2018年9月 运筹学学报 第22卷 第3期
Sept., 2018 Operations Research Transactions Vol.22 No.3

DOI: 10.15960/j.cnki.issn.1007-6093.2018.03.004

# 价格数量折扣下多阶段报童问题的在线策略*

张 永[1,†] 钟惠芬[1] 张卫国[2] 徐维军[2] 张 群[3]

**摘要** 价格数量折扣可以提高订购量，是库存决策中的一个重要因素. 特别地，当订购量达到一定水平时，价格折扣才会发生. 应用理论计算机科学兴起的弱集成算法，研究具有这种价格数量折扣的多阶段报童问题的在线策略. 弱集成算法是一种在线序列决策算法，其主要特点是不对未来输入做任何统计假设，克服了报童问题研究中需要对需求做概率假设的困难. 主要将弱集成算法应用到固定订购量的专家策略，给出了价格数量折扣下多阶段报童问题的具体在线策略；得到了该在线策略相对于最优专家策略的理论保证. 进一步将回收价值和缺货损失费引入，给出了推广的在线策略及其理论结果. 最后应用数值算例说明了给出的在线策略具有较好的竞争性能.

**关键词** 价格数量折扣，多阶段报童问题，弱集成算法，在线策略，竞争性能分析

**中图分类号** F832

**2010 数学分类号** 91G10



# On-line strategies for multi-period newsvendor problem with price quantity discount*

ZHANG Yong[1,†] ZHONG Huifen[1] ZHANG Weiguo[2]
XU Weijun[2] ZHANG Qun[3]

**Abstract** The price quantity discount can increase the order quantity, which is an important factor of inventory decision making. Particularly, price discount only occurs when the order quantity reaches a fixed level. This paper uses the Weak Aggregating Algorithm (WAA) advanced in computer science to study the multi-period newsvendor problem with this kind price quantity discount. WAA is an on-line sequential decision-making algorithm; its main advantage is that it does not make statistical assumption on future inputs, which overcomes the difficulties of having to make probability hypothesis on demand in newsvendor problem research. Mainly, this paper applies WAA to experts whose strategies are fixed order quantities to present explicit online strategy for the multi-period newsvendor problem with price quantity discount. The theoretical guarantee for

---

收稿日期：2016-04-07

* 基金项目：国家自然科学基金(No. 71501049)，2015年广州市哲学社会科学“十二五”规划课题(No. 15G29)，广东省高等学校珠江学者岗位计划(2016)

1. 广东工业大学管理学院，广州 510520; School of Management, Guangdong University of Technology, Guangzhou 510520, China

2. 华南理工大学工商管理学院，广州 510006; School of Business Administration, South China University of Technology, Guangzhou 510006, China

3. 广东外语外贸大学金融学院，广州 510006; School of Finance, Guangdong University of Foreign Studies, Guangzhou 510006, China

† 通信作者 E-mail: yq21614@163.com

the proposed online strategy is obtained compared with the best expert strategy. The salvage and shortage cost are further introduced to obtain extended online strategies and their theoretical results. The numerical examples are finally used to show the good competitive performance of the proposed online strategies.

**Keywords** price quantity discount, multi-period newsvendor problem, weak aggregating algorithm, on-line strategy, competitive performance analysis

**Chinese Library Classification** F832

**2010 Mathematics Subject Classification** 91G10

# 0 引言

报童问题是库存控制研究中一类单周期的问题, 具有较强的实践性, 广泛应用于库存管理、生产计划和股票市场等各种环境中. 面对不确定性的需求, 传统报童模型反映报童在库存成本和缺货损失之间的权衡, 求解最优订购量以使得报童获得最大化的期望收益. 在传统报童模型的基础上, 众多研究从多个方向加以扩展, 构建更加贴近现实的模型[1-6]. 报童问题的传统研究方法往往基于对需求的概率假设, 并在此基础上求解最优订购量. 当需求不服从某个概率分布或者无法获得需求的统计信息时, 传统研究方法面临着挑战. 针对这一缺点, Scarf[7]提出了当需求服从某一类分布, 仅仅知道其均值和方差的求解方法. 对于 Scarf 提出的求解方法最优性, Gallego 和 Moon[8]给出了一种简单的证明方法. Moon 和 Choi[9]考虑到库存量低于一定水平时顾客可能会犹豫止步, 给出了该情形下 Gallego 和 Moon 的扩展研究. 随后, Alfares 和 Elmorra[10]引入缺货损失费进一步推广研究了 Gallego 和 Moon 的结果.

Scarf 等众多研究学者提出的决策方法缓解了无法获得需求服从的分布时报童问题求解的困难. 但该方法仍然需要获得需求服从分布的均值和方差, 没有完全脱离统计信息的假设. 近年来理论计算机科学兴起的在线算法为库存管理[11]、在线租赁[12-14]和投资组合[15]等问题的研究提供了新的研究思路和方法. 在线算法擅长的决策是一种无风险策略, 所获得的收益是确定性的, 同时在线算法选择某一种基准策略作为比较, 提供了在某种意义下任何两种策略的性能统一度量方法. 在线算法的优点使得能将其应用到库存决策中, 即完全不对需求做任何统计假设, 仅仅根据历史需求给出即时决策, 并将这种决策方法称为在线策略. 用绩效比衡量在线策略性能, Wanger[16]研究了多阶段库存决策问题, 并推广研究可耐性产品的情形. 随后, Wanger[17]进一步用竞争比分析方法研究了成本优化下的多阶段库存决策. 张国清等[18-19]用竞争比分析法研究了单阶段报童决策问题, 分别给出了概率预期和一般预期下的风险算法. 文献[20-21]也是应用竞争比分析等在线算法研究不确定和不完全信息下的库存决策及相关问题. 库存决策的在线策略成为一个热点研究问题基于以下两个方面的原因: 一方面, 面对不确定性的决策环境, 决策者更期望获得相对较少但相对确定的收益, 而不愿意在高风险下获得比较高的收益; 另一方面, 决策者不但追求自身收益最大化, 而且更看重与同行的业绩比较.

弱集成算法 (Weak Aggregating Algorithm, WAA) 是一种集成专家意见(策略)进行预测的在线序列决策方法, 它结合每个专家的意见进行集成决策, 通过赋予每个专家权重体现对其的信任程度[22]. 集成专家意见方法的研究具有较强的实际意义, 因为现实决策中存在着众多专家意见. Levina等[11]将WAA 应用到固定订购量的专家意见, 给出了具体在线策略, 证明了该决策方法的累积收益渐近于最优专家意见的累积收益. Levina 等考虑的

是静态专家意见, 在文献[11]的基础上, Zhang等[23]进一步将 WAA 应用到动态专家意见, 研究非静止需求环境下的多阶段报童问题, 分别给出了订购量为任意实数和整数的在线策略及其理论保证, 并用数值算例说明了不同需求类型下在线策略的竞争性能. 价格数量折扣是库存决策中的一个重要因素. 报童提高订购量往往是为了付出更小的单位成本, 从而降低支出. 一般地, 当订购量达到一定水平时, 价格折扣才会发生. 考虑报纸的购买具有这种价格数量折扣, 本文继续应用 WAA 研究多阶段报童问题, 即当订购量大于某一固定值时, 报童才能以更低的价格支付单位成本. 首先应用 WAA 给出了带有这种价格数量折扣的多阶段报童问题的在线策略. 在此基础上, 进一步证明了该在线策略相对于最优专家策略具有较好的竞争性能, 给出了理论保证. 然后引入回收价值和缺货损失费将研究结果进行推广, 给出了推广的在线策略及其理论保证. 最后借助于数值算例进一步说明本文所给策略具有较强的竞争性, 能够实现较大的累积收益.

## 1 基于累积收益的弱集成算法

WAA 是理论计算机科学兴起的集成专家意见的在线序列决策算法[22]. 将其应用到报童问题时, 用收益函数来衡量每个专家意见的优劣, 以在线方式计算每个专家的收益并逐步调整对每个专家意见的信任程度. 对专家的信任程度通过对其赋予的权重体现出来. 通常收益越大的专家, 对其信任程度越大, 所赋予的权重也就越大. 用指数加权来集成专家意见, 并用学习率参数控制每个专家权重更新的幅度.

在 WAA 设置中, 有三个决策(者)体, 分别是专家决策体、在线决策者和实际决策者. 专家决策体针对具体决策问题给出预测; 在线决策者集成专家决策体中众多专家意见给出预测; 而实际决策者在专家和在线决策者决策之后给出实际的结果, 它是真正的决策者. WAA 的主要目的是提高在线决策者的竞争性能. 沿用已有的记号, 用 $\Theta$ 表示专家索引集合, 并假设定义它上的概率测度是可测的. 设在线决策者和专家决策体的预测集为 $\Gamma$, 并用 $\Omega$ 表示实际结果集合. 设 $g$ 为收益函数. 在第 $n$ 阶段, 当实际结果为 $w_n \in \Omega$, 决策者预测 $r_n \in \Gamma$ 时的收益为 $g_n = \pi(r_n, w_n)$, 专家 $\theta \in \Theta$ 预测 $r_n^\theta$ 时的收益为 $g_n^\theta = \pi(r_n^\theta, w_n)$. WAA 在初始化每个专家意见权重的基础上, 根据收益函数, 采用指数加权平均法更新每个专家意见的权重. 设 $q(\mathrm{d}\theta)$ 是在集合上的初始概率测度, 则 WAA 在每个阶段 $n$ 按照公式

$$p_n(d\theta) = \frac{{\beta_n}^{G_{n-1}^\theta} q(\mathrm{d}\theta)}{\int_\Theta {\beta_n}^{G_{n-1}^\theta} q(\mathrm{d}\theta)} \tag{1.1}$$

赋予专家策略权重, 其中 $\beta_n = \exp(\frac{1}{\sqrt{n}})$, $\frac{1}{\sqrt{n}}$ 为学习率, $G_{n-1}^\theta = \sum\limits_{i=1}^{n-1} g_i^\theta = \sum\limits_{i=1}^{n-1} \pi(r_i^\theta, w_i)$ 是专家 $\theta, \theta \in \Theta$ 在前 $n-1$ 阶段的累积收益. 根据文献 [22], WAA 的伪代码表述如下. 初始时累积收益为 0, 即 $G_0 = 0, G_0^\theta = 0, \theta \in \Theta$; 在每个决策阶段 $n = 1, 2, \cdots$

(a) 计算标准化权重 (1.1);

(b) 专家决策体给出专家意见 $r_n^\theta \in \Gamma, \theta \in \Theta$;

(c) 在线决策者集成专家意见给出决策 $r_n = \int_\Theta r_n^\theta p_n(\mathrm{d}\theta)$;

(d) 获得实际结果 $w_n \in \Omega$;

(e) 在线决策者的收益为 $g_n=\pi(r_n,w_n)$, 每个专家的收益为 $g_n^\theta=\pi(r_n^\theta,w_n),\theta\in\Theta$;

(f) 更新在线决策者和专家的累积收益 $G_n=G_{n-1}+\pi(r_n,w_n),G_n^\theta=G_{n-1}^\theta+\pi(r_n^\theta,w_n)$, $\theta\in\Theta$;

令 $N$ 表示决策的总阶段数, 当收益函数 $g$ 有界时, Levina 等给出了 WAA 实现的累积收益的理论下界[11].

**引理 1.1** 当 $-L\leqslant g\leqslant 0$, 其中 $L$ 为一固定常数, 则对于所有的 $N$ 有

$$G_N\geqslant\left(\ln\int_\Theta\exp\left(\frac{G_N^\theta}{\sqrt{N}}\right)q(\mathrm{d}\theta)-L^2\right)\sqrt{N}. \tag{1.2}$$

## 2 在线策略及其竞争性能分析

### 2.1 具体的在线策略

在多阶段报童模型中, 设单位报纸的卖出价为 $p$, 报纸的实际需求量上界为 $B$(为整数), 在第 $n$ 阶段, 报童的订购量为 $y_n$, 实际的需求量为 $d_n$, 且假设报童的订购量和实际需求量为 [0, B] 区间内的任意实数; 订购的成本具有价格数量折扣特点, 当订购量不大于 $Q(Q<B$, 为整数)时单位报纸的成本为 $c_1$, 当订购量大于 $Q$ 时单位报纸的成本为 $c_2(c_2<c_1)$, 即第 $n$ 阶段报童的收益为



$$g_n=\begin{cases}p\min(y_n,d_n)-c_1y_n, & y_n\leqslant Q;\\ p\min(y_n,d_n)-c_2y_n, & y_n>Q,\end{cases} \tag{2.1}$$

前 $n$ 阶段的累积收益为 $G_n=\sum\limits_{i=1}^{n}g_i$. 当专家 $\theta$ 的订购量为 $y_n^\theta$ 时, 其在第$n$阶段的收益为

$$g_n^\theta=\begin{cases}p\min(y_n^\theta,d_n)-c_1y_n^\theta, & y_n^\theta\leqslant Q;\\ p\min(y_n^\theta,d_n)-c_2y_n^\theta, & y_n^\theta>Q,\end{cases}$$

前 $n$ 阶段的累积收益为 $G_n^\theta=\sum\limits_{i=1}^{n}g_i^\theta$.

**定理 2.1** 将 WAA 应用到固定订购量的专家意见, 在收益函数(2.1)的基础上, 报童可得到具体的在线策略.

**证明** 考虑静态的专家意见, 也就是固定订购量的专家策略, 即每个专家 $\theta=y\in[0,B]$ 在每个阶段 $n$ 总是推荐订购量 $y_n^\theta=y$. 根据 WAA 决策步骤(c)知报童在第 $n$ 阶段的订购量为

$$y_n=\int y_n^\theta p_n(\mathrm{d}\theta)=\frac{\int_0^B y\exp\left(\frac{G_{n-1}^y}{\sqrt{n}}\right)q(\mathrm{d}y)}{\int_0^B\exp\left(\frac{G_{n-1}^y}{\sqrt{n}}\right)q(\mathrm{d}y)}, \tag{2.2}$$

当 $q(\mathrm{d}y)$ 是 $[0,B]$ 上的均匀分布时, 可采用文献[11]的求解技巧: 在第 $n$ 阶段确定订购量时, 前 $n-1$ 阶段的实际需求 $d_i,(1\leqslant i\leqslant n-1)$ 是已知的数据序列, 设其顺序统计量为 $d_{(1)},\cdots,d_{(n-1)}$, 并设 $d_{(0)}=0,d_{(n)}=B$. 对于价格数量折扣的临界值 $Q$, 存在正整数

$i_0, 0 \leqslant i_0 \leqslant n$, 满足关系式 $d_{(i_0)} \leqslant Q \leqslant d_{(i_0+1)}$. 基于上述假设, 在收益函数(2.1)的基础上有

$$\begin{aligned}\int_0^B y\exp\left(\frac{G_{n-1}^y}{\sqrt{n}}\right)\mathrm{d}y &= \sum_{k=0}^{n-1}\int_{d_{(k)}}^{d_{(k+1)}} y\exp\left(\frac{G_{n-1}^y}{\sqrt{n}}\right)\mathrm{d}y\\ &= \sum_{k=0}^{i_0-1}\int_{d_{(k)}}^{d_{(k+1)}} y\exp\left(\frac{G_{n-1}^y}{\sqrt{n}}\right)\mathrm{d}y + \int_{d_{(i_0)}}^{Q} y\exp\left(\frac{G_{n-1}^y}{\sqrt{n}}\right)\mathrm{d}y\\ &\quad + \int_{Q}^{d_{(i_0+1)}} y\exp\left(\frac{G_{n-1}^y}{\sqrt{n}}\right)\mathrm{d}y + \sum_{k=i_0+1}^{n-1}\int_{d_{(k)}}^{d_{(k+1)}} y\exp\left(\frac{G_{n-1}^y}{\sqrt{n}}\right)\mathrm{d}y\\ &= I_1 + I_2 + I_3 + I_4,\end{aligned}$$

其中

$$\begin{aligned}I_1 &= \sum_{k=0}^{i_0-1}\int_{d_{(k)}}^{d_{(k+1)}} y\exp\left(\frac{G_{n-1}^y}{\sqrt{n}}\right)\mathrm{d}y = \sum_{k=0}^{i_0-1}\int_{d_{(k)}}^{d_{(k+1)}} y\exp\left(n^{-\frac{1}{2}}\sum_{i=1}^{n-1}\left(p\min(y, d_{(i)}) - c_1 y\right)\right)\mathrm{d}y\\ &= \sum_{k=0}^{i_0-1}\exp\left(n^{-\frac{1}{2}}\sum_{i=1}^{k} p d_{(i)}\right)\int_{d_{(k)}}^{d_{(k+1)}} y\exp(yn^{-\frac{1}{2}}(p(n-k-1) - c_1(n-1)))\mathrm{d}y\\ &= \sum_{k=0}^{i_0-1}\exp\left(n^{-\frac{1}{2}}\sum_{i=1}^{k} p d_{(i)}\right) I_k^n,\end{aligned} \tag{2.3}$$

$$I_k^n = \begin{cases}\dfrac{1}{2}(d_{(k+1)}^2 - d_{(k)}^2), & p(n-k-1) = c_1(n-1);\\ \left[a_k^n\exp\left(\dfrac{y}{a_k^n}\right)(y - a_k^n)\right]\Bigg|_{d_{(k)}}^{d_{(k+1)}}, & p(n-k-1) \neq c_1(n-1), a_k^n = \dfrac{n^{\frac{1}{2}}}{p(n-k-1) - c_1(n-1)},\end{cases}$$

(2.3)式成立是由于当 $d_{(k)} \leqslant y \leqslant d_{(k+1)}$ 时, 有关系式

$$\sum_{i=1}^{n-1}(p\min(y, d_{(i)}) - c_1 y) = \sum_{i=1}^{k} p d_{(i)} + p(n-k-1)y - c_1(n-1)y.$$

采用类似的方法, 可得 $I_2 = \exp(n^{-\frac{1}{2}}\sum\limits_{i=1}^{i_0} p d_{(i)}) I_{i_0}^n$, 其中

$$I_{i_0}^n = \begin{cases}\dfrac{1}{2}(Q^2 - d_{(i_0)}^2), \quad p(n-i_0-1) = c_1(n-1);\\ \left[a_{i_0}^n\exp\left(\dfrac{y}{a_{i_0}^n}\right)(y - a_{i_0}^n)\right]\Bigg|_{d_{(i_0)}}^{Q}, p(n-i_0-1) \neq c_1(n-1), a_{i_0}^n = \dfrac{n^{\frac{1}{2}}}{p(n-i_0-1) - c_1(n-1)},\end{cases}$$

$I_3 = \exp(n^{-\frac{1}{2}}\sum\limits_{i=1}^{i_0} p d_{(i)}) I_{i_0+1}^n$, 其中

$$I_{i_0+1}^n = \begin{cases}\dfrac{1}{2}(d_{(i_0+1)}^2 - Q^2), \quad p(n-i_0-1) = c_2(n-1);\\ \left[a'^n_{i_0}\exp\left(\dfrac{y}{a'^n_{i_0}}\right)(y - a'^n_{i_0})\right]\Bigg|_{Q}^{d_{(i_0+1)}}, p(n-i_0-1) \neq c_2(n-1), a'^n_{i_0} = \dfrac{n^{\frac{1}{2}}}{p(n-i_0-1) - c_2(n-1)},\end{cases}$$

$I_4=\sum\limits_{k=i_0+1}^{n-1}\exp(n^{-\frac{1}{2}}\sum\limits_{i=1}^{k}pd_{(i)})I'^{n}_{k}$, 其中

$$I'^{n}_{k}=\begin{cases}\dfrac{1}{2}(d^2_{(k+1)}-d^2_{(k)}),\ p(n-k-1)=c_2(n-1);\\ \left[a'^{n}_{k}\exp\left(\dfrac{y}{a'^{n}_{k}}\right)(y-a'^{n}_{k})\right]\Big|_{d_{(k)}}^{d_{(k+1)}},p(n-k-1)\neq c_2(n-1),a'^{n}_{k}=\dfrac{n^{\frac{1}{2}}}{p(n-k-1)-c_2(n-1)}.\end{cases}$$

同样地有

$$\begin{aligned}\int_0^B\exp\left(\frac{G^y_{n-1}}{\sqrt{n}}\right)\mathrm{d}y&=\sum_{k=0}^{n-1}\int_{d_{(k)}}^{d_{(k+1)}}\exp\left(\frac{G^y_{n-1}}{\sqrt{n}}\right)\mathrm{d}y\\&=\sum_{k=0}^{i_0-1}\int_{d_{(k)}}^{d_{(k+1)}}\exp\left(\frac{G^y_{n-1}}{\sqrt{n}}\right)\mathrm{d}y+\int_{d_{(i_0)}}^{Q}\exp\left(\frac{G^y_{n-1}}{\sqrt{n}}\right)\mathrm{d}y\\&\quad+\int_Q^{d_{(i_0+1)}}\exp\left(\frac{G^y_{n-1}}{\sqrt{n}}\right)\mathrm{d}y+\sum_{k=i_0+1}^{n-1}\int_{d_{(k)}}^{d_{(k+1)}}\exp\left(\frac{G^y_{n-1}}{\sqrt{n}}\right)\mathrm{d}y\\&=J_1+J_2+J_3+J_4,\end{aligned}$$

其中

$$J_1=\sum_{k=0}^{i_0-1}\exp\left(n^{-\frac{1}{2}}\sum_{i=1}^{k}pd_{(i)}\right)J^n_k,\ J^n_k=\begin{cases}d_{(k+1)}-d_{(k)},\quad p(n-k-1)=c_1(n-1);\\ \left[a^n_k\exp\left(\dfrac{y}{a^n_k}\right)\right]\Big|_{d_{(k)}}^{d_{(k+1)}},\quad p(n-k-1)\neq c_1(n-1),\end{cases}$$

$$J_2=\exp\left(n^{-\frac{1}{2}}\sum_{i=1}^{i_0}pd_{(i)}\right)J^n_{i_0},\ J^n_{i_0}=\begin{cases}Q-d_{(i_0)},\quad p(n-i_0-1)=c_1(n-1);\\ \left[a^n_{i_0}\exp\left(\dfrac{y}{a^n_{i_0}}\right)\right]\Big|_{d_{(i_0)}}^{Q},\quad p(n-i_0-1)\neq c_1(n-1),\end{cases}$$

$$J_3=\exp\left(n^{-\frac{1}{2}}\sum_{i=1}^{i_0}pd_{(i)}\right)J^n_{i_0+1},\ J^n_{i_0+1}=\begin{cases}d_{(i_0+1)}-Q,\quad p(n-i_0-1)=c_2(n-1);\\ \left[a'^{n}_{i_0}\exp\left(\dfrac{y}{a'^{n}_{i_0}}\right)\right]\Big|_{Q}^{d_{(i_0+1)}},\quad p(n-i_0-1)\neq c_2(n-1),\end{cases}$$

$$J_4=\sum_{k=i_0+1}^{n-1}\exp\left(n^{-\frac{1}{2}}\sum_{i=1}^{k}pd_{(i)}\right)J'^{n}_{k},\ J'^{n}_{k}=\begin{cases}d_{(k+1)}-d_{(k)},\quad p(n-k-1)=c_2(n-1);\\ \left[a'^{n}_{k}\exp\left(\dfrac{y}{a'^{n}_{k}}\right)\right]\Big|_{d_{(k)}}^{d_{(k+1)}},\quad p(n-k-1)\neq c_2(n-1).\end{cases}$$

至此, 给出了具体的在线订购策略, 该在线策略在第 $n$ 阶段的订购量为

$$y_n=\frac{I_1+I_2+I_3+I_4}{J_1+J_2+J_3+J_4}.$$

## 2.2 竞争性能分析

对于定理 2.1 给出的价格数量折扣下多阶段报童问题的在线策略, 应用引理 1.1, 定理 2.2 给出了以最优专家策略的累积收益为基准时该策略竞争性能的理论保证.

**定理 2.2** 对于带有价格数量折扣的多阶段报童模型, 按照在线策略(2.2)进行决策报童的累积收益满足

$$G_N\geqslant\max_{y\in[0,B]}G^y_N-\left(B^2p^2+p-c_2+c_1+\ln(B\sqrt{N})\right)\sqrt{N}.\qquad(2.4)$$

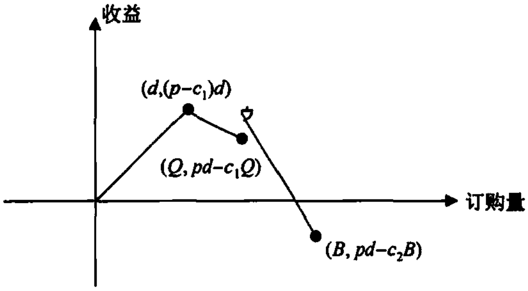
**证明** 应用引理 1.1 需要给出收益函数(2.1)的上界. 首先探讨收益与订购量的关系, 分为如下两种情况.

(a) $d \leqslant Q$ 的情形. 此种情况下收益与订购量的关系可由图 1 给出.

由图 1 可以看出, 这种情形下有关系式

$$-c_2B \leqslant g \leqslant (p-c_1)B.$$

(b) $d > Q$ 的情形. 此种情况下收益与订购量的关系可由图 2 给出.



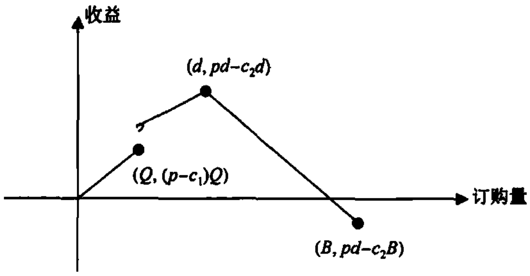




**图1** 当$d \leqslant Q$时收益与订购量之间的关系 **图2** 当$d > Q$时收益与订购量之间的关系

由图 2 可以看出, 这种情形下有关系式

$$-c_2B \leqslant g \leqslant (p-c_2)B.$$

综合以上两种情况则有

$$-c_2B \leqslant g \leqslant \max((p-c_1)B, (p-c_2)B) \leqslant (p-c_2)B,$$

即有$-c_2B \leqslant g \leqslant (p-c_2)B$. 令 $g := g-(p-c_2)B$, 则$-Bp \leqslant g \leqslant 0$.

考虑任意两个专家策略的累积收益差值, 若有两个专家的固定订购量分别为 $x$ 和 $y$, 由前面的分析知经过 $N$ 阶段之后, 这两个专家实现的累积收益 $G_N^x$ 和 $G_N^y$ 满足关系式

$$|G_N^x - G_N^y| \leqslant \max(c_1, p-c_1, c_2, p-c_2)|x-y|N \leqslant (p-c_2+c_1)|x-y|N.$$

通过任意一个专家的固定订购量策略 $y$ 的 $\frac{1}{\sqrt{N}}$ 领域给出 $\ln\int_\Theta \exp\left\{\frac{G_N^\theta}{\sqrt{N}}\right\}q(\mathrm{d}\theta)$ 的下界. 当 $|y-\theta| \leqslant \frac{1}{\sqrt{N}}$ 时, 由前面分析过程知

$$|G_N^y - G_N^\theta| \leqslant (p-c_2+c_1)N\frac{1}{\sqrt{N}} = (p-c_2+c_1)\sqrt{N},$$

所以有 $G_N^\theta \geqslant G_N^y - (p-c_2+c_1)\sqrt{N}$, 即有

$$\begin{aligned}\ln\int_\Theta \exp\left\{\frac{G_N^\theta}{\sqrt{N}}\right\}q(\mathrm{d}\theta) &= \ln\int_0^B \exp\left\{\frac{G_N^\theta}{\sqrt{N}}\right\}q(\mathrm{d}\theta)\\ &\geqslant \ln\left\{\frac{1}{B\sqrt{N}}\exp\left(\frac{G_N^y-(p-c_2+c_1)}{\sqrt{N}}\right)\right\}\\ &= \left(\frac{G_N^y}{\sqrt{N}} - p + c_2 - c_1\right) - \ln(B\sqrt{N}).\end{aligned}$$

根据引理 1.1 有

$$\begin{aligned} G_N &\geqslant \left(\ln\int_\Theta \exp\left\{\frac{G_N^\theta}{\sqrt{N}}\right\}q(\mathrm{d}\theta) - B^2p^2\right)\sqrt{N} \\ &\geqslant \left(\frac{G_N^\theta}{\sqrt{N}} - (p-c_2+c_1) - \ln(B\sqrt{N}) - B^2p^2\right)\sqrt{N} \\ &= G_N^y - \left(B^2p^2 + p - c_2 + c_1 + \ln(B\sqrt{N})\right)\sqrt{N}, \end{aligned}$$

该不等式对所有的 $y$ 成立, 故有式(2.4).

定理 2.2 中的结论 (2.4) 表明当 $N$ 越大时, 报童的累积收益 $G_N$ 越接近于最优专家意见的累积收益 $\max\limits_{y\in[0,B]} G_N^y$, 事实上, 当 $N\to\infty$ 时, 有

$$\lim_{N\to\infty}\frac{\left(B^2p^2+p-c_2+c_1+\ln(B\sqrt{N})\right)\sqrt{N}}{N}=0,$$

即有

$$\lim_{N\to\infty}\left(\frac{G_N}{N}-\frac{\max\limits_{y\in[0,B]} G_N^y}{N}\right)\geqslant 0,$$

从而知在线策略的平均收益随总阶段数 $N$ 的增大越来越接近于最优专家意见的平均收益.



## 3 扩展模型的在线策略及其竞争性能分析

### 3.1 带有回收价值的情形

在前面 2.1 节提出的模型基础上, 考虑带有回收价值的情形: 设每份报纸的回收价值为 $s(s<c)$, 则第 $n$ 阶段报童的收益为

$$g_n=\begin{cases} p\min(y_n,d_n)-c_1y_n+s[\max(y_n,d_n)-d_n], & y_n\leqslant Q;\\ p\min(y_n,d_n)-c_2y_n+s[\max(y_n,d_n)-d_n], & y_n>Q, \end{cases} \tag{3.1}$$

利用关系式 $s[\max(y_n,d_n)-d_n]=s[y_n-\min(y_n,d_n)]$, 可将式(3.1)转化为

$$g_n=\begin{cases} (p-s)\min(y_n,d_n)-(c_1-s)y_n, & y_n\leqslant Q;\\ (p-s)\min(y_n,d_n)-(c_2-s)y_n, & y_n>Q, \end{cases} \tag{3.2}$$

前 $n$ 阶段的累积收益为 $G_n=\sum\limits_{i=1}^n g_i$. 类似地, 专家 $\theta=y\in[0,B]$ 的固定订购量为 $y$ 时, 其在第 $n$ 阶段的收益为

$$g_n^y=\begin{cases} (p-s)\min(y,d_n)-(c_1-s)y, & y\leqslant Q;\\ (p-s)\min(y,d_n)-(c_2-s)y, & y>Q, \end{cases}$$

且前 $n$ 阶段的累积收益为 $G_n^y=\sum\limits_{i=1}^{n} g_i^y$.

令 $p:=p-s, c_i:=c_i-s, i=1,2$, 则收益函数(3.2)与(2.1)相同. 因此将前面求解公式(2.2)中的 $p$ 替换为 $p-s$, $c_1$替换为$c_1-s$, $c_2$ 替换为 $c_2-s$, 就得到了带有回收价值情形下的具体决策公式. 对于理论上的竞争性能分析, 同样地只需将式(2.4)中的 $p$ 替换为 $p-s$, $c_1$ 替换为 $c_1-s$, $c_2$ 替换为 $c_2-s$, 即有推论 3.1.

**推论 3.1** 对于价格数量折扣下带有回收价值的多阶段报童模型, 按照在线策略(2.2)决策的累积收益满足

$$G_N \geqslant \max_{y\in[0,B]} G_N^y-(B^2(p-s)^2+(p-s)-c_2+c_1+\ln(B\sqrt{N}))\sqrt{N}. \tag{3.3}$$

## 3.2 带有缺货损失费的情形

在实际的多阶段库存订购模型中, 当订购量小于需求时会产生一定的缺货损失费. 本小节进一步将缺货损失费引入到带有回收价值的情形; 设每份报纸的缺货损失费为 $l$, 且满足关系式 $l<p-c$, 则第 $n$ 阶段报童的收益为

$$g_n=\begin{cases} p\min(y_n,d_n)-c_1y_n+s[\max(y_n,d_n)-d_n]-l[d_n-\min(y_n,d_n)], & y_n\leqslant Q;\\ p\min(y_n,d_n)-c_2y_n+s[\max(y_n,d_n)-d_n]-l[d_n-\min(y_n,d_n)], & y_n>Q, \end{cases} \tag{3.4}$$

类似地可将 (3.4) 转化为

$$g_n=\begin{cases} (p+l-s)\min(y_n,d_n)-(c_1-s)y_n-ld_n, & y_n\leqslant Q;\\ (p+l-s)\min(y_n,d_n)-(c_2-s)y_n-ld_n, & y_n>Q, \end{cases} \tag{3.5}$$

累积收益为 $G_n=\sum\limits_{i=1}^{n} g_i$, 固定订购量为 $y$ 的专家 $\theta=y\in[0,B]$ 在第 $n$ 阶段的收益为

$$g_n^y=\begin{cases} (p+l-s)\min(y,d_n)-(c_1-s)y-ld_n, & y\leqslant Q;\\ (p+l-s)\min(y,d_n)-(c_2-s)y-ld_n, & y>Q, \end{cases}$$

累积收益为 $G_n^y=\sum\limits_{i=1}^{n} g_i^y$. 由 2.1 节中具体决策公式的推导过程和定理 2.2 的理论分析过程知具体在线策略及其竞争性能的理论保证与常数项 $ld_n$ 无关. 因此, 令

$$g_n=g_n+ld_n, \quad g_n^y=g_n^y+ld_n,$$

则报童和专家 $\theta=y\in[0,B]$ 的新收益函数为

$$g_n=\begin{cases} (p+l-s)\min(y_n,d_n)-(c_1-s)y_n, & y_n\leqslant Q;\\ (p+l-s)\min(y_n,d_n)-(c_2-s)y_n, & y_n>Q, \end{cases} \tag{3.6}$$

$$g_n^y=\begin{cases} (p+l-s)\min(y,d_n)-(c_1-s)y, & y\leqslant Q;\\ (p+l-s)\min(y,d_n)-(c_2-s)y, & y>Q, \end{cases} \tag{3.7}$$

令 $p := p + l$, 则收益函数(3.6)与(3.2)相同. 因此, 将前面求解公式(2.2)中的 $p$ 替换为 $p + l - s$, $c_1$ 替换为 $c_1 - s$, $c_2$ 替换为 $c_2 - s$, 就得到了带有缺货损失费情形下的具体决策公式. 对于理论上的竞争性能分析, 同样地只需将式(2.4)中的 $p - s$ 替换为 $p + l - s$, $c_1$ 替换为 $c_1 - s$, $c_2$ 替换为 $c_2 - s$, 即有推论 3.2.

**推论 3.2** 对于价格数量折扣下带有回收价值和缺货损失费的多阶段报童模型, 按照在线策略(2.2)决策的累积收益满足

$$G_N \geqslant \max_{y \in [0,B]} G_N^y - (B^2(p + l - s)^2 + p + l - s - c_2 + c_1 + \ln(B\sqrt{N}))\sqrt{N}. \tag{3.8}$$

## 4 数值算例

本节用数值算例进一步验证本文给出的在线策略的竞争性能. 为方便计算, 设报童的订购量和需求量为整数. 报童的订购量按照如下方法整数化: 令 $y_n$ 为按照公式 (2.2) 计算出来的第 $n$ 阶段的订购量, 则报童的订购量为

$$y_n' = \begin{cases} y_n, & y_n\text{为整数;} \\ \lceil y_n \rceil, & \text{以}y_n - \lfloor y_n \rfloor\text{概率;} \\ \lfloor y_n \rfloor, & \text{以}\lceil y_n \rceil - y_n\text{概率,} \end{cases} \tag{4.1}$$

其中$\lfloor\ \rfloor$为向下取整, $\lceil\ \rceil$ 为向上取整. 这样整数化后能够保证 $Ey_n' = y_n$. 设报童的决策天数 $N = 60$ (2 个月), 数值算例中其他固定的常数取值为: $B = 40, Q = 25, p = 10$. 在集合 $\{1, 2, \cdots, 40\}$ 随机产生了 10 次 60 个整数, 作为报童每次在 60 天面对的需求序列.

表 1 给出了不同需求下在线策略与最优专家策略的累积收益比较, 其中 AS 和 BE 表示具有价格数量折扣多阶段报童模型的在线策略和最优专家策略: AS-S 和 BE-S 表示价格数量折扣下具有回收价值的多阶段报童模型的在线策略和最优专家策略: AS-S-S 和 BE-S-S 表示价格数量折扣下具有回收价值和缺货损失费的多阶段报童模型的在线策略和最优专家策略, 其中 $c_1 = 7, c_2 = 6, s = 5, l = 3$. 从表 1 可以看出不同需求下, 在线策略 AS, AS-S 和 AS-S-S 的累积收益都接近于对应最优专家策略 BE, BE-S 和 BE-S-S 的累积收益. 具有回收价值的在线策略 AS-S 的累积收益最多, 略高于带有缺货损失费的在线策略 AS-S-S 的累积收益, 但他们的累积收益都远远高于在线策略 AS 的累积收益, 说明回收价值和缺货损失费的引入对策略累积收益的影响很大. 考虑到回收价值和缺货损失费等因素对在线策略累积收益的影响. 表 2 给出了不同参数组合下在线策略和最优专家策略的累积收益. 表 2 的数据给出了与直观一致的结论: 回收价值越大、缺货损失费越小和购买价格越小时, 在线策略及其对应最优专家策略的累积收益就越大.

图 3~5 分别给出了同一需求序列下, 在线策略 AS 与最优专家策略 BE 的日累积收益比较, 在线策略 AS-S 与最优专家策略 BE-S 的日累积收益比较, 以及在线策略 AS-S-S 与最优专家策略 BE-S-S 的日累积收益比较. 从这三个图可以看出代表在线策略 AS(AS-S 或 AS-S-S)日累积收益的实线和代表最优专家策略BE(BE-S 或 BE-S-S)日累积收益的虚线都非常接近. 图 1 波动性大, 没有图 2 和图 3 表现的平稳, 说明了不考虑回收价值和缺货损失费的的在线策略 AS 的日累积收益变化大, 而考虑这两个因素的在线策略 AS-S 和 AS-S-S 的日累积收益变化平稳. 这三个图也表明了与表 1 相一致的结论: 在线策略 AS

的累积收益远远小于在线策略 AS-S 和 AS-S-S 的累积收益, 而在线策略 AS-S 的累积收益略高于在线策略 AS-S-S 的累积收益.

**表 1** 不同需求下在线策略和最优专家策略的累积收益

| 序号 | AS | BE | AS-S | BE-S | AS-S-S | BE-S-S |
|---|---|---|---|---|---|---|
| 1 | 1 452 | 1 570 | 3 899 | 3 995 | 3 783 | 3 795 |
| 2 | 1 354 | 1 440 | 3 884 | 3 980 | 3 740 | 3 800 |
| 3 | 1 876 | 2 190 | 4 457 | 4 620 | 4 308 | 4 540 |
| 4 | 1 989 | 2 130 | 4 555 | 4 610 | 4 472 | 4 470 |
| 5 | 2 015 | 2 210 | 4 594 | 4 800 | 4 453 | 4 705 |
| 6 | 1 810 | 1 900 | 4 245 | 4 355 | 4 147 | 4 255 |
| 7 | 1 227 | 1 470 | 3 892 | 3 990 | 3 769 | 3 825 |
| 8 | 1 636 | 1 650 | 4 069 | 4 145 | 3 929 | 4 005 |
| 9 | 1 226 | 1 670 | 3 895 | 4 050 | 3 690 | 3 825 |
| 10 | 939 | 1 100 | 3 299 | 3 375 | 3 171 | 3 205 |

**表 2** 不同参数组合下在线策略和最优专家策略的累积收益

| $(c_1, c_2, s, l)$ | AS | BE | AS-S | BE-S | AS-S-S | BE-S-S |
|---|---|---|---|---|---|---|
| (7, 6, 5, 3) | 1 626 | 1 740 | 4 188 | 4 285 | 4 093 | 4 220 |
| (8, 6, 5, 3) | 1 618 | 1 740 | 4 161 | 4 285 | 4 078 | 4 220 |
| (7, 6, 4, 3) | 1 626 | 1 740 | 3 490 | 3 618 | 3 212 | 3 204 |
| (7, 6, 5, 2) | 1 626 | 1 740 | 4 188 | 4 285 | 4 127 | 4 220 |



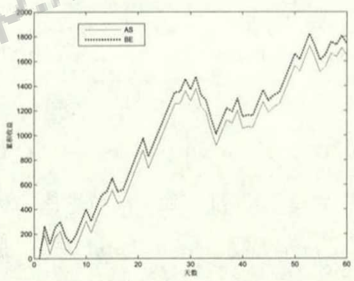



图 3 在线策略 AS 与最优专家策略 BE 的日累积收益比较

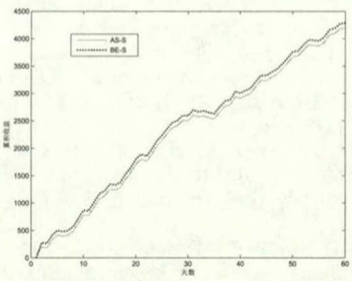



图 4 在线策略 AS-S 与最优专家策略 BE-S 的日累积收益比较

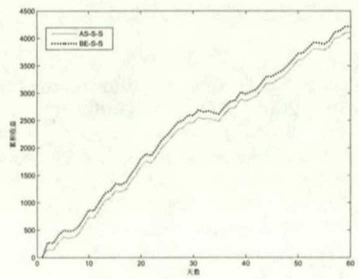



**图5** 在线策略AS-S-S 与最优专家策略BE-S-S 的日累积收益比较

## 5 结论

考虑价格数量折扣, 本文进一步应用计算机科学的弱集成算法给出了多阶段报童模型的在线策略及其理论保证. 主要基于固定订购量的专家策略给出了带有价格数量折扣的在线策略, 证明了该在线策略实现的累积收益能够与最优专家策略实现的累积收益相当. 进一步引入回收价值和缺货损失费, 将获得的在线策略及其理论结果进行了推广. 通过数值算例说明了本文给出的在线策略能够实现较好的累积收益, 追踪最优专家策略, 具有较好的竞争性能.

## 参考文献


[1] Petruzzi N C, Dada M. Pricing and newsvendor problem: A review with extension [J]. *Operations Research*, 1999, **47**(2): 183-149.

[2] Khouja M. The single-period (newsvendor) problem: Literature review and suggestions for future research [J]. *Omega*, 1999, **27**(5): 537-553.

[3] 罗春林. 基于贝叶斯信息更新的风险规避库存策略研究 [J]. 运筹学学报, 2013, **10**(1): 59-68.

[4] 黄松, 杨超, 张曦. 考虑战略顾客行为带预算约束的多产品报童问题 [J]. 中国管理科学, 2011, **19**(3): 70-78.

[5] 许民利, 李展. 基于CVaR准则具有预算约束和损失约束的报童决策 [J]. 控制与决策, 2013, **28**(11): 1614-1622.

[6] 周艳菊, 应仁仁, 陈晓红, 等. 基于前景理论的两产品报童的订货模型 [J]. 管理科学学报, 2013, **16**(11): 17-29.

[7] Scarf, H. *A Min-Max Solution of an Inventory Problem* [M]. California: Stanford University Press, 1958.

[8] Gallego G, Moon I. The distribution free newsboy problem: review and extensions [J]. *Journal of the Operational Research Society*, 1993, **44**(8): 825-834.

[9] Moon I, Choi S. Distribution free newsboy problem with balking [J]. *Journal of the Operational Research Society*, 1995, **46**(4): 537-542.

[10] Alfares H K, Elmorra H H. The distribution-free newsboy problem: extension to the shortage penalty case [J]. *International Journal of Production Economics*, 2005, **93-94**(8): 465-477.

[11] Levina T, Levin Y, McGill J, et al. Weak aggregating algorithm for the distribution-free perishable inventory problem [J]. *Operations Research Letters*, 2010, **38**(6): 516-521.

[12] El-Yaniv R, Kaniel R, Linial N. Competitive optimal on-line leasing [J]. *Algorithmica*, 1999, **25**(1): 116-140.

[13] Xu Y F, Xu W J, Li H Y. On the on-line rent-or-buy problem in probabilistic environments [J]. *Journal of Global Optimization*, 2007, **38**(1): 1-20.

[14] 张永, 张卫国, 徐维军. 可折旧设备在线租赁的随机性竞争策略 [J]. 管理科学学报, 2011, **14**(1): 69-77.

[15] Levina T, Shafer G. Portfolio selection and online learning [J]. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2008, **16**(4): 437-473.

[16] Wagner M R. Fully distribution-free profit maximization: the inventory management case [J]. *Mathematics of Operations Research*, 2010, **35**(4): 728-741.

[17] Wagner M R. Online lot-sizing problems with ordering, holding and shortage costs [J]. *Operations Research Letters*, 2011, **39**(2): 144-149.

[18] 张桂清, 徐寅峰. 概率预期下在线报童问题的最小风险策略 [J]. 中国管理科学, 2010, **18**(6): 131-137.

[19] 张桂清, 徐寅峰. 报童问题的最优竞争比策略及其风险补偿模型 [J]. 管理学报, 2011, **8**(1): 97-102.

[20] Ball M, Queyranne M. Toward robust revenue management: competitive analysis of online booking [J]. *Operations Research*, 2009, **57**(4): 950-963.

[21] Heuvel W van den, Wagelmans A P M. Worst case analysis for a general class of on-line lot-sizing heuristics [J]. *Operations Research*, 2010, **58**(1): 59-67.

[22] Kalnishkan Y, Vyugin M V. The weak aggregating algorithm and weak mixability [J]. *The Journal of Computer and System Sciences*, 2008, **74**(8): 1228-1244.

[23] Zhang Y, Vovk V, Zhang W G. Probability-free solutions to the non-stationary newsvendor problem [J]. *Annals of Operations Research*, 2014, **223**(1): 433-449.