# Ordering Policy for Perishable Commodity Using Online Algorithm with Risk-Reward

Bin Liu[1]   Wentian Cui[1] Chunlin Xin [2]

[1]School of Management, Xi'an Jiaotong University, Xi'an 710049, P. R. China
[2] School of Economics and Management, Tsinghua University, Beijing 100084, P. R. China

## Abstract

This paper studied an ordering problem under non-stationary setting. It used the online algorithm with risk-reward that is a combination of probability analysis and classical online algorithm. We obtained an online risk base stock policy whose risk was restricted to a tolerance level by designing online strategy based on forecasting. The strategy gain reward if forecasting successful, the risk is acceptable if unsuccessful. The decision-makers can design online strategy to improve performance in the light of his risk tolerance level.

**Keywords**: Perishable commodity, Online algorithm with risk-reward, Competitive ratio

## 1.  Introduction

The perishable commodity problem, also known as newsboy problem, is to find the optimal order quantity that maximizes the expected profit under uncertainty demand. If any inventory remains at the end of the period, a discount is used to sell it or it is disposed of; If the order quantity is smaller than the realized demand, the newsboy suffers some penalty because of shortage [1]. The model assumes that the demand in each period are independent and identically distributed (i.i.d.) continuous random variables.

Interest in this problem has increased the last decade since the newsboy problem was fist introduced by Within (1955)[2]. A variety of extensions to the newsboy problem have been presented and researched, one of the important extensions is to describe different states of information about demand since the optimal order quantity depends on an exact probability distribution that is different in the different situations, and lack of demand visibility is usually considered as a major challenge for efficient inventory management. Scarf (1958) assumed that only the mean $\mu$ and variance $\sigma^2$ of demand are known and derived in closed-form the optimal quantity which maximizes the expected profit against the worst possible distribution of demand [3]. Shih (1973) applied a Bayesian methodology to the newsboy problem [4]. Moon and Choi (1995) maximizes the expected profit against the worst possible distribution of demand for a distribution free newsboy problem with balking [5]. Baruch and Joseph (2005) researched the ordering policy of the risk-averse newsboy that assumed the random demand is uniformly distributed between lower bound and upper bound [6].

To design optimal ordering policy under uncertainty demand is a typical online problem in inventory management that online managers have to decide instantaneously ordering quantity to meet demand when they don't know the future information about demand. For the online problems, the traditional approach is that managers describe probability distribution as in [2]-[6] in the light of personal experience to look for the optimal policy maximizing the expected profit. However, it is very difficult for managers to describe exactly probability distribution in real situations because the demand is affected by lots of factors, real distributions in computational setting are almost always very different from any theoretical distribution assumed in an average case analysis. In such case, the optimal policy is risky, perhaps bring about bad performance. In fact, the most managers are risk averse and eager to trade some expected value to reduce the downside risk. Recently, online algorithm has been gaining recognition for being a complementary approach in the analysis of decision making under uncertainty. In the approach a player designs online strategy, in which player has no knowledge of future demand, is compared with the performance of offline strategy that has knowledge of entire future and operates optimally. This technique of evaluating an online strategy by comparing its performance to the optimal offline strategy is called competitive analysis and was first used by Sleator and Tarjan(1985)in [7]. This is not the place for a lengthy discussion of the merits of online algorithm compared with those of probability analysis; the interested reader may refer to the introductory textbook [8]. The online algorithm [13]-[16] has been used widely include foreign currency exchange, mortgage refinancing and

portfolio management [9]-[10]since it has been first introduced in list update and paging rules. In classical online algorithm, the player designs online strategy when he doesn't know absolutely future information, so the online strategy is certain and riskless.

In this paper the online algorithm with risk-reward is a combination of probability analysis and classical online algorithm. The approach assumed that online player couldn't describe probability distribution of future demand, but he could forecast the lower bound and upper bound of future demand in the light of personal experience while he beat acceptable risk level. The players design risky online strategy in the base of forecast, the performance of online strategy became better if the forecast was correct. The range of demand forecasted is different if risk preference level is different. The higher the risk preference is, the smaller the range of demand forecasted is, the better the performance of online strategy is if forecast is correct. The improvement of the performance of risky online strategy over that of the optimal riskless online strategy is looked up as the reward that player bear risk.

## 2. Online model with risk-reward

## 2.1. Problem description

In the uncertain demand setting, the managers employ base stock policy to meet future demand, the inventory system is reviewed once in a period, the replenishment is instantaneous (0 lead time). When the demand exceeds the ordering quantity in a period, excess demand is backlogged until it is subsequently filled by a delivery, but a shortage penalty $h$ is paid by the buyer to his customers for each unit of unfilled demand [11]. The remains must be disposed of, $v$ is a salvage value for each unit of inventory remaining at the end of each period. Let $D_i$ denote the demand in period $i$. Let $S_i$ be the order up to level at the beginning of period $i$, the inventory position after the order has been placed. The order quantity $q_i$, at the beginning of period $i$ should satisfy the following relationship:

$$q_i = S_i + \max\left\{0,(D_{i-1}-S_{i-1})\right\}$$

This is because the remains must be disposed of and excess demand is backlogged.

Three costs are incurred during each period which influences the order up to level $S_i$. There is an ordering cost $c$ per unit, which is paid by buyer to supplier; a loss $g = c - v$ for each unit of inventory remaining at the end of each period; and finally, a

shortage penalty $h$ associated with the failure to meet demand. The loss $g$ and shortage penalty are charged at the end of each period. The optimal ordering policy is to minimize the total cost over the duration of the process because the excess demands are backlogged and it is subsequently filled by a delivery.

We assumed that an online player didn't know probability distribution in real situations, but that he known exactly the lower bound and upper bound of future demand $[m, M]$, let $\phi = M / m$. The player designs an online ordering strategy in the light of information about demand that he knows exactly, and obtains the competitive ratio of online strategy that is riskless. Let $C_A(I)$ be the cost of the online ordering strategy $A$ on input $I$ where $I$ is an instance of the problem $\sum$. Let $C_{OPT}(I)$ be the cost of the optimal offline ordering strategy. Then the competitive ratio of a strategy $A$ on a problem $\sum$ is $r_A = \sup \dfrac{C_A(I)}{C_{OPT}(I)}$.

The less $r_A$ is, the better the performance of $A$ is, the optimal competitive ratio for a problem is $r^* = \inf\limits_A r_A$, which is riskless.

In the classical online algorithm and competitive analysis, the online player overlooks absolutely information about future demand, which has been criticized as being too conservative because of ignoring valuable information. In most cases, the managers don't seek to minimize risk, but to manage it. They would like to accept a risk level in order to potential reward. Al-Binali(1999) define the risk of an online strategy $A$ to be $r_A / r^*$, where $r_A$ is the competitive ratio of $A$ and $r^*$ is the optimal competitive ratio. From the manager's viewpoint, this measure of is the maximum opportunity cost that online strategy $A$ may incur over the optimal online strategy [12]. If $t$ is the risk tolerance of the manager (where $t \geq 1$ and higher value of $t$ denote a higher risk tolerance), then denote by $I_t = \left\{ A \mid r_A \leq tr^* \right\}$ the set of all strategy that respect the managers risk tolerance.

Within risk tolerance $t$, the online player forecast the future demand of period $i$ $D_i \in [f, F]$, let $\phi' = F / f$ ( $[f, F] \in [m, M]$ ). We define $\hat{r}_A$ to be the competitive ratio of $A$ restricted to cases when the forecast is correct: $\hat{r}_A = \sup \dfrac{C_A(I)}{C_{OPT}(I)}$, the optimal restricted ratio is $\hat{r}^* = \inf \hat{r}_A$. The reward of optimal risk strategy $A$ is $R_A = r^* / \hat{r}^*$, which is a

improvement over the optimal riskless strategy. The negative outcome of competitive ratio doesn't exceed the risk tolerance $t$ if the forecast is not correct.

## 2.2. Model

Considering a single product dynamic inventory problem, the demand in period $i$ is $D_i$ ( $i = 1, 2, \cdots n$ ), the manager employs base stock policy. Obviously, the manager always let $S_i = D_i$ if he knows the future demand in advance. We can get the optimal cost for offline problem:

$$C_{OPT}(I) = c \sum_{i=1}^{n} D_i \qquad (1)$$

For an online risk manager, he can forecast the future demand of period $i$ $D_i \in [f, F]$ within risk tolerance $t$, but can not know exact value. The manager designs online ordering strategy, let $S'$ be the order up to level. We have the cost of online strategy within risk tolerance $t$:

$$C_{S'}(I) = c \sum_{i=1}^{n} D_i + h \sum_{\substack{i=1 \\ D_i > S'}}^{n} (D_i - S') + g \sum_{\substack{i=1 \\ D_i < S'}}^{n} (S' - D_i) \qquad (2)$$

## 3. Forecast & competitive analysis

In this section, we first design online risk strategy in the base of forecasting demand, and the competitive analysis show that this strategy is optimal online ordering strategy when forecast is correct. In addition, we gain the upper bound and of range of forecasting demand that doesn't exceed the manager's acceptable risk tolerance level.

**Theorem 1** Within a certain range of future demand $[m, M]$, let $\phi = M / m$, forecasting future demand of period $i$ $D_i \in [f, F]$, let $\phi' = F / f$ ( $[f, F] \in [m, M]$ ) and designing online ordering strategy in the base of forecasting, let order up to level be

$$S' = \frac{(h + g) f F}{h f + g F} \; ; \quad \text{The competitive ratio is}$$

$1 + \dfrac{h g (\phi' - 1)}{c(h + g \phi')}$ if forecast is correct, this strategy is the optimal online strategy within forecast; the reward is

$$[1 + \frac{h g (\phi - 1)}{c(h + g \phi)}] / [1 + \frac{h g (\phi' - 1)}{c(h + g \phi')}] \; .$$

(1) the competitive ratio of strategy $S' = \dfrac{(h + g) f F}{h f + g F}$ is $1 + \dfrac{h g (\phi' - 1)}{c(h + g \phi')}$ , which is the optimal online ordering strategy.

**Proof**. For minimization problem, the offline adversary make the competitive ratio of online strategy designed by the manager as great as possible by definition, so he always select the most value $F$ of the future demand in order to maximizing shortage penalty, or select the least value $f$ to maximize the loss of inventory remaining.

First, let $D_i = F$, we have the cost of online strategy.

$$\begin{aligned} C_{S'}(I) &= \sum_{i=1}^{n} (cF + h(F - S')) \\ &= ncF + nh[F - \frac{(h + g) f F}{h f + g F}] \\ &= [1 + \frac{h g (\phi' - 1)}{c(h + g \phi')}] cnF \\ &= [1 + \frac{h g (\phi' - 1)}{c(h + g \phi')}] C_{OPT}(F) \end{aligned} \qquad (3)$$

Similarly, let $D_i = f$. We also have

$$C_{S'}(I) = [1 + \frac{h g (\phi' - 1)}{c(h + g \phi')}] C_{OPT}(f) \qquad (4)$$

For the offline adversary, the most value $F$ and the least value $f$ of demand are the best approach to make the competitive ratio as great as possible, that is, there are no other online strategies that can improve competitive ratio. The online ordering strategy $S'$ is the optimal and competitive ratio

$$r_{S'} = \sup \frac{C_{S'}(I)}{C_{OPT}(I)} \le 1 + \frac{h g (\phi' - 1)}{c(h + g \phi')} \qquad (5)$$

The proof is completed.

The equation of competitive ratio show that the less $\phi'$ is, the less the competitive ratio is. Namely, the higher the risk preference is, the better the performance of online strategy is if forecast is correct.

(2) The reward is $[1 + \dfrac{h g (\phi - 1)}{c(h + g \phi)}] / [1 + \dfrac{h g (\phi' - 1)}{c(h + g \phi')}]$

**Proof**. Recalling the proof of (1), then the manager design an online ordering strategy in the deterministic demand range $[m, M]$, $\phi = M / m$, let order up to level be

$S^* = \dfrac{(h+g)mM}{hm+gM}$ , the competitive ratio of the strategy is $1 + \dfrac{hg(\phi-1)}{c(h+g\phi)}$ , this strategy is optimal online strategy. From the definition of reward, it can be easily derived that

$$R_{S'} = [1 + \frac{hg(\phi-1)}{c(h+g\phi)}] / [1 + \frac{hg(\phi'-1)}{c(h+g\phi')}]$$

**Theorem 2** For the forecast of $D_i \in [f, F]$, the risk of strategy $S'$ doesn't exceed tolerance level $t$ if the range $\phi'$ and the upper bound $F$ satisfy the following relationship

$$\frac{h(h+g)(h+g\phi)F}{((c+h)(h+g\phi)-t(hg\phi+cg\phi+ch-hg))M} \geq (h+g\phi')$$
$$\geq \frac{g(h+g)(h+g\phi)F}{((hg\phi+cg\phi+ch-hg)t-(c-g)(h+g\phi))m}$$

**Proof** Within the forecast of $D_i \in [f, F]$ , and designing online ordering strategy, let order up to level be

$$S' = \frac{(h+g)fF}{hf+gF} \; .$$

If the forecast is incorrect, offline adversary select the most value $M$ or the least value $m$ of demand second to make the competitive ratio of risk strategy as great as possible.

In the light of the definition $r_{S'} \leq t r_{S^*}$ , the set of online strategy that respect the managers risk tolerance level $t$ . We have the following relationship about the competitive ratio of strategy $S'$ if forecast is not correct.

$$\frac{cm+g(S'-m)}{cm} \leq t(1 + \frac{hg(\phi-1)}{c(h+g\phi)}) \qquad (6)$$

$$\frac{cM+h(M-S')}{cM} \leq t(1 + \frac{hg(\phi-1)}{c(h+g\phi)}) \qquad (7)$$

| $t$ | $r^*$ | F=750 | | |
|---|---|---|---|---|
| | | $\phi'$ | $r_{S'}$ | $r^*/r_{S'}$ |
| 1.01 | 1.27 | [7.1, 9.4] | 1.185 | 1.07 |
| 1.02 | 1.27 | [6.9, 13.2] | 1.179 | 1.08 |
| 1.03 | 1.27 | [6.7, 21.3] | 1.173 | 1.083 |
| 1.04 | 1.27 | [6.6, 50.6] | 1.170 | 1.085 |

Table 1:Forecast, competitive ratio and reward.

Let $S' = \dfrac{(h+g)fF}{hf+gF}$ replace to (6), we have

$$\frac{m(c-g)(h+g\phi')+g(h+g)F}{cm(h+g\phi')} \leq t(1+\frac{hg(\phi-1)}{c(h+g\phi)}) \; , \text{ then,}$$

$$(h+g\phi') \geq \frac{g(h+g)(h+g\phi)F}{((hg\phi+cg\phi+ch-hg)t-(c-g)(h+g\phi))m} \qquad (8)$$

Similarly, Let $S' = \dfrac{(h+g)fF}{hf+gF}$ replace to (7), we also have

$$\frac{M(c+h)(h+g\phi')-h(h+g)F}{cM(h+g\phi')} \leq t(1+\frac{hg(\phi-1)}{c(h+g\phi)}), \text{ then,}$$

$$(h+g\phi') \leq \frac{h(h+g)(h+g\phi)F}{((c+h)(h+g\phi)-t(hg\phi+cg\phi+ch-hg))M} \qquad (9)$$

Therefore, the proof of theorem 2 can be derived from (8) and (9).

The theorem 1 and theorem 2 show that the manager can deal effectually with uncertain problems when he can not know probability distribution of the problem under non-stationary setting by using online algorithm with risk-reward, which make the manager not only utilizing valuable information and experience but also controlling risk within his tolerance level.

# 4. Numerical analysis

In this section, we consider one numerical example in order to assess the merits of using online algorithm with risk-reward under non-stationary demand.

To estimate the benefit of online algorithm with risk-reward, we first assume that the future demand be in the range which is enough great and is riskless, then the manager forecast the range of future demand within his risk tolerance level and design online risk strategy.

Consider a perishable commodity with purchasing cost of \$ 3.0, a salvage value of \$ 2.0, and a penalty cost of \$ 1.0 for per unit. Assume a riskless demand range [100, 1000] in the future. In the context of our model we have

$c = 3, g = 1, h = 1, M = 1000, m = 100, \phi = 10$

The optimal deterministic competitive ratio is $r^* = 1.27$ from theorem 1.

Table 1 describe the forecast of the upper bound $F$ and range $\phi'$ within different risk tolerance level $t$ , the optimal competitive ratio $r_{S'}$ under forecasting is correct, and the reward .

In numerical experiment we exam forecast range and competitive ratio within different risk tolerance level when the manager forecast that the upper bound *F* is 750. The observation that can be made from Table 1 is that the higher the risk preference is, the smaller the range forecasted is, the better the performance of online strategy is if forecast is correct.

# 5. Conclusions

The future demand is influenced by more and more factors and become extremely uncertain because of the development of technology furious competition.
Under non-stationary setting, the managers can almost not describe suitable probability distribution; in addition, classical online algorithm is too conservative. In this paper the manager employ online algorithm with risk-reward to deal with ordering problem in supply chain, which make effectually use of valuable information within acceptable risk tolerance level.

# Acknowledgement

# References

[1] M. Khouja, The Single-Period (news-vendor) Problem: Literature Review and Suggestions for Future Research, *Omega*, 27: 537-553, 1999.

[2] T. M.Within, Inventory control and price theory, *Management science*, 2(1): 61-68, 1955.

[3] K. Arrow, S. Karlins and H. Scarf, *Studies in the mathematical theory of inventory and production, Stanford*, CA: Stanford University press, p201-209, 1958.

[4] W. Shih, A note on Bayesian approach to newsboy inventory problem, *Decision science*, 4: 184-189, 1973.

[5] I. Moon and S. Choi, The distribution free newsboy problem with balking, Journal of Operational Research Society, 46: 537-542, 1995.

[6] K. Baruch and S. P. Joseph, A benchmark solution for the risk-averse newsvendor problem, *European Journal of Operational Research*，174(2): 1643-1650, 2006.

[7] D. D. Sleator and R. E. Tarjan, Amortized efficiency of list update and paging rules, *Communication of the ACM*, 28: 202-208, 1985.

[8] A. Borodin, R. El-Yaniv, *Online computation and competitive analysis*, Cambridge University Press, 1998.

[9] A. R. Karlia, M. S. Manaees and L. MeGeogh, Competitive randomized algorithms for non-uniform problem, *Algorithmica*, 11(6): 542-571, 1994.

[10] R. El-Yaniv, A. Fiat, R. M. Karp and G. Turpin, Optimal Search and One-Way Trading Online Algorithms, *Algorithmica*, 30: 101-139, 2001.

[11] L. Donald and Iglehart, Optimality of (s, S) policies in the infinite horizon dynamic inventory problem, *Management science*, 9(2): 259-267, 1963.

[12] S. Al-Binali, A risk-reward framework for the competitive analysis of financial game, Algorithmica, 25: 99-115, 1999.

[13] C. L. Xin and W. M. Ma, Scheduling for On-line taxi Problem on a real Line and Competitive Algorithms, *Proceeding of ICMLC2004 Conference, Shanghai*, pp. 3078-3083, 2004.

[14] C. L. Xin, W. M. Ma, L. L. Ding and F. L. Yi, On-line distribution system of trial-sale production and its competitive algorithms, *Proceeding of ICMLC2005 Conference, Guangzhou*, pp. 467-472, 2005.

[15] Y. F. Xu, C. L. Xin and F. L. Yi, New Results on Online Replacement Problem, *Proceeding of WINE2005 Conference, Hong Kong*, pp. 554-563, 2005.

[16] C. L. Xin, W. T. Cui and W. M. Ma, Risk Management for Online Simplified Bahncard Problem, *Proceeding of ICMLC2006 Conference, Dalian*, pp. 715-720, 2006.