

Title	自動販売機の商品品揃え最適化モデルの研究
Author(s)	根本, 学
Citation	
Issue Date	2017-03
Type	Thesis or Dissertation
Text version	author
URL	<a href="http://hdl.handle.net/10119/14180">http://hdl.handle.net/10119/14180</a>
Rights	
Description	Supervisor:平石 邦彦, 情報科学研究科, 修士

修 士 論 文

自動販売機の商品品揃え最適化モデルの研究

北陸先端科学技術大学院大学  
情報科学研究科

根 本 学

平成 29 年 3 月

修 士 論 文

# 自動販売機の商品品揃え最適化モデルの研究

指導教官 平石 邦彦

審査委員主査 平石 邦彦

審査委員 緒方 和博

審査委員 白井 清昭

北陸先端科学技術大学院大学  
情報科学研究科

1410351 根本 学

提出年月: 平成 29 年 2 月

## 概要

日本における飲料用自動販売機ビジネスでは、オペレーター企業が自動販売機への補充業務を請負い、ルートマンと呼ばれる作業員が実際の補充作業や商品交換作業を担っている。オペレーター企業においては、ルートマンによる作業の効率化と標準化、商品の品揃え改善による売上拡大と機会損失の低減の実現が経営課題となっている。

本研究では、ルートマンの作業効率化と売上拡大を目的に、自動販売機の商品品揃え改善に対する最適化モデルを提案する。本モデルでは、自動販売機に対する顧客の商品選択行動をモデル化し、部分観測マルコフ決定過程 (POMDP) の枠組みの元、将来の期待報酬を考慮した最適な品揃え改善行動をルートマンに提案する。

仮想シミュレーションによる検証の結果、認知できる情報が多い状況下においては品揃え改善効果を確認することができた。一方で現実のオペレーション業務への適応には課題が残っていることも確認できた。

# 目次

第1章	はじめに	1
第2章	先行研究	3
2.1	消費者の商品選択行動に関する研究	3
2.2	自販機に関する研究	3
第3章	品揃え改善最適化モデルの枠組み	5
3.1	自販機状態のモデル化	5
3.2	消費者の商品選択モデル	5
3.2.1	消費者行動のモデル化	5
3.2.2	商品選択行動	6
3.3	購買シミュレーション	7
3.4	品揃え改善モデルの枠組み	9
3.4.1	強化学習	9
3.4.2	部分観測マルコフ決定過程 (POMDP)	10
3.4.3	モデル化の枠組みについて	10
第4章	品揃え改善最適化モデルの提案	12
4.1	購入数の確率分布	12
4.1.1	二項分布	12
4.1.2	ポアソン二項分布	13
4.2	商品の選択確率	13
4.3	POMDPによる品揃え改善モデルの詳細	13
4.3.1	構成要素の定義	14
4.3.2	観測確率	15
4.3.3	尤度	16
4.3.4	状態遷移確率	16
4.3.5	時間遷移に伴う信頼度の推定	16
4.3.6	商品交換方策の決定	17
第5章	モデルシミュレーション	21
5.1	提案モデルと比較対照モデル	21

5.2 シミュレーションパラメーター . . . . .	22
5.3 シミュレーション結果 . . . . .	22
5.4 考察 . . . . .	26
<b>第6章 まとめ</b>	<b>31</b>
<b>付録A ポアソン二項分布</b>	<b>32</b>

# 第1章 はじめに

## 背景

日本における自販機および自動サービス機（以下、単に自販機と呼ぶ）の普及台数は約500万台である。そのうち最も割合が大きいのは飲料用自販機であり、約255万台と全体の約51%を占めている。飲料用自販機の内訳としては、清涼飲料が最も多く台数ベースで全体の約86%を占めており、以下コーヒー・ココアなどのカップ式（7%）、牛乳（6%）、酒・ビール（1%）となっている [1]。

布川ら [2] は日本における自販機ビジネスの現状についてまとめている。特に、飲料用自販機ビジネスにおける供給側のステークホルダーとして、オペレーター企業と実際に作業を行うルートマンに注目して分析をしている。

オペレーター企業は、街中やオフィスに設置されている自販機への商品補充・保守運用業務を行っており、実際の作業はルートマンと呼ばれる作業員が行っている。ルートマンは各自の担当の自販機をまわって商品の補充や商品交換、ゴミの回収作業（以下、まとめて巡回作業と呼ぶ）などを行うが、補充のタイミングや商品の品揃え交換の対象に関して一定の範囲内での個人の裁量が認められている。

オペレーター企業においては、ルートマンが行う巡回作業の効率化と標準化、および商品の品揃え改善による売上拡大や売切れに伴う機会損失の低減が経営課題となっている。

このうち巡回作業の効率化と標準化については、業務システムの進歩により、システムが自動的に自販機の売上を予測しルートマンに対して適正な巡回タイミングを指示することによって、一定レベルの精度で実現化されつつある [3]。

一方で、商品の品揃え改善手法の効率化や標準化の取り組みはまだあまりなされていない。これは、「1台1台の売上が小さい」「商品の種類が多い」「商品売上の傾向は個々の自販機の設置状況や利用する顧客の属性や趣向によって大きく左右される」という飲料用自販機特有の性質による部分が多いと考えられる。

## 本研究の目的

以上のような背景を踏まえ、本研究では自販機商品の品揃えを改善し売上拡大を実現する最適化モデルの開発を行う。具体的には、以下の点について研究を行う。

- 自販機における顧客の商品選択行動をモデル化する。

- ルートマンに対して品揃え改善を提案し、売上拡大を実現できる品揃え改善モデルを開発する。
- 仮想の売上データを対象に品揃え改善モデルを適用したシミュレーションを実施し、モデルの効果を検証する。

## 研究の効果

この研究の成果を実際のオペレーション業務に適用することができるようになると、顧客・オペレーター企業、ルートマンの三者にとって以下のようなメリット・効果が期待される。

- 顧客：自分の好む商品を欲しい時に購入できる。欲しい商品の売切れが少なくなる。
- オペレーター企業：売上拡大および機会損失の低減といったサービスレベルの維持・向上が図られる。また、ルートマン作業の最適化や標準化といった効果も期待できる。
- ルートマン：改善提案に従うことで個人の経験や能力に依存しない一定レベルの作業を皆が行えるようになる。また失敗を減らすことができる。



## 第2章 先行研究

### 2.1 消費者の商品選択行動に関する研究

マーケティング分野における多種商品に対する消費者の商品選択行動に関する近年の代表的な研究としては、佐藤ら [4][5][6] の研究がある。小売店の販売実績データ（POS データ）を元に消費者の商品選択行動をベイジアンモデリングで形式化したものである。また藤田ら [7] は、飲食店を対象に、消費者の商品に対する効用をロジットモデルで形式化し、POS データから統計的手法で推計したパラメーターを用いて需要予測を試みている。他に、CVS 店舗を対象に、商品欠品時の消費者行動をマルチエージェントモデルでシミュレーションした松村ら [8] の研究がある。

いずれの研究も小売店・飲食店の販売実績データに基づく統計的なモデル構築とその実証に関する研究であるが、決まった品揃え（小売店における棚割り）に対する消費者の選択行動の予測に留まっている。店舗の売上拡大に向けた商品の品揃え改善にまで踏み込んだ研究はまだ少数である。

自販機は小売店とは異なり、販売実績が販売時点では分からない（巡回訪問時に初めてデータ取得ができる）ことや、設置環境が一台一台異なるため統一的なパラメーター設定が難しいことなどに特徴がある。また、品揃え交換のタイミングや対象商品の選定については、ルートマン自身の適切な判断が求められる。このため自販機の商品選択行動のモデル化にはこれらの先行研究とは異なるアプローチが必要である。

### 2.2 自販機に関する研究

自販機に関する研究はいくつかの分野で行われているが、ここでは主にオペレーター・ルートマンの業務における諸問題に関する研究を取り上げる。

第一に、複数の自販機への効率的な巡回作業を実現するための在庫配送計画問題に関する研究である。主にオペレーションズリサーチの側面から多くの研究がなされており、代表的なものとして宮本ら [10] がある。

第二に、自販機へのコラム最適化問題があげられる。コラムとは自販機内における商品の保管場所のことであり、限られたコラムへの適切な商品の割当てと在庫管理により、売上の拡大や売切れの削減が実現できる。自販機コラム割当て問題を組合せ最適化問題として定式化した宮本らの研究 [11][12] がある。これらの研究では総コスト最小化や売切れ削減を目的関数として整数計画問題を解く試みを行なっている。一方、消費者の購買需

要がポアソン過程に従うものとしてモデル化した竹内らの研究 [13] がある。最適化の目的関数として長期的な利益率を採用している点が宮本らとは異なる。また、Anupindi ら [9] は売切れが発生した際に消費者が代替商品を選択するような行動をモデル化した需要推定法を提案している。

このように、自販機のコラム最適化問題における既存研究では、一定環境のもとコスト最小化や利益最大化を目的とした最適化手法が提案されている。しかし、先述のように自販機は一台一台の設置状況が異なる上、季節や天候、消費者の嗜好により好まれる商品が常に変化するという特徴がある。したがって、オペレーター企業やルートマンの売上拡大や効率化の実現には、変動する設置環境や消費者需要に逐次対応できるようなコラム最適化（品揃え改善）手法の提案が求められる。

## 第3章 品揃え改善最適化モデルの枠組み

先行研究を参考に、設置環境や消費者需要の変化に対応し常に適切な品揃え改善案を提案できるような最適化モデルを提案する。また、提案モデル実証のための仮想シミュレーションの枠組みを検討した。

### 3.1 自販機状態のモデル化

まず、設置環境の変化のモデル化として、以下のような自販機状態と状態変化を仮定する。

- 自販機は時刻  $t$  においてある状態  $S_j$  にある。
- ルートマンが巡回するごとに、状態  $S_j$  からある遷移確率で他の状態  $S_{j'}$  に遷移する。
- 自販機状態は種々の属性で記述されるが、その属性が「固定／変化する」か、およびルートマンから「観測可能／観測不可能」かで分類した。本研究ではその中から以下の3つを自販機の属性として採用することとした。
  - － 固定で、観測可能なもの：設置場所 ... オフィス、屋外、学校の3パターン
  - － 変化し、観測可能なもの：外部要因 ... 気温、高／中／低の3パターン
  - － 変化し、観測不可能なもの：顧客属性 ... 男女比、8:2 / 5:5 / 2:8 の3パターン

自販機状態のモデル化イメージを図3.1に示した。

### 3.2 消費者の商品選択モデル

次に、自販機に対する消費者行動と商品選択行動を以下のようにモデル化した。

#### 3.2.1 消費者行動のモデル化

本研究では、自販機で商品を購入しようとする消費者（顧客）の行動を以下のようにモデル化する。

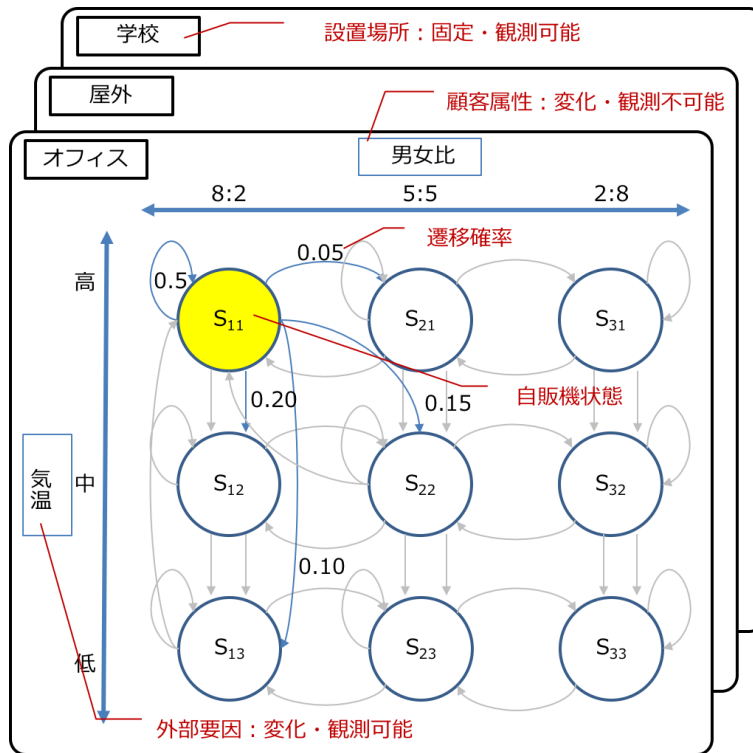


図 3.1: 自販機状態

- 顧客の属性は性別のみ（男性または女性）とする。
- 顧客は自販機で商品を購入する際に、購入しようとする商品の一つを選択する。
- 購入しようとする商品は、顧客の性別や設置場所、気温によって確率的に変動する。
- 購入しようとする商品が品揃えにあり、かつ在庫もある場合、実際に購入する。
- 購入しようとする商品が品揃えになかったり、あっても売切れであったりした場合、購入しない（代替商品の購入はしない）。

### 3.2.2 商品選択行動

本研究では、顧客の商品選択行動に多項ロジットモデル [4] を採用した。

#### 多項ロジットモデル

多項ロジットモデルは、複数の選択可能な商品がある状態において消費者がそのうち一つを選択する確率を商品の効用値のロジットを用いて表すモデルである。即ち、消費者は各商品に対する効用値から導かれる選択確率に従って、購入しようとする商品を決定す

る。商品ごとの効用値は、顧客の性別と自販機の3状態に依存するパラメーターで記述するものとした。

多項ロジットモデルを用いると、ある顧客  $k$  の商品  $A$  に対する効用値は以下のように定式化できる。

$$V_{Ajk} = V_A^0 + S_k^M (V_A^M + \beta_A^M T_j) + S_k^F (V_A^F + \beta_A^F T_j) \quad (3.1)$$

ここで、

$S_k^M / S_k^F$  :  $k$  が男性 / 女性の場合=1、女性 / 男性の場合=0

$V_A^0$  : 効用値定数 (男女共通)

$V_A^M / V_A^F$  : 効用値定数 (男性 / 女性別項)

$\beta_A^M / \beta_A^F$  : 気温の効用値への気温寄与係数 (男性 / 女性)

$T_j$  : 自販機状態  $S_j$  における気温パラメーター。

高 (*high*)  $\rightarrow 1$ , 中 (*middle*)  $\rightarrow 0$ , 低 (*low*)  $\rightarrow -1$

である。例えば自販機で購入可能な商品が  $A_1, A_2, \dots, A_M$  の  $M$  種類であるとき、商品  $A_i (1 \leq i \leq M)$  の選択確率は以下のように表せる。

$$P_{A_ijk} = \frac{\exp(V_{A_ijk})}{\sum_{m=1}^M \exp(V_{A_mjk})} = \frac{\exp(V_{A_ijk})}{\exp(V_{A_1jk}) + \exp(V_{A_2jk}) + \dots + \exp(V_{A_Mjk})} \quad (3.2)$$

この定式化において、現在品揃えにない商品についても効用値および選択確率を設定することが特徴である。これにより、消費者が選択した商品が品揃えになれば商品を購入しない、という事象を再現することが可能となる。

### 3.3 購買シミュレーション

前節までの自販機および消費者行動のモデル化に従い、ある自販機状態における顧客の商品選択購買シミュレーションを形式化した。

まず、本研究において以下の前提を仮定する。

#### 商品

1台の自販機に補充可能な商品はA~Jの10種類とする。

#### 商品コラム

1台の自販機には、商品を補充可能な商品コラムが6コラムのみ存在する。1コラムには、商品の種類によらず20本の在庫が補充可能である。ただし、1種類の商品を複数のコラムに割り当てることは可能とする (複数コラム商品)。

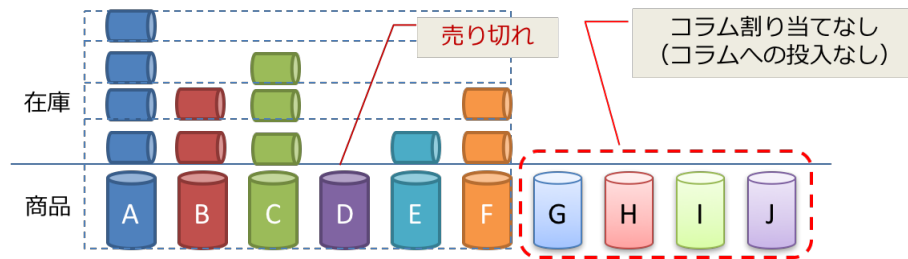


図 3.2: コラム割当てと商品在庫

### 購買シミュレーション

まず自販機状態  $S_j$  において、ルートマンは自販機への品揃えを決定する。例えば A~J のうち A~F の 6 商品を選択したものとする。自販機へは A~F の 6 商品を 1 コラムずつ、20 本ずつ補充する。G~J は補充しない。各商品の在庫は、

$$Z_A = \dots = Z_F = 20, Z_G = \dots = Z_J = 0$$

である。

次に、ルートマンが商品を補充してから次の補充を行うまでを 1 回の販売期間と考え、以下のフローに従ってシミュレーションを実行する。

1. 自販機状態をある遷移確率のもと遷移させる。遷移後の状態  $S_j$  に従って男女比・気温を決定する（設置場所は固定）。
2. 販売期間の間に顧客  $N$  人が1人ずつ自販機を訪れ商品の購入を試みるものと考ええる。
3. 自販機の販売数  $R = 0$  とする。
4.  $k(1 \leq k \leq N)$  番目の顧客の性別を、 $S_j$  における男女比 ( $p_j^M : p_j^F$ ) を元にランダムに決定する。
5. 決定した性別と  $S_j$  における設置場所・気温から、顧客  $k$  の商品 A ~ J に対する効用値  $V_{Ajk} \sim V_{Jjk}$  を計算する。
6. 効用値から選択確率  $P_{Ajk} \sim P_{Jjk}$  を求め、それに従ってランダムに購入商品 ( $X$ ) を決定する。
7.  $X$  が品揃えにあり、かつ在庫  $Z_X > 0$  場合：顧客  $k$  が商品  $X$  を購入したものとする。自販機の販売数  $R \leftarrow R + 1$ 、 $Z_X \leftarrow Z_X - 1$  とする。
8.  $X$  が品揃えにない、または品揃えにあっても在庫  $Z_X = 0$  場合：顧客  $k$  は何も購入しない。
9.  $k \leftarrow k + 1$  とし、 $k \leq N$  である限り、4に戻って繰り返す。
10.  $N$  人の購入が終了した後、今回の補充作業における報酬  $r =$  販売数  $R$  である。

図 3.2 はこのシミュレーションの途中における商品在庫のイメージである。

## 3.4 品揃え改善モデルの枠組み

前節までに定義した購買シミュレーションを対象として、ルートマンが取るべき最適な品揃え改善を実現できるようなモデルを提案する。本研究では、変化する自販機状態に逐次対応できること、および自販機状態の一部属性が観測できない点に着目し、強化学習と部分観測マルコフ決定過程 (POMDP) を最適化モデルの枠組みとして採用した。

### 3.4.1 強化学習

機械学習の一手法で、学習主体者（エージェント）が環境との相互作用により情報の獲得と行動を選択する学習法である [14][15]。エージェントは、(1) 環境から状態  $s$  を観測

し、(2) $s$ に基づいて行動 $a$ を起こし、(3)その結果として状態 $s'$ へ遷移するとともに、報酬 $r$ を得るといったサイクルを繰り返しながら学習とタスクを実行する。環境の状態遷移と、観測値・エージェントの行動と報酬との関係を図3.3に示した。

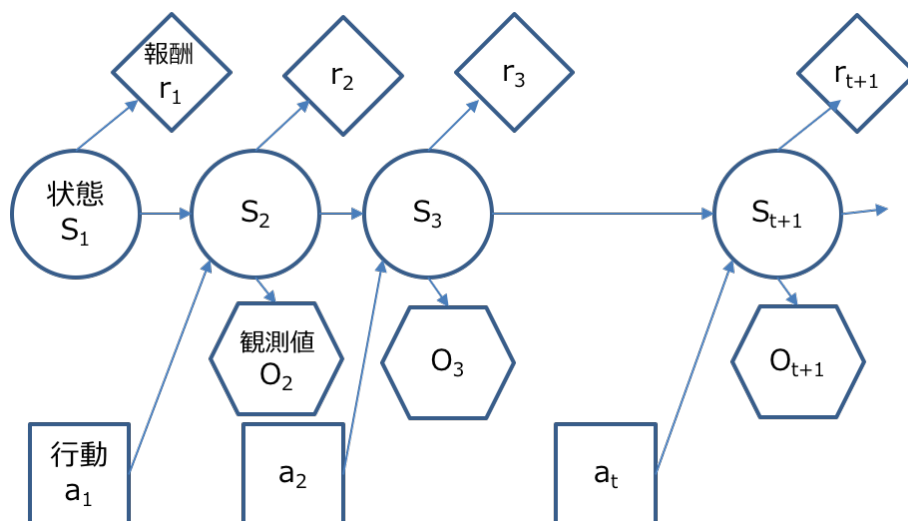


図 3.3: 強化学習による状態変化とエージェントの関係

### 3.4.2 部分観測マルコフ決定過程 (POMDP)

部分観測マルコフ決定過程 (Partially Observable Markov Decision Process: POMDP) とは、マルコフ決定過程 (MDP) を、エージェントの状態観測の不確実性を付加することで拡張したものである [16]。強化学習の枠組みと組み合わせることで不確実な環境下における学習と制御を実現する手法として、音声認識や制御工学の分野などにおいて活用が研究されている [17][18]。

### 3.4.3 モデル化の枠組みについて

品揃え改善モデルの実装に向け、今回の課題に強化学習や POMDP を適用する枠組みについて整理した。

これまでの定式化・モデル化において、シミュレーションの過程を左右するパラメーターが明らかになっている。つまり、大きく分けて「A：商品選択モデルのパラメーター」と「B：自販機状態の遷移確率」の2つである。

これらの情報をルートマンが認知できるか否かは、最適化モデルの成否に大きく関わってくる。例えば、Aのパラメーターを知っている場合、どの商品がどの程度売れるかを把握することができる。またBの遷移確率を認知している場合、現在の状態から次にどの状態へ遷移するかを推定することができ、その状態に合わせた品揃えが可能となる。しか



しいずれも現実のルートマンには直接的に知り得ない情報であり、過去の売上データなどから推察するしかない。

そこで、A や B の情報それぞれについてルートマンが「わかっている／わかっていない」かで場合分けを行い、その組み合わせによってモデルに必要な手法について整理した(表 3.1)。

表中のうち、今回の研究では、情報が最も多く通常の POMDP として形式化できる①の枠組みを対象に品揃え改善最適化モデルの構築を行うこととした。

一部の情報を欠く②・③・④は今後の研究課題となるが、それらのモデル構築時は実データを用いたり実装済みの①と比較するなどしてモデル評価の検証を行うこととなる。

		A：商品選択モデルのパラメーター	
		わかっている	わかっていない
B：自販機状態の遷移確率	わかっている	① 通常の POMDP として形式化	③ 多項ロジットのパラメーター推定 + 強化学習
	わかっている	② POMDP + 強化学習	④ ②+③の枠組みで実現

表 3.1: 品揃え改善最適化モデルの枠組み

## 第4章 品揃え改善最適化モデルの提案

本章では品揃え改善最適化モデルの具体的な定式化を行う。

### 4.1 購入数の確率分布

#### 4.1.1 二項分布

顧客数が  $N$ , 男女比が  $p^M : p^F, p^M + p^F = 1$  のとき、各商品の購入確率を男女別に以下のように定義する。

$$\begin{aligned} \text{A を買う確率： 男性} \cdots P_A^M, \text{ 女性} \cdots P_A^F \\ \text{B を買う確率： 男性} \cdots P_B^M, \text{ 女性} \cdots P_B^F \\ \vdots \\ \text{J を買う確率： 男性} \cdots P_J^M, \text{ 女性} \cdots P_J^F \end{aligned}$$

ここで、ある顧客が購入する商品 A~J を選択する事象は互いに独立であると仮定する。これにより、

$$P_A^M + P_B^M + \cdots + P_J^M = 1, P_A^F + P_B^F + \cdots + P_J^F = 1 \quad (4.1)$$

である。

以降、10種類の商品を代表して商品 A の購入確率について定式化を行う。B~J についても同様である。

もし商品 A の購入確率が男女で等しい場合、つまり  $P_A^M = P_A^F = P_A$  の場合、確率変数  $X_A$  を商品 A の購入数と定義すると、 $N$  人の顧客により商品 A が  $l$  個購入される確率  $\Pr(X_A = l)$  は、二項分布に従う。

$$\Pr(X_A = l) = {}_N C_l P_A^l (1 - P_A)^{N-l} \quad (4.2)$$

$$\text{期待値： } E[X_A] = NP_A, \text{ 分散： } \sigma_A^2 = NP_A(1 - P_A) \quad (4.3)$$

### 4.1.2 ポアソン二項分布

一方、商品 A の購入確率が顧客 1 人 1 人異なる場合を考える。つまり、 $N$  人の顧客の購入確率  $P_{A1}, P_{A2}, \dots, P_{AN}$  が一定でない場合であるが、このとき  $\Pr(X_A = l)$  はポアソン二項分布に従う（付録 A 参照）。

本研究では、購入確率は顧客の性別により決定される、と仮定している。付録 A より、男性顧客数： $N^M = Np^M$ 、女性顧客数： $N^F = Np^F$  における、商品 A に対するポアソン二項分布は以下ようになる。

$$\Pr(X_A = l) = \begin{cases} (1 - P_A^M)^{N^M} \times (1 - P_A^F)^{N^F}, & l = 0 \\ \frac{1}{l} \sum_{i=1}^l (-1)^{i-1} \Pr(X_A = l - i) T(i), & l > 0 \end{cases} \quad (4.4)$$

$$T(i) = N^M \left( \frac{P_A^M}{1 - P_A^M} \right)^i + N^F \left( \frac{P_A^F}{1 - P_A^F} \right)^i \quad (4.5)$$

$$\text{期待値} : E[X_A] = E_A^M + E_A^F = N^M P_A^M + N^F P_A^F \quad (4.6)$$

$$\text{分散} : \sigma_A^2 = \sigma_A^{M^2} + \sigma_A^{F^2} = N^M P_A^M (1 - P_A^M) + N^F P_A^F (1 - P_A^F) \quad (4.7)$$

## 4.2 商品の選択確率

(3.1) より、顧客  $k$ 、時刻  $t$  における、商品 A に対する効用値は、

$$k \text{ が男性の場合} : V_{A_t}^M = V_A + V_A^M + \beta_A^M T_t \quad (4.8)$$

$$k \text{ が女性の場合} : V_{A_t}^F = V_A + V_A^F + \beta_A^F T_t \quad (4.9)$$

となり、商品 A の時刻  $t$  における選択確率は、

$$k \text{ が男性の場合} : P_{A_t}^M = \frac{\exp(V_{A_t}^M)}{\exp(V_{A_t}^M) + \exp(V_{B_t}^M) + \dots + \exp(V_{J_t}^M)} \quad (4.10)$$

$$k \text{ が女性の場合} : P_{A_t}^F = \frac{\exp(V_{A_t}^F)}{\exp(V_{A_t}^F) + \exp(V_{B_t}^F) + \dots + \exp(V_{J_t}^F)} \quad (4.11)$$

となる。

## 4.3 POMDP による品揃え改善モデルの詳細

Kaelbling ら [16] に従い、本研究における POMDP による品揃え改善モデルの詳細について記述する。

### 4.3.1 構成要素の定義

#### 状態

時刻  $t$  における自販機状態:  $s_t = \{ \text{男女比}, \text{気温} \} \in S$ ,  $S$ : 存在しうる状態集合とする。  
3.1 節の仮定により、

- 男女比は  $\{8:2, 5:5, 2:8\}$  の 3 通り。時刻  $t$  においてルートマンには未知。
- 気温は  $\{high, middle, low\}$  の 3 通り。時刻  $t$  においてルートマンには既知。

である。よって、状態集合  $S$  には  $3 \times 3 = 9$  通りの状態が存在することになる。

#### 行動

時刻  $t$  における商品変更後の品揃えを、行動  $a_t = \{t \text{ における品揃え} \}$  とする。自販機が 6 コラムであるため、10 種類の商品から 6 商品を選ぶ品揃えの組み合わせは  ${}_{10}C_6 = 210$  通り。複数コラム商品を許す場合は  ${}_{10}H_6 = 5005$  通りである。

#### 方策

時刻  $t$  における品揃え、つまり  $a_{t-1} \rightarrow a_t$  の変更方針を商品交換方策と呼び、 $\pi^i$  で表す。各時刻において、商品交換方策は表 4.1 に掲げる方策集合  $\Pi$  の 8 方策のうちいずれか一つから選択するものとする。

#### 観測値

時刻  $t$  における商品別売上を、POMDP における観測値  $o_t = \{t \text{ における商品別売上} \}$  とする。

#### 報酬

時刻  $t$  における商品別売上の合計を、POMDP における報酬  $r_t = \{t \text{ における商品別売上の合計} \}$  とする。

表 4.1: 方策集合 II

方策	内容
$\pi^0$	何もしない
$\pi^1$	コラム内商品のうち最も効用値の低い商品と、コラム外商品のうち最も効用値の高い商品を入れ替える
$\pi^2$	コラム内商品のうち最も効用値の低い商品と、コラム外商品のうちランダムで一商品を入れ替える
$\pi^3$	コラム内商品のうちランダムで一商品と、コラム外商品のうち最も効用値の高い商品を入れ替える
$\pi^4$	コラム内商品のうちランダムで一商品と、コラム外商品のうちランダムで一商品を入れ替える
$\pi^5$	一コラム商品のうち最も効用値の高い商品を一コラム増やし、コラム内商品のうち最も効用値の低い商品を撤去する
$\pi^6$	複数コラム商品の内、最も効用値の低い商品を一コラム減らし、コラム外商品のうち最も効用値の高い商品と入れ替える
$\pi^7$	コラム内商品のうち最も効用値の低い2商品と、コラム外商品のうち最も効用値の高い2商品を入れ替える

## 信頼度

時刻  $t$  において、ルートマンが現在の自販機状態が  $s_t$  であると推定している割合を信頼度と呼び、

$$b_t(s_t) = b_t(s_t = \{ \text{男女比, 気温} \}) \quad (4.12)$$

とする。このとき、全ての状態に対する信頼度の合計は1となる。つまり、

$$\sum_{s_t \in S} b_t(s_t) = 1 \quad (4.13)$$

### 4.3.2 観測確率

時刻  $t$  における観測値  $o_t$  を得る確率  $O$  は、時刻  $t-1$  における品揃え  $a_{t-1}$  と時刻  $t$  における自販機状態に依存する。

$$O(s_t, a_{t-1}, o_t) = O(s_t = \{ \text{男女比, 気温} \}, a_{t-1} = \{ t-1 \text{ における品揃え} \}, o_t = \{ \text{商品別売上} \}) \quad (4.14)$$

これは、 $t-1$  における品揃え  $a_{t-1}$  と  $t$  における状態  $s_t$  である条件下で  $o_t$  を得る条件付き確率に等しい。つまり、

$$O(s_t, a_{t-1}, o_t) = \Pr(o_t | s_t, a_{t-1}) \quad (4.15)$$

### 4.3.3 尤度

各時刻において得られた観測値から、自販機状態に対する尤度を計算できる。例えば時刻  $t$  において、状態  $s_t = \{8 : 2, high\}$  で商品別売上（商品 A:  $X_A = 20$ , 商品 B:  $X_B = 10$ , 商品 C:  $X_C = 15$ , ..., 商品 F:  $X_F = 7$ ）を得る尤度は、以下のようなになる。

$$\begin{aligned} L(s_t, o_t) &= L_t(s_t = \{8 : 2, high\}, o_t = \{X_A = 20, X_B = 10, X_C = 15, \dots, X_F = 7\}) \\ &= \Pr^{\{8:2,high\}}(X_A = 20) \times \Pr^{\{8:2,high\}}(X_B = 10) \times \Pr^{\{8:2,high\}}(X_C = 15) \\ &\quad \times \dots \times \Pr^{\{8:2,high\}}(X_F = 7) \end{aligned} \quad (4.16)$$

### 4.3.4 状態遷移確率

時刻  $t \rightarrow t + 1$  における状態遷移確率を以下のように表す。

$$\begin{aligned} T(s_{t+1}|s_t, a_t) &= T(s_{t+1}|s_t) (\because \text{品揃えは状態遷移に寄与しない}) \\ &= T(s_{t+1} = \{\text{男女比, 気温}\} | s_t = \{\text{男女比, 気温}\}) \\ &= T(s_{t+1} = \{\text{男女比}\} | s_t = \{\text{男女比}\}) \times T(s_{t+1} = \{\text{気温}\} | s_t = \{\text{気温}\}) \end{aligned} \quad (4.17)$$

ここで、男女比と気温の状態遷移は独立であると仮定している。

### 4.3.5 時間遷移に伴う信頼度の推定

時刻  $t$  における状態  $s_t$  に対し、 $s_t = s, s \in S$  である全ての  $s$  に対する信頼度  $b_t(s)$  が得られていたとき、時刻  $t + 1$  における状態  $s_{t+1}$  がある状態  $s'$  である信頼度は、 $t + 1$  における観測値  $o_t$  をもとに以下の遷移式で求められる。

$$\begin{aligned} b_{t+1}(s') &= \Pr(s'|o_t, a_t, b_t) \\ &= \frac{\Pr(o_t|s', a_t, b_t) \Pr(s'|a_t, b_t)}{\Pr(o_t|a_t, b_t)} \\ &= \frac{\Pr(o_t|s', a_t) \sum_{s \in S} \Pr(s'|a_t, b_t) \Pr(s|a_t, b_t)}{\Pr(o_t|a_t, b_t)} \\ &= \frac{O(s', a_t, o_t) \sum_{s \in S} T(s'|s) b_t(s)}{\Pr(o_t|a_t, b_t)} \end{aligned} \quad (4.18)$$

また時刻  $t$  において観測が得られているとき、(4.18) を逐次適用することにより、時刻  $t + n (n \geq 0)$  における状態が  $s'$  である信頼度を推定することができる。

$$\begin{aligned}
b_{t+n}(s') &= \sum_{s \in S} T(s'|s) b_{t+n-1}(s) \\
&= \sum_{s \in S} \{T(s'|s)\}^2 b_{t+n-2}(s) \\
&\vdots \\
&= \sum_{s \in S} \{T(s'|s)\}^n b_t(s) \\
&= \sum_{s \in \{\text{男女比}\}} \{T(s'|s)\}^n b_t(s) \quad (\because \text{時刻 } t \text{ での状態 } s \text{ のうち } \{\text{気温}\} \text{ は既知})
\end{aligned} \tag{4.19}$$

#### 4.3.6 商品交換方策の決定

各時刻において、得られた観測値から推定した信頼度より商品交換方策を決定する。

以降本節では、時刻  $t = 0$  における観測値と品揃えにより時刻  $t = 1$  で得られた観測値を前提として、 $t = 1$  における商品交換方策の決定プロセスを例に説明を行う。

1.  $t = 0$  で 気温 = *middle* を観測

$b_0(s_0)$ :

気温は既知なので、未知である男女比の3状態に対する信頼度を推定する。 $t = 0$  では何も情報がないため、3つの男女比の状態について同一の信頼度を推定する。つまり、

$$b_0(s_0 = \{8 : 2, middle\}) = b_0(s_0 = \{5 : 5, middle\}) = b_0(s_0 = \{2 : 8, middle\}) = \frac{1}{3} \tag{4.20}$$

2.  $t = 1$  で 気温 = *high*、観測値 (= 売上)  $o_1$  を得る。

$b_1(s_1)$ :

男女比の3状態について推定する。(4.18) の遷移式に (4.20) を適用して、

$$\begin{aligned}
b_1(s_1 = \{8 : 2, high\}) &= \frac{L_1(s_1 = \{8 : 2, high\}, o_1) \sum_{s \in S} T(s_1 = \{8 : 2, high\} | s) b_0(s)}{\sum_{s' \in S} b_1(s')} \\
&= \frac{L(s_1 = \{8 : 2, high\}, o_1)}{\sum_{s' \in S} b_1(s')} \\
&\quad \times \{T(\{8 : 2, high\} | \{8 : 2, middle\}) \times b_0(\{8 : 2, middle\}) \\
&\quad + T(\{8 : 2, high\} | \{5 : 5, middle\}) \times b_0(\{5 : 5, middle\}) \\
&\quad + T(\{8 : 2, high\} | \{2 : 8, middle\}) \times b_0(\{2 : 8, middle\})\}
\end{aligned} \tag{4.21}$$

$$\begin{aligned}
b_1(s_1 = \{5 : 5, high\}) &= \frac{L_1(s_1 = \{5 : 5, high\}, o_1) \sum_{s \in S} T(s_1 = \{5 : 5, high\} | s) b_0(s)}{\sum_{s' \in S} b_1(s')} \\
&= \frac{L(s_1 = \{5 : 5, high\}, o_1)}{\sum_{s' \in S} b_1(s')} \\
&\quad \times \{T(\{5 : 5, high\} | \{8 : 2, middle\}) \times b_0(\{8 : 2, middle\}) \\
&\quad \quad + T(\{5 : 5, high\} | \{5 : 5, middle\}) \times b_0(\{5 : 5, middle\}) \\
&\quad \quad + T(\{5 : 5, high\} | \{2 : 8, middle\}) \times b_0(\{2 : 8, middle\})\}
\end{aligned} \tag{4.22}$$

$$\begin{aligned}
b_1(s_1 = \{2 : 8, high\}) &= \frac{L_1(s_1 = \{2 : 8, high\}, o_1) \sum_{s \in S} T(s_1 = \{2 : 8, high\} | s) b_0(s)}{\sum_{s' \in S} b_1(s')} \\
&= \frac{L(s_1 = \{2 : 8, high\}, o_1)}{\sum_{s' \in S} b_1(s')} \\
&\quad \times \{T(\{2 : 8, high\} | \{8 : 2, middle\}) \times b_0(\{8 : 2, middle\}) \\
&\quad \quad + T(\{2 : 8, high\} | \{5 : 5, middle\}) \times b_0(\{5 : 5, middle\}) \\
&\quad \quad + T(\{2 : 8, high\} | \{2 : 8, middle\}) \times b_0(\{2 : 8, middle\})\}
\end{aligned} \tag{4.23}$$

### 3. $t = 1$ における商品交換方策の決定

$t = 1$  における商品交換方策  $\pi_1^n (n = 0, 1, \dots, 7) \in \Pi$  を、将来得られる期待報酬を元に決定する。まず、 $\pi_1^n : a_0 \rightarrow a_1^n$  を選択した時の  $t = 2$  における報酬  $r_2$  の期待値を算出する。

(例)  $\pi_1^1 : a_0 \rightarrow a_1^1$  を選択した場合

- $t = 2$  における期待報酬は、(4.19) を用いて、

$$\begin{aligned}
E_{\pi_1^1}[r_2] &= \sum_{s' \in \{\text{男女比, 気温}\}} b_2(s') r_2(s', a_1^1) \\
&= \sum_{s' \in \{\text{男女比, 気温}\}} \left\{ \sum_{s \in \{\text{男女比}\}} T(s' | s) b_1(s) \right\} r_2(s', a_1^1)
\end{aligned} \tag{4.24}$$

ここで、

$o_2(s', a_1^1)$  : 状態  $s'$ , 品揃え  $a_1^1$  における商品別の売上期待値  $E_{s', a_1^1}[X_A], \dots, E_{s', a_1^1}[X_J]$  より算出 (4.25)

$r_2(s', a_1^1)$  :  $o_2(s', a_1^1)$  より算出 (4.26)  
 $\Rightarrow r_2(s', a_1^1) = E_{s', a_1^1}[X_A] + \dots + E_{s', a_1^1}[X_J]$



- 次に、 $t = 2$  で  $\pi_2^i (i = 0, 1, \dots, 7)$  を選択した場合を考える。即ち、品揃え  $a_{1,2}^{1,i}$

$$\pi_1^1 \rightarrow \pi_2^i : a_0 \rightarrow a_1^1 \rightarrow a_{1,2}^{1,0}, a_{1,2}^{1,1}, \dots, a_{1,2}^{1,7} \quad (4.27)$$

における報酬  $r_3$  の期待値を各々算出する。 $t = 2$  の時と同様に、

$$\begin{aligned} E_{\pi_1^1, \pi_2^i}[r_3] &= \sum_{s' \in \{\text{男女比, 気温}\}} b_3(s') r_3(s', a_{1,2}^{1,i}) \\ &= \sum_{s' \in \{\text{男女比, 気温}\}} \left\{ \sum_{s \in \{\text{男女比, 気温}\}} T(s'|s) b_2(s) \right\} r_3(s', a_{1,2}^{1,i}) \\ &= \sum_{s' \in \{\text{男女比, 気温}\}} \left\{ \sum_{s \in \{\text{男女比}\}} \{T(s'|s)\}^2 b_1(s) \right\} r_3(s', a_{1,2}^{1,i}) \end{aligned} \quad (4.28)$$

$r_3(s', a_{1,2}^{1,i})$  は  $o_3(s', a_{1,2}^{1,i})$  より計算できる。 $E_{\pi_1^1, \pi_2^i}[r_3]$  は方策  $\pi_2^i$  の選び方により 8 通りである。

- 以降同様に、 $t = 4, 5$  における報酬の期待値を計算する。即ち、 $t = 3$  における方策  $\pi_3^j$  による  $t = 4$  のときの期待値  $E_{\pi_1^1, \pi_2^i, \pi_3^j}[r_4]$  ( $8^2$  通り)、 $t = 4$  における方策  $\pi_4^k$  による  $t = 5$  のときの期待値  $E_{\pi_1^1, \pi_2^i, \pi_3^j, \pi_4^k}[r_5]$  ( $8^3$  通り) である。
- 以上より、時刻  $t = 1$  において方策  $\pi_1^1$  を選択した場合に期待される将来の最大報酬が計算できる。即ち、 $t = 5$  における期待報酬が最大になるような方策を  $t = 4$  において選択し、 $t = 4$  における期待報酬が最大になるような方策を  $t = 3$  において選択し、 $t = 3$  における期待報酬が最大になるような方策を  $t = 2$  において選択した場合の期待報酬である。つまり、方策  $\pi_1^1$  を選択した場合の期待報酬  $V_{t=1}(\pi_1^1)$  は、

$$\begin{aligned} V_{t=1}(\pi_1^1) &= E_{\pi_1^1}[r_2] + \gamma \max_{\pi_2^i \in \Pi} \left\{ E_{\pi_1^1, \pi_2^i}[r_3] \right. \\ &\quad \left. + \gamma \max_{\pi_3^j \in \Pi} \left\{ E_{\pi_1^1, \pi_2^i, \pi_3^j}[r_4] + \gamma \max_{\pi_4^k \in \Pi} \left\{ E_{\pi_1^1, \pi_2^i, \pi_3^j, \pi_4^k}[r_5] \right\} \right\} \right\} \end{aligned} \quad (4.29)$$

となる。ここで、将来の期待報酬に対しては割引率  $\gamma (0 < \gamma < 1)$  を掛けることにより、より直近の期待報酬を重視するようにした。

- 上記のような手順で、方策  $\pi_1^0, \dots, \pi_1^7$  による期待平均報酬  $V_{t=1}(\pi_1^0), \dots, V_{t=1}(\pi_1^7)$  を算出する。
- $V_{t=1}(\pi_1^n)$  が最大となる方策を  $t = 1$  における商品交換方策  $\pi_1^n$  として決定する。

4. 時刻  $t = 2$  における観測値を得る。  $t = 1$  と同様に、
- 信頼度  $b_2(s)$  ( $\{ \text{気温} \}$  が既知なので3通り) を算出する。
  - $t = 2, 3, 4, 5$  における方策  $\pi_2^n, \pi_3^i, \pi_4^j, \pi_5^k$  を想定し、これに基づく期待平均報酬  $V_{t=2}(\pi_2^n), n = 0, \dots, 7$  を算出する。
  - $V_{t=2}(\pi_2^n)$  が最大となる方策を  $t = 2$  における商品交換方策  $\pi_2^n$  として決定する。
5. 以降、時刻  $t$  ごとに同様に繰り返す。

# 第5章 モデルシミュレーション

前章までの枠組みやモデル化をベースとし、複数のモデルを対象として仮想パラメータによる品揃え改善モデルの数値シミュレーションを実施する。

## 5.1 提案モデルと比較対照モデル

今回の研究での提案モデルについて複数のパターンでシミュレーションを実施し、精度と有効性検証を行う。また評価の比較対照となるようなモデルも用意し同様のシミュレーションを実施、提案モデルとの比較を行う。

### 提案モデル

期待報酬の算出式 (4.29) において、いくつの時間ステップ先まで報酬の期待値を和に含めるかで複数のモデルを設定した。(4.29) では、 $t = 1$  における期待報酬  $V_{t=1}(\pi_1^1)$  に対して4つ先の時間ステップ  $t = 5$  までの期待報酬を算出し和を求めているが、推定する将来の時間ステップ数を減らして和をとったものも設定し別の提案モデルとした。つまり以下の4モデルである。

モデル1 1つ先の時間ステップでの期待報酬  $E_{\pi_1^1}[r_2]$  までを和に含める。

モデル2 2つ先の時間ステップでの期待報酬  $E_{\pi_1^1, \pi_2^i}[r_3]$  までを和に含める。

モデル3 3つ先の時間ステップでの期待報酬  $E_{\pi_1^1, \pi_2^i, \pi_3^j}[r_4]$  までを和に含める。

モデル3 4つ先の時間ステップでの期待報酬  $E_{\pi_1^1, \pi_2^i, \pi_3^j, \pi_4^k}[r_5]$  までを和に含める。  
(式 (4.29) の通り)

### 比較対照モデル

比較対照モデルとして、以下の2つの単純でヒューリスティックな手法を採用した。

モデル0 初期の品揃え（商品 A~F の6種類）のまま一切変更しない。

モデル0' 毎回、方策  $\pi^1$  を行う。つまり (4.8), (4.9) における効用値の低い商品と高い商品を入れ替える。このとき、自販機状態としての男女比は常に初期状態 (5:5) であるとの仮定をとっている。

## 5.2 シミュレーションパラメーター

本研究において使用したシミュレーションパラメーターについて具体的な数値を示す。

### 共通パラメーター

全モデルに共通なパラメーターとして、顧客数  $N = 100$ 、時間ステップ数=100とした。前述の通り、設置場所はオフィス、屋外、学校の3種類である。また、(4.29)における割引率  $\gamma = 0.9$ とした。

### 状態遷移確率

男女比の遷移確率、および気温の遷移確率を表5.1~5.4に示した。

設置場所の特性を考慮して、屋外の場合は男女比の変動が大きくなるよう他の状態への遷移確率を高くした。一方、学校の場合は変動が小さいものと考え逆に現在の状態に留まる確率を高くした。またオフィスはその中間の値を採用してある。

### 商品別効用値パラメーター

商品の選択確率に関わる商品別効用値パラメーターを、表5.5~5.8に示した。商品A~Jには、想定する種類（コーヒーや緑茶など）とICE/HOTの区分を設定しており、各商品別のパラメーターの値にそのイメージを反映してある（コーヒーは男性の方が効用値が高い、気温が高いとICEの方がより売れる、など）。

表5.5は設置場所によらない共通パラメーターで、商品のベースの効用値に相当する。表5.6~5.8は設置場所の特性を考慮したパラメーター値とした。ここで、 $\beta_M / \beta_F$ は効用値への気温寄与係数であり、この値が正の場合、気温が高いとより売れる商品であることを表している。

以上の条件で、各モデルについて3種類の設置場所に対し10回ずつシミュレーションを実施した。

## 5.3 シミュレーション結果

比較対照モデル2モデル、提案モデル4モデルについて、設置場所別に10回ずつのシミュレーションを試行し、各々100時間ステップ分の売上の結果を得た。時間ステップごとの自販機状態と売上と売切れの経過の例を設置場所別に図5.1に示した。

各モデルについて、主に100時間ステップ全体における売上の平均と標準偏差、売切れの平均を集計し評価指標として比較した。各モデルの10回ずつの指標の分布を図5.2~

表 5.1: 男女比の遷移確率:オフィス

$T(s_{t+1} = \{\text{男女比}\}   s_t = \{\text{男女比}\})$	$s_{t+1} = \{8:2\}$	$s_{t+1} = \{5:5\}$	$s_{t+1} = \{2:8\}$
$s_t = \{8:2\}$	0.6	0.3	0.1
$s_t = \{5:5\}$	0.25	0.5	0.25
$s_t = \{2:8\}$	0.15	0.35	0.5

表 5.2: 男女比の遷移確率:屋外

$T(s_{t+1} = \{\text{男女比}\}   s_t = \{\text{男女比}\})$	$s_{t+1} = \{8:2\}$	$s_{t+1} = \{5:5\}$	$s_{t+1} = \{2:8\}$
$s_t = \{8:2\}$	0.45	0.35	0.2
$s_t = \{5:5\}$	0.3	0.4	0.3
$s_t = \{2:8\}$	0.2	0.3	0.5

表 5.3: 男女比の遷移確率:学校

$T(s_{t+1} = \{\text{男女比}\}   s_t = \{\text{男女比}\})$	$s_{t+1} = \{8:2\}$	$s_{t+1} = \{5:5\}$	$s_{t+1} = \{2:8\}$
$s_t = \{8:2\}$	0.8	0.15	0.05
$s_t = \{5:5\}$	0.15	0.7	0.15
$s_t = \{2:8\}$	0.1	0.15	0.75

表 5.4: 気温の遷移確率

$T(s_{t+1} = \{\text{気温}\}   s_t = \{\text{気温}\})$	$s_{t+1} = \{high\}$	$s_{t+1} = \{middle\}$	$s_{t+1} = \{low\}$
$s_t = \{high\}$	0.35	0.5	0.15
$s_t = \{middle\}$	0.2	0.6	0.2
$s_t = \{low\}$	0.25	0.45	0.3

表 5.5: 商品別効用値パラメーター：共通

商品	A	B	C	D	E
種類	コーヒー	コーヒー	カフェオレ	緑茶	紅茶
ICE/HOT	ICE	HOT	ICE	ICE	ICE
$V^0$	1.0	0.5	0.0	1.0	0.5

商品	F	G	H	I	J
種類	栄養ドリンク	炭酸	ミネラルウォーター	紅茶	スポーツドリンク
ICE/HOT	ICE	ICE	ICE	HOT	ICE
$V^0$	-0.5	0.5	1.0	-0.5	0.5

表 5.6: 商品別効用値パラメーター：オフィス

商品	A	B	C	D	E
種類	コーヒー	コーヒー	カフェオレ	緑茶	紅茶
ICE/HOT	ICE	HOT	ICE	ICE	ICE
$V^M$	1.0	0.5	-0.5	1.0	0.5
$V^F$	0.5	0.5	0.5	1.0	1.0
$\beta^M$	0.0	-0.5	0.0	0.0	0.0
$\beta^F$	0.5	-0.5	-0.5	0.0	0.5

商品	F	G	H	I	J
種類	栄養ドリンク	炭酸	ミネラルウォーター	紅茶	スポーツドリンク
ICE/HOT	ICE	ICE	ICE	HOT	ICE
$V^M$	0.0	0.5	0.0	-1.0	0.5
$V^F$	-1.0	0.0	0.0	0.5	0.0
$\beta^M$	0.0	1.0	0.0	-1.0	0.0
$\beta^F$	0.0	0.5	0.0	-0.5	0.5

表 5.7: 商品別効用値パラメーター：屋外

商品	A	B	C	D	E
種類	コーヒー	コーヒー	カフェオレ	緑茶	紅茶
ICE/HOT	ICE	HOT	ICE	ICE	ICE
$V^M$	0.5	0.5	-1.0	0.5	0.0
$V^F$	-0.5	-1.0	0.0	0.5	1.0
$\beta^M$	0.5	-0.5	0.5	0.0	1.0
$\beta^F$	0.5	-1.0	1.0	0.5	0.5

商品	F	G	H	I	J
種類	栄養ドリンク	炭酸	ミネラルウォーター	紅茶	スポーツドリンク
ICE/HOT	ICE	ICE	ICE	HOT	ICE
$V^M$	-0.5	0.5	-0.5	-1.0	1.0
$V^F$	-1.0	0.0	0.0	0.5	0.5
$\beta^M$	0.0	1.0	0.5	-0.5	0.5
$\beta^F$	0.5	1.0	0.0	-1.0	1.0

表 5.8: 商品別効用値パラメーター：学校

商品	A	B	C	D	E
種類	コーヒー	コーヒー	カフェオレ	緑茶	紅茶
ICE/HOT	ICE	HOT	ICE	ICE	ICE
$V^M$	0.0	-1.0	0.0	1.0	0.5
$V^F$	-0.5	-1.0	0.5	1.0	1.0
$\beta^M$	0.0	-1.0	-0.5	0.0	0.5
$\beta^F$	0.5	-1.0	0.0	-0.5	0.0

商品	F	G	H	I	J
種類	栄養ドリンク	炭酸	ミネラルウォーター	紅茶	スポーツドリンク
ICE/HOT	ICE	ICE	ICE	HOT	ICE
$V^M$	-0.5	0.5	0.0	-1.0	1.0
$V^F$	-1.0	0.0	-0.5	0.5	0.5
$\beta^M$	0.0	1.0	0.0	-0.5	1.0
$\beta^F$	0.0	0.5	0.5	-1.0	0.5

5.4 で示した。さらに、設置場所別モデル別に結果をサマリーしたものが表 5.9～5.10 である。

なおここでの「売切れ」数は、コラムに割り当てた商品が在庫切れで売切れが発生してしまい購入したい商品が購入できなかった顧客数を集計している。最初からコラムに割り当てられていないため購入できなかった商品についてはカウントしていない。

## 5.4 考察

表 5.9～5.11 から、各モデルの品揃え改善能力について評価を行なった。

比較対照モデル 0 と比較して、提案モデル 1～4 は売上数が平均で 6～20 程度高くなっている。特に学校での売上向上が大きい結果となった。

また比較対照モデル 0' と比較しても、提案モデル 1～4 は売上数の平均が 1～2 程度高く、売上の向上が見られた。設置場所別による売上の平均値に大きな差は見られなかった。屋外と学校では売上の標準偏差が 2 程度減少した一方、売切れ数が 0.2～0.5 程度増加した。

提案モデル 1～4 間で比較すると、全体としては大きな売上向上は見られなかったものの、唯一屋外のみ、モデル 1 とモデル 2～4 の間で 0.5 前後の向上が見られた。しかし、提案モデル 2～4 間では、推定時間ステップ数を増加しても売上や売切れの値についての向上は見られず、逆に売上が減少するケースも存在した。

以上より、本シミュレーションにより以下のような考察を得た。

- 比較対照モデル 0' のような単純な品揃え改善手法でも、全く品揃えを替えないよりは売上の向上が期待できる。
- 比較対照モデル 0' 自体に比較的高い品揃え改善能力があると考えられるが、そこから提案モデルを採用することにより若干ながらさらなる品揃え改善が期待できる。
- しかし、そこからさらに提案モデルにおいて数ステップ先の期待報酬までを含める品揃え改善を試みても、今回の条件・パラメーターではあまり効果がないことがわかった。

本研究では、比較対照モデル 0' として方策  $\pi^1$ （効用値の高低に従って商品を入れ替える）を常に行うモデルとして採用した。この手法で比較的高い品揃え改善能力が見られたことから、顧客の商品選択行動を左右するパラメーターとしての効用値が既知であることが、品揃え改善に有利であることが推察される。今後の研究課題として、より現実に近い状況へ品揃え改善モデルを適応させる際には、効用値パラメーターが未知の状況における対応が重要になるものと考えられる。



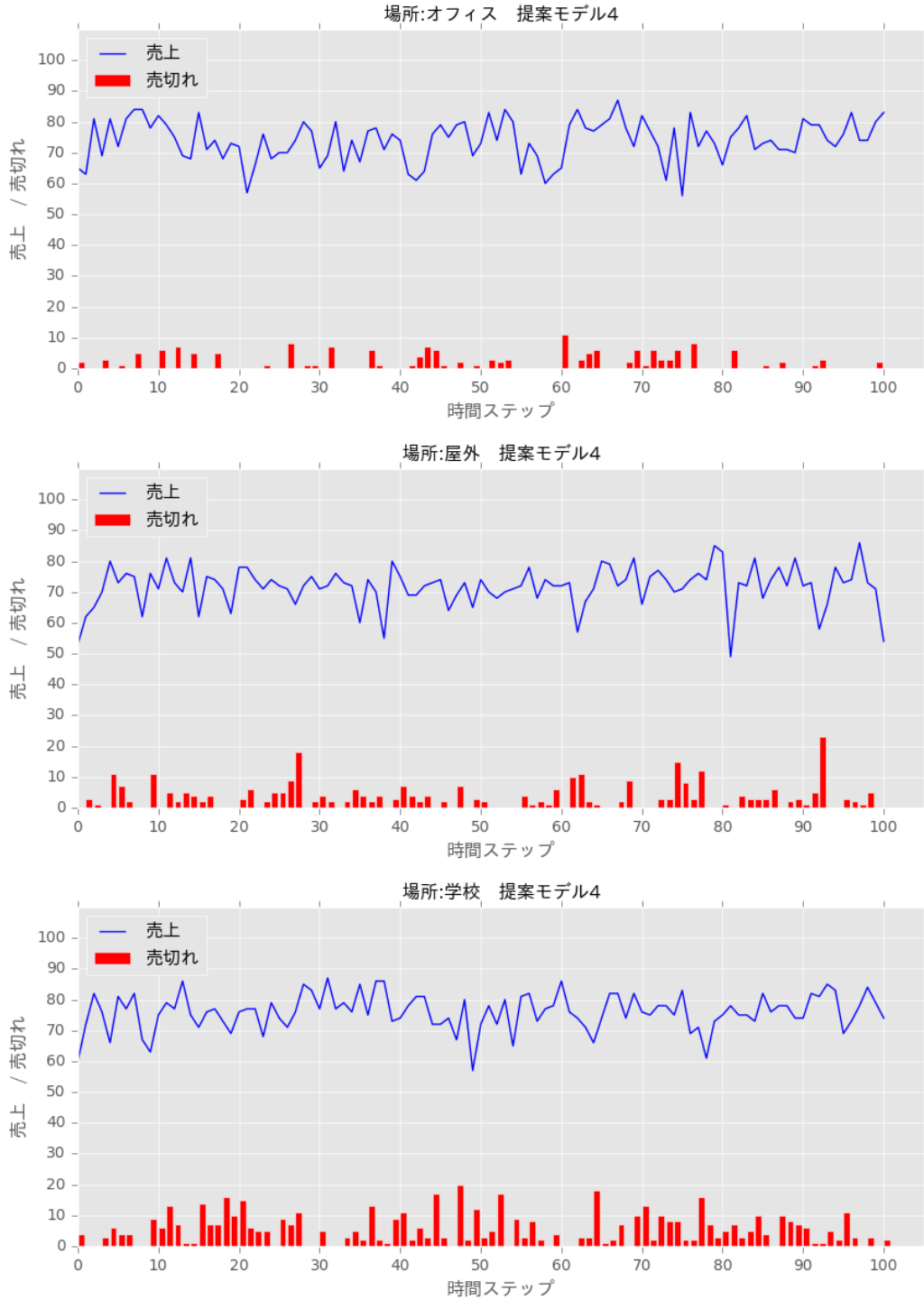


図 5.1: シミュレーション結果の例

表 5.9: 結果サマリー：オフィス

モデル名	推定する 時間ステップ数	売上平均	売上標準偏差	売切れ平均
モデル 0	-	67.150	5.404	1.149
モデル 0'	-	72.323	6.766	1.352
モデル 1	1	73.782	6.662	1.571
モデル 2	2	73.898	6.418	1.502
モデル 3	3	73.678	6.386	1.522
モデル 4	4	73.652	6.600	1.483

表 5.10: 結果サマリー：屋外

モデル名	推定する 時間ステップ数	売上平均	売上標準偏差	売切れ平均
モデル 0	-	54.312	5.742	0.931
モデル 0'	-	70.611	8.668	2.636
モデル 1	1	71.863	6.744	3.391
モデル 2	2	72.368	6.902	3.225
モデル 3	3	72.224	7.194	3.366
モデル 4	4	72.353	6.911	3.343

表 5.11: 結果サマリー：学校

モデル名	推定する 時間ステップ数	売上平均	売上標準偏差	売切れ平均
モデル 0	-	55.431	7.962	2.598
モデル 0'	-	73.598	7.712	5.062
モデル 1	1	75.399	5.634	5.482
モデル 2	2	75.585	6.031	5.241
モデル 3	3	75.290	5.849	5.224
モデル 4	4	75.645	6.181	5.493

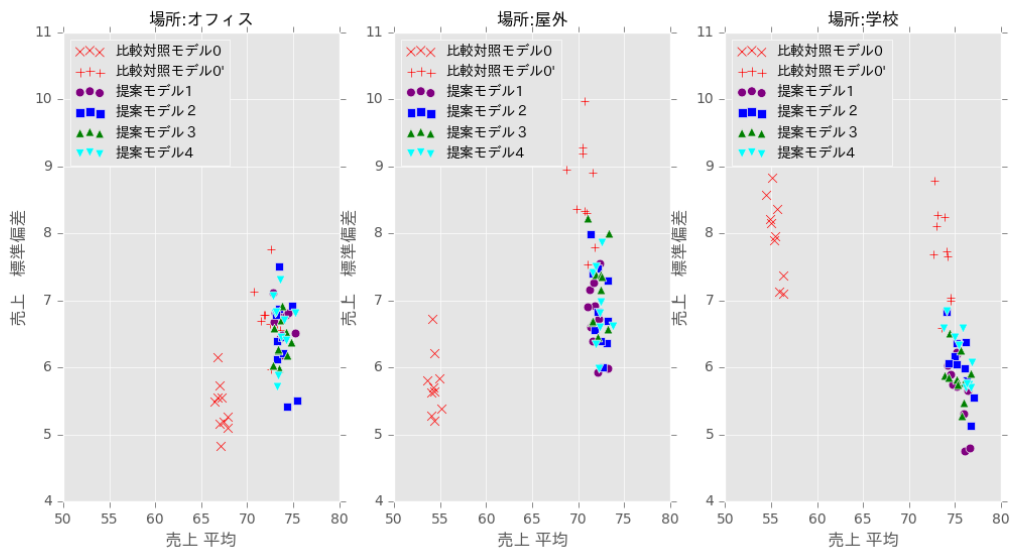


図 5.2: 売上の平均と標準偏差の分布

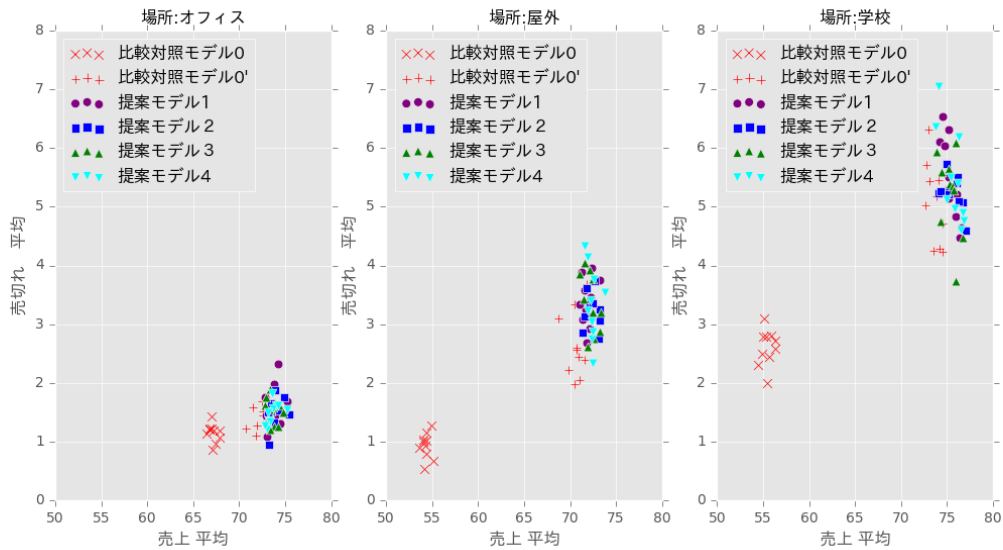


図 5.3: 売上の平均と売切れの平均の分布

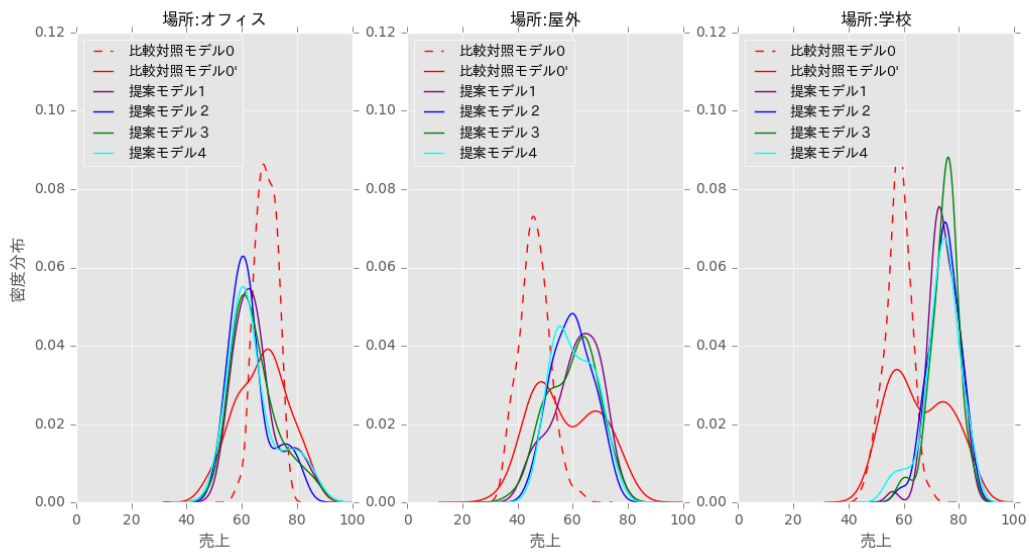


図 5.4: 売上の密度分布

## 第6章 まとめ

本研究では、自販機のオペレーター企業による補充・巡回作業を対象に、ルートマンに対して設置環境や消費者需要の変化に応じた適切な品揃え改善案を提案できるような最適化モデルの提案を試みた。消費者の商品選択行動を多項ロジットモデルで定式化した上で、POMDPの枠組みを利用し、自販機の状態が未知で変化する状況でも将来得られるであろう期待報酬を推定し改善提案を行えるモデルを開発した。

仮想パラメーターを用いたシミュレーションの結果、商品の効用値と自販機状態の遷移確率が既知であれば、比較的十分な品揃え改善を実現することができた。

しかし、本研究の条件下では、より単純でヒューリスティックな手法と比較して大きな改善を実現することはできなかった。特に将来の期待報酬を推定する手法の効果について検証することができず、モデル全体の有用性を実証するまでには至らなかった。

今後の課題として、より情報量が少ない状況においても対応できるモデルへの発展が必要と考える。つまり、表3.1における②・③・④の環境下での枠組みのモデル化と開発を行い、再シミュレーション下での効果検証を行いたいと考えている。

また、本研究はあくまで仮想パラメーターを用いたシミュレーション検証にとどまっている。さらなる研究課題としては、現実の自販機オペレーター企業の協力の元、実際の販売データ・業務データを利用して、提案モデルが現実の業務に適用できるかの検証が不可欠である。最終的には本研究で提案したモデルをオペレーター企業に提供し、実際のオペレーション業務に適用いただくことで業務改善に貢献するところまでを目標としたいと考えている。

## 付録A ポアソン二項分布

独立だがそれぞれの成功確率が同一でない複数のベルヌーイ試行を考える。この独立なベルヌーイ試行の和が従う確率分布を、ポアソン二項分布 (Poisson binomial distribution) と呼ぶ [19]。

確率変数  $X_i, i \in \{1, 2, \dots, n\}$  を、 $n$  個の独立なベルヌーイ試行に従うものとし、 $p_i$  を  $i$  番目のベルヌーイ試行の成功確率とする。このとき、すべての  $i$  について、 $0 \leq p_i \leq 1$  であり、

$$X_i \in \{1, 0\}, \Pr(X_i = 1) = p_i, \Pr(X_i = 0) = 1 - p_i \quad (\text{A.1})$$

である。確率変数  $X = \sum_{i=1}^n X_i$  を、 $n$  回の試行のうちの成功回数の和を表す確率変数とする。このとき、 $n$  回のうち  $k$  回成功する確率は次のように表せる。

$$\Pr(X = k) = \sum_{\alpha \in F_k} \prod_{i \in \alpha} p_i \prod_{j \in \alpha^C} (1 - p_j) \quad (\text{A.2})$$

ここで、 $F_k$  は  $\{1, 2, 3, \dots, n\}$  から選べるすべての  $k$  要素部分集合の族である。例えば、 $n = 3$  であれば  $F_2 = \{\{1, 2\}, \{1, 3\}, \{2, 3\}\}$  である。また  $\alpha^C$  は  $\alpha$  の補集合を表し、 $\alpha^C = \{1, 2, 3, \dots, n\} \setminus \alpha$  である。このとき、 $\Pr(X = k)$  は以下のように再帰的に求めることができる [20][21]。

$$\Pr(X = k) = \begin{cases} \prod_{i=1}^N (1 - p_i), & k = 0 \\ \frac{1}{k} \sum_{i=1}^k (-1)^{i-1} \Pr(X = k - i) T(i), & k > 0 \end{cases} \quad (\text{A.3})$$

ただし、

$$T(i) = \sum_{j=1}^N \left( \frac{p_j}{1 - p_j} \right)^i \quad (\text{A.4})$$

である。また、 $X$  の期待値と分散は、各ベルヌーイ分布に対する期待値と分散の和として求まる。

$$\text{期待値 : } E[X] = \sum_{i=1}^N p_i \quad (\text{A.5})$$

$$\text{分散 : } \sigma^2 = \sum_{i=1}^N p_i (1 - p_i) \quad (\text{A.6})$$

# 謝辞

本論文の作成にあたり、主指導教員である平石邦彦教授にはいつも貴重な時間を割いていただき、丁寧な御指導と厚いご助言を賜りました。この場を借りまして心よりの感謝を申し上げます。大変ありがとうございました。

また、副テーマ指導教員としてご指導いただきました白井清昭准教授、審査委員として貴重なご意見をいただきました緒方和博教授、上原隆平教授の各先生方にも改めて感謝の意を表します。

勤務先である株式会社インテージテクスフィア、株式会社インテージ、および株式会社クロスコンパス インテリジェンスの上司・同僚の皆さんからは、大学院に通うにあたっての理解や多くの支援・激励をいただきました。

そして、仕事と研究の両立に理解と協力を示していただき、いつも大きな心の支えになってくれている家族に、改めて深い感謝の意を表します。

最後に、2度の大学院進学に理解をいただき、長い間、常に有形・無形の支援をいただいていたものの、その支援への恩返しを果たせないまま一昨年の春に亡くなった実父に、最大限の感謝の意を表します。

## 参考文献

- [1] 一般社団法人日本自動販売機工業会, 自販機普及台数及び年間自販金額 2015 年(平成 27 年)版, 2016.
- [2] 布川 博士, 佐藤 究, 飲料自販機ビジネスにおけるルートマンの作業分析, 日本感性工学会論文誌, Vol.15, No.4, pp.471-478, 2016.
- [3] 日経 BP 社, “成果出す業務革新の現場 販売予測システム アサヒ飲料 自販機 1 台 3.2 分の補充時間短縮 営業車との往復が 1 回で作業完結”, 日経情報ストラテジー, Vol. 17, No, 12, pp.118-121, 2009.
- [4] 佐藤 忠彦, 樋口 知之, ビックデータ時代のマーケティング - ベイジアンモデリングの活用, 講談社, 2013.
- [5] 佐藤 忠彦, マーケティングにおける結果データ動的活用のためのベイジアンモデリング, オペレーションズ・リサーチ: 経営の科学, Vol. 55, No. 1, pp.25-30, 2010.
- [6] 佐藤忠彦, 樋口知之, 動的個人モデルによる消費者来店行動の解析, 日本統計学会誌, Vol. 38, No. 1, pp.1-19, 2008.
- [7] 藤田 真理奈, 河本 健, 相菌 敏子, 荒 宏視, 多品種商品を扱う店舗の需要予測のための商品選択予測手法の構築, 電子情報通信学会技術研究報告, Vol. 112, No. 466(LOIS2012-79), pp.59-64, 2013
- [8] 松村 直樹, 和泉 潔, 山田 健太, POS データに基づく欠品時の顧客行動を考慮した小売店舗の購買シミュレーション, 人工知能学会論文誌, Vol. 31, No. 2, pp.F-F13.1-8F, 2016.
- [9] R.Anupindi, M.Dada, S.Gupta, Estimation of Consumer Demand with Stock-Out Based Substitution: An Application to Vending Machine Products, Marketing Science, Vol. 17, No. 4, pp.406-423, 1998.
- [10] 宮本 裕一郎, 久保 幹雄, 自動販売機に対する在庫配送計画の事例, Journal of the Operations Research Society of Japan, Vol. 44, No. 4, pp.378-389, 2001.
- [11] 宮本 裕一郎, 久保 幹雄, 自動販売機補充問題に対する組合せ最適化アプローチ, 東京商船大学研究報告. 自然科学, Vol. 51, pp.27-33, 2000.



- [12] 伊藤 志保, 久保 幹雄, スイ キウン, 宮本 裕一郎, 自動販売機コラム割当問題, 日本オペレーションズ・リサーチ学会秋季研究発表会アブストラクト集, pp.22-23, 2000.
- [13] 竹内 俊子, 伊藤 一, 福地 純一郎, 自動販売機コラム割当最適化問題: 需要がポアソン過程に従う場合, 学習院大学経済論集, Vol. 49, No. 1, pp.47-52, 2012.
- [14] R.S.Sutton, A.G.Barto, (三上 貞芳, 皆川 雅章 訳), 強化学習, 森北出版, 2000.
- [15] 電気学会 GA・ニューロを用いた学習法とその応用調査専門委員会, 学習とそのアルゴリズム ニューラルネットワーク・遺伝アルゴリズム・強化学習. POD版, 森北出版, 2012.
- [16] Leslie Pack Kaelbling, Michael L. Littman, Anthony R. Cassandra, Planning and acting in partially observable stochastic domains, ArtificialIntelligence, Vol. 101, pp.99-134, 1998,
- [17] 木村 元, Kaelbling Leslie Pack, 部分観測マルコフ決定過程下での強化学習 (<特集>強化学習), 人工知能学会誌, Vol. 12, No. 6, pp.822-830, 1997.
- [18] 南 泰浩, 部分観測マルコフ決定過程に基づく対話制御, 日本音響学会誌, Vol. 67, No. 10, pp482-487, 2011.
- [19] Y. H. Wang, On the number of successes in independent trials, Statistica Sinica, Vol. 3, No. 2, pp.295-312, 1993.
- [20] B. K. Shah, On the distribution of the sum of independent integer valued random variables, The American Statistician, Vol. 27, No. 3, pp.123-124, 1973.
- [21] Xiang-Hui Chen, Arthur P. Dempster Jun S. Liu, Weighted finite population sampling to maximize entropy, Biometrika, Vol. 81, No. 3, pp.457-469, 1994.