中国应用数学通报
Bulletin of Chinese Applied Mathematics
2026 年第 04 卷第 02 期
Vol. 04, No. 02, 2026

研究性论文 http://www.scifootprint.com/BCAM/index.html

退货情形下部分延期交付库存系统的最优订货策略

苏 欣[1,2,*]，杨建奎[1,2]

(1. 北京邮电大学 数学科学学院，北京 100876；
2. 数学与信息网络教育部重点实验室(北京邮电大学)，北京 100876)

摘要：在随机库存管理研究中，当发生缺货时，管理者可通过延期交付的方式满足愿意等待的顾客需求，但现有研究较少考虑退货行为对订货策略的影响。事实上，大量的退货能够显著影响库存水平。本文针对有退货情形的库存系统，综合考虑顾客的退货行为以及缺货时顾客耐心的随机性，证明了在零交货期条件下，基本库存策略为最优策略，并给出其解析解。有趣的是，最优库存水平可能为负值。

关键词：库存管理；退货；部分延期交付；最优策略

DOI：10.48014/bcam.20260212001

引用格式：苏欣，杨建奎．退货情形下部分延期交付库存系统的最优订货策略[J]．中国应用数学通报，2026，4(2)：29-42.

1 引言

随着我国零售渠道的快速扩张，商品退货已逐渐演变为供应链运营中的常态性环节。国家统计局数据显示[1]，2025 年 1—11 月份全国网上零售额达 14.46 万亿元，同比增长 9.1%，其中实物商品网上零售额 11.82 万亿元，占社会消费品零售总额比重约 25.9%，线上渠道在居民消费中的重要性持续提升。在此背景下，退货政策的制度化也进一步强化了退货的普遍性，以国家市场监管总局发布的《网络购买商品七日无理由退货暂行办法》为例[2]，线上交易退货具备了明确依据，这使得退货在消费者决策与平台运营中更加常见。与此同时，退货率在部分品类中呈现出显著的高位特征。以服饰类商品销售情况为例，根据网经社电子商务研究中心数据[3]，电商服饰领域退货率长期居高不下，女装退货率可达 50%～60%，直播带货情境下甚至出现 80%及以上的极端退货现象。大量且波动的退货不仅抬升了逆向物流与再处理成本，也将改变有效需求与库存动态，从而对库存控制策略提出更高要求。因此，如何在高退货常态下制定行之有效的订货策略，已成为亟待解决的重要问题。

零交货期的周期性盘点系统作为供应链管理中的基础模型，其最优策略的求解始终是学术研究的重点问题。自 Harris[4] 提出经典经济订货量(Economic Order Quantity，简写为 EOQ)模型以来，随机库存系统的研究在 Arrow 与 Dvoretzky[5,6] 等学者的推动下，逐渐建立起以动态规划为核心的理论框架。Karlin 与 Zipkin[7,8] 进一步证明，在延期交付和需求损失模型中，基本库存策略具有最优性。随后，研究者围绕该策略展开系统性分析：Janakiraman 与 Roundy[9] 指出基本库存策略下长期平均成本对库存水平呈凸性；Janakiraman 等人[10] 对基本库存策略下完全积压系统与完全损失系统的成本进行了对比分析。然而，上述研究普遍基于顾客对缺货行为反应为确定性的假设，即需求

* 通讯作者 Corresponding author：苏欣，2023111631@bupt.cn

收稿日期：2026-02-12； **录用日期**：2026-03-24； **发表日期**：2026-06-28

基金项目：本项研究得到了国家自然科学基金项目(项目编号：12571344)的资助。

要么完全延期交付，要么完全流失，未能刻画现实中顾客耐心程度的随机性。实际上，顾客行为往往介于这两种极端情形之间：部分顾客愿意等待下一周期交付，而其余顾客则选择放弃等待并取消订单[11]。在考虑顾客随机耐心的部分延期交付库存系统中，相关研究仍较为有限，其中 Andrew 等人[12]在零交货期条件下引入顾客随机耐心行为，并给出其最优策略的解析解。

此外，考虑到退货现象在供应链管理实践中普遍存在[13]，大量研究从不同角度探讨了退货因素对库存决策及供应链利润的影响。Gokbayrak 和 Kayis[14]假设某一周期的退货量依赖于前一周期的销售数量，并给出相应的最优策略；Braz 等人[15]分析了闭环供应链中长鞭效应的成因与缓解因素，指出高退货率有助于缓解长鞭效应；Narang 和 Shankar[16]采用双重差分法研究了购物应用程序的使用对退货量产生的影响。尽管上述文献从不同层面揭示了退货的形成机制及其影响，但相关研究多聚焦于退货行为本身，而没有人对顾客的随机耐心与退货行为共同纳入考虑。同时，现有文献通常仅从定性层面探讨退货因素对库存水平的影响，本文则在此基础上，给出了最优订货量的显式解析解。

本文构建并分析了一个同时考虑顾客耐心随机性与退货行为的库存模型，允许顾客在面临库存短缺时表现出不同的等待态度，并且将退货行为纳入考虑。值得注意的是，本模型允许退货数量大于需求数量。假设退货经简单处理后可以重新销售，本文得到了最优库存水平的显示解。有趣的是，这个最优解满足的方程与 Andrew 等人[12]给出的无退货情形下的最优策略形式上是一致的。不同的是，无退货情形下的最优库存水平始终为正，而考虑退货因素后，最优库存水平 s 则可能取负值。这体现了当退货能够在概率意义上大量补充库存时，允许库存水平为负值是一种理性的最优选择。这一结果为实际供应链管理中处理大量退货与顾客随机耐心的策略制定提供了理论依据。

2 模型与假设

我们考虑一个周期性盘点的单产品类型的库存系统。设 I_i 为第 i 个周期初决策者所持有的库存。D_i 为第 i 期的需求，其均值为 d，假定$\{D_i:i\geqslant 1\}$为独立同分布的离散型非负随机变量序列。定义 R_i 为第 i 期被退回的商品数，其均值为 r。当发生退货时，原销售交易所取得的销售收入需全额退还。假设所有退回的商品经简单处理后均可重新出售。本文不考虑因退货产生的附加成本（如退货运费、检测翻新费用、二次销售折价损失以及客户服务成本等）。

考虑到品牌官方专卖店的运营特点，消费者可以将从任何授权渠道购买的同品牌产品退回至该店，故假设$\{R_i:i\geqslant 1\}$为独立同分布的离散型非负整数值随机变量序列，且与$\{D_i:i\geqslant 1\}$相互独立。

本文仅考虑订货即时交付的情形，即订货交货期为 0。设第 i 周期的订货量为 Q_i。在本文中，订货量 Q_i 是每周期唯一的决策变量，其他系统参数均为预先给定的常量或随机变量。

记 U_i 为本周期因库存不足而未被满足的需求量。在每周期末，系统将评估所有未被满足的需求，每个未被满足的需求会以概率 p 等待下一周期交付，或以概率$(1-p)$选择放弃等待。当 $0<p<1$ 时，称此情形为 P 系统。特别地，当 $p=1$ 时，所有未满足的需求均延期交付，我们称之为 B 系统；当 $p=0$ 时，所有未满足的需求均放弃等待，我们称之为 L 系统。

为刻画上述特征，我们引入伯努利随机变量序列$\{\xi_{i;l}:i\in N^+,l\geqslant 1\}$。其中，当第 i 周期第 l 个未被满足的需求有耐心等待延期交付时，$\xi_{i;l}=1$，否则 $\xi_{i;l}=0$。由定义可知 $\xi_{i;l}$ 的期望为 p。由此，第 i 周期为满足的需求量为

$$U_i=(B_{i-1}+D_i-I_i-Q_i-R_i)^+ \tag{2-1}$$

记 $B_i=\sum_{l=1}^{U_i}\xi_{i;l}=\xi_{i;[1,U_i]}$，$L_i=U_i-\xi_{i;[1,U_i]}$。$B_i$ 表示第 i 个周期未满足且继续等待的需求数量，L_i 表示第 i 个周期未满足且最终放弃等待的需求数量。默认地，$B_0=L_0=0$。特别地，在 B 系统中，$B_i=U_i$，$L_i=0$；在 L 系统中，$B_i=0$，$L_i=U_i$。假定$\{\xi_{i;l}:i\in N^+,l\geqslant 1\}$与$\{R_i:i\geqslant 1\}$、$\{D_i:i\geqslant 1\}$相互独立。

下面，我们按步骤给出库存系统的状态更新过程：

(S1)在第 i 周期初，决策者可以观测到当前库存 I_i，以及从上一周期未满足且继续等待的需求数

量 B_{i-1}。此时,定义第 i 期的净库存

$$N_i = I_i - B_{i-1} \tag{2-2}$$

注意到 $I_i \cdot B_{i-1} = 0$。这是因为,如果当前周期存在现有库存,即 $I_i > 0$,这意味着上一周期的需求已全部满足,不会有遗留的未完成的订单,即 $B_{i-1} = 0$。

(S2)第 i 周期开始后,决策者根据 I_i 和 B_{i-1},确定本周期的订购数量 Q_i。

(S3)利用现有库存 I_i、本周期的订货量 Q_i 和被退回的商品数量 R_i,来满足当期需求 D_i 和上一周期未满足且继续等待的需求 B_{i-1}。因此,本周期的实际销售量由本周期的可用供给量($I_i + Q_i + R_i$)和总需求量($B_{i-1} + D_i$)中的较小值决定,即($I_i + Q_i + R_i$)$\wedge$($B_{i-1} + D_i$),且

$$I_{i+1} = (I_i + Q_i + R_i - B_{i-1} - D_i)^+ \tag{2-3}$$

当可用供给量($I_i + Q_i + R_i$)大于或等于总需求量($B_{i-1} + D_i$)时,意味着当期的需求不仅能全部交付,还可能会产生剩余库存。这部分剩余库存将结转至下一周期继续使用,并产生相应的库存持有成本。本周期末的剩余库存,即下一周期初的现有库存 $I_{i+1} = (I_i + Q_i + R_i - B_{i-1} - D_i)^+ = I_i + Q_i + R_i - B_{i-1} - D_i$。由于本周期没有未交付的需求,所以 $U_i = (B_{i-1} + D_i - I_i - Q_i - R_i)^+ = 0$。

当可用供给量($I_i + Q_i + R_i$)小于总需求量($B_{i-1} + D_i$)时,本周期会出现需求无法完全交付的情况,未交付的需求 $U_i = (B_{i-1} + D_i - I_i - Q_i - R_i)^+ = B_{i-1} + D_i - I_i - Q_i - R_i$。本周期的剩余库存,即下一周期期初的现有库存 $I_{i+1} = (I_i + Q_i + R_i - B_{i-1} - D_i)^+ = 0$。

(S4)在第 i 周期末,如果存在未交付的需求,则每个未交付的需求 $\xi_{i;l}$(其中 $l = 1, 2, \cdots, U_i$)会以概率 p 等待下一周期交付,或以概率($1-p$)放弃等待,取消订单。于是,周期未满足且继续等待的需求数量

$$B_i = \sum_{l=1}^{U_i} \xi_{i;l} = \xi_{i;[1,U_i]} \tag{2-4}$$

默认地,$U_i = 0$ 意味着 $B_i = 0$。

(S5)在第($i+1$)周期初,决策者可以观测到本周期未满足且继续等待的需求数量 B_i,以及其目前所持有的库存

$$\begin{aligned} I_{i+1} &= (I_i + Q_i + R_i - B_{i-1} - D_i)^+ \\ &= (N_i + Q_i + R_i - D_i)^+ \end{aligned} \tag{2-5}$$

这是由上周期末的剩余库存结转而来的。进一步地,第($i+1$)期的净库存为

$$\begin{aligned} N_{i+1} &= I_{i+1} - B_i \\ &= (N_i + Q_i + R_i - D_i)^+ - \xi_{i;[1,U_i]} \\ &= (N_i + Q_i + R_i - D_i)^+ \\ &\quad - \xi_{i;[1,(D_i - N_i - Q_i - R_i)+]} \end{aligned}$$

基于上述状态,系统在第($i+1$)个周期继续按上述规则更新。

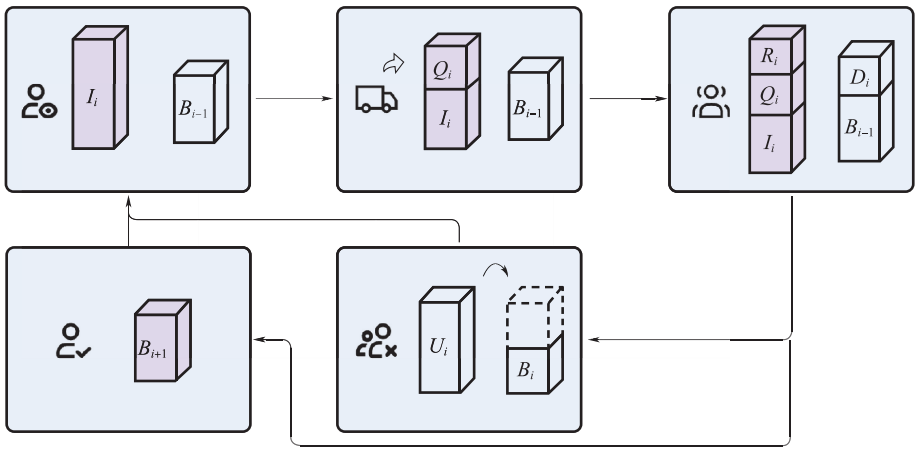

图 1 库存流转示意图

Fig. 1 Inventory flow diagram

定义 $X_i = (I_i, B_{i-1})$第 i 周期的状态,这是一个二维非负整数型随机变量。其中,第一个分量是第 i 周期初的现有库存 I_i,第二个分量是第 $i-1$ 周期延至第 i 期交付的需求 B_{i-1}。特别地,初始状态

$X_1=(I_1,B_0)$。在第 i 期,决策者根据本周期的状态 X_i,即当前现有的库存 I_i 和上一周期未满足且继续等待的需求 B_{i-1},决定本周期的订购数量 Q_i。我们将整个时间序列上的订购决策序列$(Q_1,Q_2,\cdots)$记为 π,称之为订购策略。所有订购策略 π 组成的集合记为 A。

根据上述库存系统的动态描述,第 i 期状态 $X_i=(I_i,B_{i-1})$完整地涵盖了前 $i-1$ 期所有库存变动的累积效应,即系统状态序列$\{X_i:i\geqslant 1\}$是一个马尔可夫链,它的状态空间为 $X_i=\{(I_i,B_{i-1}):I_i\cdot B_{i-1}=0\}\subseteq\mathbb{N}^2$。

又注意到 I_i 和 B_{i-1} 均为非负整数,根据定义(2-2)及

$$I_i\cdot B_{i-1}=0 \tag{2-6}$$

可得

$$(N_i)^+=(I_i-B_{i-1})^+=I_i \tag{2-7}$$

$$(N_i)^-=(I_i-B_{i-1})^-=B_{i-1} \tag{2-8}$$

即状态空间可以用 N_i 来完全表示。进一步地,状态空间从$\mathbb{N}^2$缩减为$\mathbb{N}$。

在每个运营周期内,系统的利润由产品销售收入扣除销售退款和库存持有成本构成。定义 u 为单位产品的销售价格,h 为单位产品的持有成本。在策略 $\pi\in\mathcal{A}$下,第 i 期的实际销售量为$(I_i+Q_i+R_i)\wedge(B_{i-1}+D_i)$,第 i 期接收的退货量为R_i,期末结转库存量为 $I_{i+1}=(I_i+Q_i+R_i-B_{i-1}-D_i)^+$。因此,在给定 N_i 的条件下,第 i 期的期望利润为:

$$\begin{aligned}C_i^\pi(N_i)=\mathbb{E}_{N_i}[&u\cdot(I_i+Q_i+R_i)\wedge(B_{i-1}+D_i)-\\&u\cdot R_i-h\cdot(I_i+Q_i+R_i-\\&B_{i-1}-D_i)^+]\end{aligned}$$

令初始状态 $N_1=x$,策略 π 下的长期平均利润表示如下

$$C^\pi(u,h;x)=\liminf_{k\to\infty}\frac{1}{k}\mathbb{E}\Big(\sum_{i=1}^{k}C_i^\pi(N_i)\mid N_1=x\Big)$$

本文旨在求解最优库存策略 $\pi^*\in A$,使系统长期平均利润最大化。

3 交货期为零时 P 系统的最优策略

首先介绍基本库存策略的概念。

具有水平为 s 的基本库存策略是一种经典的库存控制策略,它的执行机制如下:在每一周期,决策者通过计算目标水平 s 与当前净库存 N_i^s 的差来决定补货量,当且仅当差额为正时进行补货。在这一策略下,每期订购数量 Q_i^s 的表达式为 $Q_i^s=(s-N_i^s)^+$。基本库存策略的核心思想是通过动态调整订购量,将系统的净库存 N_i^s 维持在一个预设的固定水平 s。

Zipkin[5]已经证明,当交货期为 0 时,基本库存策略对于 L 系统和 B 系统都是最优的。下面证明 P 系统也有类似的结论。

引理 3.1 P 系统在策略 $\pi=(Q_1,Q_2,\cdots)$下的状态更新过程可以由以下递推方程描述:

$N_i=N_{i-1}+Q_{i-1}+R_{i-1}-D_{i-1}+(U_{i-1}-\xi_{i-1;[1,U_{i-1}]})$其中 N_1 是初始状态。

证明 对于 $i\geqslant 2$,分两种情况进行讨论。

Ⅰ. 当 $N_{i-1}+Q_{i-1}+R_{i-1}-D_{i-1}\leqslant 0$ 时,

$$\begin{aligned}&(D_{i-1}-N_{i-1}-Q_{i-1}-R_{i-1})^+=\\&\quad D_{i-1}-N_{i-1}-Q_{i-1}-R_{i-1}\end{aligned} \tag{3-1}$$

且由(2-5)式可得

$$I_i=0 \tag{3-2}$$

由式(2-1)及式(2-2)可得

$$\begin{aligned}U_{i-1}&=(B_{i-2}+D_{i-1}-I_{i-1}-Q_{i-1}-R_{i-1})^+\\&=(D_{i-1}-N_{i-1}-Q_{i-1}-R_{i-1})^+\end{aligned} \tag{3-3}$$

所以

$N_i=I_i-B_{i-1}=-B_{i-1}$(由(2-1)式和(3-2)式可得)

$$\begin{aligned}&=(N_{i-1}+Q_{i-1}+R_{i-1}-D_{i-1})+(D_{i-1}-\\&\quad N_{i-1}-Q_{i-1}-R_{i-1})-B_{i-1}\\&=N_{i-1}+Q_{i-1}+R_{i-1}-D_{i-1}+\\&\quad(D_{i-1}-N_{i-1}-Q_{i-1}-R_{i-1})^+-B_{i-1}\end{aligned}$$

(由(3-2)式可得)

$$\begin{aligned}&=N_{i-1}+Q_{i-1}+R_{i-1}-D_{i-1}+\\&\quad(U_{i-1}-\xi_{i-1;[1,U_{i-1}]})\end{aligned}$$

(由(3-3)式和(2-4)式可得)

Ⅱ. 当 $N_{i-1}+Q_{i-1}+R_{i-1}-D_{i-1}>0$ 时,由式(2-1)及式(2-2)可得

$$U_{i-1}=(D_{i-1}-N_{i-1}-Q_{i-1}-R_{i-1})^+=0$$

进一步地,由式(2-4),有 $B_{i-1}=0$,从而 $N_i=I_i-B_{i-1}=I_i\geqslant 0$

所以

$N_i=I_i=(N_{i-1}+Q_{i-1}+R_{i-1}-D_{i-1})^+$(由

(2-5)式可得)

$$=N_{i-1}+Q_{i-1}+R_{i-1}-D_{i-1}$$
$$=N_{i-1}+Q_{i-1}+R_{i-1}-D_{i-1}+U_{i-1}-B_{i-1}$$

综合Ⅰ,Ⅱ两种情况,引理得证。

我们将使用消失折扣方法来求解长期平均利润方程的最优策略 π^*。为此,考虑最优策略 π^* 下具有折扣因子 α 的长期平均利润方程:

$$C^{\alpha}(u,h;x)=\max_{\pi\in\mathcal{A}}\mathbb{E}\left\{\sum_{i=1}^{\infty}\alpha^{i-1}C_i^{\pi}(N_i)\mid N_1=x\right\}$$
$$=\max_{\pi\in\mathcal{A}}\mathbb{E}\left\{\sum_{i=1}^{\mathbb{E}}\alpha^{i-1}\mathbb{E}_{N_i}[u\cdot(I_i+Q_i+R_i)\wedge(B_{i-1}+D_i)-u\cdot R_i-h\cdot(I_i+Q_i+R_i-B_{i-1}-D_i)^+]\mid N_1=x\right\} \tag{3-4}$$

引理 3.2 最优策略 π^* 下具有折扣因子 α 的长期平均利润方程可以有如下表示形式:

$$C^{\alpha}(u,h;x)=\max_{\pi\in\mathcal{A}}\mathbb{E}\left\{\sum_{i=1}^{\infty}\alpha^{i-1}[\mathbb{E}_{N_i}(u\cdot[(Q_i+R_i)\wedge(N_i)^-+(N_i+Q_i+R_i)^+\wedge D_i]-u\cdot R_i-h\cdot(N_i+Q_i+R_i-D_i)^+)]\mid N_1=x\right\} \tag{3-5}$$

证明 由(2-2)式可知 $N_i=I_i-B_{i-1}$。进一步对比(3-4)式和(3-5)式的差异,可以发现,要证原命题只需证明下式成立

$$(I_i+Q_i+R_i)\wedge(B_{i-1}+D_i)=(Q_i+R_i)\wedge(N_i)^-+(N_i+Q_i+R_i)^+\wedge D_i$$

分两种情况进行讨论。

Ⅰ.当 $B_{i-1}=0$ 时,由(2-2)式可得

$$(Q_i+R_i)\wedge(N_i)^-=(Q_i+R_i)\wedge(I_i)^-=(Q_i+R_i)\wedge 0=0 \tag{3-6}$$
$$N_i=I_i-B_{i-1}=I_i\geqslant 0 \tag{3-7}$$

于是

$$(I_i+Q_i+R_i)\wedge(B_{i-1}+D_i)=(I_i+Q_i+R_i)\wedge D_i=0+(I_i+Q_i+R_i)^+\wedge D_i=(Q_i+R_i)\wedge(N_i)^-+(N_i+Q_i+R_i)^+\wedge D_i$$

最后一个等号由(3-6)、(3-7)可得。

Ⅱ.当 $B_{i-1}>0$ 时,由(2-6)式可得 $I_i=0$。于是 $(I_i+Q_i+R_i)\wedge(B_{i-1}+D_i)=(Q_i+R_i)\wedge(B_{i-1}+D_i)$,也就是要证明 $(Q_i+R_i)\wedge(B_{i-1}+D_i)=(Q_i+R_i)\wedge(N_i)^-+(N_i+Q_i+R_i)^+\wedge D_i$。

当 $Q_i+R_i<B_{i-1}$ 时,

$$(Q_i+R_i)\wedge(B_{i-1}+D_i)=Q_i+R_i=Q_i+R_i+0\wedge D_i$$
$$=(Q_i+R_i)\wedge B_{i-1}+(Q_i+R_i-B_{i-1})^+\wedge D_i$$
$$=(Q_i+R_i)\wedge(N_i)^-+(Q_i+R_i+I_i-B_{i-1})^+\wedge D_i$$

(由(2-8)式及 $I_i=0$ 可得)

$$=(Q_i+R_i)\wedge(N_i)^-+(N_i+Q_i+R_i)^+\wedge D_i$$

(由(2-2)式可得)

当 $Q_i+R_i\geqslant B_{i-1}$ 时,进一步分两种子情况讨论:

Ⅱ.1 若 $Q_i+R_i-B_{i-1}>D_i$,则有

$$(Q_i+R_i)\wedge(B_{i-1}+D_i)=B_{i-1}+D_i$$
$$=(Q_i+R_i)\wedge B_{i-1}+(Q_i+R_i-B_{i-1})^+\wedge D_i$$
$$=(Q_i+R_i)\wedge(N_i)^-+(N_i+Q_i+R_i)^+\wedge D_i$$

(与Ⅱ.1 的情形同理)

Ⅱ.2 若 $Q_i+R_i-B_{i-1}\leqslant D_i$,则有

$$(Q_i+R_i)\wedge(B_{i-1}+D_i)=Q_i+R_i$$
$$=B_{i-1}+(Q_i+R_i-B_{i-1})$$
$$=(Q_i+R_i)\wedge B_{i-1}+(Q_i+R_i-B_{i-1})^+\wedge D_i$$
$$=(Q_i+R_i)\wedge(N_i)^-+(N_i+Q_i+R_i)^+\wedge D_i$$

(与Ⅱ.1 的情形同理)

综合Ⅰ,Ⅱ两种情况,原命题得证。

给定 x,Q,定义第 1 周期的单期利润 $L(x,Q)$ 为

$$L(x,Q)=\mathbb{E}(u\cdot[(Q+R_1)\wedge x^-+(x+Q+R_1)^+\wedge D_1]-u\cdot R_1-h\cdot(x+Q+R_1-D_1)^+) \tag{3-8}$$

回顾式(2-1)、(2-2)、(2-4)、(2-5),我们有

$$N_{i+1}=I_{i+1}-B_i=(N_i+Q_i+R_i-D_i)^+-\xi_{i;[1,U_i]}=(N_i+Q_i+R_i-D_i)^+-\xi_{i;[1,(D_i-N_i-Q_i-R_i)^+]}$$

于是可以给出长期平均利润方程 $C^{\alpha}(u,h;x)$ 相应的动态规划方程

$$C^{\alpha}(u,h;x)=\sup_{Q\in N^+}\{\mathcal{L}(x,Q)+\alpha\,\mathbb{E}[C^{\alpha}(u,h;(x+Q+R_1-D_1)^+-\xi_{1;[1,(D_1-x-Q-R_1)^+]})]\} \tag{3-9}$$

引理 3.3 动态规划方程(3-9)式可以等价地

写为如下形式：

$$\begin{aligned}\mathcal{C}^{\alpha}(u,h;x)-ux^{-}=\sup_{Q\in N^{+}}\{&-u\cdot\mathbb{E}[(x+Q+R_1)^{-}]+\\&\mathbb{E}(u\cdot[(x+Q+R_1)^{+}\wedge D_1]-\\&u\cdot R_1-h\cdot(x+Q+R_1-\\&D_1)^{+})+au\,\mathbb{E}[((x+Q+R_1-\\&D_1)^{+}-\xi_{1;[1,(D_1-x-Q-R_1)^{+}]})^{-}]+\\&\alpha\,\mathbb{E}[\mathcal{C}^{\alpha}(u,h;(x+Q+R_1-\\&D_1)^{+}-\xi_{1;[1,(D_1-x-Q-R_1)^{+}]})-\\&u((x+Q+R_1-D_1)^{+}-\\&\xi_{1;[1,(D_1-x-Q-R_1)^{+}]})^{-}]\}\end{aligned}\tag{3-10}$$

证明 我们首先证明下式成立

$$(Q+R_1)\wedge x^{-}-x^{-}=-(x+Q+R_1)^{-}$$

由于 Q、R_1 非负，所以

$(Q+R_1)\geqslant x^{-}\Leftrightarrow x$ 非负，或 x 是负数且 $(Q+R_1)\geqslant -x$

$x^{-}\geqslant(Q+R_1)\Leftrightarrow x$ 为负，且 $-x\geqslant(Q+R_1)$

于是有

$$(Q+R_1)\wedge x^{-}-x^{-}=\begin{cases}x^{-}-x^{-}=0 & \text{当}(Q+R_1)\geqslant x^{-}\text{时}\\ Q+R_1-x^{-}=x+Q+R_1 & \text{当 } x^{-}\geqslant(Q+R_1)\text{时}\end{cases}$$

$$-(x+Q+R_1)^{-}=\begin{cases}0 & \text{当}(Q+R_1)\geqslant x^{-}\text{时}\\ -[-(x+Q+R_1)] & \\ =x+Q+R_1 & \text{当 } x^{-}\geqslant(Q+R_1)\text{时}\end{cases}$$

从而我们证明了

$$(Q+R_1)\wedge x^{-}-x^{-}=-(x+Q+R_1)^{-}$$

于是有

$$\begin{aligned}\mathbb{E}[(Q+R_1)\wedge x^{-}]-x^{-}&=\mathbb{E}[(Q+R_1)\wedge x^{-}-x^{-}]\\&=-\mathbb{E}[(x+Q+R_1)^{-}]\end{aligned}$$

从而

$$\begin{aligned}\mathcal{L}(x,Q)-ux^{-}=&\mathbb{E}(u\cdot[(Q+R_1)\wedge x^{-}+\\&(x+Q+R_1)^{+}\wedge D_1]-u\cdot R_1-\\&h\cdot(x+Q+R_1-D_1)^{+})-\\&ux^{-}\text{（由(3-8)式可得）}\\=&-u\cdot\mathbb{E}[(x+Q+R_1)^{-}]+\\&\mathbb{E}(u\cdot[(x+Q+R_1)^{+}\wedge D_1]-\\&u\cdot R_1-h\cdot(x+Q+R_1-\\&D_1)^{+})\end{aligned}\tag{3-11}$$

$$\begin{aligned}\mathcal{C}^{\alpha}(u,h;x)-ux^{-}=\sup_{Q\in N^{+}}\{&\mathcal{L}(x,Q)+\alpha\,\mathbb{E}[C^{\alpha}(u,h;\\&(x+Q+R_1-D_1)^{+}-\\&\xi_{1;[1,(D_1-x-Q-R_1)^{+}]})]\}-\\&ux^{-}\text{（由(3-9)式可得）}\\=\sup_{Q\in N^{+}}\{&-u\cdot\mathbb{E}[(x+Q+R_1)^{-}]+\\&\mathbb{E}(u\cdot[(x+Q+R_1)^{+}\wedge D_1]-\\&u\cdot R_1-h\cdot(x+Q+R_1-\\&D_1)^{+})+au\,\mathbb{E}[((x+Q+\\&R_1-D_1)^{+}-\\&\xi_{1;[1,(D_1-x-Q-R_1)^{+}]})^{-}]+\\&\alpha\,\mathbb{E}[\mathcal{C}^{\alpha}(u,h;(x+Q+R_1-\\&D_1)^{+}-\xi_{1;[1,(D_1-x-Q-R_1)^{+}]})-\\&u((x+Q+R_1-D_1)^{+}-\\&\xi_{1;[1,(D_1-x-Q-R_1)^{+}]})^{-}]\}\\&\text{（由(3-11)式可得）}\end{aligned}$$

命题得证。

引理 3.4 令

$$\mathcal{C}^{\alpha,o}(u,h;x)=\mathcal{C}^{\alpha}(u,h;x)-ux^{-}\tag{3-12}$$

$$\begin{aligned}\mathcal{L}^{o}(x)=&-u\cdot\mathbb{E}[(x+R_1)^{-}]+\\&\mathbb{E}[u\cdot((x+R_1)^{+}\wedge D_1)-u\cdot R_1-\\&h(x+R_1-D_1)^{+}+\alpha u\cdot((x+R_1-D_1)^{+}-\\&\xi_{1;[1,(D_1-R_1-x)^{+}]})^{-}]\end{aligned}\tag{3-13}$$

则动态规划方程可以进一步写为

$$\begin{aligned}\mathcal{C}^{\alpha,o}(u,h;x)=\sup_{y\geqslant x}\{&\mathcal{L}^{o}(y)+\alpha\,\mathbb{E}[\mathcal{C}^{\alpha,o}(u,h;\\&(y+R_1-D_1)^{+}-\xi_{1;[1,(D-y-R)^{+}]})]\}\end{aligned}$$

证明 由 $\mathcal{L}^{o}(x)$ 的定义可得

$$\begin{aligned}\mathcal{L}^{o}(x+Q)=&-u\cdot\mathbb{E}[(x+Q+R_1)^{-}]+\\&\mathbb{E}(u\cdot[(x+Q+R_1)^{+}\wedge D_1]-\\&u\cdot R_1-h\cdot(x+Q+R_1-D_1)^{+})+\\&au\,\mathbb{E}[((x+Q+R_1-D_1)^{+}-\\&\xi_{1;[1,(D_1-R_1-x-Q)^{+}]})^{-}]\end{aligned}$$

由 $\mathcal{C}^{\alpha,o}(u,h;x)$ 的定义可得

$$\begin{aligned}\mathcal{C}^{\alpha,o}(u,h;(x+Q+R_1-D_1)^{+}-&\\\xi_{1;[1,(D_1-R_1-x-Q-R_1)^{+}]})=&\mathcal{C}^{\alpha}(u,h;(x+Q+R_1-\\&D_1)^{+}-\xi_{1;[1,(D_1-R_1-x-Q)^{+}]})-\\&u((x+Q+R_1-D_1)^{+}-\\&\xi_{1;[1,(D_1-R_1-x-Q)^{+}]})^{-}\end{aligned}$$

从而

$$\mathcal{C}^{\alpha,o}(u,h;x)=\mathcal{C}^{\alpha}(u,h;x)-ux^{-}$$

$$
\begin{aligned}
&= \sup_{Q\in N^+}\{-u\cdot\mathbb{E}[(x+Q+R_1)^-]+\\
&\quad \mathbb{E}(u\cdot[(x+Q+R_1)^+\wedge D_1]-\\
&\quad u\cdot R_1-h\cdot(x+Q+R_1-D_1)^+)+\\
&\quad au\,\mathbb{E}[((x+Q+R_1-D_1)^+-\\
&\quad \xi_{1;[1,(D_1-R_1-x-Q)^+]})^-]+\\
&\quad \alpha\,\mathbb{E}[\mathcal{C}^{\alpha}(u,h;(x+Q+R_1-D_1)^+-\\
&\quad \xi_{1;[1,(D_1-R_1-x-Q)^+]})-\\
&\quad u((x+Q+R_1-D_1)^+-\\
&\quad \xi_{1;[1,(D_1-R_1-x-Q)^+]})^-]\}\\
&\quad (\text{由(3-10)式可得})\\
&= \sup_{Q\in N^+}\{\mathcal{L}^{o}(x+Q)+\alpha\,\mathbb{E}[\mathcal{C}^{\alpha,o}(u,h;\\
&\quad (x+Q+R_1-D_1)^+-\\
&\quad \xi_{1;[1,(D_1-R_1-x-Q)^+]})]\}
\end{aligned}
$$

令 $y=x+Q$,从而 $Q\in N^+$ 等价于 $y\geqslant x$,则上式等价于

$$
\begin{aligned}
\mathcal{C}^{\alpha,o}(u,h;x)=\sup_{y\geqslant x}\{&\mathcal{L}^{o}(y)+\\
&\alpha\,\mathbb{E}[\mathcal{C}^{\alpha,o}(u,h;(y+R_1-D_1)^+-\\
&\xi_{1;[1,(D_1-y-R_1)^+]})]\}
\end{aligned}
$$

我们的目标是求解上述的动态规划方程,即求解$\mathcal{C}^{\alpha,o}(u,h;x)$的最大值点。为此,我们首先回顾拟凹函数的定义。

定义 设 S 是$\mathbb{R}^n$ 中的非空凸集,$f(x)$是定义在 S 上的实值函数。对于$\forall x,y\in S,\lambda\in(0,1)$,若满足

$$f(\lambda x+(1-\lambda)y)\geqslant\min\{f(x),f(y)\}$$

则称 $f(x)$为拟凹函数。

引理 3.5 考虑一元函数的情形,已知 $f(x)$是定义在区间 J 上的拟凹函数,$g(x)$是定义在区间 I 上的单调不减函数,且$g(I)\subseteq J$,则 $h(x)=f(g(x))$是拟凹函数。

证明 因为$g(x)$是定义在区间 I 上的单调不减函数,所以对于$\forall x_1,x_2\in I,\lambda\in(0,1)$,有

$$
\begin{aligned}
\max\{g(x_1),g(x_2)\}&\geqslant g(\lambda x_1+(1-\lambda)x_2)\\
&\geqslant\min\{g(x_1),g(x_2)\}
\end{aligned}
$$

所以,$\exists\mu\in(0,1)$使得

$$g(\lambda x_1+(1-\lambda)x_2)=\mu g(x_1)+(1-\mu)g(x_2)$$

因为 $f(x)$是定义在区间 J 上的拟凹函数,所以对于$g(x_1),g(x_2)\in J,\mu\in(0,1)$有

$$
\begin{aligned}
f(g(\lambda x_1+(1-\lambda)x_2))&=f(\mu g(x_1)+\\
&\quad(1-\mu)g(x_2))\\
&\geqslant\min\{f(g(x_1)),\\
&\quad f(g(x_2))\}
\end{aligned}
$$

于是,我们证明了对于$\forall x_1,x_2\in I,\in(0,1)$ $f(g(\lambda x_1+(1-\lambda)x_2))\geqslant\min\{f(g(x_1)),f(g(x_2))\}$

即 $h(x)=f(g(x))$是拟凹函数。

引理 3.6 对于任意的 $x\in X$,都有

$$
\begin{aligned}
&((x+R_1-D_1)^+-\xi_{1;[1,(D_1-R_1-x)^+]})^-\\
&\quad=\xi_{1;[1,(D_1-R_1-x)^+]}
\end{aligned}
$$

证明 分两种情况进行讨论。

当 $D_1-R_1\geqslant x$ 时,

$$
\begin{aligned}
&((x+R_1-D_1)^+-\xi_{1;[1,(D_1-R_1-x)^+]})^-\\
&\quad=(0-\xi_{1;[1,(D_1-R_1-x)^+]})^-\\
&\quad=\xi_{1;[1,(D_1-R_1-x)^+]}
\end{aligned}
$$

当 $D_1-R_1<x$ 时,

$$
\begin{aligned}
&((x+R_1-D_1)^+-\xi_{1;[1,(D_1-R_1-x)^+]})^-\\
&\quad=(x+R_1-D_1-0)^-\\
&\quad=0=\xi_{1;[1,(D_1-R_1-x)^+]}
\end{aligned}
$$

引理得证。

引理 3.7 对于任意的 $x\in X$,都有

$$\mathbb{E}[(x+R_1)^+\wedge D_1]=d-\mathbb{E}[(D_1-(x+R_1)^+)^+]$$

证明 同引理 3.3 开头的证明,容易得到

$$
\begin{aligned}
&D_1\wedge(-x-R_1)^--(-x-R_1)^-\\
&\quad=-(D_1-x-R_1)^-
\end{aligned}
$$

又由 $a^-=(-a)^+$可得

$$
\begin{aligned}
D_1\wedge(x+R_1)^+=&-(-D_1+x+R_1)^+\\
&+(x+R_1)^+
\end{aligned}
$$

分两种情况进行讨论。

Ⅰ. 当$-D_1+x+R_1\geqslant0$ 时

$$
\begin{aligned}
&\mathbb{E}[((x+R_1)^+\wedge D_1)]\\
&\quad=\mathbb{E}[-(-D_1+x+R_1)^++(x+R_1)^+]\\
&\quad=\mathbb{E}(D_1)=d-\mathbb{E}[(D_1-x-R_1)^+]\\
&\quad=d-\mathbb{E}[(D_1-(x+R_1)^+)^+]
\end{aligned}
$$

Ⅱ. 当$-D+x+R_1<0$ 时,进一步分两种子情况讨论

若 $x+R_1\leqslant0$

$$
\begin{aligned}
\mathbb{E}[((x+R_1)^+\wedge D_1)]&=\mathbb{E}(0\wedge D_1)=0\\
&=d-\mathbb{E}(D_1)
\end{aligned}
$$

$$=d-\mathbb{E}[(D_1-0)^+]$$
$$=d-\mathbb{E}[(D_1-(x+R_1)^+)^+]$$

若 $x+R_1>0$

$$\mathbb{E}[((x+R_1)^+\wedge D_1)]=\mathbb{E}(x+R_1)$$
$$=d-d+\mathbb{E}(x+R_1)$$
$$=d-\mathbb{E}(D_1-x-R_1)$$
$$=d-\mathbb{E}[(D_1-(x+R_1)^+)^+]$$

综合Ⅰ,Ⅱ两种情况,原命题得证。

引理 3.8 式(3-13)关于 x 具有离散凹性,且最大值点为

$$x_0=arg\max_{x\in\mathbb{N}}\mathcal{L}^o(x)$$
$$=\min\left\{x\in\mathbb{Z}:\mathbb{P}(D_1-R_1\leqslant x)\geqslant\frac{(1-\alpha p)u}{(1-\alpha p)u+h}\right\}$$

证明 由引理 3.6,式(3-13)可整理成如下形式

$$\mathcal{L}^o(x)=-u\,\mathbb{E}[(x+R_1)^-]+\mathbb{E}[u\cdot((x+R_1)^+\wedge D_1)-u\cdot R_1-h\cdot(x+R_1-D_1)^++\alpha u\cdot\xi_{1;[1,(D_1-R_1-x)^+]}]$$

由二项分布的期望公式可知

$$\mathbb{E}[\xi_{1;[1,(D_1-R_1-x)^+]}|D_1,R_1]=p\cdot(D_1-R_1-x)^+$$

进而

$$\mathcal{L}^o(x)=-u\,\mathbb{E}[(x+R_1)^-]+\mathbb{E}[u\cdot((x+R_1)^+\wedge D_1)-u\cdot R_1-h(x+R_1-D_1)^++\alpha pu\cdot(D_1-R_1-x)^+]$$

由引理 3.7,式(3-13)可进一步整理得

$$\mathcal{L}^o(x)=-u\,\mathbb{E}[(x+R_1)^-]+ud-ur+\mathbb{E}[-u(D_1-(x+R_1)^+)^+-h(x+R_1-D_1)^++\alpha pu\cdot(D_1-R_1-x)^+]\tag{3-14}$$

定义一阶差分

$$\Delta\mathcal{L}^o(x)=\mathcal{L}^o(x+1)-\mathcal{L}^o(x),x\in\mathbb{Z}$$

令 $Z=D_1-R_1$,容易证明,对于任意整数 t,有

$$(t+1)^--t^-=-\mathbf{1}_{\{t\leqslant-1\}}\tag{3-15}$$
$$(t-1)^+-t^+=-\mathbf{1}_{\{t\geqslant1\}}\tag{3-16}$$
$$(t+1)^+-t^+=\mathbf{1}_{\{t\geqslant0\}}\tag{3-17}$$

其中$\mathbf{1}_{\{.\}}$为指示函数。

常数项 $ud-ur$ 差分为 0,下面分别计算式(2-14)中四个含 x 部分的差分。

(1)$-u\,\mathbb{E}[(x+R_1)^-]$的差分

由(3-15)可得

$$(x+1+R_1)^--(x+R_1)^-=-\mathbf{1}_{\{x+R_1\leqslant-1\}}=-\mathbf{1}_{\{R_1\leqslant-x-1\}}$$

因此

$$-u\,\mathbb{E}([(x+1+R_1)^--(x+R_1)^-])=-u\,\mathbb{E}(-\mathbf{1}_{\{R_1\leqslant-x-1\}})=u\,\mathbb{P}(R_1\leqslant-x-1)\tag{3-18}$$

(2)$\mathbb{E}[u(D_1-(x+R_1)^+)^+]$的差分

令 $\Phi(x)=(D_1-(x+R_1)^+)^+$

当 $R_1\leqslant-x-1$ 时,$x+R_1\leqslant-1$ 且 $x+1+R_1\leqslant0$,故$(x+R_1)^+=(x+1+R_1)^+=0$,于是 $\Phi(x+1)-\Phi(x)=0$。

当 $R_1\geqslant-x$ 时,$(x+R_1)^+=x+R_1$,$(x+1+R_1)^+=x+1+R_1$。

令 $t=D_1-(x+R_1)$,则由(3-16)可得

$$\Phi(x+1)-\Phi(x)=(t-1)^+-t^+=-\mathbf{1}_{\{t\geqslant1\}}=-\mathbf{1}_{\{D_1\geqslant x+1+R_1\}}=-\mathbf{1}_{\{R_1\geqslant-x,Z\geqslant x+1\}}$$

因此

$$-u(\mathbb{E}[\Phi(x+1)]-\mathbb{E}[\Phi(x)])=u\,\mathbb{P}(R_1\geqslant-x,Z\geqslant x+1)\tag{3-19}$$

(3)$\mathbb{E}[h(x+R_1-D_1)^+]$的差分

由(3-17)可得$(x+1-Z)^+-(x-Z)^+=\mathbf{1}_{\{x-Z\geqslant0\}}=\mathbf{1}_{\{Z\geqslant x\}}$

因此

$$-h(\mathbb{E}[(x+1-Z)^+-(x-Z)^+])=-h\,\mathbb{P}(Z\leqslant x)\tag{3-20}$$

(4)$\mathbb{E}[\alpha pu(Z-x)^+]$的差分

由(3-16)可得$(Z-(x+1))^+-(Z-x)^+=-1_{\{Z\geqslant x+1\}}$

因此

$$\alpha pu\,\mathbb{E}[(Z-(x+1))^+-(Z-x)^+]=-\alpha pu\,\mathbb{P}(Z\geqslant x+1)\tag{3-21}$$

由(3-18)-(3-21)相加得到对于任意 $x\in\mathbb{Z}$,有

$$\Delta\mathcal{L}^o(x)=u\,\mathbb{P}(R_1\leqslant-x-1)+u\,\mathbb{P}(R_1\geqslant-x,Z\geqslant x+1)-h\,\mathbb{P}(Z\leqslant x)-$$

$$\alpha pu\ \mathbb{P}(Z\geqslant x+1) \tag{3-22}$$

注意到

$$\begin{aligned}\mathbb{P}(Z\geqslant x+1)&=\mathbb{P}(D_1\geqslant x+R_1+1)\\&=\sum_{r=0}^{\mathbb{P}}\mathbb{P}(R_1=r)\,\mathbb{P}(D_1\\&\geqslant x+r+1)\end{aligned}$$

当 $r\leqslant -x-1$ 时,$x+r+1\leqslant 0$,由 $D_1\geqslant 0$ 可得 $\mathbb{P}(D_1\geqslant x+r+1)=1$,于是 $\sum_{r=0}^{-x-1}\mathbb{P}(R_1=r)\,\mathbb{P}(D_1\geqslant x+r+1)=\mathbb{P}(R_1\leqslant -x-1)$。

当 $r\geqslant -x$ 时,$\sum_{-x}^{\infty}\mathbb{P}(R_1=r)\,\mathbb{P}(D_1\geqslant x+r+1)=\mathbb{P}(R_1\geqslant -x,D_1\geqslant x+r+1)$。因此

$$\begin{aligned}\mathbb{P}(Z\geqslant x+1)&=\sum_{r=0}^{\infty}\mathbb{P}(R_1=r)\,\mathbb{P}(D_1\geqslant x+r+1)\\&=\sum_{r=0}^{-x-1}\mathbb{P}(R_1=r)\,\mathbb{P}(D_1\\&\geqslant x+r+1)+\sum_{-x}^{\infty}\mathbb{P}(R_1=r)\,\mathbb{P}(D_1\\&\geqslant x+r+1)\\&=\mathbb{P}(R_1\leqslant -x-1)+\mathbb{P}(R_1\geqslant -x,\\&\quad D_1\geqslant x+R_1+1)\end{aligned} \tag{3-23}$$

将(3-23)代入到(3-22)中可得

$$\begin{aligned}\Delta\mathcal{L}^o(x)&=u\ \mathbb{P}(Z\geqslant x+1)-h\ \mathbb{P}(Z\leqslant x)-\\&\quad\alpha pu\ \mathbb{P}(Z\geqslant x+1)=u(1-\alpha p)\\&\quad\mathbb{P}(Z\geqslant x+1)-h\ \mathbb{P}(Z\leqslant x)\end{aligned}$$

令 $F_Z(x)=\mathbb{P}(Z\leqslant x)$,则$\mathbb{P}(Z\geqslant x+1)=1-F_Z(x)$

从而

$$\Delta\mathcal{L}^o(x)=(1-\alpha p)u-[(1-\alpha p)u+h]F_z(x) \tag{3-24}$$

定义二阶差分

$$\Delta^2\mathcal{L}^o(x)=\Delta\mathcal{L}^o(x+1)-\Delta\mathcal{L}^o(x)=-[(1-\alpha p)u+h][F_z(x+1)-F_z(x)]$$

注意到

$$F_z(x+1)-F_z(x)=\mathbb{P}(Z=x+1)\geqslant 0$$

且$(1-\alpha p)u+h\geqslant 0$,

所以$\Delta^2\mathcal{L}^o(x)\leqslant 0$

因此,$\mathcal{L}^o(x)$具有离散凹性。进一步地,$\mathcal{L}^o(x)$的最大值点 x_0 满足

$$x_0=\min\{x\in\mathbb{Z}:\Delta\mathcal{L}^o(x)\leqslant 0\}$$

将(3-24)代入 $\Delta\mathcal{L}^o(x)\leqslant 0$ 得

$$F_z(x)\geqslant\frac{(1-\alpha p)u}{(1-\alpha p)u+h}$$

故$\mathcal{L}^o(x)$的最大值点为

$$x_0=min\left\{x\in\mathbb{Z}:\mathbb{P}(D_1-R_1\leqslant x)\geqslant\frac{(1-\alpha p)u}{(1-\alpha p)u+h}\right\}$$

注 1 在连续情形下,我们同样能够得到完全一致的研究结论。

引理 3.9 $\mathcal{C}^{\alpha,o}(u,h;x)=\sup_{y\geqslant x}\{\mathcal{L}^o(y)+\alpha\ \mathbb{E}[\mathcal{C}^{\alpha,o}(u,h;(y+R_1-D_1)^+-\xi_{1;[1,(D_1-y-R_1)^+]})]\}$与$\mathcal{L}^o(x)$的全局最大值点相同,都是

$$s^{*,\alpha}=arg\max_{x\in\mathbb{N}}\mathcal{L}^o(x)$$

证明 已知

$$\begin{aligned}\mathcal{L}^o(x)=&-u\ \mathbb{E}[(x+R_1)^-]+ud-ur+\\&\mathbb{E}[-u(D_1-(x+R_1)^+)^+-\\&h(x+R_1-D_1)^++\alpha pu\cdot\\&(D_1-R_1-x)^+]\end{aligned}$$

考虑由下式定义的函数序列,对于 $k\geqslant 1$,有

$$\begin{aligned}\mathcal{G}^{\alpha,(k-1)}(x)=&\mathcal{L}^o(x)+\alpha\ \mathbb{E}[\mathcal{C}^{\alpha,(k-1)}(u,h;\\&(x+R_1-D_1)^+-\\&\xi_{1;[1,(D_1-x-R_1)^+]})],\\&\text{且}\mathcal{C}^{\alpha,(0)}(u,h;x)\equiv 0\end{aligned}$$

$$\mathcal{C}^{\alpha,(k)}(u,h;x)=\max_{y\geqslant x}\{\mathcal{G}^{\alpha,(k-1)}(y)\}$$

我们用数学归纳法证明$\mathcal{G}^{\alpha,(k)}(\cdot)$是拟凹的,并且具有全局最大值点 $s^{*,\alpha}=arg\ \max_{x\in\mathbb{N}}\mathcal{L}^o(x)$。

当 $k=1$ 时,

$$\begin{aligned}\mathcal{G}^{\alpha,(0)}(x)=&\mathcal{L}^o(x)+\alpha\ \mathbb{E}[\mathcal{C}^{\alpha,(0)}(u,h;\\&(x+R_1-D_1)^+-\\&\xi_{1;[1,(D_1-x-R_1)^+]})]=\mathcal{L}^o(x)\end{aligned}$$

所以$\mathcal{G}^{\alpha,(0)}(x)$是拟凹的,且具有全局最大值点 $s^{*,\alpha}=arg\ \max_{x\in\mathbb{E}}\mathcal{L}^o(x)$。

假设$\mathcal{G}^{\alpha,(k-1)}(x)$拟凹,并且具有全局最大值点 $s^{*,\alpha}$ 成立,我们期望证明$\mathcal{G}^{\alpha,(k)}(x)$也是拟凹的,并且具有相同的全局最大值点 $s^{*,\alpha}$。

对于$\mathcal{C}^{\alpha,(k)}(u,h;x)=\max_{y\geqslant x}\{\mathcal{G}^{\alpha,(k-1)}(y)\}$,当 $x\leqslant s^{*,\alpha}$ 时,$\max_{y\geqslant x}\{\mathcal{G}^{\alpha,(k-1)}(y)\}=\mathcal{G}^{\alpha,(k-1)}(s^{*,\alpha})$;当 $x\geqslant s^{*,\alpha}$ 时,由拟凹函数的性质可知$\mathcal{G}^{\alpha,(k-1)}(x)$单调不增,所以 $\max_{y\geqslant x}\{\mathcal{G}^{\alpha,(k-1)}(y)\}=\mathcal{G}^{\alpha,(k-1)}(x)$。于是我们有

$$\mathcal{C}^{\alpha,(k)}(u,h;x)=\max_{y\geqslant x}\{\mathcal{G}^{\alpha,(k-1)}(y)\}$$
$$=\begin{cases}\mathcal{G}^{\alpha,(k-1)}(s^{*,\alpha}) & \text{当 } x\leqslant s^{*,\alpha} \text{ 时}\\ \mathcal{G}^{\alpha,(k-1)}(x) & \text{当 } x\geqslant s^{*,\alpha} \text{ 时}\end{cases}$$

可以发现，当 $x\leqslant s^{*,\alpha}$ 时，$\mathcal{C}^{\alpha,(k)}(u,h;x)$ 是一个常数；当 $x\geqslant s^{*,\alpha}$ 时，$\mathcal{C}^{\alpha,(k)}(u,h;x)$ 是拟凹函数并且单调不增。于是，我们证明了，当 $\mathcal{G}^{\alpha,(k-1)}(x)$ 是拟凹函数且具有全局最大值点 $s^{*,\alpha}$ 时，$\mathcal{C}^{\alpha,(k)}(u,h;x)$ 也是拟凹函数，并且具有相同的全局最大值点 $s^{*,\alpha}$。

考虑 $\mathcal{C}^{\alpha,(k)}(u,h;(x+R_1-D_1)^+-\xi_{1;[1,(D_1-x-R_1)^+]})$。易证 $(x+R_1-D_1)^+-\xi_{1;[1,(D_1-x-R_1)^+]}$ 关于 x 单调不减。由引理 3.5 可得，$\mathcal{C}^{\alpha,(k)}(u,h;(x+R_1-D_1)^+-\xi_{1;[1,(D_1-x-R_1)^+]})$ 也是拟凹函数。

于是，我们证明了

$$\mathcal{G}^{\alpha,(k)}(x)=\mathcal{L}^{o}(x)+\alpha\,\mathbb{E}[\mathcal{C}^{\alpha,(k)}(u,h;(x+R_1-D_1)^+-\xi_{1;[1,(D_1-x-R_1)^+]})]$$

是拟凹函数，并且有全局最大值点 $s^{*,\alpha}$。

综上，我们用数学归纳法证明了 $\mathcal{G}^{\alpha,(k)}(x)$ 是拟凹的，并且具有全局最大值点 $s^{*,\alpha}$。进一步地，$\mathcal{C}^{\alpha,(k)}(x)$ 也是拟凹函数，并且具有相同的全局最大值点 $s^{*,\alpha}$。

根据 Bellman 提出的最优性原理[17]（参考 3.3 节），可得

$$\begin{aligned}\mathcal{C}^{\alpha,o}(u,h;x)&=\sup_{y\geqslant x}\{\mathcal{L}^{o}(y)+\alpha\,\mathbb{E}[\mathcal{C}^{\alpha,o}(u,h;(y+R_1-D_1)^+-\xi_{1;[1,(D_1-y-R_1)^+]})]\}\\&=\lim_{k\to\infty}\mathcal{C}^{\alpha,(k)}(\cdot)\end{aligned}$$

这意味着 $\mathcal{C}^{\alpha,o}(u,h;x)$ 是拟凹的，并且具有全局最大值点 $s^{*,\alpha}$。命题得证。

进一步地，直接应用刘克、曹平[18]中的定理 4.1 可得

引理 3.10 对于折扣准则，存在平稳策略 $\pi^*\in\mathcal{A}$ 和折扣因子 $0\leqslant\alpha_0<1$，使得对任意的 $\alpha\in(\alpha_0,1)$，价值函数都满足

$$V_\alpha(\pi^*)=\sup_{\pi\in\Pi}V_\alpha(\pi)$$

进一步地，这个平稳策略 π^* 是平均准则在 A 中的最优策略。

所以，存在 $s^*\in S$ 和折扣因子 $0\leqslant\alpha_0<1$，使得当任意的 $\alpha\in(\alpha_0,1)$ 时，都满足

$$\mathcal{C}^{\alpha,o}(u,h;s^*)=\sup_{s\in S}\mathcal{C}^{\alpha,o}(u,h;s)$$

并且具有水平 s^* 的基本库存策略就是平均准则在 S 中的最优策略，即当 $\alpha=1$ 时，无折扣时的长期平均利润方程 $\mathcal{C}^{\pi}(u,h;x)$ 的最大值点为

$$s^*=\min\left\{x\in\mathbb{Z}:\mathbb{P}(D_1-R_1\leqslant x)\geqslant\frac{(1-p)u}{(1-p)u+h}\right\}\tag{3-25}$$

定理 3.1（基本库存策略的最优性） 对于零交货期的 P 系统，具有水平 s^*（3-25）的基本库存策略是最优的。即，每周期的最优订货量为 $Q_i=(s^*-N_i)^+$，s^* 的值由（3-25）给出。

注 2 观察上式发现，当 $D_1\geqslant R_1$ 时，退货满足部分需求后，应对剩余需求的最优库存水平（D_1-R_1）正好与 Andrew 等人[12]给出的无退货情形下的最优库存水平 $s^*=\min\left\{x\in\mathbb{Z}:\mathbb{P}(D_1\leqslant x)\geqslant\frac{(1-p)u}{(1-p)u+h}\right\}$ 形式上保持一致，此时，最优库存水平的取值始终为非负值；当 $D_1<R_1$ 时，不同于传统认知，最优库存水平 s^* 可能为负值。该结果直观地表明，当退货能够在概率意义上大量补充库存时，决策者更倾向于暂不订货或更少订货。

4 数值实验

在本节中，我们通过数值实验分析退货情形下部分延期交付库存系统的订货策略，以评估前文所提出的基本库存策略的实际性能。在所有数值实验中，我们统一设定单位库存持有成本 $h=1$，并将单位销售价格 u 从较低水平 $u=4$ 逐步提高至较高水平 $u=64$，以考察销售价格变化对最优库存决策及系统利润的影响。

需求随机变量 D_i 与退货随机变量 R_i 的分布组合如表中所示，分别考虑泊松分布与二项分布下的多种参数设定。针对未满足需求的耐心程度，本文重点比较两种典型情形：高耐心情形（$p=0.7$）和低耐心情形（$p=0.3$）。

我们进一步比较考虑退货情形与忽略退货情形下所得库存决策及系统利润的差异，以强调将退货因素显式纳入模型的重要性。表 3 和表 4 分别给出了考虑或忽略退货因素时，最优库存水平及长期平均利润的比较。其中，$(s^*,\mathcal{C}^{s^*}(u,h;x))$ 分别表示由解析解计算得到的最优库存水平及对应的长期

平均利润,$(s', C^{s'}(u,h;x))$表示在忽略退货因素的前提下,依据不含退货因子的解析解所得到的库存水平及其对应的长期平均利润。通过对比这两组结果,可以直观刻画忽略退货行为所带来的决策偏差及其对系统绩效的影响。

表 1 基础库存策略的最优库存水平与每期平均利润(h=1,p=0.7)

Table 1 Optimal inventory level and average profit per period under the basic inventory strategy (h = 1,p = 0.7)

		$(s^*,\bar{s}^*)$					$C^{\bar{s}^*}(u,h;x)$				
		$u=4$	$u=8$	$u=16$	$u=32$	$u=64$	$u=4$	$u=8$	$u=16$	$u=32$	$u=64$
D_i	$P(10)$	(7,8)	(9,8)	(10,10)	(12,11)	(13,14)	24.86	51.72	106.57	217.16	441.21
R_i	$P(3)$										
D_i	$B(10,0.5)$	(3,3)	(4,4)	(5,5)	(6,5)	(6,6)	10.23	21.68	45.15	92.51	188.11
R_i	$B(10,0.2)$										
D_i	$P(10)$	(3,3)	(5,4)	(7,6)	(8,8)	(10,10)	7.99	18.75	41.57	88.31	182.95
R_i	$P(7)$										
D_i	$B(10,0.5)$	(1,0)	(2,1)	(3,2)	(4,3)	(5,4)	1.41	4.66	11.98	27.36	58.81
R_i	$B(10,0.4)$										
D_i	$P(10)$	(−3,−4)	(0,−1)	(2,3)	(4,4)	(5,6)	−3.59	−1.40	4.87	19.63	50.03
R_i	$P(9)$										
D_i	$B(10,0.5)$	(−1,−2)	(−1,0)	(1,1)	(2,2)	(3,3)	−2.19	−0.95	2.27	9.51	24.38
R_i	$B(10,0.45)$										

表 2 基础库存策略的最优库存水平与每期平均利润(h=1,p=0.3)

Table 2 Optimal inventory level and average profit per period under the basic inventory strategy (h = 1,p = 0.3)

		$(s^*,\bar{s}^*)$					$C^{\bar{s}^*}(u,h;x)$				
		$u=4$	$u=8$	$u=16$	$u=32$	$u=64$	$u=4$	$u=8$	$u=16$	$u=32$	$u=64$
D_i	$P(10)$	(9,9)	(11,11)	(12,12)	(13,14)	(15,14)	23.43	50.25	105.02	216.48	440.02
R_i	$P(3)$										
D_i	$B(10,0.5)$	(4,4)	(5,5)	(6,6)	(6,6)	(7,7)	9.52	20.99	44.40	91.92	187.39
R_i	$B(10,0.2)$										
D_i	$P(10)$	(5,5)	(7,6)	(9,8)	(10,10)	(11,10)	6.39	17.10	39.89	86.59	181.29
R_i	$P(7)$										
D_i	$B(10,0.5)$	(2,2)	(3,3)	(4,3)	(5,4)	(5,5)	0.52	3.80	11.09	26.58	57.94
R_i	$B(10,0.4)$										
D_i	$P(10)$	(1,1)	(2,3)	(4,4)	(6,6)	(8,8)	−5.87	−3.54	3.06	17.57	48.94
R_i	$P(9)$										
D_i	$B(10,0.5)$	(−1,0)	(1,1)	(2,2)	(3,3)	(4,4)	−3.15	−1.96	1.25	8.37	23.43
R_i	$B(10,0.45)$										

表 1 和表 2 分别给出了在不同顾客耐心水平下,最优基本库存水平及其对应的长期平均利润。其中,s^* 表示由前文解析表达式 $s^* = \min\left\{x \in \mathbb{Z}: \mathbb{P}(D_1 - R_1 \leqslant x) \geqslant \frac{(1-p)u}{(1-p)u+h}\right\}$ 计算得到的最优库存水平;$\bar{s}^*$ 表示通过数值模拟搜索得到、使得系统平均利润最大的库存水平;$C^{s^*}(u,h;x)$则表示在策略 s^* 下系统每周期的平均利润。从整体上看,解析解 s^* 与数值模拟得到的最优库存水平 $\bar{s}^*$ 在绝大多数情形下高度一致。这表明本文给出的最优库存水平解析解具有良好的数值稳定性。

观察表中数据可知，当退货分布 R_i 的期望水平提高时，最优库存水平 s^* 整体呈现下降趋势。这是由于退货在系统中起到了随机库存补充的作用，使得管理者在决策时无需维持过高的初始库存水平。这一现象在不同需求分布下均表现一致，说明考虑退货因素对库存决策具有显著的缓冲效应。其次，对比表 1 与表 2 可以发现，当未满足需求的耐心程度降低时，最优库存水平 s^* 明显上升。这是因为在低耐心情形下，未满足的需求更可能直接流失，从而显著提高了缺货的机会成本。相应地，决策者倾向于通过提高库存水平来降低需求流失的风险。相反，在高耐心情形下，部分需求可以延期至下一周期完成交付，结合本文中订货交货期为 0 这一设定，决策者维持相对较低的库存水平是直观的。此外，随着销售价格 u 的提高，最优库存水平 s^* 整体呈现上升趋势，同时系统的长期平均利润近似呈线性增长。这反映了当单位销售收益占主导地位时，增加库存在经济效益上更具优势。值得注意的是，最优库存水平的取值可能为负。这表明：当退货能够在概率意义上大量补充库存时，决策者更倾向于暂不订货或更少订货。

表 3　考虑/忽略退货因素下库存水平及长期平均利润的比较(h=1,p=0.7)

Table 3　Comparison of inventory levels and long-run average profit with and without considering returns(h=1,p=0.7)

		$(s^*, C^{s^*}(u,h;x))$					$(s', C^{s'}(u,h;x))$				
		$u=4$	$u=8$	$u=16$	$u=32$	$u=64$	$u=4$	$u=8$	$u=16$	$u=32$	$u=64$
D_i	$P(10)$	(7,24.86)	(9,51.72)	(10,106.57)	(12,217.16)	(13,441.21)	(10,24.11)	(12,50.38)	(13,105.49)	(14,216.97)	(15,439.02)
R_i	$P(3)$										
D_i	$B(10,0.5)$	(3,10.23)	(4,21.68)	(5,45.15)	(6,92.51)	(6,188.11)	(5,9.61)	(6,20.81)	(6,44.74)	(7,92.06)	(8,186.78)
R_i	$B(10,0.2)$										
D_i	$P(10)$	(3,7.99)	(5,18.75)	(7,41.57)	(8,88.31)	(10,182.95)	(10,3.82)	(12,13.75)	(13,36.89)	(14,83.66)	(15,179.34)
R_i	$P(7)$										
D_i	$B(10,0.5)$	(1,1.41)	(2,4.66)	(3,11.98)	(4,27.36)	(5,58.81)	(5,−1.42)	(6,1.68)	(6,9.73)	(7,24.57)	(8,55.61)
R_i	$B(10,0.4)$										

表 4　考虑/忽略退货因素下库存水平及长期平均利润的比较(h=1,p=0.3)

Table 4　Comparison of inventory levels and long-run average profit with and without considering returns(h=1,p=0.3)

		$(s^*, C^{s^*}(u,h;x))$					$(s', C^{s'}(u,h;x))$				
		$u=4$	$u=8$	$u=16$	$u=32$	$u=64$	$u=4$	$u=8$	$u=16$	$u=32$	$u=64$
D_i	$P(10)$	(9,23.43)	(11,50.25)	(12,105.02)	(13,216.48)	(15,440.02)	(12,22.36)	(13,49.42)	(15,103.57)	(16,214.75)	(17,437.73)
R_i	$P(3)$										
D_i	$B(10,0.5)$	(4,9.52)	(5,20.99)	(6,44.40)	(6,91.92)	(7,187.39)	(6,8.74)	(7,19.84)	(7,43.77)	(8,91.00)	(8,187.08)
R_i	$B(10,0.2)$										
D_i	$P(10)$	(6,6.39)	(7,17.10)	(9,39.89)	(10,86.59)	(11,181.29)	(12,1.79)	(13,12.88)	(15,35.03)	(16,81.89)	(17,176.82)
R_i	$P(7)$										
D_i	$B(10,0.5)$	(2,0.52)	(3,3.80)	(4,11.09)	(5,26.58)	(5,57.94)	(6,−2.35)	(7,0.60)	(7,8.59)	(8,23.61)	(8,55.61)
R_i	$B(10,0.4)$										

一个最为显著的现象是：当退货水平较高时，忽略退货因素所得到的库存水平 s' 明显偏大，而由此产生的利润损失也更为显著。具体而言，在当退货分布由 $P(3)$ 增至 $P(7)$，或由 $B(10,0.2)$ 增至 $B(10,0.4)$ 的情形下，可以观察到 s' 相对于 s^* 的偏差明显扩大。相应地，过高的库存水平会带来额外的持有成本，使得长期平均利润 $C^{s'}(u,h;x)$ 相比于 $C^{s^*}(u,h;x)$ 出现更明显的下降。此外，对比表 3 与表 4 还可以发现，当顾客耐心程度降低时，忽略退货所造成的负面影响进一步加剧。在低耐心情形下，

未满足需求更可能直接流失,此时若仍基于不含退货的模型进行决策,不仅库存水平偏差更大,而且利润损失更加显著。这表明,在顾客放弃行为较为普遍的环境中,同时考虑退货与顾客耐心对库存决策尤为关键。最后,随着销售价格 u 的提高,考虑退货与忽略退货两种情形下的利润差距整体呈现缩小趋势。这是因为当销售收益在利润结构中占主导地位时,库存决策偏差所引致的相对损失被较高的销售收益部分抵消。然而,即便在高价格区间,忽略退货所带来的库存偏高现象仍然存在,说明退货因素在库存系统中的影响并不会随着价格上升而完全消失。

综上所述,数值结果表明:在存在退货与顾客放弃行为的库存系统中,若在决策时忽略退货因素,不仅会导致库存水平系统性偏高,而且可能造成显著的长期利润损失。这一现象在退货水平较高或顾客耐心较低的情形下尤为突出。

利益冲突:作者声明没有利益冲突。

参考文献(References)

[1] 国家统计局. 2025 年 1—11 月份社会消费品零售总额数据[EB/OL]. 北京:国家统计局,2025-12-15[2025-12-16].

[2] 国家市场监督管理总局. 网络购买商品七日无理由退货暂行办法[EB/OL]. 北京:国家市场监督管理总局,2017-01-06[2025-12-16].

[3] 网经社电子商务研究中心. 2024 年 Q4 中国电商平台商家投诉数据报告[EB/OL]. 杭州:网经社,2025-01[2025-12-16].

[4] Harris F. How many parts to make at once[J]. Factory,1913,10:135-136,152.
https://doi.org/10.1287/opre.38.6.947

[5] Arrow K J,Harris T E,Marschak J. Optimal inventory policy[J]. Econometrica,1951,19:250-272.
https://doi.org/10.2307/1906813

[6] Dvoretzky A,Kiefer J,Wolfowitz J. The inventory problem[J]. Econometrica,1952,20:187-222.
https://doi.org/10.2307/1907847

[7] Karlin S. Inventory models of the Arrow-Harris-Marschak type with time lag[M]. In:Studies in the Mathematical Theory of Inventory and Production. Stanford:Stanford University Press,1958.

[8] Zipkin P H. Foundations of inventory management[M]. New York:McGraw-Hill/Irwin,2000.

[9] Janakiraman G,Roundy R O. Lost-sales problems with stochastic lead times:Convexity results for base-stock policies[J]. Operations Research,2004,52(5):795-803.
https://doi.org/10.1287/opre.1040.0130

[10] Janakiraman G,Seshadri S,Shanthikumar J G. A comparison of the optimal costs of two canonical inventory systems[J]. Operations Research,2007,55(5):866-875.
https://doi.org/10.1287/opre.1070.0463

[11] Corsten D,Gruen T W. Stock-outs cause walkouts[J]. Harvard Business Review,2004,82(5):26-28.

[12] Andrew E B L,Wei Z,Zhang H. Partial Backorder Inventory System:Asymptotic Optimality and Demand Learning[R/OL]. 2024-04-15[2026-03-11].
http://dx.doi.org/10.2139/ssrn.4772763

[13] Islam M T,Huda N. Reverse logistics and closed-loop supply chain of waste electrical and electronic equipment:A comprehensive literature review[J]. Resources,Conservation and Recycling,2018,137:48-75.
https://doi.org/10.1016/j.resconrec.2018.05.026

[14] Gökbayrak E,Kayis E. Single item periodic review inventory control with sales dependent stochastic return flows[J]. International Journal of Production Economics,2023,255:108699.
https://doi.org/10.1016/j.ijpe.2022.108699

[15] Braz A C,De Mello A M,Gomes L A D,et al. The bullwhip effect in closed-loop supply chains:A systematic literature review[J]. Journal of Cleaner Production,2018,202:376-389.
https://doi.org/10.1016/j.jclepro.2018.08.042

[16] Narang U,Shankar V. Mobile app introduction and online and offline purchases and product returns[J]. Marketing Science,2019,38(5):756-772.
https://doi.org/10.1287/mksc.2019.1169

[17] Bellman R. Dynamic programming[M]. Princeton:Princeton University Press,1957.

[18] 刘克,曹平. 马尔可夫决策过程理论与应用[M]. 北京:科学出版社,2015.

Optimal Ordering Policy for Inventory Systems with Partial Delayed Delivery under Return Scenarios

SU Xin[1,2,*], YANG Jiankui[1,2]

(1. School of Mathematical Sciences, Beijing University of Posts and Telecommunications, Beijing 100876, China;
2. Key Laboratory of Mathematics and Information Networks Ministry of Education
(Beijing University of Posts and Telecommunications), Beijing 100876, China)

Abstract: In stochastic inventory management studies, managers can fulfill the demand of customers willing to wait through replenishment and delayed fulfillment when stockouts occur. However, the impact of product returns on ordering policies has received limited attention. In fact, a large volume of product returns can significantly affect inventory levels. This paper investigates an inventory system with product returns, comprehensively considering customer return behavior and the randomness of customer patience during stockouts. It is proved that the base-stock policy is optimal under the condition of zero lead time, and the analytical solution is derived. Interestingly, the optimal inventory level may be negative.

Keywords: Inventory management; returns; partial backlogging; optimal policy

DOI: 10.48014/bcam.20260212001

Citation: SU Xin, YANG Jiankui. Optimal ordering policy for inventory systems with partial delayed delivery under return scenarios[J]. Bulletin of Chinese Applied Mathematics, 2026, 4(2): 29-42.

Copyright © 2026 by author(s) and Science Footprint Press Co., Limited. This article is open accessed under the CC-BY License (http://creativecommons.org/licenses/by/4.0/).