

Data-Driven Algorithms for Two-Location Inventory Systems

Zijun Zhong¹, Mingyang Yuan² and Zhou He^{1,3,4,5,*} 

¹ School of Economics and Management, University of Chinese Academy of Sciences, 3 Zhongguancun Nanyitiao, Beijing 100190, China; zhongzijun21@mailsucas.edu.cn

² China Academy of Information and Communications Technology, 52, Hua Yuan Bei Road, Beijing 100191, China

³ MOE Social Science Laboratory of Digital Economic Forecasts and Policy Simulation at UCAS, 3 Zhongguancun Nanyitiao, Beijing 100190, China

⁴ Key Laboratory of Big Data Mining and Knowledge Management, Chinese Academy of Sciences, 80 Zhongguancun East Road, Beijing 100190, China

⁵ International Society for the Systems Sciences, Ashland, TN 41101, USA

* Correspondence: hezhou@ucas.ac.cn

Abstract: In this paper, we consider a multiperiod, two-location inventory system with unknown demand distributions and perishable products. Products can be transshipped from the location with excess inventory to the other with excess demand to better fulfill customer demand. The demand distributions are assumed to follow a family of parametric distributions and can only be learned on the fly. To address the challenge, we propose a data-driven inventory management algorithm called DD2LI that achieves a good performance in terms of regret. This algorithm, DD2LI, employs maximum likelihood estimation to approximate the unknown parameter and determines the order quantity based on these estimations. In addition, we emphasize a key assumption that tightens regret bound. Finally, we test the effectiveness of our proposed algorithm by conducting numerical experiments for two scenarios.

Keywords: data-driven algorithm; transshipment; inventory management; regret analysis



Citation: Zhong, Z.; Yuan, M.; He, Z. Data-Driven Algorithms for Two-Location Inventory Systems. *Systems* **2024**, *12*, 153. <https://doi.org/10.3390/systems12050153>

Academic Editor: Sitalakshmi Venkatraman

Received: 27 March 2024

Revised: 17 April 2024

Accepted: 28 April 2024

Published: 29 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Inventory management research is a significant branch of operations research that addresses the challenges businesses face in meeting customer demand while minimizing costs associated with excess and insufficient inventory. The traditional newsvendor model has long been a cornerstone of inventory management theory, providing a framework for single-location inventory decisions under uncertainty. However, in today's globalized and interconnected business environment, companies often operate across multiple locations, each with its own unique demand dynamics and inventory [1]. For instance, a clothing store chain may operate multiple outlets in a city [2]. In this way, the store chain can carry out transshipment if one store has excess inventory and another is experiencing a stockout [3]. In such a system, efficient management of inventory across multiple locations is vital to maintaining customer satisfaction and operational efficiency.

The current literature on inventory management and transshipment primarily presupposes a known demand distribution, hence the inventory managers can derive optimal inventory levels and transshipment strategies [3,4]. Yet in practice, the specific form of the demand distribution is generally unknown, and inventory decisions must be made based on limited demand samples. This serves the starting point for research on data-driven inventory management strategies that can adapt to unknown demand patterns. When demand distribution is unknown a priori, inventory managers must learn the demand on the fly and make inventory decisions by past demand data [5,6]. Despite the prevalence of data-driven inventory management studies, current studies mainly focus on single-location inventory decisions.

Given the aforementioned challenges, this paper considers a multiperiod, two-location inventory system with perishable products. The firm can incur a cost to transship excess inventory in one location to the other experiencing stockout. The model with full demand information has been explored in [3,4]. Our work differs from previous research by assuming the demand distributions are unknown and by proposing data-driven algorithms tailored for the two-location inventory system. In this paper, the demand distributions in the two locations are assumed to follow a parametric family of distributions and are unknown a priori. Inventory decisions can only be made by past demand data. The objective of our paper is to design an adaptive algorithm for two-location inventory systems and to establish the theoretical performance of the algorithm.

Next, we summarize the main results of our work. First, we propose a parametric data-driven algorithm called DD2LI, specifically designed for the two-location inventory system with transshipment. This algorithm utilizes past demand data, employs maximum likelihood estimation to approximate the unknown demand parameters, and determines the order quantity based on the estimations. Second, we characterize the regret bound of our proposed algorithm based on asymptotics of the maximum likelihood estimator. The regret bound of our algorithm is $O(\sqrt{T})$ and can be strengthened to $O(\log T)$ with an additional assumption that the optimal order quantity is Lipschitz continuous with respect to parameters. Finally, the proved regret bound is validated through numerical experiments.

There are two main contributions in this paper. First, we design adaptive inventory control algorithms for two-location inventory systems with unknown demand distributions. This algorithm holds significant practical value, given the common scenario of unknown demand in real-world applications. We also hope that our study will offer valuable insights to store managers tasked with inventory management across multiple locations. Second, our paper is among the first to explore data-driven inventory management strategies for problems with more than one location. The novelty of our work lies in the performance analysis of DD2LI, where we bound the difference between expected cost functions using demand parameters. We provide proof that our proposed algorithm achieves an optimal convergence rate.

The remainder of this paper is organized as follows. The literature is reviewed in Section 2. In Section 3, we outline the classic standard newsvendor model and the two-location inventory system with transshipment. In Section 4, we present the formulation of a multiperiod, two-location inventory system. In Section 5, we propose a data-driven algorithm called DD2LI, and we present its performance analysis in the section Performance Analysis. In Section 6, we present results of numerical experiments under two scenarios. We conclude our paper in Section 7.

2. Literature Review

We review the relevant literature from two aspects: two-location inventory systems with transshipment and data-driven inventory management.

2.1. Two-Location Inventory Systems with Transshipment

First, our paper relates to the literature on two-location inventory systems with transshipment. Ref. [7] first introduced the concept of inventory sharing across multiple locations. It is demonstrated that, when the demand at each location follows an independent normal distribution, inventory sharing can reduce the system's total cost to $O(1/\sqrt{n})$, where n is the number of locations, without considering transportation costs. Ref. [7] serves as a foundation for expanding the standard newsvendor model to incorporate multilocation extensions. An important research stream is the study of inventory transshipment. For a comprehensive review, please refer to [8]. When multiple locations can adjust inventory through transshipment, Ref. [9] studied a centralized transshipment network of n newsvendors, establishing the optimality of a base-stock policy. Refs. [10–13] conducted research on optimal strategies for two-location production and sales systems, considering various settings and using different demand and supply models. Ref. [14] examined single-warehouse

multilocation inventory systems and proved that there are five possible optimal strategies. Ref. [4] considered the impact of the manufacturer's pricing on multi-retailer systems, finding that the manufacturer's profit is significantly influenced by whether the manufacturer is the price setter or price taker in the presence of retailer inventory sharing. Recent literature has examined general multilocation systems. Ref. [15] investigated multi-warehouse, multi-store systems with no external replenishment, designing asymptotically optimal policy via Lagrangian relaxation. Ref. [16] applied distributionally robust optimization to analyze a multilocation newsvendor network with demand ambiguity. The authors applied a moment-based uncertainty set and derived inventory levels that minimize the worst-case expected cost. Ref. [17] studied a multilocation inventory system with an additive or multiplicative random yield. They found the comparison between centralization and decentralization hinges significantly on demand uncertainty.

Another important stream explores the strategic behavior of decentralized newsvendors. Ref. [3] investigated the behavior of two newsvendors maximizing profits through mutual transshipment, deriving equilibrium order quantities and transshipment prices. Ref. [18] explored the impact of demand information asymmetry and designed an optimal information coordination mechanism. Ref. [19] developed a two-stage game model to examine the inventory and end-of-season transshipment decisions between competing retailers. Ref. [20] examined the impact of inventory transshipment network structures on transshipment equilibria. Additionally, from a behavioral perspective, Ref. [21] studied the transshipment equilibrium between bounded rational newsvendors. Ref. [22] found that overconfidence could undermine the benefits brought by transshipment. Ref. [23] discovered from the practices of procurement managers that they tend to order less when inventory sharing is involved. Based on this observation, the authors in [23] constructed a behavioral model to explain this phenomenon.

The existing literature on inventory transshipment primarily assumes that the underlying demand distribution is known. We complement this line of research by developing data-driven inventory control strategies for a two-location inventory system with transshipment when the demand distribution is unknown a priori.

2.2. Data-Driven Inventory Management

Our work is also related to data-driven inventory management, particularly focusing on the parametric stream. The data-driven research on inventory management can primarily be categorized into two parts: one assumes that the demand distribution falls within a specific parametric family, while the other approach does not make such an assumption. Our work belongs to the former category. Early studies in this category are mainly based on Bayesian dynamic programming. Ref. [24] assumed that the demand distribution was generated from a certain parametric family with unknown parameters, but the posterior distribution of the parameters could be obtained from demand samples, thereby modeling inventory decisions as a Bayesian Markov decision process. Refs. [25,26] examined cases where the demand distribution follows a specific parametric family. Ref. [26] provides explicit solutions for inventory control strategies when the demand distribution follows the Weibull distribution. Ref. [27] studied nonstationary demand process. Nevertheless, methods based on Bayesian Markov decision processes only guarantee the convergence of the algorithm, providing no performance analysis compared to full information benchmarks. The methods' performances are solely described through limited numerical experiments, offering only a partial understanding of their effectiveness.

Recent research on parametric methods utilizes maximum likelihood estimation (MLE) and concentration inequalities to obtain the regret bound. Ref. [5] studied a joint pricing and inventory management problem for perishable products with changing demand distributions. The authors in [5] considered that the random error in the demand distribution follows an exponential family distribution and then used the maximum likelihood estimation to estimate the parameters of the exponential family distribution. Refs. [28,29] investigated joint pricing and inventory management with limited price adjustments and

parametric demand. Ref. [29] provided the concentration properties of the maximum likelihood estimator for censored demand. In addition, there are other research studies that employ different underlying methods. Ref. [30] studied the network revenue management problem, assuming a parameterized prior distribution for the demand. They designed an algorithm based on Thompson sampling by calculating the Bayesian posterior distribution from demand samples. Ref. [31] introduced operational statistics, which integrate demand estimation and inventory optimization.

Contrary to existing studies that mainly investigate single-location inventory problems, our work contributes to the literature by proposing a parametric learning algorithm tailored for two-location inventory systems. Moreover, we have successfully established the performance of our algorithm and validated its effectiveness through numerical experiments.

3. Preliminaries

3.1. The Standard Newsvendor Model

This section provides a review of the setup and solution to the standard newsvendor model. For a more comprehensive review of the newsvendor model, readers may refer to [32,33].

The newsvendor model assumes that the demand distribution for a product is stochastic, represented by the random variable D . At the beginning of the period, the risk-neutral newsvendor must decide on the order quantity y . The product has a fixed unit retail price of p , a unit purchase cost of c (with $p > c$), and, for simplicity, the salvage cost is normalized to zero.

For a real number x , let x^+ represent $\max\{x, 0\}$. The expected profit of the newsvendor, denoted as $\Pi(y)$, is a function of the order quantity y . It is defined as follows:

$$\Pi(y) = p\mathbb{E} \min\{y, D\} - cy.$$

When D is a continuous random variable, Π is differentiable. Setting the derivative of Π to zero, it can be found that the optimal order quantity y^* satisfies the equation

$$F(y^*) = \frac{p-c}{p},$$

which shows that the optimal order quantity is the $(p-c)/p$ quantile of the demand distribution.

3.2. Two-Location Inventory System with Transshipment

To fulfill demand efficiently and maximize profits, a firm operating two outlets can carry out transshipments whenever a stockout occurs in one location, while excess stock is available at the other location.

We follow the framework established by [4]. Consider two retailers designated as i and j (where $j = 3 - i$). The two retailers are operated by a single firm. In the event of a stockout at one location, excess demand can be met through transshipment if another location has surplus inventory. We represent the random demand in the two locations as D_1 and D_2 . It is assumed that the demands are continuous random variables, each characterized by differentiable cumulative distribution functions F_1 and F_2 , as well as continuous probability density functions f_1 and f_2 .

A central planner decides the order quantity y_1 and y_2 for the two locations. The product has a unit price p and a unit purchase c (with $p > c$), and transshipment can be carried out at a cost of τ .

The transshipment quantity between the two retailers is the minimum of excess demand and demand shortage. It can be expressed as $\min\{(D_1 - y_1)^+ + (D_2 - y_2)^+, (y_1 - D_1)^+ + (y_2 - D_2)^+\}$.

The central planner aims to maximize the per-period expected profit:

$$\begin{aligned}
 Q(y_1, y_2) &= \mathbb{E}[p(\min\{D_1, y_1\} + \min\{D_2, y_2\})] - c(y_1 + y_2) \\
 &\quad + \mathbb{E}[(p - \tau) \min\{(D_1 - y_1)^+ + (D_2 - y_2)^+, (y_1 - D_1)^+ + (y_2 - D_2)^+\}] \\
 &= -\mathbb{E}\{(p - c - \tau)(D_1 + D_2 - y_1 - y_2)^+ + \tau[(D_1 - y_1)^+ + (D_2 - y_2)^+]\} \\
 &\quad + (p - c)\mathbb{E}[D_1 + D_2].
 \end{aligned} \tag{1}$$

The expected profit comprises three components: the expected revenue generated from direct sales at both locations, minus the associated costs, and the additional revenue obtained through transshipment.

The equality (1) follows from the observation that

$$\begin{aligned}
 &\min\{(D_1 - y_1)^+ + (D_2 - y_2)^+, (y_1 - D_1)^+ + (y_2 - D_2)^+\} \\
 &= \min\{D_1 + D_2, y_1 + y_2\} - (\min\{D_1, y_1\} + \min\{D_2, y_2\}),
 \end{aligned}$$

and the fact that $\min\{a, b\} = a - (a - b)^+$ and $a = b - (b - a)^+ + (a - b)^+$.

Taking partial derivatives of Q with respect to y_1, y_2 , we obtain

$$\begin{aligned}
 \frac{\partial Q}{\partial y_1} &= p - c - (p - c - \tau)F_{12}(y_1 + y_2) - \tau F_1(y_1), \\
 \frac{\partial Q}{\partial y_2} &= p - c - (p - c - \tau)F_{12}(y_1 + y_2) - \tau F_2(y_2)
 \end{aligned}$$

and

$$\nabla^2 Q = \begin{bmatrix} -(p - c - \tau)f_{12}(y_1 + y_2) - \tau f_1(y_1) & -(p - c - \tau)f_{12}(y_1 + y_2) \\ -(p - c - \tau)f_{12}(y_1 + y_2) & -(p - c - \tau)f_{12}(y_1 + y_2) - \tau f_2(y_2) \end{bmatrix},$$

where and F_{12}, f_{12} represent the distribution function and probability density function of $D_1 + D_2$, respectively.

Following the analysis in [4,9], we claim that $Q(y_1, y_2)$ is jointly concave in (y_1, y_2) , and the profit-maximizing central planner chooses optimal order quantity (y_1^*, y_2^*) satisfying the first order condition if the distributions of D_1 and D_2 are known:

$$\frac{\partial Q}{\partial y_1}(y_1^*, y_2^*) = 0, \quad \frac{\partial Q}{\partial y_2}(y_1^*, y_2^*) = 0.$$

4. Problem Formulation

In this section, we consider a T -period two-location inventory system with parametric demand distributions.

For each period $t = 1, 2, \dots, T$, we denote D_t^1, D_t^2 as the demand distributions for the two locations and d_t^1, d_t^2 as their realizations. For ease of exposition, we introduce the random vector $\mathbf{D}_t = (D_t^1, D_t^2)$ and its realization $\mathbf{d}_t = (d_t^1, d_t^2)$. Specifically, we assume \mathbf{D} is parameterized by $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^k$, where Θ is a compact subset of \mathbb{R}^k and represents the domain of $\boldsymbol{\theta}$. We denote this parameterized demand distribution as $\mathbf{D}(\boldsymbol{\theta})$ and its corresponding probability density function as $f(\mathbf{d}; \boldsymbol{\theta})$.

We make the following assumptions about the demand distributions.

Assumption 1. The random vector \mathbf{D}_t is independently and identically distributed (i.i.d.) across time period t .

Assumption 2. The family of distributions $\{f(\mathbf{d}; \boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta\}$ is identifiable: the probability function $f(\mathbf{d}; \boldsymbol{\theta}_1) \neq f(\mathbf{d}; \boldsymbol{\theta}_2)$ for $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_2$.

Assumption 3. The Fisher information matrix $I(\theta)$ is bounded and positive definite. $I(\theta)$ is defined by

$$[I(\theta)]_{ij} = \mathbb{E} \left[-\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log \mathcal{L}^\theta(\mathbf{D}) \right],$$

where $\mathcal{L}^\theta(\mathbf{D})$ is the likelihood function.

Assumption 1 assumes the demand process is stationary. This assumption is common in inventory literature [6,34,35]. Assumption 2 assumes the family of parametric demand distributions under consideration is identifiable, ensuring that each parameter vector uniquely determines the corresponding probability distribution. Identifiability is an important concept in mathematical statistics, and it has been extensively employed in the literature [36]. Assumption 3 ensures the convergence of the maximum likelihood estimator (MLE). Notably, exponential family distributions inherently satisfy both Assumptions 2 and 3. Consequently, the analysis presented in our paper applies to a wide range of common parametric distribution families.

At the beginning of each period t , the central planner decides the order quantities $\mathbf{y}_t = (y_t^1, y_t^2)$ for both outlets. The central planner has no knowledge of the true underlying demand distribution a priori but can rely on historical demand data and make adaptive inventory decisions based on the available information.

In this work, we consider only perishable products, implying there are no inventory carryovers across periods. Therefore, in each period t , the central planner collects the profit $(p - c - \tau)(d_t^1 + d_t^2 - y_1 - y_t^2)^+ + \tau[(d_t^1 - y_t^1)^+ + (d_t^2 - y_t^2)^+] + (p - c)(d_t^1 + d_t^2)$. We define $Q(\mathbf{y}_t, \theta)$ as the per-period expected profit function when the order quantity is \mathbf{y}_t and when the demand parameter is θ :

$$Q(\mathbf{y}_t, \theta) = -\mathbb{E}_{D \sim D(\theta)} \left\{ (p - c - \tau)(D_t^1 + D_t^2 - y_t^1 - y_t^2)^+ + \tau[(D_t^1 - y_t^1)^+ + (D_t^2 - y_t^2)^+] \right\} + (p - c)\mathbb{E}[D_t^1 + D_t^2]. \tag{2}$$

Let $\{\mathcal{H}_t : t \geq 0\}$ represent the sequence of filtrations generated by demand data and decisions accumulated up to time t . Precisely, \mathcal{H}_t is defined as the sigma algebra $\sigma(\mathbf{y}_k, \mathbf{d}_k : k = 1, 2, \dots, t)$ with $\mathcal{H}_0 = \emptyset$. A feasible policy ϕ is a sequence of functions $\mathbf{y}_t = \phi(\mathcal{H}_{t-1})$, which maps the historical information to current inventory decisions.

According to the analysis in Section 3.2, if the underlying demand distribution is known, then there exists an order quantity \mathbf{y}^* that maximizes the per-period profit, i.e., $\mathbf{y}^* = \arg \max_{\mathbf{y}} Q(\mathbf{y}, \theta)$. In the system with perishable products, the optimal policy is a myopic policy ϕ^* that sets the order quantity as $\mathbf{y}^* = (y_1^*, y_2^*)$. We refer to \mathbf{y}^* as the clairvoyant optimal solution and ϕ^* as the clairvoyant optimal policy.

To measure the performance of data-driven policy ϕ , we use regret as the criterion, which is defined as the expected total profit loss incurred by ϕ when compared to the clairvoyant optimal policy ϕ^* :

$$R_T(\phi) = Q(\mathbf{y}^*, \theta) - \mathbb{E} \left[\sum_{t=1}^T Q(\mathbf{y}_t, \theta) \right]. \tag{3}$$

Regret is a widely adopted metric in the literature on data-driven inventory management [6,29,37]. The metric quantifies the reduced profit caused by a lack of demand information. A lower regret value indicates a more effective policy. Thus, the central planner's goal is to devise an algorithm that minimizes regret.

Finally, we make the following assumptions about the problem.

Assumption 4. The clairvoyant optimal order quantity \mathbf{y}^* is upper-bounded by $M > 0$: $y_1^* \leq M$, $y_2^* \leq M$.

Assumption 5. The profit function $Q(\mathbf{y}, \theta)$ is Lipschitz continuous with respect to θ : there exists a constant $L_1 > 0$ such that $|Q(\mathbf{y}, \theta_1) - Q(\mathbf{y}, \theta_2)| \leq L_1 \|\theta_1 - \theta_2\|_2$ for $\theta_1, \theta_2 \in \Theta$.

Assumption 6. $\mathbf{y}^*(\theta)$ is Lipschitz continuous with respect to θ : there exists a constant $L_2 > 0$ such that $\|\mathbf{y}^*(\theta_1) - \mathbf{y}^*(\theta_2)\|_2 \leq L_2 \|\theta_1 - \theta_2\|_2$ for $\theta_1, \theta_2 \in \Theta$, where $\mathbf{y}^*(\theta)$ is the optimal order quantity given the demand parameter θ .

Assumption 4 is mild. It claims the clairvoyant optimal order quantity is upper-bounded, which can be met with common demand distributions. This assumption facilitates our theoretical analysis. Assumption 5 assumes that the per-period profit function exhibits Lipschitz continuity with respect to demand parameters. This assumption suggests that an accurate estimation of the unknown demand parameter can lead to a close approximation of the actual per-period profit function. Assumption 6 assumes the optimal order quantity is Lipschitz continuous with respect to demand parameters. This assumption implies that minor changes in the parameter will not result in significant variations in the optimal order quantity. Similar assumptions can be found in [5,29].

While Assumption 6 is not an absolute necessity for our algorithm, it plays a pivotal role in improving the regret bound. We remark that, although Assumption 6 is generally satisfied in most situations, there exist degenerate cases where it may not hold true by the following example.

Example 1. Consider a two-location system where one location's demand is equal to zero and where the demand distribution in the other location is a Bernoulli distribution with a cumulative distribution function (CDF)

$$F(0) = (p - c)/p + \theta, \quad F(1) = 1,$$

where θ is the parameter under consideration. For a small enough $\epsilon > 0$, the slight change in θ from $\theta = -\epsilon$ to $\theta = \epsilon$ will lead to a shift in its optimal order quantity, from 1 to 0.

5. Data-Driven Inventory Control Algorithms

Without any knowledge of the true underlying distribution of D_t a priori, we aim to find a provably good, adaptive, data-driven inventory control policy that makes the total expected system profits close to the optimal strategy. In this section, we introduce the data-driven algorithm DD2LI (Data-Driven Two-Location Inventory Management Algorithm) and prove its theoretical performance.

Detailed steps of DD2LI are presented below.

DD2LI: Data-Driven Two-Location Inventory Management Algorithm

Step 0. (Initialization.) In period $t = 1$, order an initial quantity $\mathbf{y}_1 = \mathbf{0}$ or select any other permissible value. Collect demand \mathbf{d}_1 and initialize the demand dataset $S = \{\mathbf{d}_1\}$ to be the available demand data up to decision time.

For each period $t = 2, 3, \dots, T$, repeat the following steps:

Step 1. (Maximum Likelihood Estimation.) Given the past demand dataset S , compute the maximum likelihood estimation of parameter θ :

$$\hat{\theta}_{t-1} = \arg \max_{\theta \in \Theta} \sum_{i=1}^{t-1} \log f(\mathbf{d}_i; \theta). \quad (4)$$

Step 2. (Order Quantity Optimization.) Using the estimated parameter $\hat{\theta}_{t-1}$ from Step 1, compute $\mathbf{y}_t = \arg \max_{\mathbf{y}} Q(\mathbf{y}, \hat{\theta}_{t-1})$, and order \mathbf{y}_t for the current period t .

Step 3. (Demand Data Update.) Observe the realized demand \mathbf{d}_t , and update the demand dataset S by adding the new data point: $S \leftarrow S \cup \{\mathbf{d}_t\}$.

Throughout the algorithm, we maintain a set containing past demand data that can be used to make adaptive inventory decisions. In Step 1, given the available demand data, we compute the MLE of the unknown demand parameter constrained on the compact parameter set Θ . In Step 2, we use the MLE estimator to obtain an empirically optimal order quantity and implement it. Lastly, in Step 3, we update the demand dataset.

Performance Analysis

In this subsection, we analyze the performance of our proposed algorithm. To facilitate our analysis, we initially introduce several lemmas that serve as useful tools in our investigation.

Lemma 1 is the direct result of Theorem 36.3 in [36].

Lemma 1. *Defining $\check{\theta}_t$ as the unconstrained maximum likelihood estimator given demand data $\{d_1, d_2, \dots, d_t\}$:*

$$\check{\theta}_t = \arg \max_{\theta} \sum_{i=1}^t \log f(d_i; \theta),$$

then there exists the positive constant $K_1, K_2 > 0$, such that, for $\epsilon > 0$,

$$\mathbb{P}(\sqrt{t} \|\check{\theta}_t - \theta\|_2 \geq \epsilon) \leq K_1 e^{-K_2 \epsilon^2}. \tag{5}$$

It is well-known in mathematical statistics that maximum likelihood estimators are asymptotically normal under proper conditions. Additionally, Lemma 1 provides a further concentration inequality pertaining to MLE, which plays a vital role in performance analysis.

Corollary 1. *For $\epsilon > 0$, the projection onto Θ of the maximum likelihood estimator $\hat{\theta}_t$ defined in (4) satisfies*

$$\mathbb{P}(\|\hat{\theta}_t - \theta\|_2 \geq \epsilon) \leq K_1 e^{-tK_2 \epsilon^2}.$$

Proof. Note $\|\check{\theta}_t - \theta\|_2 \geq \|\hat{\theta}_t - \theta\|_2 \Rightarrow \mathbb{P}(\|\hat{\theta}_t - \theta\|_2 \geq \epsilon) \leq \mathbb{P}(\|\check{\theta}_t - \theta\|_2 \geq \epsilon)$. Replacing ϵ with $\sqrt{t}\epsilon$ in (5) yields the result. \square

Below, we introduce a lemma that is widely used in probability. Lemma 2 states the expectation of X can be obtained by the integral of its survival function.

Lemma 2. *Suppose a continuous random variable $X \sim F(\cdot)$ is nonnegative; if $\mathbb{E}[X] < \infty$, then*

$$\mathbb{E}[X] = \int_0^\infty [1 - F(x)] dx.$$

Proof. We relegate the detailed proof of Lemma 2 to Appendix A.1. \square

Theorem 1 below presents one of the main results in our paper.

Theorem 1. *With Assumption 6, there exists some positive constant $A_1 > 0$ such that, for $T \geq 1$,*

$$R_T = Q(\mathbf{y}^*, \theta) - \mathbb{E} \left[\sum_{t=1}^T Q(\mathbf{y}_t, \theta) \right] \leq A_1 \log T,$$

where \mathbf{y}_t is the order quantity generated by algorithm DD2LI.

Proof. First, we apply Taylor’s expansion to $Q(\mathbf{y}_t, \theta)$ at the maximizer $\mathbf{y} = \mathbf{y}^*$,

$$Q(\mathbf{y}_t, \theta) = Q(\mathbf{y}^*, \theta) + \langle \nabla_{\mathbf{y}} Q(\mathbf{y}^*, \theta), \mathbf{y}_t - \mathbf{y}^* \rangle + \langle \mathbf{y}_t - \mathbf{y}^*, \nabla_{\mathbf{y}}^2 Q(\xi \mathbf{y}_t + (1 - \xi) \mathbf{y}^*, \theta) (\mathbf{y}_t - \mathbf{y}^*) \rangle, \quad \xi \in [0, 1].$$

Since $Q(\mathbf{y}, \boldsymbol{\theta})$ is maximized at $\mathbf{y} = \mathbf{y}^*$, we have $\nabla_{\mathbf{y}}Q(\mathbf{y}^*, \boldsymbol{\theta}) = \mathbf{0}$.

In addition, given Assumption 4, we claim that ∇^2Q is continuous and hence upper-bounded in a compact set. Therefore, we get the inequality that there exists a constant $K_3 > 0$ such that

$$|Q(\mathbf{y}_t, \boldsymbol{\theta}) - Q(\mathbf{y}^*, \boldsymbol{\theta})| \leq K_3\|\mathbf{y}_t - \mathbf{y}^*\|_2^2 \leq L_2K_3\|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}\|_2^2. \tag{6}$$

The second inequality follows from Assumption 6.

Next, we define an event \mathcal{E}_1 by

$$\mathcal{E}_1 = \left\{ \|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}\|_2 < \sqrt{\frac{\log t}{K_2t}} \right\}. \tag{7}$$

According to Corollary 1, we have

$$\mathbb{P}(\mathcal{E}_1) > 1 - \frac{K_1}{t}, \quad \mathbb{P}(\mathcal{E}_1^c) \leq \frac{K_1}{t}.$$

Now, we decompose $\mathbb{E}[\|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}\|_2^2]$ by event \mathcal{E}_1 :

$$\begin{aligned} \mathbb{E}[\|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}\|_2^2] &= \mathbb{P}(\mathcal{E}_1)\mathbb{E}[\|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}\|_2^2 \mid \mathcal{E}_1] + \mathbb{P}(\mathcal{E}_1^c)\mathbb{E}[\|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}\|_2^2 \mid \mathcal{E}_1^c] \\ &\leq \mathbb{E}[\|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}\|_2^2 \mid \mathcal{E}_1] + \frac{K_1}{t}(\text{diam } \Theta)^2 \end{aligned} \tag{8}$$

$$\leq \int_0^{\log t / (K_2t)} K_1 e^{-tK_2\epsilon} d\epsilon + \frac{K_1}{t}(\text{diam } \Theta)^2 \tag{9}$$

$$\leq \frac{K_4}{t}. \tag{10}$$

The inequality (8) follows from $\mathbb{P}(\mathcal{E}_1) \leq 1, \mathbb{P}(\mathcal{E}_1^c) \leq K_1/t$ and $\|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}\|_2 \leq \text{diam } \Theta$. Here, $\text{diam } \Theta = \max_{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \Theta} \|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|_2$ is the diameter of the set Θ . The inequality (9) is due to

Lemma 2 and Corollary 1. By inequality (10), we claim the result is bounded by $O(1/t)$.

Combining (6) and (10) yields

$$\begin{aligned} Q(\mathbf{y}^*, \boldsymbol{\theta}) - \mathbb{E} \left[\sum_{t=1}^T Q(\mathbf{y}_t, \boldsymbol{\theta}) \right] &\leq \sum_{t=1}^T \mathbb{E}[|Q(\mathbf{y}^*, \boldsymbol{\theta}) - Q(\mathbf{y}_t, \boldsymbol{\theta})|] \\ &\leq L_2K_3 \sum_{t=1}^T \mathbb{E}[\|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}\|_2^2] \\ &\leq L_2K_3K_4 \sum_{t=1}^T \frac{1}{t}. \end{aligned}$$

Recall a well-known result in mathematical analysis that the sequence

$$f(n) = \sum_{k=1}^n \frac{1}{k} - \log n$$

converges to Euler’s constant as $n \rightarrow \infty$, which indicates $\sum_{t=1}^T \frac{1}{t} \sim \log T$. Hence, we claim the theorem holds. \square

Theorem 1 states the regret bound is $O(\log T)$ if Assumption 6 holds. According to [37], it will never be possible to find a policy with a regret smaller than $O(\log T)$. Hence, our proposed algorithm can achieve the best possible regret bound.

In contrast, for cases where Assumption 6 does not hold, we have the following theorem:

Theorem 2. Without Assumption 6, there exists some positive constant $A_2 > 0$ such that, for $T \geq 1$,

$$R_T = Q(\mathbf{y}^*, \boldsymbol{\theta}) - \mathbb{E} \left[\sum_{t=1}^T Q(\mathbf{y}_t, \boldsymbol{\theta}) \right] \leq A_2 \sqrt{T},$$

where \mathbf{y}_t is the order quantity generated by algorithm DD2LI.

Proof. First, we decompose the per-period regret $\mathbb{E}[Q(\mathbf{y}^*, \boldsymbol{\theta}) - Q(\mathbf{y}_t, \boldsymbol{\theta})]$ into three parts:

$$\begin{aligned} \mathbb{E}[Q(\mathbf{y}^*, \boldsymbol{\theta}) - Q(\mathbf{y}_t, \boldsymbol{\theta})] &= \mathbb{E}[Q(\mathbf{y}^*, \boldsymbol{\theta}) - Q(\mathbf{y}^*, \hat{\boldsymbol{\theta}}_t)] + \mathbb{E}[Q(\mathbf{y}^*, \hat{\boldsymbol{\theta}}_t) - Q(\mathbf{y}_t, \hat{\boldsymbol{\theta}}_t)] \\ &\quad + \mathbb{E}[Q(\mathbf{y}_t, \hat{\boldsymbol{\theta}}_t) - Q(\mathbf{y}_t, \boldsymbol{\theta})]. \end{aligned} \tag{11}$$

Since $Q(\cdot, \hat{\boldsymbol{\theta}}_t)$ is maximized at $\mathbf{y} = \mathbf{y}_t$, the second part in (11) is

$$\mathbb{E}[Q(\mathbf{y}^*, \hat{\boldsymbol{\theta}}_t) - Q(\mathbf{y}_t, \hat{\boldsymbol{\theta}}_t)] \leq 0. \tag{12}$$

For the first and third parts, by Assumption 5 we have

$$\mathbb{E}[Q(\mathbf{y}^*, \boldsymbol{\theta}) - Q(\mathbf{y}^*, \hat{\boldsymbol{\theta}}_t)] + \mathbb{E}[Q(\mathbf{y}_t, \hat{\boldsymbol{\theta}}_t) - Q(\mathbf{y}_t, \boldsymbol{\theta})] \leq 2L_1 \mathbb{E}[\|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}\|_2]. \tag{13}$$

Recall the definition of \mathcal{E}_1 in (7). Now, we decompose $\mathbb{E}[\|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}\|_2]$ by event \mathcal{E}_1 :

$$\begin{aligned} \mathbb{E}[\|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}\|_2] &= \mathbb{P}(\mathcal{E}_1) \mathbb{E}[\|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}\|_2 \mid \mathcal{E}_1] + \mathbb{P}(\mathcal{E}_1^c) \mathbb{E}[\|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}\|_2 \mid \mathcal{E}_1^c] \\ &\leq \mathbb{E}[\|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}\|_2 \mid \mathcal{E}_1] + \frac{K_1}{t} \text{diam } \Theta \\ &\leq \int_0^{\sqrt{\log t / (K_2 t)}} K_1 e^{-tK_2 \epsilon^2} d\epsilon + \frac{K_1}{t} \text{diam } \Theta \\ &\leq \frac{K_5}{\sqrt{t}}. \end{aligned} \tag{14}$$

Since $\int_0^\infty e^{-t\epsilon^2} = \frac{\sqrt{\pi}}{2\sqrt{t}}$, the integral in (14) can be bounded by $O(1/\sqrt{t})$. By (15), we claim the result is bounded by $O(1/\sqrt{t})$.

Combining (11), (12), (13), and (15) yields

$$\begin{aligned} Q(\mathbf{y}^*, \boldsymbol{\theta}) - \mathbb{E} \left[\sum_{t=1}^T Q(\mathbf{y}_t, \boldsymbol{\theta}) \right] &\leq 2L_1 \sum_{t=1}^T \mathbb{E}[\|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}\|_2] \\ &\leq 2L_1 K_5 \sum_{t=1}^T \frac{1}{\sqrt{t}} \\ &\leq 4L_1 k_5 \sqrt{T}, \end{aligned}$$

where the last inequality follows from the fact $\sum_{t=1}^T \frac{1}{t} \leq 2\sqrt{T}$. \square

By comparing the proofs of Theorems 1 and 2, it is shown that Assumption 6 serves to strengthen the continuity between per-period profit function and the parameter (see the difference between (6) and (13)). Specifically, the inclusion of Assumption 6 allows for a more precise estimation of the per-period profit function compared to scenarios where this assumption is not applicable.

In addition, we remark on the analogy between Assumption 6 in our paper and Assumption (iii) in [37]. As discussed in [38], the absence of a key assumption about the optimal quantity and the parameter will lead the cumulative regret to change from $O(\log T)$ to $O(\sqrt{T})$.

Overall, in this section, we have proposed the algorithm DD2LI and proved its regret bound. The algorithm’s regret bound is at least $O(\sqrt{T})$ under proper conditions. Further-

more, with a stronger assumption about the optimal quantity’s continuity with respect to the parameter, we are able to further enhance the regret bound to $O(\log T)$. These regret bounds demonstrate that our algorithm is close to the optimal strategy on average when the time periods are large enough.

6. Numerical Experiments

In this section, we conduct numerical experiments for two scenarios. Following [39], we measure the performance of a learning algorithm by the percentage of relative regret defined as

$$\frac{R_T}{T \cdot Q(\mathbf{y}^*, \boldsymbol{\theta})} \times 100\%.$$

In both scenarios, we set unit price $p = 10$, unit cost $c = 3$, and unit transshipment cost $\tau = 1$.

In the first scenario, we consider that the demand distributions in the two locations follow exponential distributions $\text{Exp}(\lambda_1)$ and $\text{Exp}(\lambda_2)$. Hence, the demand parameter $\boldsymbol{\theta} = (\lambda_1, \lambda_2)$.

Given the demand data up to period t , the parameters λ_1, λ_2 can be estimated by

$$\hat{\lambda}_{1,t} = \frac{t}{\sum_{k=1}^t d_k^1}, \quad \hat{\lambda}_{2,t} = \frac{t}{\sum_{k=1}^t d_k^2}.$$

For two parameter sets, we conduct $N = 100$ runs and compute the average. The results are shown in Figures 1 and 2.

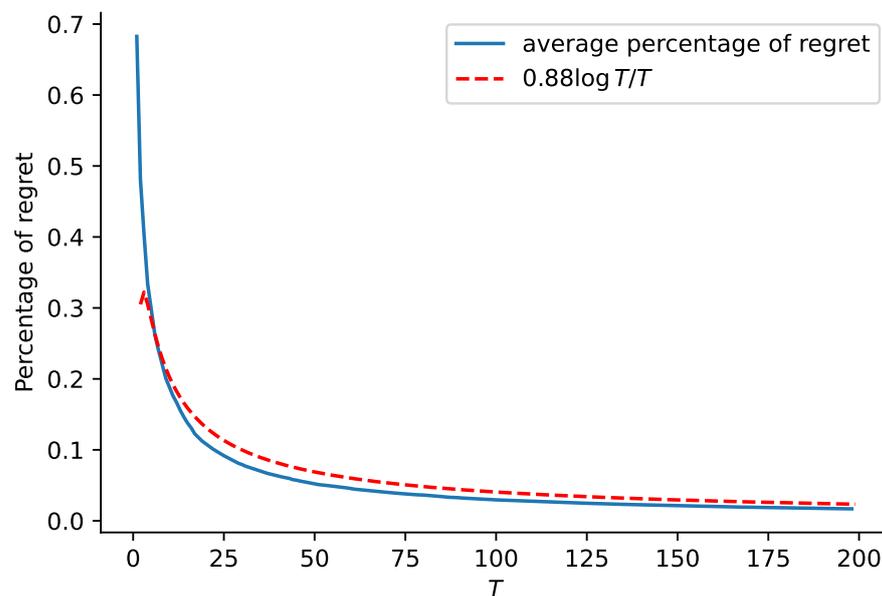


Figure 1. Curve of average percentage of regret in Scenario 1 with $\lambda_1 = 2$ and $\lambda_2 = 3$.

The blue curves in Figure 1 and Figure 2 show the average percentage of regret for $(\lambda_1, \lambda_2) = (2, 3)$ and $(\lambda_1, \lambda_2) = (0.5, 0.2)$, respectively, while the red dashed curves in the two figures represent function $\frac{0.88 \log T}{T}$ and $\frac{0.7 \log T}{T}$, respectively. This comparison shows that the regret of our proposed algorithm is close to the rate of $O(\log T)$.

In the second scenario, we consider that the demand distributions in the two locations follow a multivariate normal distribution with mean vector (μ_1, μ_2) and covariance matrix $\begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$. Hence, the parameter $\boldsymbol{\theta} = \left((\mu_1, \mu_2), \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \right)$. Note that, in this scenario, the demand distributions in the two locations may be correlated.

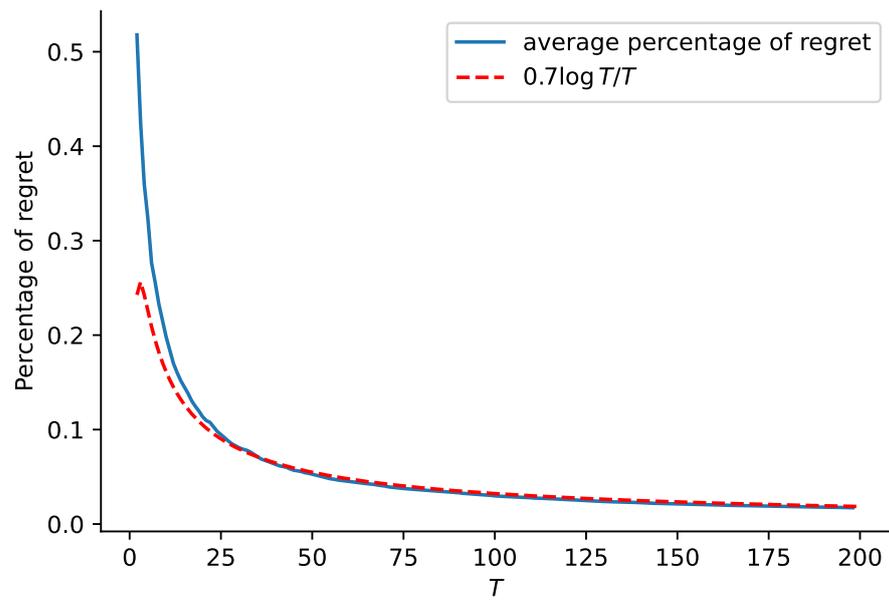


Figure 2. Curve of average percentage of regret in Scenario 1 with $\lambda_1 = 0.5$ and $\lambda_2 = 0.2$.

Given the demand data up to period t ($t \geq 2$), the parameters can be estimated by

$$\hat{\mu}_{1,t} = \frac{\sum_{k=1}^t d_k^1}{t}, \quad \hat{\mu}_{2,t} = \frac{\sum_{k=1}^t d_k^2}{t}$$

$$\hat{\sigma}_{1,t}^2 = \frac{\sum_{k=1}^t (d_k^1 - \hat{\mu}_{1,t})^2}{t-1}, \quad \hat{\sigma}_{2,t}^2 = \frac{\sum_{k=1}^t (d_k^2 - \hat{\mu}_{2,t})^2}{t-1}, \quad \rho = \frac{\sum_{k=1}^t (d_k^1 - \hat{\mu}_{1,t})(d_k^2 - \hat{\mu}_{2,t})}{t-1}$$

For two parameter sets, we conduct $N = 100$ runs and compute the average. The results are shown in Figures 3 and 4.

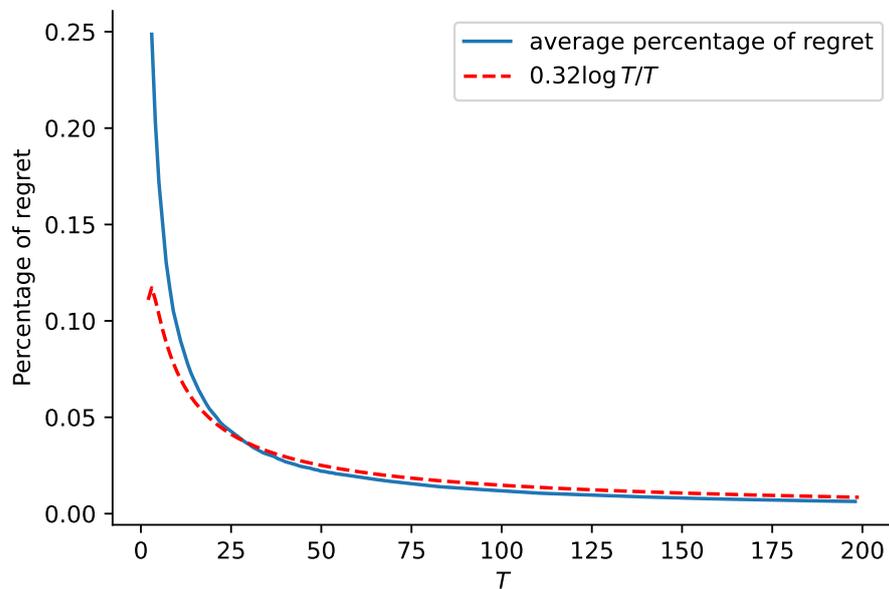


Figure 3. Curve of average percentage of regret in Scenario 2 with $\mu_1 = 8, \mu_2 = 12, \sigma_1^2 = 4, \sigma_2^2 = 9$, and $\rho = 0.1$.

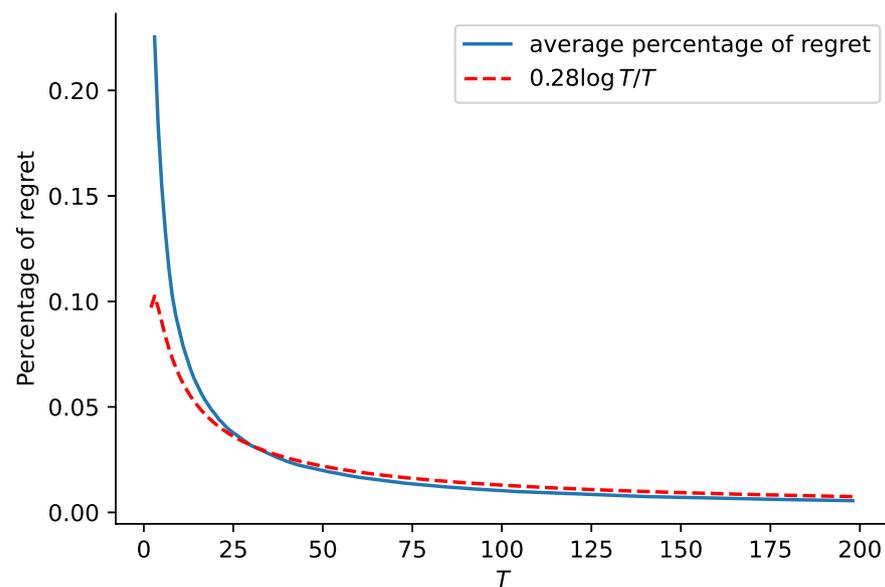


Figure 4. Curve of average percentage of regret in Scenario 2 with $\mu_1 = 5, \mu_2 = 6, \sigma_1^2 = 1, \sigma_2^2 = 4$, and $\rho = -0.5$.

The blue curves in Figure 1 and Figure 2 show average percentage of regret for $\mu_1 = 8, \mu_2 = 12, \sigma_1^2 = 4, \sigma_2^2 = 9, \rho = 0.1$ and $\mu_1 = 5, \mu_2 = 6, \sigma_1^2 = 1, \sigma_2^2 = 4, \rho = -0.5$, respectively, while the red dashed curves in the two figures represent functions $\frac{0.32 \log T}{T}$ and $\frac{0.28 \log T}{T}$, respectively. This comparison shows that the regret of our proposed algorithm is close to the rate of $O(\log T)$, and it is also applicable to correlated demands.

Based on the outcomes of four experiments, we can also conclude that Assumption 6 holds when demand follows exponential distributions or Gaussian distributions.

7. Conclusions

When a firm operates multiple stores, to better fulfill customer demand and make more profit, transshipment can be carried out if one location has excess inventory while another is experiencing a stockout. Despite the extensive research on transshipment between newsvendors, most assume the demand distribution is known a priori.

In this work, we introduce a data-driven inventory management algorithm for a multiperiod, two-location inventory system with perishable products and unknown demand distributions, which are assumed to follow a family of parametric distributions. The proposed algorithm, called DD2LI, uses past demand data to make adaptive inventory decisions. By using maximum likelihood estimation to estimate the unknown parameters, the algorithm determines the order quantity based on these estimations. We successfully derive the regret bound of the proposed algorithm under proper assumptions, which shows the algorithm is close to the optimal strategy on average. Additionally, we emphasize the assumption that the continuity of optimal order quantity with respect to parameters plays a key role in a tighter regret bound. Finally, to validate the effectiveness of our proposed algorithm, we conduct numerical experiments in two distinct scenarios.

There are several future research directions. First, future studies can design algorithms for two-location inventory systems when demand data are censored. Note that the demand censoring may be more intricate in two-location inventory systems with transshipment than in the single-location newsvendor problem. Second, while our work considers the stationary demand process, it is worth exploring algorithms that handle shifting demand. Third, our work only considers two locations, which restrict its use in practice. Future research can try to design algorithms for more complex inventory networks.

Author Contributions: Conceptualization, Z.Z. and M.Y.; methodology, Z.Z.; formal analysis, Z.Z.; investigation, Z.H.; writing—original draft preparation, Z.Z.; writing—review and editing, Z.H.; supervision, Z.H.; project administration, M.Y.; funding acquisition, M.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (grant number 71932002), the China Academy of Information and Communications Technology, the Fundamental Research Funds for the Central Universities (grant numbers E1E40810X2 and E2ET0808X2), the Youth Innovation Promotion Association CAS (grant number 110800EAG2), the MOE Social Science Laboratory of Digital Economic Forecasts and Policy Simulation at UCAS, and the Weiqiao Guoke Joint Laboratory at UCAS.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Appendix A.1. Proof of Lemma 2

Proof. Because the cumulative distribution function gives the probability of a random variable being smaller than a given value, we can obtain $F_X(x) = \Pr(X \leq x)$, and then we have $1 - F_X(x) = \Pr(X > x)$ such that $\int_0^\infty (1 - F_X(x)) dx = \int_0^\infty \Pr(X > x)$. By the probability density function, it can be written as follows:

$$\begin{aligned} \int_0^\infty (1 - F_X(x)) dx &= \int_0^\infty \int_x^\infty f_X(z) dz dx \\ &= \int_0^\infty \int_0^z f_X(z) dx dz \\ &= \int_0^\infty f_X(z) \int_0^z dx dz \\ &= \int_0^\infty z \cdot f_X(z) dz. \end{aligned}$$

Then, we apply the definition of the expectation and obtain

$$\int_0^\infty (1 - F_X(x)) dx = \int_0^\infty z \cdot f_X(z) dz = \mathbb{E}[X].$$

This completes the proof. \square

References

1. Noh, J. Reinforcement Learning for Optimizing Can-Order Policy with the Rolling Horizon Method. *Systems* **2023**, *11*, 350. [[CrossRef](#)]
2. Griffin, E.C.; Keskin, B.B.; Allaway, A.W. Clustering retail stores for inventory transshipment. *Eur. J. Oper. Res.* **2023**, *311*, 690–707. [[CrossRef](#)]
3. Rudi, N.; Kapur, S.; Pyke, D.F. A Two-Location Inventory Model with Transshipment and Local Decision Making. *Manag. Sci.* **2001**, *47*, 1668–1680. [[CrossRef](#)]
4. Dong, L.; Rudi, N. Who Benefits from Transshipment? Exogenous vs. Endogenous Wholesale Prices. *Manag. Sci.* **2004**, *50*, 645–657. [[CrossRef](#)]
5. Keskin, N.B.; Li, Y.; Song, J.-S. Data-Driven Dynamic Pricing and Ordering with Perishable Inventory in a Changing Environment. *Manag. Sci.* **2022**, *68*, 1938–1958. [[CrossRef](#)]
6. Huh, W.T.; Rusmevichientong, P. A nonparametric asymptotic analysis of inventory planning with censored demand. *Math. Oper. Res.* **2009**, *34*, 103–123. [[CrossRef](#)]
7. Eppen, G.D. Note—Effects of Centralization on Expected Costs in a Multi-Location Newsboy Problem. *Manag. Sci.* **1979**, *25*, 498–501. [[CrossRef](#)]
8. Paterson, C.; Kiesmüller, G.; Teunter, R.; Glazebrook, K. Inventory models with lateral transshipments: A review. *Eur. J. Oper. Res.* **2011**, *210*, 125–136. [[CrossRef](#)]

9. Robinson, L.W. Optimal and approximate policies in multiperiod, multilocation inventory models with transshipments. *Oper. Res.* **1990**, *38*, 278–295. [[CrossRef](#)]
10. Hu, X.; Duenyas, I.; Kapuscinski, R. Existence of Coordinating Transshipment Prices in a Two-Location Inventory Model. *Manag. Sci.* **2007**, *53*, 1289–1302. [[CrossRef](#)]
11. Yang, J.; Qin, Z. Capacitated Production Control with Virtual Lateral Transshipments. *Oper. Res.* **2007**, *55*, 1104–1119. [[CrossRef](#)]
12. Hu, X.; Duenyas, I.; Kapuscinski, R. Optimal Joint Inventory and Transshipment Control Under Uncertain Capacity. *Oper. Res.* **2008**, *56*, 881–897. [[CrossRef](#)]
13. Zhao, H.; Ryan, J.K.; Deshpande, V. Optimal Dynamic Production and Inventory Transshipment Policies for a Two-Location Make-to-Stock System. *Oper. Res.* **2008**, *56*, 400–410. [[CrossRef](#)]
14. Wee, K.E.; Dada, M. Optimal Policies for Transshipping Inventory in a Retail Network. *Manag. Sci.* **2005**, *51*, 1519–1533. [[CrossRef](#)]
15. Miao, S.; Jasin, S.; Chao, X. Asymptotically Optimal Lagrangian Policies for MultiWarehouse, Multi-Store Systems with Lost Sales. *Oper. Res.* **2022**, *70*, 141–159. [[CrossRef](#)]
16. Govindarajan, A.; Sinha, A.; Uichanco, J. Distribution-free inventory risk pooling in a multilocation newsvendor. *Manag. Sci.* **2021**, *67*, 2272–2291. [[mns.2020.3719CrossRef](#)] [[CrossRef](#)]
17. Xiao, L.; Wang, C. Multi-location newsvendor problem with random yield: Centralization versus decentralization. *Omega* **2023**, *116*, 102795. [[CrossRef](#)]
18. Yan, X.; Zhao, H. Decentralized inventory sharing with asymmetric information. *Oper. Res.* **2011**, *59*, 1528–1538. [[CrossRef](#)]
19. Fu, Q.; Liu, L.; Shang, W. Bilateral transshipment between competing retailers. *Nav. Res. Logist.* **2023**, *70*, 509–521. [[CrossRef](#)]
20. Fang, X.; Cho, S.H. Stability and Endogenous Formation of Inventory Transshipment Networks. *Oper. Res.* **2014**, *62*, 1316–1334. [[CrossRef](#)]
21. He, Q.; Shi, T.; Xu, F.; Qiu, W. Decentralized Inventory Transshipments with Quantal Response Equilibrium. *Systems* **2023**, *11*, 357. [[CrossRef](#)]
22. Li, J.; Li, M.; Zhao, X. Transshipment between overconfident newsvendors. *Prod. Oper. Manag.* **2021**, *30*, 2803–2813. [[CrossRef](#)]
23. Zhao, H.; Xu, L.; Siemsen, E. Inventory sharing and demand-side underweighting. *Manuf. Serv. Oper. Manag.* **2021**, *23*, 1217–1236. [[CrossRef](#)]
24. Ding, X.; Puterman, M.L.; Bisi, A. The censored newsvendor and the optimal acquisition of information. *Oper. Res.* **2002**, *50*, 517–527. [[CrossRef](#)]
25. Lariviere, M.A.; Porteus, E.L. Stalking information: Bayesian inventory management with unobserved lost sales. *Manag. Sci.* **1999**, *45*, 346–363. [[CrossRef](#)]
26. Bisi, A.; Dada, M.; Tokdar, S. A Censored-Data Multiperiod Inventory Problem with Newsvendor Demand Distributions. *Manu. Ser. Oper. Manag.* **2011**, *13*, 525–533. [[CrossRef](#)]
27. Bensoussan, A.; Çakanyıldırım, M.; Sethi, S.P. A multiperiod newsvendor problem with partially observed demand. *Math. Oper. Res.* **2007**, *32*, 322–344. [[CrossRef](#)]
28. Chen, B.; Chao, X. Parametric demand learning with limited price explorations in a backlog stochastic inventory system. *IIE Trans.* **2019**, *51*, 605–613. [[24725854.2018.1538594CrossRef](#)] [[CrossRef](#)]
29. Chen, B.; Chao, X.; Wang, Y. Technical Note—Data-Based Dynamic Pricing and Inventory Control with Censored Demand and Limited Price Changes. *Oper. Res.* **2020**, *68*, 1445–1456. [[CrossRef](#)]
30. Ferreira, K.J.; Simichi-Levi, D.; Wang, H. Online network revenue management using thompson sampling. *Oper. Res.* **2018**, *66*, 1586–1602. [[CrossRef](#)]
31. Liyanage, L.; Shanthikumar, J. A practical inventory control policy using operational statistics. *Oper. Res. Lett.* **2005**, *33*, 341–348. [[CrossRef](#)]
32. Petruzzi, N.C.; Dada, M. Pricing and the Newsvendor Problem: A Review with Extensions. *Oper. Res.* **1999**, *47*, 183–194. [[CrossRef](#)]
33. Qin, Y.; Wang, R.; Vakharia, A.J.; Chen, Y.; Seref, M.M.H. The newsvendor problem: Review and directions for future research. *Eur. J. Oper. Res.* **2011**, *213*, 361–374. [[CrossRef](#)] [[CrossRef](#)]
34. Shi, C.; Chen, W.; Duenyas, I. Technical note—Nonparametric data-driven algorithms for multiproduct inventory systems with censored demand. *Oper. Res.* **2016**, *64*, 362–370. [[CrossRef](#)]
35. Chen, B.; Chao, X. Dynamic inventory control with stockout substitution and demand learning. *Manag. Sci.* **2020**, *66*, 5108–5127. [[CrossRef](#)]
36. Borokov, A.A. *Mathematical Statistics*, 1st ed.; Routledge: London, UK, 1999; pp. 211–215.
37. Besbes, O.; Muharremoglu, A. On Implications of Demand Censoring in the Newsvendor Problem. *Manag. Sci.* **2013**, *59*, 1407–1424. [[CrossRef](#)]
38. Chen, X.; Jasin, S.; Shi, C. *The Elements of Joint Learning and Optimization in Operations Management*, 1st ed.; Springer Nature: Cham, Switzerland, 2022; pp. 273–279.
39. Chen, B. Data-driven inventory control with shifting demand. *Prod. Oper. Manag.* **2021**, *30*, 1365–1385. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.