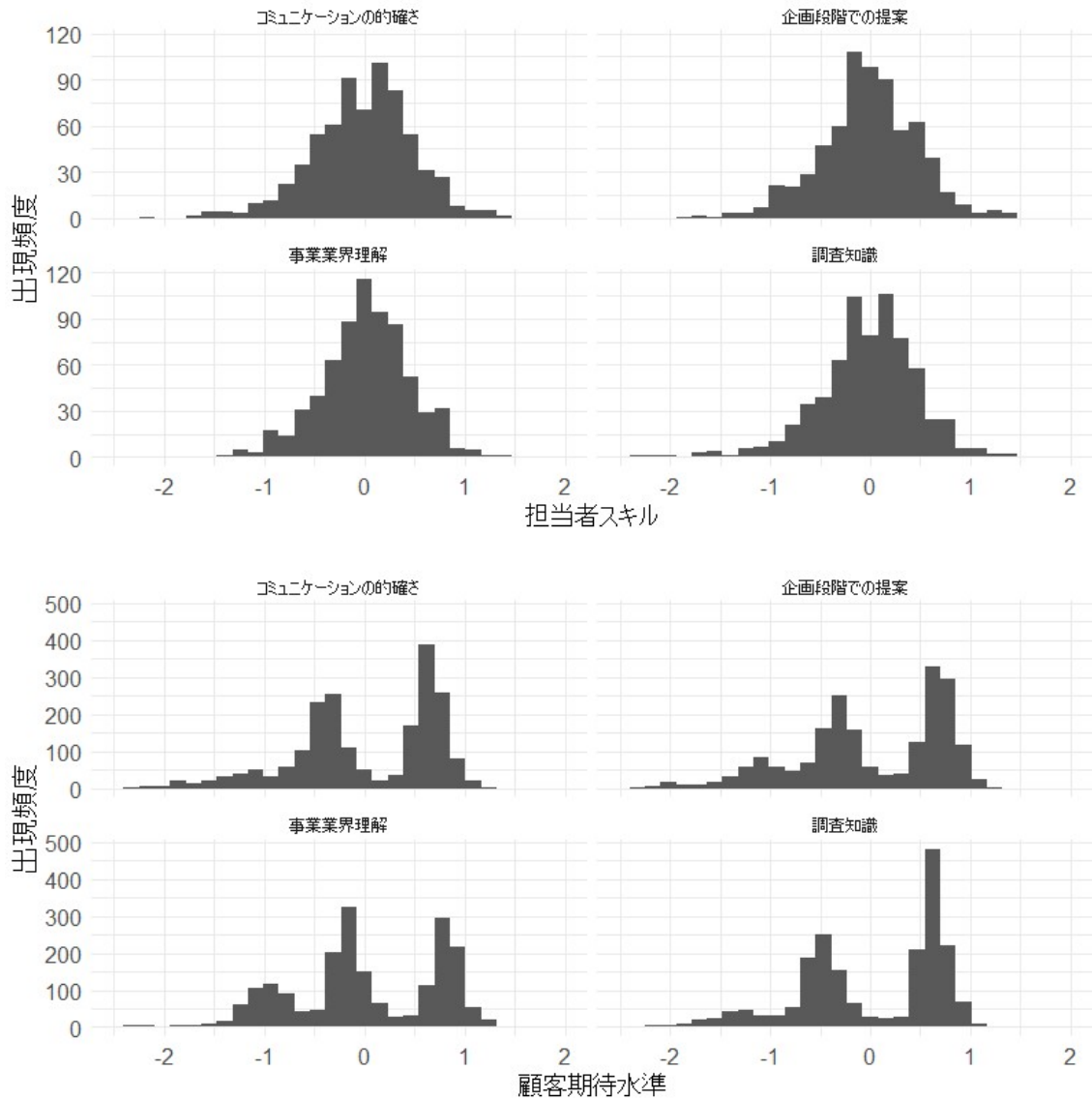


図 15 項目別担当者スキル・顧客期待水準の推定値分布

上記の項目別推定値を元に、顧客と担当者を非階層クラスタリング(k-means)によって得られたクラスタの中心とその解釈を表 12 と表 13 に示す。担当者については各項目が連動して高いクラスタ低いクラスタがある他に、コミュニケーション力によってクラスタが分かれる傾向があり、コミュニケーション力の属人性の高さを伺わせる。顧客については、全体的に高い低いクラスタの他に、コミュニケーション、業界理解、提案力など、期待するポイントが異なるクラスタが見られることから、顧客の期待の多様性を示していると言える。

表 12 項目別スキル推定を用いた担当者クラスタリング結果

	人数 シェア	事業業界 理解	調査知識	企画段階 での提案	コミュニケ ーション	要約
1	15%	-0.5	-0.5	-0.5	-0.5	全体的に低スキル
2	13%	0.5	0.7	0.5	0.6	全体的に高スキル
3	17%	0.1	-0.2	0.0	-0.4	平均的だがコミュニケーション 苦手
4	27%	0.2	0.2	0.3	0.2	全体的に平均よりやや上
5	25%	-0.1	0.0	-0.2	0.1	平均的
6	3%	-0.6	-1.5	-1.0	-1.1	非常に低スキル

表 13 項目別スキル推定を用いた顧客クラスタリング結果

	顧客数 シェア	事業業界 理解	調査知識	企画段階 での提案	コミュニケ ーション	要約
1	9%	-1.0	-1.4	-1.2	-1.3	非常に期待が高い
2	12%	0.1	0.0	0.6	-0.5	コミュニケーション重視
3	17%	-0.4	0.2	0.2	0.6	業界理解重視
4	24%	-0.5	-0.6	-0.6	-0.4	全体的に期待高い
5	27%	0.8	0.6	0.7	0.6	全体的に期待低い
6	12%	0.4	0.4	-0.4	0.0	提案力重視

5.8 考察

実際の顧客満足度アンケートデータを用いた担当者スキル・顧客期待水準の推定と予測精度検証の結果、本論文で提案した3つのモデルの中では「提案モデル」が相対的には高い予測精度を示した。これは、ポアソン回帰モデルや順序付きプロビットモデルは推定用データを用いたパラメータ推定完了後はパラメータを更新しないのに対し、提案モデルではテストデータの予測の過程においても逐次的なパラメータの更新が行える優位性によるものと考えられる。また、推定にかかる時間も提案モデルの逐次的な推定手法はMCMCを用いる手法に比べ大幅に短縮できるため、効率的であるといえる。

ただし、提案モデルの正解率62%であり、データの約60%を満足度5が占める偏りを考慮すると、全てを満足度5と予測した場合と大差がなく、高い水準とは言えない。この背景として、今回のデータは2018年の1カ年分を前後2つに分けて推定用データとテストデータとしたが、1カ年の後半に初めて出現する新規顧客や受注頻度の低い顧客が数多く存在することが影響していると考えられる。これはビジネスの構造上ロングテールな顧客を対象とする限りは避けられない課題であると言えるが、理想としては新規顧客に対する満足度予測であっても、人員が固定されている担当者スキルの推定が逐次的に進むことによって、一定の精度を担保できることが望ましい。しかし今回

の実証分析では、担当者スキルの関与回数が予測精度に与える影響は有意とは言えない結果であったため、担当者スキルの更新が進んだとしても予測精度の向上は期待できない。

以上から、本論文で用いた案件顧客満足度アンケートからの担当者スキル・顧客期待水準の推定と顧客満足度アンケート結果予測については、提案モデルによる逐次的な推定と予測が相対的には優位であると言えるが、その正解率はデータ全体で6割、一定以上の関与回数がある顧客に限定しても7割程度と高くはない精度であるため、実用で予測値をそのまま鵜呑みにして信頼することは難しい。ただし、案件の品質が顧客の最低条件を下回り顧客満足度が低くなるリスクを回避するという目的においては、結果3を「低い満足度」と仮定すると、予測値別の実測横比を参考に「予測値が5であれば、結果3になる可能性はごく低い安全」、「予測値が4であれば結果3の可能性が出てくるため、要注意」、「予測値が3であれば、結果5を獲得できる可能性が低い」といったように、案件担当者の任命時の目安として、もしくは任命後の案件運用時のフォロー必要性の目安として運用することが可能であると考えられる。

予測精度について別の要因としては、モデルが現実と十分にフィットしていない可能性が考えられる。特に担当者スキルの影響力の低さを鑑みるに、担当者のスキルのモデル化に課題があるように思われる。1つには、複数人担当の案件の成果を担当者ごとの成果の平均としている点について、現実には、最もスキルの高い担当者が全体の成果を自身のレベルまで引き上げる、もしくは最もスキルの低い担当者がボトルネックとなって成果を引き下げるといった事態が起こりうる。また、個々の担当者が案件の成果に対して担う重みが一律ではなく、重み付き平均になるといったことも考えられる。これをモデルに反映させるには、例えば推定スキルの高い担当者ほど重要な役割を担うと考え、推定スキルをチーム内平均の重みとして使うといった方策が挙げられる。

もう1つには「スキル」といってもビジネスにおけるスキルは多面的であり、例えば調査の設計力、顧客とのコミュニケーション力、提案力、資料作成力など、様々な要素が含まれる。1つの案件を複数人で担当する際、多くの場合はこれらの多面的なスキルをチームで分担するため、担当者全員を1つのスキルで評価するという本論文のモデルは、やや単純化が過剰であるとも考えられる。一方で顧客の側も、相対する担当者のどのスキルを特に評価するかといった好みがあり、1つの満足度という指標では測りきれない可能性がある。

上記の検証として5.7節において、顧客満足アンケートで聴取された項目別の推定を行い担当者や顧客をクラスタリングした結果、担当者はコミュニケーション力によるクラスタの分割、顧客はコミュニケーション力、提案力、業界知識など多様な項目によるクラスタの分割が見られた。ここからも、一義的なスキル・期待水準だけでなく、顧客の多様な期待ニーズに応えられる担当者のマッチングが、最終的な満足度の予測精度に影響を及ぼすものと考えられる。

6 結論

6.1 まとめ

本論文では、顧客-担当者型サービスビジネスにおける顧客満足度(顧客満足度)の構造的な課題の解決を目的として、顧客満足度アンケートデータを用いた顧客期待水準と担当者スキル推定の手法を提案し、顧客満足度の予測精度についての評価を行った。具体的には、期待-不一致モデルを元にした顧客と担当者のマッチング問題をゲームのプレイヤースキル推定とマッチングの問題に置き換え、ネットワーク対戦ゲームのプレイヤースキルレーティングアルゴリズム TrueSkill を参考として、顧客期待水準と担当者スキルの推定モデルを構築した。このモデルを単純な回帰モデル(ポアソン回帰モデル、順序プロビットモデル)と比較し、実際の顧客満足度アンケートデータを用いた実証実験により、予測精度と計算速度の面から優位性を示した。しかし実用上は、データに含まれる新規顧客の影響もあり予測精度が高いとは言えないため、本論文で提案した手法は、極端に低い顧客満足度を回避する目安としての運用にとどまると結論する。

6.2 今後の課題と展望

今後の課題として、第一に顧客満足度の予測精度の向上が挙げられる。本論文の実証分析では、顧客の推定有無が予測精度に大きく寄与するという結果であったが、これは新規の顧客が多いビジネスにおいては精度向上が難しいことを意味する。一方で現実的には新規の顧客についても事前情報からある程度の「予測水準のあたり」をつけて担当者を検討することがあるため、例えば顧客の事業規模、業界、競合の利用有無、案件の価格などといった情報によって、顧客の期待水準の事前分布を変動させる仕組みをモデルに導入できれば、新規顧客についての予測精度を向上できる可能性があると考えられる。

また本論文で用いた顧客満足度アンケートでは総合的な満足度以外にも複数の要素(コミュニケーション力、企画提案力、プレゼンテーション力など)についての満足度を聴取しているため、顧客期待水準や担当者スキルをそれぞれの要素ごとに推定し、それぞれが総合満足度にどのような影響を与えるかといった構造を推定することで、満足度の予測精度向上が見込められると思われる。

今後の展望としては、案件に対する担当者の提案システムが挙げられる。本論文で提案した顧客満足度の予測モデルを用いると、期待水準の高い顧客案件にスキルの低い担当者を任命すると顧客満足度が低くなるのが事前に予測可能である。実用シーンではその場合、代わりに高いスキルを有する担当者を任命しなければならないが、その際に「多数存在する担当者候補のうち、誰を担当させるのが最適か」という問題が生じる。ここでは、交代した担当者は別の案件からは外れることになり、交代させられた担当についても別途自身のスキルに適した案件に任命するといった点を考慮して、組織全体の案件の顧客満足度が最大となるような最適化が求められる。理想的には、日々の案件の積み重ねにより担当者スキルと顧客期待水準を逐次更新しつつ、新規の案件については全体最適となるような適任の担当者を提示するシステムが望まれる。

また、満足度予測精度の向上という課題が達成された後には、一般的に「売上」や「担当案件数」といった成果で測られることの多い「顧客-担当者型サービス」における担当者の評価について、「結果として顧客の満足度につながるスキル」という新たな軸を加えることができると考える。これは顧客満足度が企業の財務的指標の先行指標である [4]という仮定に基づけば、担当者スキルは顧客満足度のさらに先行指標であるため、売上や案件数といった短期的な指標ではなく、将来の経営状況の予測指標として重要視され得るものとする。またこのスキルを前述の項目別に推定することで、担当者自身にとっては強みや弱みの把握に役立てられるほか、企業研修の前後のスキル変動によって研修の効果を検証する、メンバーのスキル変動を時系列で観察することでマネージャーを評価する、採用においては一定の試用期間でスキルを推定し本採用の可否を判断するなど、ビジネス上の活用範囲は多岐に渡ると考えられる。

謝辞

本研究活動および本論文の執筆にあたり格別のご指導をいただきました、滋賀大学大学院データサイエンス研究科の岩山幸治先生に心より感謝申し上げます。統計学については門外漢であった自分の初歩的な質問や相談にも常に親切に快く応じていただき、非常に多くの物事を学ばせていただきました。また論文執筆にあたっては内容の精査、修正に多くのお時間を割いていただき、多大なるご助力をいただくとともに、たいへんご迷惑をおかけいたしました。心より恐縮に思うとともに、この上なくありがたく感じております。

また修士1年次には指導教員として、修士2年次は副指導教員としてご指導いただきました河本薫先生、修士2年次に副指導教員をお引き受けいただきました姫野哲人先生に感謝いたします。河本先生には紆余曲折した私の研究テーマの方向性を定めていただくとともに、ご専門として岩山先生をご紹介いただいたことを特にありがたく存じております。

そして、私を滋賀大学大学院への社会人留学として送り出すとともに、本研究で使用したデータの利用を許可いただいた株式会社マクロミルの皆様に感謝いたします。特に私を候補として推挙いただいた橋本厚司さん、論文執筆中の業務を分担してくださった技術研究グループの皆様に御礼申し上げます。

最後に、本研究活動と論文執筆にあたり心身ともに支えとなってくれた家族に感謝いたします。

参考文献

- [1] R. L. Oliver, "A cognitive model of the antecedents and consequences of satisfaction decisions," *Journal of Marketing Research*, vol. 17, pp. 460-469, 1980.
- [2] E. W. Anderson and M. W. Sullivan, "The antecedents and consequences of customer satisfaction for firms," *Marketing Science*, vol. 12, no. 2, pp. 125-143, 1993.
- [3] F. F. Reichheld and W. E. Sasser, "Zero defections: Quality comes to services," *Harvard Business Review*, vol. 68, no. 5, pp. 105-111, 1990.
- [4] T. S. Gruca and L. L. Rego, "Customer Satisfaction, Cash Flow, and Shareholder Value," *Journal of Marketing*, vol. 69, no. 3, pp. 115-130, 2005.
- [5] R. L. Oliver, *Satisfaction: A Behavioral Perspective on the Consumer*, McGraw-Hill, 1997, pp. 98-131.
- [6] 小野謙司, 顧客満足[CS]の知識, 日本経済新聞出版社, 2010.
- [7] R. Herbrich, T. Minka and T. Graepel, "TrueSkill: A Bayesian Skill Rating System," in *Proceedings of the 19th international conference on neural information processing systems*, MIT Press, 2006, pp. 569-576.
- [8] 劉兵 and 神山進, "テーマパークにおける顧客満足の構造分析," *彦根論叢*, no. 397, pp. 38-53, 2013.
- [9] 南知恵子 and 小川孔輔, "日本版顧客満足度指数(JCSI)のモデル開発とその理論的な基礎," *マーケティングジャーナル*, vol. 30, no. 1, pp. 4-19, 2010.
- [10] A. E. Elo, *The rating of chess players: Past and present.*, Acro Publishing, 1978.
- [11] T. Minka., "A family of algorithms for approximate Bayesian inference.," in *PhD Dissertation*, Massachusetts Institute of Technology, 2001.
- [12] 西山慶彦, 新谷元嗣, 川口大司, 奥井亮, 計量経済学, 有斐閣, 2019.
- [13] 山本勲, 実証分析のための計量経済学, 中央経済社, 2015.
- [14] 馬場真哉, RとStanではじめるベイズ統計モデリングによるデータ分析入門, 講談社, 2019.
- [15] M. D. Hoffman and A. Gelman, "The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo," *Journal of Machine Learning Research*, no. 15, pp. 1593-1623, 2014.

[16] 須山敦志, ベイズ深層学習, 講談社, 2019.

[17] N. L. Johnson, S. Kotz , N. Balakrishnan, Continuous Univariate Distributions, Volume 1, 2nd Ed., Wiley, 1994, p. Section 10.1.