

レベニューマネジメントと需要関数の学習

佐藤 公俊

販売環境における不確実性の高まりや消費者の価値観の多様化により、商品・サービスの需要予測はますます困難になっている。近年では、収益最大化のための手法であるレベニューマネジメント分野において、販売から得られるデータを用いて需要関数を逐次的に推測し、価格を調整する研究に注目が集まっている。本稿では、そのようなダイナミックプライシングモデルを多腕バンディット問題として扱った研究を数値例とともに紹介する。

キーワード：レベニューマネジメント、ダイナミックプライシング、多腕バンディット問題

1. はじめに

今日のコロナ禍による経済的リスクに加えてウクライナ危機による地政学的リスクが需要と価格の関係に不確実性を増大させた。特に航空・鉄道などの輸送産業およびホテル・レストラン・イベントなどのサービス産業ではコロナ禍において甚大な影響を受けている。今後の需要回復期に向けて中長期的な視点で収益を上げるためにレベニューマネジメントの活用が求められている。レベニューマネジメント (RM: revenue management, 収益管理) は企業収益を最大にするために、さまざまな商品やサービスの供給容量を管理し、販売時の価格付けに関する経営戦略を策定するための理論である [1]。RM モデルでは、同一商品を価格差別化することで複数の料金クラスに分け、総供給量 (在庫量) をいかに各クラスに配分するかという問題を扱う。ダイナミックプライシング (DP: dynamic pricing) は、リアルタイムに得られる情報に基づき、いかに販売価格を調整するかという問題を扱う。RM モデルと DP モデルともに目的は収益最大化であるが、RM モデルでは価格を所与として「量」を制御するのに対して DP モデルでは「価格 (構成)」を制御することで量も間接的に制御するため、Revenue Management and Pricing と呼ばれることが多い。

近年、さまざまな業界において DP の有用性が報告されている。たとえば、フロリダ州に本社を置く Holiday Retirement では、介護付き住宅の賃料を決定する問題に対して価格決定モデルを適用した結果、9% (8,800 万ドル) 増収したことが報告されている [2]。そのほかに、フランスのレンタカー会社である Europcar [3] や

ホテル経営企業である Wyndham Destinations のパッケージレンタル [4]、マイアミに本拠を置くクルーズ会社である Carnival Cruise Line [5] などにおいても収益の改善効果が報告されている。

DP の理論は Gallego and van Ryzin [6] 以降、構築されていった。DP モデルにおける重要な問題の一つは需要関数の決定である。十分な量の過去の販売データが蓄積されている商品・サービスに対しては、需要モデルの選択とそのパラメータを推定しやすいため、需要関数を既知と仮定した DP モデルを応用できる。しかし、過去データの少ない新商品・新サービスおよび市場環境の変化が激しい商品などに対しては、販売後に得られる情報をもとに需要関数を逐次的に学習し、意思決定の質を高めることが有効であると予想される。このため、近年では、未知の需要関数の学習を考慮した価格決定の研究が注目されている。先行研究として、在庫制約の下でパラメトリックな需要関数を扱った研究 [7, 8] や、ノンパラメトリックな需要モデルを扱った研究 [9, 10]、複数商品を扱った研究 [11, 12] などが挙げられる。これらの研究は、価格の候補を複数のアームと見做した多腕バンディット問題として価格決定問題を捉え、未知の需要関数を学習するために広範囲の価格を設定する探索 (exploration) と収益を増加のために最適価格を設定する活用 (exploitation) を繰り返し、長期的な収益最大化を図ることを目的とする。さらに、近年では、個々の顧客情報に基づき価格を設定する (カスタマイズドプライシング) 研究へと拡張されている [13]。

本稿では、パラメトリックな需要関数の学習を考慮した DP モデルの基礎モデルとアルゴリズムを紹介する。さらに、DP に対する消費者の不公平感の観点から価格の上限値について感度分析を行う。

さとう きみとし

神奈川大学工学部経営工学科

〒 221-8686 神奈川県横浜市神奈川区六角橋 3-27-1

k-sato@kanagawa-u.ac.jp

2. DP モデル

ダイナミックプライシング (DP) は在庫数や将来の需要予測値、他社価格、顧客個々の購買行動などの観測情報に基づき、動的に販売価格を変えて販売する手法である。DP は設定する際に候補となる価格の集合が離散値であるか連続値であるかにより、assortment optimization, dynamic price adjustment, continuous pricing の三つのメカニズムに分類される [14]。以下では、continuous pricing を紹介する。

ある企業は $C (> 0)$ 単位の商品を期間 $t = 1, 2, \dots, T$ に販売する。 $t = 1$ は販売期間の初期、 $t = T$ は販売期間の最終期を表す。1 期間に顧客が多くとも 1 人購入に訪れるものと仮定し、その到着率を $m (> 0)$ とする。顧客が商品に対して支払っても良いと考える最大価格 (支払い意思額) を確率変数 X によって表し、確率変数のパラメータを ϑ とする。各期の期首で企業は価格 $p \in [p, \bar{p}]$ を決定する。ここで、 p と \bar{p} はそれぞれ価格の下限値と上限値である。顧客の支払い意思額が販売価格よりも高いときに商品は購入されるため、購入率を $d_\vartheta(p) := \Pr(X \geq p)$ と定義する。したがって、需要率はパラメータ $\lambda_\theta(p) = md_\vartheta(p)$ のポアソン過程に従う。ここで、到着率と購入率は独立と仮定する。需要率に含まれるパラメータ $\theta := (m, \vartheta)$ は未知であり、販売後に得られるデータより推定する。

各期において、(i) 顧客が訪れて商品を購入する、(ii) 顧客は訪れたが商品を購入しない、(iii) 顧客が訪れない、のいずれかの事象が起こるため、第 t 期の期首で在庫量が c 単位 ($c = 0, \dots, C$) であるときの t 期から T 期までの最大総期待収益 $V_t^\theta(c)$ は以下のように定式化される。

$$\begin{aligned} V_t^\theta(c) &= \sup_{p \in [p, \bar{p}]} [md_\vartheta(p)(p + V_{t+1}^\theta(c-1)) \\ &\quad + m(1 - d_\vartheta(p))V_{t+1}^\theta(c) + (1 - m)V_{t+1}^\theta(c)] \\ &= \sup_{p \in [p, \bar{p}]} [\lambda_\theta(p)(p - \Delta V_{t+1}^\theta(c)) + V_{t+1}^\theta(c)]. \end{aligned} \quad (1)$$

ここで、 $\Delta V_{t+1}^\theta(c) = V_{t+1}^\theta(c) - V_{t+1}^\theta(c-1)$ は在庫 1 単位あたりの限界価値を表す。商品の陳腐化性と希少性により、境界条件をそれぞれ $V_{T+1}^\theta(c) = 0, c = 1, \dots, C$ 、と $V_t^\theta(0) = 0, t = 1, \dots, T$ 、と仮定すれば、各期における最適政策は $t = T$ から $t = 1$ まで後ろ向きに解くことで得られる。もし、収益関数 $r_\theta(p) := \lambda_\theta(p)p$ が p について凹関数ならば、最適価格が存在し、それを

$$p_t^{*\theta}(c) := \operatorname{argmax}\{p \in [\bar{p}, p] \mid \lambda_\theta(p)(p - \Delta V_{t+1}^\theta(c))\} \quad (2)$$

と定義する。ここで、 $\Delta V_{t+1}^\theta(c)$ を 1 単位余分に保有することの機会費用と見做せば、初期在庫量 C が十分大のとき ($T \ll C$)、容量制約がないのと同じ状態なので、機会費用はゼロと仮定することができる。したがって、(2) 式は t に依存しないため、一定価格による販売が最適価格となる。

ここでは、先行研究 [15–17] と同様に支払い意思額 X がパラメータ α の指数分布に従うとき、 $\vartheta = \alpha$ より、 $d_\alpha(p) = \exp\{-\alpha p\}$ であり、最適価格は次式のよりに支払い意思額の期待値と限界価値の和となる。

$$p_t^{*\theta}(c) = \frac{1}{\alpha} + \Delta V_{t+1}^\theta(c). \quad (3)$$

ここで、需要パラメータは $\theta = (m, \alpha)$ である。また、最大総期待収益は

$$V_t^\theta(c) = \frac{m}{\alpha} e^{-1-\alpha \Delta V_{t+1}^\theta(c)} + V_{t+1}^\theta(c) \quad (4)$$

となる。4 節では、(3) 式および (4) 式を用いてシミュレーションを行う。

3. パラメータの学習と DP アルゴリズム

需要率 $\lambda_\theta(p)$ のパラメータ θ が未知であるとき、販売データを用いて逐次的にパラメータを学習する。需要モデルの学習を考慮した DP の先行研究はパッシブ学習、アクティブ学習、強化学習に分類される。本節では、パッシブ学習とアクティブ学習による DP を解説する。学習においてはさまざまなデータの活用が考えられるが、ここでは販売価格データと販売価格に対応した購入有無のデータのみを用いる。

パッシブ学習による販売政策は販売期間の初めにそれまでの販売データを用いて需要パラメータを推定し、(2) 式より得られる最適政策を次の販売期間に適用する方法である (図 1)。需要パラメータの推定精度が低ければ、それに基づく最適政策の質も低くなるため、次のパラメータの更新まで収益の増加は期待できない。需要パラメータの推定においては、過去に設定された価格のばらつきが大きいほど、データの質が高まり推定精度は高まる。そのため、アクティブ学習による販売政策では、最適政策とは異なる任意の価格を設定することでデータの質を高め、パラメータの推定精度を改善すること (探索, exploration)、そのパラメータを用いて (2) 式より得られる最適政策を適用すること (活用, exploitation) を各期に繰り返す (図 2)。

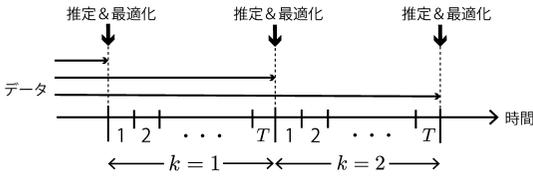


図1 パッシブ学習のイメージ

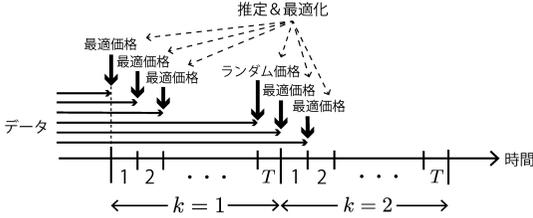


図2 アクティブ学習のイメージ

パラメータの推定は最尤法を用いる。商品の販売開始から販売終了までの販売期間 $t = 1, \dots, T$ を1エピソードと呼び、エピソードを添え字 k で表す。エピソード $k = 1$ の販売開始時点 $t = 1$ から経過した期間数を s とする。このとき、エピソード k の第 t 期の期首までに得られる販売データを $\mathcal{D}_s := \{(p_s, d_s), s = 1, 2, \dots, T(k-1) + t - 1\}$ とする。ここで、 p_s は s 期の販売価格、 d_s は s 期に商品が売れた場合は1、そうでない場合は0をとる変数である。ここで、販売期間の途中で在庫切れとなる場合 ($c = 0$)、その時点以降の販売価格を $p_s = \infty$ 、販売データを $d_s = 0$ とする。確率変数 D_s を成功率 $\lambda_\theta(p_s)$ のベルヌーイ分布に従う確率変数とすれば、 $\{d_s\}$ はベルヌーイ分布から無作為抽出された購買履歴のサンプル列を表す。このとき、第 s 期の需要パラメータの推定値 θ_s は以下の最大化問題の解として求められる。

$$\max_{\theta_s} \log \mathcal{L}(\theta_s | \mathcal{D}_s). \quad (5)$$

ここで、

$$\mathcal{L}(\theta_s | \mathcal{D}_s) = \prod_{j=1}^s [\lambda_{\theta_s}(p_j)]^{d_j} [1 - \lambda_{\theta_s}(p_j)]^{1-d_j} \quad (6)$$

である。

需要関数の学習を考慮したDPでは、パラメータの推定と価格最適化を繰り返すために、複数の販売期間から得られる総収益と、真のパラメータを適用したときの総収益との差によって最適政策を評価する。真のパラメータを $\theta_0 = (\Lambda_0, \vartheta_0)$ とし、意思決定者には観測できない値であるとする。また、エピソード数が K であるときの総期待収益は $\mathbb{E}_\pi [\sum_{s=1}^{KT} p_s^\pi D_s^\pi]$ であり、

これを最大化する価格政策 π を求める問題を考える。この問題は決定変数が連続値をとる多腕バンディット問題である。

最適政策の良さを評価するためにリグレット (Regret) を用いる。残り k 回のエピソードに価格政策 π を適用したときのリグレットは以下のように定義される。

$$\text{Regret}_k^\pi = k \cdot V_1^{\theta_0, 0}(C) - \mathbb{E}_\pi \left[\sum_{s=1}^{Tk} p_s^\pi D_s^\pi \right]. \quad (7)$$

ここで、 $V_1^{\theta_0, 0}(C)$ は真のパラメータ θ_0 の下で (1) 式より得られる1エピソードの最大総期待収益である。リグレットは最適政策を適用しなかったことによる期待損失額であり、リグレットを最小化することと長期的な総期待収益を最大化することは等価である。最適政策を求める方法として、den Boer and Zwart [8] のアルゴリズムを紹介する。

DP アルゴリズム

- $s = 1$ のとき：初期の需要パラメータ θ_1 を任意に定め、定数 $\epsilon \in (0, (\bar{p} - \underline{p})/4)$ を選択する。
 - $s \geq 2$ のとき：過去データ \mathcal{D}_{s-1} を用いて、(5) より需要パラメータ θ_s を推定する。
 - $c_s = 0$ のとき、 $p_s = \bar{p}$ とする。
 - $c_s > 0$ のとき、(2) 式より、 $p^* := p_s^{\theta_s}(c_s)$ を求める。次の (i), (ii), (iii) を満たすならば、 $p_s \in [p^* - 2\epsilon, p^* + 2\epsilon] \cup [p, \bar{p}]$ とする。
 - (i) $|p_i - p_j| < \epsilon, i, j = 1, \dots, s, S_i = S_j = S_s,$
 - (ii) $|p_i - p^*| < \epsilon, i = 1, \dots, s, S_i = S_s,$
 - (iii) $c = 1$ または、 $s = T$ 。
- もし、(i), (ii), (iii) を満たさなければ、 $p_s = p^*$ とする。

ここで、 $S_i = 1 + \lfloor (i-1)/T \rfloor$ は i 期が属するエピソードの番号を表す。条件 (i)-(iii) より、このアルゴリズムでは販売期間内の価格変動が小さく (条件 (i), (ii)) かつ残在庫が1または販売期間の最終期 (条件 (iii)) ならば、探索 (exploration) を行い、販売価格がある区間からランダムに選択される。それ以外の期間では最適政策が適用される。また、パラメータ ϵ によって、最適政策の活用と需要パラメータの探索のバランスが調整される。den Boer and Zwart [8] はリグレットの上界が $\text{Regret}(K) = O(\log^2(K))$ となることを示している。すなわちエピソード数 K が増えるにつれて $1/K^2$ 程度に比例してリグレットは減少していくことを意味する。

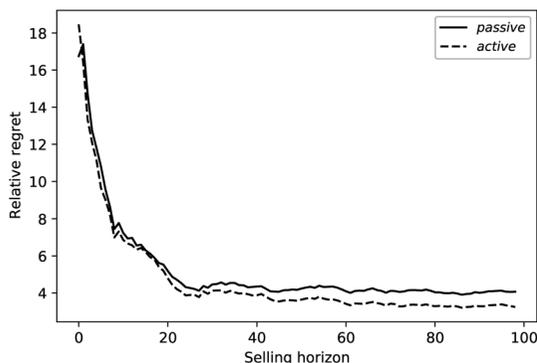


図3 エピソードとリグレットの関係

4. シミュレーション

4.1 パラメータの設定とリグレットの評価指標

本節では、支払い意思額の確率分布に指数分布を仮定した(3)式を用いて、パッシブ学習とアクティブ学習による販売を比較する。エピソード数を $K = 100$ 、1エピソードあたりの販売期間数を $T = 20$ 、初期在庫量を $C = 10$ とする。また、価格の上限値を $\bar{p} = 5$ 、下限値を $p = 1$ とし、真の到着率を $m_0 = 0.75$ 、真の需要率パラメータを $\alpha_0 = 0.4$ とする。到着率の初期値 $0.7m_0$ 、需要率パラメータの初期値 $1.5\alpha_0$ とする。3節のDPアルゴリズムでは、毎期にパラメータの推定と価格の最適化を行うが、ここでは計算時間の短縮のためにエピソードの終了時にのみパラメータの推定と価格の最適化を行う。推定と最適化の時期はパッシブ学習と同様であるが、アクティブ学習ではエピソード内に探索を行う点が異なる。エピソード k におけるリグレットの評価を以下の式を用いて評価する。

$$R_p(k) = 100 \times \frac{kV_1^{\theta,0}(C) - \sum_{s=1}^{T_k} p_s d_s}{kV_1^{\theta,0}(C)}. \quad (8)$$

100回のシミュレーションを実施し、評価値の平均値を用いてパッシブ学習とアクティブ学習を比較する。その平均値を $\bar{R}_p(k) = (1/n) \sum_{i=1}^n R_{p,i}(k)$ とする。ここで、 $R_{p,i}(k)$ は i 回目のシミュレーションの評価値である。

図3はそれぞれの学習を適用したときの各エピソード k におけるリグレット評価値の平均 $\bar{R}_p(k)$ を示している。エピソード数が増えるにつれて、リグレットの評価値が収束していることが確認できる。また、アクティブ学習の方が評価値が低いため、真のパラメータの下での総収益に近い結果が得られている。

図4は100エピソードにわたる販売価格のサンプル

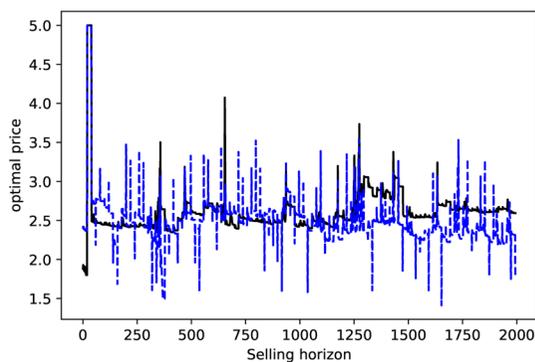


図4 最適価格のサンプルパス

パスを示している(1エピソードは20期間のため、横軸は0から2,000となる)。アクティブ学習では、価格の探索時に価格が大きく変動している。

4.2 上限価格の感度分析

4.1節より、アクティブ学習における探索は収益拡大に有効であるが、販売価格の変動は大きくなる。消費者が価格に不公平感を抱くのは、消費者が同じ時期に同じ商品・サービスに対して他者よりも高い価格を支払った場合に多いことが知られている[1, 18]。消費者の不公平感の高まりはロイヤルティの高い顧客を失う可能性があるため、(i) 基準価格および下限・上限価格の提示、(ii) 自社と他社の価格差に対応する同額保証や最低価格保証、(iii) 異時点間の価格差に対応する購入延期オプション、などの方策が実務において採用されている。ここでは、上限価格に着目し、価格変動とリグレットの関係を考察する。

表1はエピソードが $k = 100$ のときのリグレットの平均評価値 $\bar{R}_p(100)$ と需要パラメータの平均評価値 $\bar{\alpha}_p(100)$ 、 $\bar{m}_p(100)$ を示している。ここで、 k 回目のエピソードにおける需要パラメータの評価指標として、 $\alpha_p(k) = 100 \times (\alpha - \alpha_0)/\alpha_0$ 、 $m_p(k) = 100 \times (m - m_0)/m_0$ を用いた。 $\bar{\alpha}_p(k)$ および $\bar{m}_p(k)$ は100回のシミュレーションの平均値である。表1より、上限価格が低いとき ($\bar{p} = 4, 4.5$) は、パッシブ学習の方がアクティブ学習よりもリグレットが低いことがわかる。これは、上限価格が低いために探索する価格の範囲が狭く、パラメータの予測精度が低いことと、上限価格が低いために平均販売価格が低下することが起因している。一方、上限価格が高いとき ($\bar{p} = 6.5, 7$) もまたパッシブ学習の方がリグレットが低い。この場合には、価格の範囲が広いためにパラメータの予測精度は高まるが、価格が高いために販売の機会損失が生じるためである。以上より、アクティブ学習が収益改

表 1 顧客の到着情報がない場合

上限価格	パッシブ学習			アクティブ学習		
	\hat{R}_p	$\hat{\alpha}_p$	\hat{m}_p	\hat{R}_p	$\hat{\alpha}_p$	\hat{m}_p
4.0	4.33	22.75	22.03	4.35	22.94	21.64
4.5	4.14	21.08	20.18	4.42	22.01	20.63
5.0	4.48	20.22	19.84	4.05	20.32	19.46
5.5	4.82	19.90	20.46	4.61	18.39	17.63
6.0	4.07	22.29	22.05	3.25	15.29	15.40
6.5	4.25	18.33	18.35	4.44	18.57	18.22
7.0	4.82	20.64	20.78	5.24	19.67	19.43

表 2 顧客の到着情報がある場合

上限価格	パッシブ学習			アクティブ学習		
	\hat{R}_p	$\hat{\alpha}_p$	\hat{m}_p	\hat{R}_p	$\hat{\alpha}_p$	\hat{m}_p
4.0	0.724	2.629	1.044	0.874	2.609	1.048
4.5	0.776	2.603	1.046	0.907	2.649	1.048
5.0	0.754	2.596	1.046	0.873	2.613	1.049
5.5	0.784	2.589	1.046	0.858	2.645	1.048
6.0	0.788	2.600	1.046	0.910	2.678	1.048
6.5	0.781	2.595	1.046	0.956	2.711	1.048
7.0	0.790	2.596	1.046	0.951	2.676	1.049

善において有効となる上限価格の幅が存在することが確認できる。この範囲内で上限価格を適切に定め、消費者に公開することで、価格上昇の可能性に伴う消費者の不安の軽減と企業の収益増加の双方にとって有益であるといえる。

4.3 顧客の到着データの追加

表 1 より、学習方法によらず需要パラメータの評価値 $\hat{\alpha}_p$ および \hat{m}_p は 20% 前後であり、真の値と推定値の間に誤差が見受けられる。パラメータの推定に用いた販売データ $d_s = 0$ は第 s 期に購入がなかったことを表すが、このデータからは s 期に顧客が訪れなかったのか、顧客は訪れたが購入しなかったのかを判断することができない。もし、商品・サービスがオンラインで販売されており、顧客のアクセスログが利用可能であれば、顧客が購入に訪れたかどうかを正確に知ることができるため、需要パラメータをより精度良く推定できるだろう。そこで、本節では、顧客のアクセスの有無に関するデータを追加し、上限価格の感度分析を行う。

顧客が訪れた場合に $y_s = 1$ 、訪れない場合に $y_s = 0$ をとる成功率 m のベルヌーイ分布を考える。 s 期に $y_s = 0$ かつ $d_s = 0$ ならば、顧客が訪れなかったことがわかり、 $y_s = 1$ かつ $d_s = 0$ ならば、顧客が訪れたものの購入しなかったことがわかる。エピソード k の第 t 期までに得られる履歴 (p_s, d_s, y_s) , $s = 1, 2, \dots, T(k-1) + t - 1$ を用いると、対数尤度関数は

$$\begin{aligned} \log \mathcal{L}(\theta_s | \mathcal{D}_s) &= \log \left(\prod_{j=1}^s m^{\mathbb{1}_{\{y_j=1\}}} (1-m)^{\mathbb{1}_{\{y_j=0\}}} \right. \\ &\quad \left. \times [d(p)]^{\mathbb{1}_{\{y_j=1, d_j=1\}}} \cdot [1-d(p)]^{\mathbb{1}_{\{y_j=1, d_j=0\}}} \right) \end{aligned} \quad (9)$$

となる。したがって、最尤推定量は以下のように求められる。

$$\begin{aligned} \hat{m}_s &= \frac{1}{s-1} \sum_{j=1}^s y_j, \\ \hat{\alpha}_s &= \operatorname{argmax}_{\alpha} \sum_{j \in \{1, \dots, s\}} [d_t \log(d(p)) \\ &\quad + (1-d_t) \log(1-d(p))]. \end{aligned} \quad (10)$$

表 2 より、到着データの追加によって、需要パラメータの評価値が大きく改善されていることが見受けられる。それに伴い、リグレットの評価値も改善している。特に、上限価格の大きさによらずパッシブ学習の方がアクティブ学習よりもリグレットの評価値が低い。

5. おわりに

本稿では、レベニューマネジメントの最近のテーマである需要関数の学習を考慮したダイナミックプライシングについて解説した。いくつかの仮定の下で、アクティブ学習による探索が収益改善に有効であることをシミュレーションを通して確認した。また、web サイトへのアクセスデータなどの顧客の到着データが利用できる場合には探索を行わないパッシブ学習の方が有効であることを確認した。アクティブ学習による探索行動の結果、価格が大きく変動すると消費者の不公平感を招く恐れがあるため、パッシブ学習による DP は企業と消費者の双方にとって良い方策であるといえる。先行研究の多くはリグレット上界に関する研究に留まり、消費者視点からの研究はまだ十分にされていないといえる。このため、今後は消費者の不公平感を緩和するための方策に着目した研究の進展も必要であると思われる。

謝辞 本研究は JSPS 科研費基盤研究 (C)20K04976 の助成を受けたものである。

参考文献

[1] 佐藤公俊, 澤木勝茂, 『レベニューマネジメント—収益管理の基礎からダイナミックプライシングまで—』, 共立出版,

2020.

- [2] A. Kuyumcu, U. Yildirim, A. Hyde, S. Shan-berger, K. Hsiao, S. Donahoe, S. Wu, M. Murray and M. B. Maron, “Revenue management delivers significant revenue lift for Holiday Retirement,” *Interfaces*, **48**, pp. 7–23, 2018.
- [3] J. Guillen, P. Ruiz, U. Dellepiane, L. Maccarrone, R. Maccioni, A. Pinzuti and E. Procacci, “Europcar integrates forecasting, simulation, and optimization techniques in a capacity and revenue management system,” *INFORMS Journal on Applied Analytics*, **49**, pp. 40–51, 2019.
- [4] Y. Wang, K. McGuire, J. Terbush, M. Towns and C. K. Anderson, “Choice-based dynamic pricing for vacation rentals,” *INFORMS Journal on Applied Analytics*, **51**, pp. 450–462, 2021.
- [5] J. Beck, J. Harvey, K. Kaylen, C. Sala, M. Urban, P. Vermeulen, N. Wilken, W. Xie, D. Ilescu and P. Mital, “Carnival optimizes revenue and inventory across heterogenous cruise line brands,” *INFORMS Journal on Applied Analytics*, **51**, pp. 26–41, 2021.
- [6] G. Gallego and G. van Ryzin, “Optimal dynamic pricing of inventories with stochastic demand over finite horizon,” *Management Science*, **40**, pp. 1241–1259, 1994.
- [7] J. Broder and P. Rusmevichientong, “Dynamic pricing under a general parametric choice model,” *Operations Research*, **60**, pp. 965–980, 2012.
- [8] A. V. den Boer and B. Zwart, “Dynamic pricing and learning with finite inventories,” *Operations Research*, **63**, pp. 965–978, 2015.
- [9] O. Besbes and A. Zeevi, “Blind network revenue management,” *Operations Research*, **60**, pp. 1537–1550, 2012.
- [10] Z. Wang, S. Deng and Y. Ye, “Close the gaps: A learning-while-doing algorithm for single-product revenue management problems,” *Operations Research*, **62**, pp. 318–331, 2014.
- [11] O. Besbes and A. Zeevi, “Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms,” *Operations Research*, **57**, pp. 1407–1420, 2009.
- [12] K. J. Ferreira, D. Simchi-Levi and H. Wang, “Online network revenue management using Thompson sampling,” *Operations Research*, **66**, pp. 1586–1602, 2018.
- [13] X. Chen, Z. Owen, C. Pixton and D. Simchi-Levi, “A statistical learning approach to personalization in revenue management,” *Management Science*, **68**, pp. 1923–1937, 2021.
- [14] M. D. Wittman and P. P. Belobaba, “Dynamic pricing mechanisms for the airline industry: A definitional framework,” *Journal of Revenue and Pricing Management*, **18**, pp. 100–106, 2019.
- [15] K. Sato and K. Sawaki, “A continuous-time dynamic pricing model knowing the competitor’s pricing strategy,” *European Journal of Operational Research*, **229**, pp. 223–229, 2013.
- [16] K. Sato, “Price trends and dynamic pricing in perishable product market consisting of superior and inferior firms,” *European Journal of Operational Research*, **274**, pp. 214–226, 2019.
- [17] K. Sato, “Dynamic pricing with automated purchase-reservation algorithms,” *Journal of Revenue and Pricing Management*, **20**, pp. 33–41, 2021.
- [18] L. Xia and K. B. Monroe, “Is a good deal always fair? Examining the concepts of transaction value and price fairness,” *Journal of Economic Psychology*, **31**, pp. 884–894, 2010.