DATA-DRIVEN TWO-STAGE INVENTORY PROBLEM

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF

MASTER OF SCIENCE

IN

INDUSTRIAL ENGINEERING

By

Simay Ayça Çolak

September 2021

DATA-DRIVEN TWO-STAGE INVENTORY PROBLEM
By Simay Ayça Çolak
September 2021

We certify that we have read this thesis and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Nesim K. Erkip(Advisor)

Alper Şen

Cem İyigün

Approved for the Graduate School of Engineering and Science:

Ezhan Karaşan
Director of the Graduate School

ABSTRACT

DATA-DRIVEN TWO-STAGE INVENTORY PROBLEM

Simay Ayça Çolak
M.S. in Industrial Engineering
Advisor: Nesim K. Erkip
September 2021

In this thesis, we consider two-stage newsvendor problem where the decision maker selling a seasonal product only uses the historical demand information in her decisions. In our setting, there are two decisions to be made: the order quantity, and a marked-down price. We decide on how many products to order for the first stage, as well as how to set a marked-down price for remaining unsold inventory in the second stage. To solve the problem considered, data-driven models which do not require any distributional assumption are provided. Specifically, we propose six data-driven methods that solve the problem hierarchically in addition to another method which finds the order quantity and the marked-down price for the remaining inventory simultaneously by using a mixed integer linear program. We generate the data from selected demand distributions and divide it into a training data and a testing data. The generated data is a function of the way that decisions were made historically. We make a definition of the relevancy level based on what decisions the data depends on. We conduct a numerical study to evaluate: (a) the effect of data relevancy, (b) the effect of training data size, (c) the performance of proposed models. We investigate the performances of proposed models in three ways: (1) comparison the best model with the worst one, (2) comparison with respective expected values, (3) comparison with respective the inverse coefficient of variation. Lastly, we measure how many times one model is the best among testing samples and compare models based on their performances.

Keywords: newsvendor, discount pricing, data-driven models, sample average approximation, mixed-integer linear programming formulation, regression analysis.

ÖZET

VERİYE DAYALI İKİ AŞAMALI ENVANTER PROBLEMİ

Simay Ayça Çolak
Endüstri Mühendisliği, Yüksek Lisans
Tez Danışmanı: Nesim K. Erkip
September 2021

Bu tezde, sezonluk bir ürün satan bir karar vericinin kararlarında yalnızca geçmiş talep bilgilerini kullandığı iki aşamalı gazeteci çocuk problemi ele alınmıştır. Bu modelde, verilmesi gereken iki karar vardır: sipariş miktarı ve indirimli fiyat. İlk aşamada ne kadar ürün sipariş edileceğine ve ikinci aşamada kalan satılmayan envanter için nasıl bir indirimli fiyat belirleneceğine karar verilir. Ele alınan problemi çözmek için, herhangi bir dağılım varsayımına ihtiyaç duymayan veriye dayalı modeller sunulmuştur. Spesifik olarak, sorunu hiyerarşik olarak çözen veriye dayalı altı tane yöntem ve karışık tamsayı lineer programlama modeli kullanarak sipariş miktarını ve kalan envanter için indirimli fiyatı aynı anda bulan başka bir yöntem önerilmiştir. Veriler seçilen talep dağılımlarından oluşturulmuştur. Bu veriler eğitim verisi ve test verisi olmak üzere ikiye bölünmüştür. Üretilen veriler, geçmişte alınan kararların bir fonksiyonudur. Verilerin hangi kararlara bağlı olduğuna göre alaka düzeyi tanımlanmıştır. Veri alaka düzeyinin etkisini, eğitim verisi boyutunun etkisini, ve önerilen modellerin performanslarını değerlendirmek için sayısal bir çalışma yürütülmüştür. Önerilen modellerin performansları üç farklı karşılaştırma ile değerlendirilmiştir: (1) en iyi modeli en kötü olanla karşılaştırma, (2) beklenen değerlerle karşılaştırma, (3) ters varyasyon katsayısı ile karşılaştırma. Son olarak, modelleri performanslarına göre karşılaştırmak için bir modelin test örneklerinde kaç kez en iyi olduğu da dikkate alınmıştır.

Anahtar sözcükler: gazeteci çocuk, veriye dayalı modeller, indirimli fiyatlandırma, örnek ortalama yaklaşımı, karışık tamsayı lineer programlama formülasyonu, regresyon analizi.

Acknowledgement

First and foremost, I would like to express my deepest appreciation to my advisor, Prof. Nesim K. Erkip. The completion of this thesis could not have been possible without him pushing me to do my best. From beginning to end, our enjoyable discussions of our study not only taught me a lot, but also broadened my perspective. He always inspired me to become a hard working, open-minded and creative researcher. Words cannot express the gratitude I feel for unfailing encouragement, everlasting patience and invaluable guidance he provided throughout my study. I am truly honored to have worked under his supervision and I consider myself lucky to have the opportunity to benefit from his wisdom and immense expertise.

I am also grateful to Prof. Alper Şen and Prof. Cem İyigün for accepting to read and review this thesis. Their comments and suggestions have been invaluable.

From the bottom of my heart, I would like to thank my mother Hatice Çolak, my father Hakan Çolak, and my sister Aylin Çolak for their endless love, lifetime support and everlasting belief in me. They are always there for me. I am forever indebted to them who enabled me to achieve all my accomplishments with their timely encouragement and unconditional trust. I would like to dedicate this thesis to my precious family.

I am grateful to my friends Alkım Başar and Gözde Taşoluk for their unique motivation and moral support. I am also thankful to Özlem Mahmutoğulları, Gül Çulhan, Pelin Keşrit, Deniz Emre, İrem Gürsesli for their great friendship.

I would like to thank Ahmet Özban, Elif Koşok and her family who have given me valuable support whenever I need since the very first day I moved to Ankara.

Last but not the least, I am also very grateful to all my professors at Bilkent University Department of Industrial Engineering. I feel extremely lucky to have a chance to learn from their wisdom. I also wish to thank Asst. Prof. Özlem Karsu and Prof. Bahar Yetiş Kara for their unwavering support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The newsvendor problem, which Porteus provides a summary of in [1], is a well known single period inventory model where the decision maker optimizes the order quantity of a product having a short selling season such as fashion or seasonal product. Historically, the newsvendor problem was first formulated by Whitin [2] in 1955 and it has long been a key component of studies in operations management. As it forms the basis in many supply chain models, it is applied successfully in a wide variety of business areas. Determining stock levels of perishable products, choosing spare parts at the end of production process of a product, staff sizing in a service business, airline overbooking of flights, capacity management are some applications of the newsvendor problem. See [3] for a historical background and [4] for more application areas. In the framework of all these problems, the decision of order quantity is made before (or without) knowing what the random demand will be. Furthermore, the optimal solution to this simple problem is selected in a way that maximizes the expected profit in the presence of uncertain demand.

Conventionally, the newsvendor problem is often solved analytically under the assumption that the demand distribution is known [5]. However, placing an accurate order is a challenge for the retailers selling the short life cycle products since the demand for such products is stochastic in nature. As more and more data becomes available, data-driven methods are applied to inventory management and thereby the retailers have benefited from it to make better ordering decisions and increase their profits. Data-driven approaches are approaches where the decisions are made based on historical data and without demand distribution assumption (see the agenda by Nguyen et al. [6] for more about the areas of supply chain management that data analytic is used). In data-driven approaches, feature data can be also available. Features about the data are exogenous variables which are predictors of the demand.

When the demand distribution is ambiguous, Saghafian and Tomlin propose an approach which incorporates the available data with partial distributional information to update an estimation made about possible demand distributions dynamically in [7]. A data-driven approach, however, can be used when the decision maker has the only historical data of a fashion product and some features about it. Moreover, the retailers sometimes do not have access to the historical data if a new product will be sold. In such case, sales record of the similar products sold before may help them to obtain accurate forecasts for the new product. To illustrate, Gallien et al. develop a forecasting module which uses historical sales of the products similar with the new product to estimate a demand distribution in [8]. They consider two-period problem where they update the initial allocation and withdrawal decisions in observation period by using a data-driven approach.

To sum up, the traditional newsvendor model has evolved to include the concept of the data-driven modelling in addition to the several contexts that it is useful in. There are a number of papers that study the data-driven newsvendor model: e.g., Ban and Rudin [9], Bertsimas and Kallus [10], Beutel and Minner [11], Levi et al. [12], Huber et al. [13]. These works consider either the single period or a repeated newsvendor model. Furthermore, in the setting of those papers the only decision variable to be determined is the order quantity.

Considering the lack of two period retail model in this line of work, the specific problem that we are interested in is a data-driven newsvendor problem with recourse action. Two-stage data-driven newsvendor problem and data-driven newsvendor problem with recourse action can be used interchangeably. Motivated by this retail inventory management problem, data-driven approaches that do not need any assumption of specific demand distribution are presented in this thesis. We consider a single fashion product to be sold through two stages (see Figure 1.1). In our setting, there are two decisions to be made: the order quantity, and a marked-down price. We decide on how many products to order for the season, as well as how to set a marked-down price for remaining unsold inventory as the second period problem. At the end of the second stage, unsold products are salvaged. For both stages, the only accessible information is the historical demand data although the optimal decision in all of the papers mentioned above is predicted from past demand observations and feature data such as weather and seasonality. Our problem setting is similar to the framework mentioned in [8], but it is different in terms of decision variables and methodology. Gallien et al. [8] study a two-period problem close to ours because it involves the incorporation of forecast updating and an optimization model capturing demand learning. In contrast, we achieve the order quantity and the marked-down price decisions from data-driven approaches without estimating the demand distribution.

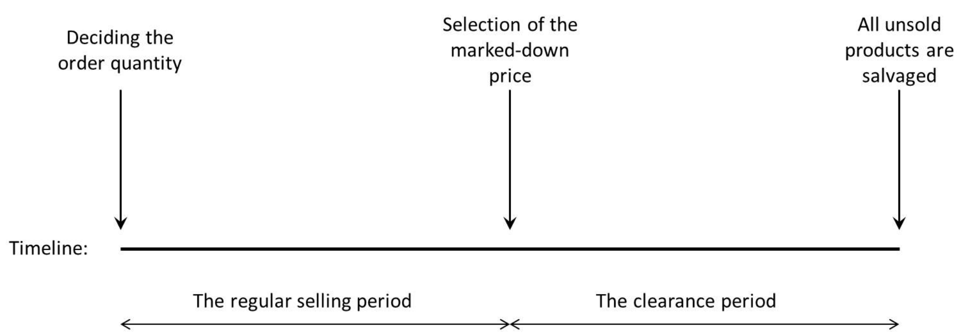

Figure 1.1: Timeline of two-stage inventory problem

For the sake of simplicity, it is convenient to restrict ourselves to a set of marked-down prices consisting of one high and one low price since the marked-down price should be between the selling price of the first-stage and the salvage value at the end of the second-stage. In contrast to the traditional newsvendor setting, we decide on the new price instead of assuming a standard salvage value. As the salvage value of leftovers is quite low, selling unsold seasonal product at a marked-down price at the end of a selling season makes sense. The inventory problem that we investigate can in fact be viewed as an extension of the classical newsvendor problem to the framework that includes quantity order and marked-down price decisions which are made by data-driven approaches. The salvage (or clearance) price is also considered as a decision variable in [14] and [15].

There are other papers related to the problem that we considered. The paper by Aviv et al. [16] is also relevant to our work because of its pricing problem at the second period. Similarly, they also consider a retailer selling a product during two periods. The seller's pricing problem for the second period in [16] is written by using first-period sales and remaining quantity as well as the consumers' policy parameter. However, there are some parts of this article that are not incoherent with our work. First, they consider the strategic customers for the newsvendor problem. We do not separate customers as myopic and strategic. Also, they have an assumption that customers who are willing to buy are Poisson distributed with a mean. In our problem, we do not make any assumption about the true demand distribution. Aflaki et al. [17] also study a problem where a fashion product is sold at a high price in the first period and it is offered at a low price in the second period. Their problem, however, is quite different than ours as they analyze the customer behavior and do not apply any data-driven approach. Furthermore, Herrero et al. [18] consider a newsvendor problem with two decision variables, which are price and stock quantity. Although this work may be relevant to ours due to consideration of two decision variables for a single-period newsvendor problem. These papers are studies which assume a demand distribution. Unlike these studies, we focus on an newsvendor problem with recourse action where the historical demand information is used and there is no assumption about demand distribution.

1.1 Outline of The Approaches

Under the assumption that the only accessible information is the historical data, we solve the problem of how to deal with both two periods of a fashionable product by using a data-driven approach. There are six methods that we proposed to solve the data-driven two-stage newsvendor problem hierarchically. The second stage decision variable, which is the marked down price, is first found given the leftover at the end of the first stage. Based on the solution of the second period problem, the first stage decision variable, which is the order quantity, is obtained. To do so, three data-driven approaches are presented for the second period problem: We applied an algorithm based on sample average approximation (SAA) to estimate expected profit as well as the marked-down price for the leftovers. As an alternative to our algorithm based on SAA, linear regression model and quadratic regression model are used to solve the second period problem. Two data-driven approaches are developed to solve the first period problem: We provide an algorithm based on sample average approximation and a mixed integer linear program that works with the historical demand data. For each approach which solves the first period problem, the result of any second period method can be used. By combining these methods, we have six approaches to decide on two decision variables. In addition to that, we also propose a mixed integer linear program which finds the order quantity and the marked-down price for the remaining inventory simultaneously. Our approaches do not require any assumption about the demand distribution. Instead, the historical data is used as an input. The data is a function of the way that decision were made historically. We make a definition of the relevancy level based on what decisions the data depends on. Three different relevancy level is considered and the effect of the relevancy level of data is examined numerically.

In addition to, we conduct a numerical study to evaluate the performance of proposed models and compare them with each other as well as the effect of the training data size. To do so, we generate random data set from selected demand distributions and divide it into a training set and test set (see Figure 1.2). The training data set (or in-sample data) is used for training the models and estimating the parameters of the models. The test data set (or out-of sample data) is used to validate the trained models. After training the models, the estimated parameters which are the order quantity and the marked-down price decisions, are used with the observations in the test data to predict the objective function of data-driven models. According to the average profit value obtained from the testing data, a discussion is made whether a data-driven model performs well.

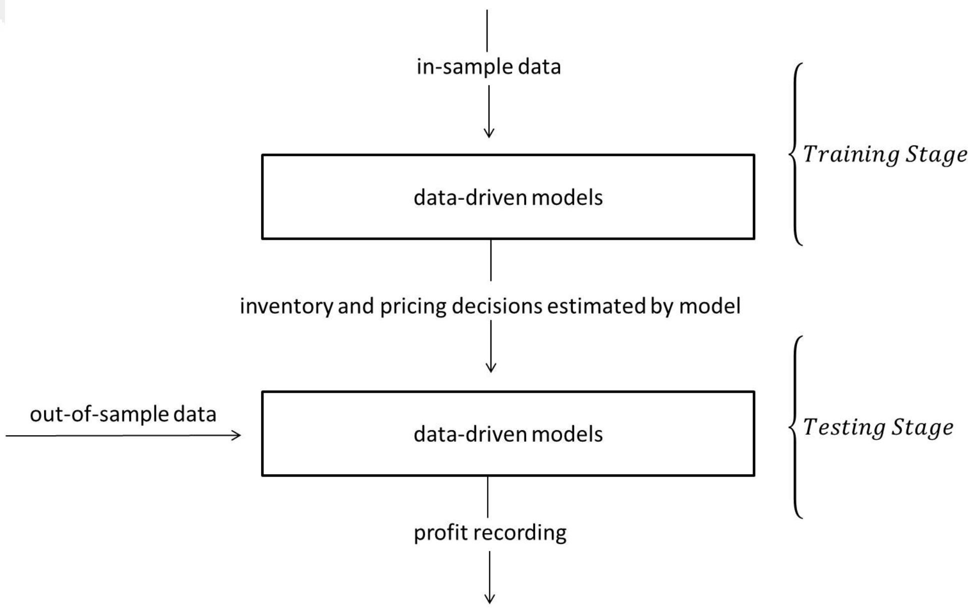

Figure 1.2: Implementation of data-driven models

As a result of this study, we find the optimal order quantity and also show that there exist a threshold value which describes how to choose the marked-down price for leftovers: The decision maker introduces the product at a high price as long as the remaining inventory is less than the threshold value; otherwise, the product is offered at a low price through the second period, i.e. clearance period. Hence, it can be said that the decision variables are the order quantity and the threshold that is used in marked-down price decision. In addition to, we derive a number of measurements to show the importance of solving the second period problem. We also observe that there are sets of parameters where the value of the second period problem is highlighted.

Our Contribution

This thesis makes three main contributions. First, we propose data-driven models for the two-period newsvendor problem defined. To the best of our knowledge, this work is the first to study two-stage newsvendor model within a non-parametric data-driven framework consisting of the initial order quantity and the marked-down price decisions. As a result, the challenge of deciding on the amount to order and the marked-down price for a fashion product with unknown demand can be effectively addressed by properly utilizing the historical demand data available from previous selling seasons. Second, we compare the performances of data-driven models. In addition to that, we numerically address the importance of size of the training data as well as the relevancy level of data.

The remainder of this thesis is organized as follows. An overview of the literature relevant to our problem is provided in the next Section 1.2. We present a comparison of literature and contributions of this work. Section 2 contains the problem description and some analytical results about the problem assuming known demand distributions. We give a description of the proposed data-driven approaches in Section 3. In Section 4, data generation process and numerical experiments are introduced. Section 4 includes performance measures defined to evaluate the potential benefit of the second period problem. The comparison of proposed methods is given as well. This thesis is concluded with our findings and directions of further research in Section 5.

1.2 Literature Review

As stated before, numerous extensions of the classical newsvendor problem have been proposed based on the basic formulation throughout the years. Within this literature, we mention its data-driven extension that are particularly relevant to our setting.

With the growing availability of data, the new data-driven approaches for the inventory models have been developed by researchers in the recent years. Existing approaches for the data-driven newsvendor problem can be classified into two categories which are parametric and non-parametric approaches. In the parametric approaches some distributional assumptions are required whereas in the setting of the non-parametric newsvendor problem the decision maker does not know the distribution family to which the true demand belongs . Instead, it is assumed that the demand data is a random sample drawn from some distribution with unknown cumulative distribution function [12]. The most well known parametric approach is the Bayesian approach which is applied by Lariviere and Porteus [19] to the newsvendor problem. The assumption is made that the distribution of demand is known in the Bayesian approach (e.g., Scarf et al. [20]).

Furthermore, one part of the non-parametric data-driven models is the separated estimation and optimization approaches. It is a two-step procedure consisting of estimating the demand distribution and optimizing the optimal order quantity. Using machine learning or time series approaches, the aim is to find a demand model which fits the past data well in the estimation step. Time series approaches are more traditional than the machine learning approaches such as neural networks and support vector machines. In the recent papers by Barrow et al. [21] and Crone et al. [22], it is stated that machine learning techniques are applicable with large data sets for forecasting tasks. Depending on the demand forecast, the purpose of the optimization step is to make an appropriate inventory decision to maximize profit.

Moreover, the other component of non-parametric approaches is the approaches that integrating the demand estimation and optimization into a single model. It is called data-driven (or distribution-free) approach as it is not based on any theoretical demand distributions (see Bertsimas and Thiele [23]). Liyanage and Shanthikumar [24] were the first to introduce an approach called operational statistics, which integrates demand estimation and optimization. Another approach in the literature to solve the data-driven newsvendor problem is sample average approximation (SAA) which has a wide range of applications (Kleywegt et al. [25]). The SAA method is a popular useful tool for the two-stage optimization problems that are hard to solve computationally even though the underlying distribution is known [12]. We therefore apply SAA approach to the data-driven two-stage newsvendor problem. In SAA approach, true distribution is replaced by the empirical data and the profit averaged over the empirical distribution is maximized.

Recently, Beutel and Minner [11], Ban and Rudin [9], Huber et al. [13] use the data-driven approach to solve the single product newsvendor problem. Müller et al. [26] is the first to use this data-driven approach to multi product newsvendor problem with uncertain demand by using artificial neural network technique. Also, Huber et al. [13] provide data-driven approaches integrating alternative machine learning methods such as artificial neural networks and decisiton trees into an optimization model. In the recent paper by Oroojlooyjadid et al. [27], artificial neural networks is also used to propose an algorithm in the same concept and the multi-feature newsvendor problem is considered. The traditional model to decision making under uncertainty is maximizing the worst case profit, called the maximin approach. It is applied by Scarf [28] and Gallego et. al. [29] to drive the maximin optimal order quantity for the newsvendor problem by assuming that mean and variance are given. Furthermore, Perakis and Roels [30] use the minimax regret approach that works with partial information about the demand distribution and determine the inventory levels that minimize the newsvendor's maximum regret of not acting optimally. Moreover, Lau and Lau [31] propose a non-parametric model based on the product limit method which is a form of the Kaplan-Meier estimator. This method also does not require any prior distributional assumptions. Huh et al. [32] solve the data-driven inventory model by using the well-known Kaplan-Meier estimator which is another non-parametric method.

Bertsimas and Thiele [33], [34] find the solution to the inventory planning problem by proposing a linear program working directly with demand data. Beutel and Minner [11] extend this data-driven approach which does not require any distributional assumptions and directly determine safety stocks from historical demand data with several explanatory variables. Sampling-based method is another approach applied by Levi et al. [35] to determine the optimal ordering quantities. Levi et al. [12] improve the bounds without features provided by Levi et al. [35] for the single-period. Like Beutel and Minner [11], Ban and Rudin [9] also assume that the order quantity is the linear function of the explanatory features and propose algorithms based on empirical risk minimization principle and Kernel-weights optimization. They drive the performance bounds which are the generalizations of the bounds of Levi et al. [12] in the data-driven inventory decisions.

In a departure from the existing data-driven newsvendor frameworks, we consider a two-stage model that not only optimizes the order quantity, but also determines the marked-down price, which is of important strategy for the retailers. Other major aspect that departs our approaches from earlier frameworks is that we find a way to implement a data-driven approach to the two-stage newsvendor problem. By introducing a new concept, we provide six different data-driven approaches which apart from ordering decision also incorporate the marked-down price decision. The papers by Bertsimas and Thiele [33], Beutel and Minner [11], Sachs and Minner [36] are the related studies to a part of our work since they use a linear programming formulation to solve the data-driven newsvendor problem. The optimal order quantity is found as a linear function of exogenous variables whose coefficients are determined by a linear programming model. In the paper by Sachs and Minner [36], the approach developed in [11] is extended to study the problem with censored demand observations. In contrast to these works, we provide a mixed integer programming that works with only the historical demand data for each period to find the order quantity and the marked-down price. Although we do not use feature data, our approaches can be extended such that feature data is also considered to estimate the decision variables. To conclude, our approach is distinguished from these in several ways, including in having two decision variables and in the consideration of two stages as well as in the implementation of data-driven approaches.

Chapter 2

Problem Definition

In this section we consider the newsvendor problem with recourse action under the assumption that the demand distributions are known. As the salvage value of leftovers is quite low, offering unsold products at a marked-down price makes sense to clear them at the end of a selling season. Therefore, we extend the classical newsvendor problem in a way that the retailer has two periods to sell the product and not only decides the order quantity of the product, but also decides the new marked down price of the product for the second period. The marked-down price will be selected from a set containing a high price and a low price. Although a generalization for more than two possible marked down prices might be made, we consider two possible marked-down prices throughout this thesis. As a solution to the second period of the problem considered, we find that there exist a threshold which describes the range of the selection for both marked-down prices. Therefore, we find the order quantity as well as the threshold value to solve the problem considered.

The remainder of this chapter is structured as follows. We first define the formulation of the two-stage newsvendor problem in Section 2.1. Section 2.2 provides the solution of the second period problem and some assumptions about the problem. Before solving the problem with proposed data-driven approaches in Section 3, we start our study by analyzing the problem under the assumption that the demand distributions of the periods are known. Under full knowledge of the demand distribution, we developed the total profit function of the newsvendor problem with recourse action by using the second period solution in Section 2.3. Also, we obtain an insight into the properties of the expected total profit function by analysing the second derivation. Finally, the way of how to find the exact solutions to the newsvendor problem with recourse action is given. The distributions that we consider in this section are Poisson, Triangular and Uniform distributions.

2.1 The Newsvendor Problem with Recourse Action

In the classical version of the newsvendor problem, one sells a single fashion product in a short season. In the presence of uncertain demand, the aim of the vendor is to order a quantity of the product, Q, that maximizes the expected value of the total profit. The decision of how much to order is made once and prior to the beginning of the selling season. Each unit sold during the period produces a revenue of p. Considering the situation faced by the decision maker, overage occurs if the amount ordered exceeds demand rate which is denoted by ξ. Unsold products have a salvage value q at the end of the selling season, and a per unit purchasing cost of v is assessed on each unit ordered. On the other hand, there is an unsatisfied demand which causes lost sales if the order quantity is under stocked. Therefore, an underage or shortage cost, B, is charged for each demand that cannot be met. This is also called stock-out or penalty cost.

Considering the economic consequences of ordering too much or too little, the vendor has to decide the order quantity that maximizes the expected profit and hence minimizes overstocking and understocking costs.

The profit of order Q and demand ξ is as follows.

$$\Pi(Q,\xi) = \begin{cases} -Qv + p\xi + q(Q-\xi), & \text{if } Q \geq \xi \\ -Qv + pQ - B(\xi - Q), & \text{if } Q \leq \xi \end{cases} \tag{2.1}$$

It is known that optimal solution Q^* satisfies $F(Q^*) = \frac{p-v+B}{p-q+B}$ unless the demand distribution is unknown. (See [1] and [5] for more about classical newsvendor.)

In the two-stage newsvendor problem, there are two periods in which the demand indicated by ξ_1 occurs in the first period and the demand indicated by ξ_2 occurs in the second period respectively. We assume, without loss of generality, that lengths of each period are fixed and demand for the second period is price-dependent and stochastic. Let Q be the amount of order quantity to be purchased and I be the remaining inventory from the first period or equivalently initial inventory level of the second period, i.e. $I = \max\{0, (Q - \xi_1)\}$. The retailer decides on how many products to order for the regular season without knowing the actual demand of both periods $(\xi_1, \xi_2(p_2))$. Subsequently, the retailer makes a one shot decision of how to set a marked-down price, p_2, for remaining unsold inventory as the second period problem. Thus, there is one more decision variable that account for marked-down price in our setting. Having decided on the marked-down price whose expected profit is called the salvage value, the aim is to determine the order quantity that maximizes the total expected profit from period 1 and 2.

Recall that there is a cost, v, associated with every purchased unit and a penalty, B, for every unmet demand in the first period. The seller offers the fashion product for a fixed price $p_1 > v$ during the first period and has the opportunity to offer its remaining inventory for a marked-down price $p_2 < p_1$ in the second period. There are two possible marked-down prices considered in the second period problem: a high price p_2^H and a low price p_2^L. If, at the end of the second period, there is an inventory left, they can be salvaged at a price of $q < v$ per product. The demand ξ_1 for the first period is a nonnegative random variable with density function f_{ξ_1} and cumulative distribution function F_{ξ_1}. Given p_2 for the second period, $f_{\xi_2|p_2}(.)$ and $F_{\xi_2|p_2}(.)$ denote the density function and cumulative distribution function of the demand ξ_2 respectively.

The uncertain demand of the first period as well as the expected profit function in the second period are taken into account for the decision-making process of the quantity order. The two-stage newsvendor problem is as follows.

$$\max_{Q \geq 0, p_2 \in \{p_2^H, p_2^L\}} \mathbb{E}[\Pi(Q)] \tag{2.2}$$

The objective is to maximize the expected profit of the first-period decision plus the expected profit of the second-period recourse decision. As mentioned before, the expected profit Π^2 earned from the second period can be considered as a salvage value in the regular season and hence Π can be written as given below [1].

$$\Pi(Q, \xi_1) = \begin{cases} -Qv + p_1\xi_1 + \mathbb{E}[\Pi^2(Q - \xi_1)], & \text{if } Q \geq \xi_1 \\ -Qv + p_1 Q - B(\xi_1 - Q), & \text{if } Q \leq \xi_1 \end{cases} \tag{2.3}$$

and for a continuous demand ξ_1, the expected total profit function is

$$\mathbb{E}[\Pi(Q, \xi_1)] = \int_0^Q \Big(- Qv + p_1 x + \mathbb{E}[\Pi^2(Q - x, \xi_2)|p_2] \Big) f_{\xi_1}(x)dx \\ + \int_Q^\infty \Big(- Qv + p_1 Q - B(x - Q) \Big) f_{\xi_1}(x)dx \tag{2.4}$$

The discrete non negative demand can also be assumed to express the above functions. For discrete case, the equation in (2.4) becomes

[1] An alternative expression is $\Pi(Q, \xi_1) = -Qv + p_1 \min(\xi_1, Q) - B\max(\xi_1 - Q, 0) + \mathbb{E}[\Pi^2(\max(Q - \xi_1, 0))]$

$$
\mathbb{E}[\Pi(Q,\xi_1)] = \sum_{k=0}^{Q}\Big(-Qv + p_1 k + \mathbb{E}[\Pi^2(Q-k,\xi_2)|p_2]\Big) f_{\xi_1}(k)
$$
$$
+ \sum_{k=Q+1}^{\infty}\Big(-Qv + p_1 Q - B(k-Q)\Big) f_{\xi_1}(k)
$$

The expected second period profit function in (2.3) comprises of two expressions. The first expression is for the case when the leftover is more than the demand. It indicates that the demand of the customer is met and the inventory in excess of demand is disposed of at salvage value q. In the second expression, demand exceeds the leftover inventory, and hence all the leftover items are sold. Thus,

$$
\Pi^2(I,\xi_2(p_2)) = \begin{cases} p_2\xi_2(p_2) + q(I-\xi_2(p_2)), & \text{if } I \geq \xi_2(p_2) \\ p_2 I, & \text{if } I \leq \xi_2(p_2) \end{cases} \tag{2.5}
$$

is the random second period profit of the marked-down price p_2, inventory level I and demand $\xi_2(p_2)$.

Since the demand has not been realized at the beginning of the second period, the seller cannot observe the actual profit. Also, the second period profit depends on the order quantity and the random demand. The decision maker, therefore, has to determine the marked-down price that maximizes the expected profit and then use the expected profit of the second period as a salvage value for the first period. To do so, it is necessary to solve the following second period problem:

$$
\max_{p^L < p^H} \mathbb{E}[\Pi^2(I)] \tag{2.6}
$$

where

$$\mathbb{E}[\Pi^2(I,\xi_2)|p_2^i] = \int_0^I \Big(p_2^i y + q(I-y)\Big) f_{\xi_2|p_2^i}(y)dy$$
$$+ \int_I^\infty p_2^i I f_{\xi_2|p_2^i}(y)dy \quad \forall i = H, L \tag{2.7}$$

The function $\mathbb{E}[\Pi^2(I_1,\xi_2)|p_2]$ can also be expressed for a discrete demand. If the demand is discrete, the equation in (2.7) is written as follows.

$$\mathbb{E}[\Pi^2(I,\xi_2)|p_2^i] = \sum_{k=0}^{I} \Big(p_2^i k + q(I-k)\Big) f_{\xi_2|p_2^i}(k)$$
$$+ \sum_{k=I+1}^{\infty} p_2^i I f_{\xi_2|p_2^i}(k) \quad \forall i = H, L$$

2.2 The Expected Profit Function of The Second Stage

In this subsection, we give analytical results for the second period problem. We assumed that demand for the second period is price-dependent ($\xi_2(p_2)$). In other words, the demand rate is assumed to be a linearly decreasing function of the marked-down selling price. Hence, the demand that will occur when the low price is selected is expected to be higher than the demand that will occur when the high price is selected. Second of all, the expected profit of the high price and the low price are calculated separately in order to decide on the new price of the remaining inventory. The recourse decision is made between these two prices, depending on which one gives larger expected profit. Equivalently, the marked-down price can be decided for each remaining inventory level by looking at the difference between the expected profit of the high price and the low price. In cases where the difference is non negative, the high price is selected. Otherwise, the low price is selected. Lastly, it can always be negative which means that the low price is selected for all leftovers. On the other hand, it can always be positive which means that the high price will always be chosen.

When the retailer has I inventory level and determines the marked-down price p_2 for that inventory level, the expected profit of the second period for continuous demand is as follows.

$$\mathbb{E}[\Pi^2(I)|p_2^i] = \int_0^I (yp_2^i + q(I-y))f_{\xi_2|p_2^i}(y)dy + \int_I^\infty Ip_2^i f_{\xi_2|p_2^i}(y)dy \quad \forall i = H, L \tag{2.8}$$

$$\begin{aligned}
\mathbb{E}[\Pi^2(I)|p_2^i] &= (p_2^i - q)\int_0^I yf_{\xi_2|p_2^i}(y)dy + qI\int_0^I f_{\xi_2|p_2^i}(y)dy \\
&+ Ip_2^i \int_I^\infty f_{\xi_2|p_2^i}(y)dy \quad \forall i = H, L \\
&= (p_2^i - q)\Big(\underbrace{\int_0^\infty yf_{\xi_2|p_2^i}(y)dy}_{\mu^i} - \int_I^\infty yf_{\xi_2|p_2^i}(y)dy\Big) \\
&+ qI\Big(\underbrace{\int_0^\infty f_{\xi_2|p_2^i}(y)dy}_{1} - \int_I^\infty f_{\xi_2|p_2^i}(y)dy\Big) \\
&+ Ip_2^i \int_I^\infty f_{\xi_2|p_2^i}(y)dy \quad \forall i = H, L \\
&= (p_2^i - q)\mu^i - (p_2^i - q)\int_I^\infty yf_{\xi_2|p_2^i}(y)dy + qI \\
&\underbrace{-qI\int_I^\infty f_{\xi_2|p_2^i}(y)dy + Ip_2^i\int_I^\infty f_{\xi_2|p_2^i}(y)dy}_{(p_2^i - q)I\int_I^\infty f_{\xi_2|p_2^i}(y)dy} \\
&= qI + (p_2^i - q)\Big(\mu^i - \int_I^\infty (y - I)f_{\xi_2|p_2^i}(y)dy\Big) \quad \forall i = H, L
\end{aligned} \tag{2.9}$$

For discrete demand, the function $\mathbb{E}[\Pi^2(I_1)|p_2]$ in (2.9) is expressed as follows.

$$\mathbb{E}[\Pi^2(I)|p_2^i] = qI + (p_2^i - q)\Big(\mu^i - \sum_{k=I}^{\infty}(k - I)f_{\xi_2|p_2^i}(k)\Big) \quad \forall i = H, L$$

An alternative expression for the expected second period profit function is shown in Appendix A.

When the recourse decision is made between two prices, there exists a threshold, I^t, under the following assumptions:

1. $p_2^L < p_2^H$
2. $\mathbb{P}\{\xi_2 \leq I^t | p_2^L\} \leq \mathbb{P}\{\xi_2 \leq I^t | p_2^H\}$
3. $\mathbb{E}[\xi_2 | p_2^H] \leq \mathbb{E}[\xi_2 | p_2^L]$

and threshold is neither zero nor infinite (i.e.,$I^t \notin \{0, \infty\}$) if the following conditions are satisfied:

$$\mathbb{P}\{\xi_2 \geq 1 | p_2^H\} p_2^H + \mathbb{P}\{\xi_2 = 0 | p_2^H\} q \geq \mathbb{P}\{\xi_2 \geq 1 | p_2^L\} p_2^L + \mathbb{P}\{\xi_2 = 0 | p_2^L\} q \tag{2.10}$$

$$p_2^H \mathbb{E}[\xi_2 d | p_2^H] + q \Bigg[\mathbb{E}[\xi_2 | p_2^L] - \mathbb{E}[\xi_2 | p_2^H] \Bigg] \leq p_2^L \mathbb{E}[\xi_2 | p_2^L] \tag{2.11}$$

The threshold value I^t is defined as the point where the expected profit difference is zero for continuous demand. Hence,

$$\begin{aligned}
&\mathbb{E}[\Pi^2(I^t) | p_2^H] - \mathbb{E}[\Pi^2(I^t) | p_2^L] = 0 \\
&\left[q I^t + (p_2^H - q) \left(\mu^H - \int_{I^t}^{\infty} (y - I^t) f_{\xi_2 | p_2^H}(y) dy \right) \right] \\
-&\left[q I^t + (p_2^L - q) \left(\mu^L - \int_{I^t}^{\infty} (y - I^t) f_{\xi_2 | p_2^L}(y) dy \right) \right] = 0
\end{aligned} \tag{2.12}$$

For discrete demand, it is unlikely that any values of the expected profit difference exactly equal to zero as the distribution function increases by jumps. Therefore, the threshold value I^t is defined as the point where the difference, whose value is non-negative at the beginning, becomes negative. Hence, threshold value is the smallest inventory level that satisfies the following condition.

$$\begin{aligned}
&\mathbb{E}[\Pi^2(I^t)|p_2^H] - \mathbb{E}[\Pi^2(I^t)|p_2^L] < 0 \\
&\left[qI^t + (p_2^H - q)\left(\mu^H - \sum_{I^t}^{\infty}(y - I^t)f_{\xi_2|p_2^H}(y)dy\right)\right] \\
-&\left[qI^t + (p_2^L - q)\left(\mu^L - \sum_{I^t}^{\infty}(y - I^t)f_{\xi_2|p_2^L}(y)dy\right)\right] < 0
\end{aligned} \tag{2.13}$$

If the difference is greater than or equal to zero, the highest price p_2^H is chosen. If the difference is less than zero, then the lowest price p_2^L is selected. Hence, we can write our decision rule for choosing marked down price is as follows.

- The high price p_2^H is determined if $I < I^t$.
- The low price p_2^L is determined as the new price if $I \geq I^t$,

Given I^t , the solution of the second period is shown below.

$$\mathbb{E}[\Pi^2(I)] = \begin{cases} \mathbb{E}[\Pi^2(I)|p_2^H], & \text{if } I < I^t \\ \mathbb{E}[\Pi^2(I)|p_2^L], & \text{if } I \geq I^t \end{cases} \tag{2.14}$$

Extreme Cases:

When $I \to \infty$, the difference approaches

$$\lim_{I\to\infty} \mathbb{E}[\Pi^2(I)|p_2^H] - \mathbb{E}[\Pi^2(I)|p_2^L] = (p_2^H - q)\mathbb{E}[\xi_2|p_2^H] - (p_2^L - q)\mathbb{E}[\xi_2|p_2^L]$$

If the condition

$$p_2^H\mathbb{E}[\xi_2|p_2^H] + q\left[\mathbb{E}[\xi_2|p_2^L] - \mathbb{E}[\xi_2|p_2^H]\right] \leq p_2^L\mathbb{E}[\xi_2|p_2^L]$$

does not hold, then the solution of the second period can be shown as follows.

$$\mathbb{E}[\Pi^2(I)] = \mathbb{E}[\Pi^2(I)|p_2^H] \quad \forall I$$

If the condition

$$\mathbb{P}\{\xi_2 \geq 1|p_2^H\}p_2^H + \mathbb{P}\{\xi_2 = 0|p_2^H\}q \geq \mathbb{P}\{\xi_2 \geq 1|p_2^L\}p_2^L + \mathbb{P}\{\xi_2 = 0|p_2^L\}q$$

does not hold, then the solution of the second period can be shown as follows.

$$\mathbb{E}[\Pi^2(I)] = \mathbb{E}[\Pi^2(I)|p_2^L] \quad \forall I$$

We solve the second period problem when the demand is described by Poisson distribution. In addition to that, the second period problem is also solved when demand is Uniform and Triangular distributed, respectively.

The Second Period Problem for Poisson Demand

In our study, Poisson distribution is one of the choices for modeling demand uncertainty. We show how to find optimal order quantity and the marked down price when the demand is assumed to be Poisson which is a discrete probability distribution. Recall that the marked-down price is determined from the discrete set consisting p_2^H, and p_2^L. Thus, it is assumed that $p_2^H \leq p_2^L$ and demand is decreasing in p_2. The demand for p_2^H follows Poisson distribution with rate λ_H whereas the demand for p_2^L is assumed to be a Poisson distributed random variable with rate λ_L. The probability density function of Poisson distribution is given in Appendix F.2.

The expected profit of the second period when the retailer has I inventory level is written as below.

$$\mathbb{E}[\Pi^2(I)|p_2^i] = \sum_{k=0}^{I} \big[p_2^i k \frac{e^{-\lambda_j}\lambda_j^k}{k!} + q(I-k)\frac{e^{-\lambda_j}\lambda_j^k}{k!}\big] + \sum_{k=I+1}^{\infty} p_2^i I \frac{e^{-\lambda_j}\lambda_j^k}{k!} \quad \forall i,j = H,L$$

$$= p_2^i \sum_{k=0}^{I} k\frac{e^{-\lambda_j}\lambda_j^k}{k!} + qI\sum_{k=0}^{I} \frac{e^{-\lambda_j}\lambda_j^k}{k!} - q\sum_{k=0}^{I} k\frac{e^{-\lambda_j}\lambda_j^k}{k!}$$

$$+ p_2^i I\big(1 - \sum_{k=0}^{I} \frac{e^{-\lambda_j}\lambda_j^k}{k!}\big) \quad \forall i,j = H,L$$

$$= p_2^i I + \sum_{k=0}^{I} ((p_2^i - q)(k-I))\frac{e^{-\lambda_j}\lambda_j^k}{k!} \quad \forall i,j = H,L \tag{2.15}$$

When the marked-down price decision is made, the difference

$$\mathbb{E}[\Pi^2(I)|p_2^H] \; - \; \mathbb{E}[\Pi^2(I)|p_2^L]$$

is taken into account. The expected profit difference function is as follows:

$$\mathbb{E}[\Pi(I)|p_2^H] - \mathbb{E}[\Pi(I)|p_2^L] = (p_2^H - q)\sum_{k=0}^{I} k\frac{e^{-\lambda_H}\lambda_H^k}{k!} + qI\sum_{k=0}^{I} \frac{e^{-\lambda_H}\lambda_H^k}{k!}$$

$$+ p_2^H I\big(1 - \sum_{k=0}^{I} \frac{e^{-\lambda_H}\lambda_H^k}{k!}\big) - (p_2^L - q)\sum_{k=0}^{I} k\frac{e^{-\lambda_L}\lambda_L^k}{k!} - qI\sum_{k=0}^{I} \frac{e^{-\lambda_L}\lambda_L^k}{k!}$$

$$- p_2^L I\big(1 - \sum_{k=0}^{I} \frac{e^{-\lambda_L}\lambda_L^k}{k!}\big)$$

$$= p_2^H\big(\sum_{k=0}^{I} k\frac{e^{-\lambda_H}\lambda_H^k}{k!} - I\sum_{k=0}^{I} \frac{e^{-\lambda_H}\lambda_H^k}{k!}\big) - q\big(\sum_{k=0}^{I} k\frac{e^{-\lambda_H}\lambda_H^k}{k!} - \sum_{k=0}^{I} k\frac{e^{-\lambda_L}\lambda_L^k}{k!}\big)$$

$$- p_2^L\big(\sum_{k=0}^{I} k\frac{e^{-\lambda_L}\lambda_L^k}{k!} - I\sum_{k=0}^{I} \frac{e^{-\lambda_L}\lambda_L^k}{k!}\big) + qI\big(\sum_{k=0}^{I} \frac{e^{-\lambda_H}\lambda_H^k}{k!} - \sum_{k=0}^{I} \frac{e^{-\lambda_L}\lambda_L^k}{k!}\big)$$

$$+ I(p_2^H - p_2^L)$$

$$= \sum_{k=0}^{I} \Big\{ p_2^H k\frac{e^{-\lambda_H}\lambda_H^k}{k!} - p_2^H I\frac{e^{-\lambda_H}\lambda_H^k}{k!} - qk\frac{e^{-\lambda_H}\lambda_H^k}{k!} + qk\frac{e^{-\lambda_L}\lambda_L^k}{k!} - p_2^L k\frac{e^{-\lambda_L}\lambda_L^k}{k!}$$

$$+ p_2^L I\frac{e^{-\lambda_L}\lambda_L^k}{k!} + qI\frac{e^{-\lambda_H}\lambda_H^k}{k!} - qI\frac{e^{-\lambda_L}\lambda_L^k}{k!} \Big\} + I(p_2^H - p_2^L)$$

$$
\begin{aligned}
&= \sum_{k=0}^{I} \Big\{ (k-I)p_2^H \frac{e^{-\lambda_H}\lambda_H^k}{k!} - (k-I)p_2^L \frac{e^{-\lambda_L}\lambda_L^k}{k!} \\
&- (qk - qI)\big(\frac{e^{-\lambda_H}\lambda_H^k}{k!} - \frac{e^{-\lambda_L}\lambda_L^k}{k!}\big) \Big\} + I(p_2^H - p_2^L) \\
&= \sum_{k=0}^{I} \Big\{ (k-I)\big(p_2^H \frac{e^{-\lambda_H}\lambda_H^k}{k!} - p_2^L \frac{e^{-\lambda_L}\lambda_L^k}{k!}\big) - q(k-I)\big(\frac{e^{-\lambda_H}\lambda_H^k}{k!} - \frac{e^{-\lambda_L}\lambda_L^k}{k!}\big) \Big\} \\
&+ I(p_2^H - p_2^L) \\
&= \sum_{k=0}^{I} \Big\{ (I-k)\Bigg(q\big(\frac{e^{-\lambda_H}\lambda_H^k}{k!} - \frac{e^{-\lambda_L}\lambda_L^k}{k!}\big) - \big(p_2^H \frac{e^{-\lambda_H}\lambda_H^k}{k!} - p_2^L \frac{e^{-\lambda_L}\lambda_L^k}{k!}\big)\Bigg) \Big\} \\
&+ I(p_2^H - p_2^L)
\end{aligned}
$$

We know that $p_2^L \leq p_2^H$ and $\lambda_H \leq \lambda_L$ as the demand for higher price is less than the demand for the low price of the fashion product. In other words, $p_2^L = \alpha p_2^H$ where $\alpha \in [0,1]$ and $\lambda_H = \beta\lambda_L$ where $\beta \in [0,1]$. Therefore, we can write the following equation:

$$
\begin{aligned}
&\sum_{k=0}^{I} \Big\{ (I-k)\Bigg(q\big(\frac{e^{-\lambda_H}(\lambda_H)^k}{k!} - \frac{e^{-\frac{\lambda_H}{\beta}}(\frac{\lambda_H}{\beta})^k}{k!}\big) - \big(p_2^H \frac{e^{-\lambda_H}(\lambda_H)^k}{k!} - \alpha p_2^H \frac{e^{-\frac{\lambda_H}{\beta}}(\frac{\lambda_H}{\beta})^k}{k!}\big)\Bigg) \Big\} \\
&+ I(p_2^H - \alpha p_2^H) \geq 0
\end{aligned}
$$

The algorithm which can be found in Appendix C.1, is used to solve the second period problem with Poisson distributed demand. The threshold value I^t is found as the smallest inventory level that satisfies the condition defined above.

Up to now, we developed an expression for the expected profit function of the second period based on the assumption that the demand is described by a discrete probability distribution. Since working with continuous demand distribution is also desirable, we find the optimal solution to the second period by assuming continuous demand. Let $f_{\xi_1}(x)$ and $F_{\xi_1}(x)$ denote the probability distribution and cumulative distribution function of ξ_1, respectively. Also, we will assume that the demand $\xi_2(p_2)$ is a non-negative random variable with density function $f_{\xi_2|p_2}(y)$ and cumulative distribution function $F_{\xi_2|p_2}(y)$.

We use Uniform distribution and Triangular distribution to model the demand, respectively.

The Second Period Problem for Uniform Demand

First of all, let's assume that

$$\xi_{2|p_2^H} \sim U(a^H, b^H)$$

and

$$\xi_{2|p_2^L} \sim U(a^L, b^L)$$

where $\mu_2^H = \frac{a^H+b^H}{2} < \mu_2^L = \frac{a^L+b^L}{2}$.

For the Uniform distributed demand, the threshold $I_1^{(t)}$ is found by setting the expected profit difference equation to zero. As shown in Appendix D, the solution to the second period problem is as follows.

If $0 \leq I^t \leq b^H$,

$$\mathbb{E}[\Pi^2(I)] = \begin{cases} -I^2\left(\frac{p_2^H-q}{2b^H}\right) + Ip_2^H, & \text{if } I < I^t \text{and } 0 \leq I \leq b^H \\ -I^2\left(\frac{p_2^L-q}{2b^L}\right) + Ip_2^L, & \text{if } I \geq I^t \text{and } 0 \leq I \leq b^L \\ Iq + (p_2^L - q)\frac{b^L}{2}, & \text{if } I \geq I^t \text{and } b^L < I \end{cases} \tag{2.16}$$

If $b^H \leq I^t \leq b^L$,

$$\mathbb{E}[\Pi^2(I)] = \begin{cases} -I^2\left(\frac{p_2^H - q}{2b^H}\right) + Ip_2^H, & \text{if } I < I^t \text{and } 0 \leq I \leq b^H \\ Iq + (p_2^H - q)\frac{b^H}{2}, & \text{if } I < I^t \text{and } b^H < I \\ -I^2\left(\frac{p_2^L - q}{2b^L}\right) + Ip_2^L, & \text{if } I \geq I_1^{(t)} \text{and } 0 \leq I \leq b^L \\ Iq + (p_2^L - q)\frac{b^L}{2}, & \text{if } I \geq I^t \text{and } b^L < I \end{cases} \tag{2.17}$$

If $b^L \leq I^t$,

$$\mathbb{E}[\Pi^2(I)] = \begin{cases} -I^2\left(\frac{p_2^H - q}{2b^H}\right) + Ip_2^H, & \text{if } I < I^t \text{and } 0 \leq I \leq b^H \\ Iq + (p_2^H - q)\frac{b^H}{2}, & \text{if } I < I^t \text{and } b^H < I \\ Iq + (p_2^L - q)\frac{b^L}{2}, & \text{if } I \geq I^t \text{and } b^L < I \end{cases} \tag{2.18}$$

The Second Period Problem for Triangular Distributed Demand

For the Triangular distributed demand, let's assume that

$$\xi_{2|p_2^H} \sim Trian(a^H, c^H, b^H)$$

and

$$\xi_{2|p_2^L} \sim Trian(a^L, c^L, b^L)$$

where $a^H = a^L = 0 < c^H < b^H < c^L < b^L$. Suppose that $I \geq 0$, and $\mu^H = \frac{a^H + b^H + c^H}{2} < \mu^L = \frac{a^L + b^L + c^H}{2}$. Likewise, the threshold value I^t is found firstly, then the solution is derived as follows (see Appendix E for details and the case where $a^H = a^L = 0 < c^H < c^L < b^H < b^L$).

If $0 \leq I^t \leq c^H$,

$$\mathbb{E}[\Pi^2(I)] = \begin{cases} \left[-\frac{(p_2^H-q)}{3b^Hc^H}\right]I^3 + p_2^H I, & \text{if } I < I^t \text{and } 0 \leq I \leq c^H \\ \left[-\frac{(p_2^L-q)}{3b^Lc^L}\right]I^3 + p_2^L I, & \text{if } I \geq I^t \text{and } 0 \leq I \leq c^L \\ \left[\frac{p_2^L-q}{3(b^L-c^L)b^L}\right]I^3 - \frac{p_2^L-q}{b^L-c^L}I^2, & \text{if } I \geq I^t \text{and } c^L < I \leq b^L \\ qI + (p_2^L - q)\mu^L & \text{if } I \geq I^t \text{and } b^L < I \end{cases} \tag{2.19}$$

If $c^H \leq I^t \leq b^H$,

$$\mathbb{E}[\Pi^2(I)] = \begin{cases} \left[-\frac{(p_2^H-q)}{3b^Hc^H}\right]I^3 + p_2^H I, & \text{if } I < I^t \text{and } 0 \leq I \leq c^H \\ \left[\frac{p_2^H-q}{3(b^H-c^H)b^H}\right]I^3 - \frac{p_2^H-q}{b^H-c^H}I^2, & \text{if } I < I^t \text{and } c^H \leq I \leq b^H \\ \left[-\frac{(p_2^L-q)}{3b^Lc^L}\right]I^3 + p_2^L I, & \text{if } I \geq I^t \text{and } 0 \leq I \leq c^L \\ \left[\frac{p_2^L-q}{3(b^L-c^L)b^L}\right]I^3 - \frac{p_2^L-q}{b^L-c^L}I^2, & \text{if } I \geq I^t \text{and } c^L < I \leq b^L \\ qI + (p_2^L - q)\mu^L & \text{if } I \geq I^t \text{and } b^L < I \end{cases} \tag{2.20}$$

If $b^H \leq I^t \leq c^L$,

$$\mathbb{E}[\Pi^2(I)] = \begin{cases} \left[-\frac{(p_2^H-q)}{3b^Hc^H}\right]I^3 + p_2^H I, & \text{if } I < I^t \text{and } 0 \leq I \leq c^H \\ \left[\frac{p_2^H-q}{3(b^H-c^H)b^H}\right]I^3 - \frac{p_2^H-q}{b^H-c^H}I^2, & \text{if } I < I^t \text{and } c^H \leq I \leq b^H \\ qI + (p_2^H - q)\mu^H & \text{if } I < I^t \text{and } b^H < I \\ \left[-\frac{(p_2^L-q)}{3b^Lc^L}\right]I^3 + p_2^L I, & \text{if } I \geq I^t \text{and } 0 \leq I \leq c^L \\ \left[\frac{p_2^L-q}{3(b^L-c^L)b^L}\right]I^3 - \frac{p_2^L-q}{b^L-c^L}I^2, & \text{if } I \geq I^t \text{and } c^L < I \leq b^L \\ qI + (p_2^L - q)\mu^L & \text{if } I \geq I^t \text{and } b^L < I \end{cases} \tag{2.21}$$

If $c^L \leq I^t \leq b^L$,

$$\mathbb{E}[\Pi^2(I)] = \begin{cases} \left[-\frac{(p_2^H-q)}{3b^Hc^H}\right]I^3 + p_2^H I, & \text{if } I < I^t \text{and } 0 \leq I \leq c^H \\ \left[\frac{p_2^H-q}{3(b^H-c^H)b^H}\right]I^3 - \frac{p_2^H-q}{b^H-c^H}I^2, & \text{if } I < I^t \text{and } c^H \leq I \leq b^H \\ qI + (p_2^H - q)\mu^H & \text{if } I < I^t \text{and } b^H < I \\ \left[\frac{p_2^L-q}{3(b^L-c^L)b^L}\right]I^3 - \frac{p_2^L-q}{b^L-c^L}I^2, & \text{if } I \geq I^t \text{and } c^L < I \leq b^L \\ qI + (p_2^L - q)\mu^L & \text{if } I \geq I^t \text{and } b^L < I \end{cases} \tag{2.22}$$

If $b^L \leq I^t$,

$$\mathbb{E}[\Pi^2(I)] = \begin{cases} \left[-\frac{(p_2^H-q)}{3b^Hc^H}\right]I^3 + p_2^H I, & \text{if } I < I^t \text{and } 0 \leq I \leq c^H \\ \left[\frac{p_2^H-q}{3(b^H-c^H)b^H}\right]I^3 - \frac{p_2^H-q}{b^H-c^H}I^2, & \text{if } I < I^t \text{and } c^H \leq I \leq b^H \\ qI + (p_2^H - q)\mu^H & \text{if } I < I^t \text{and } b^H < I \\ qI + (p_2^L - q)\mu^L & \text{if } I \geq I^t \text{and } b^L < I \end{cases} \tag{2.23}$$

2.3 The Expected Profit Function of The First Stage

By applying this decision rule for the second period, the following total profit function can be rewritten:

$$\Pi(Q, \xi_1) = \begin{cases} -Qv + p_1\xi_1 + \mathbb{E}[\Pi(I)|p_2^H], & \text{if } 0 \leq I < I^t \\ -Qv + p_1\xi_1 + \mathbb{E}[\Pi(I)|p_2^L], & \text{if } I \geq I^t \\ -Qv + p_1 Q - B(\xi_1 - Q), & \text{if } I \leq 0 \end{cases} \tag{2.24}$$

An alternative expression for the equation (2.24) is given in Appendix B. Next, we can write the following equation when the demand is assumed to be discrete

$$
\begin{aligned}
\mathbb{E}[\Pi(Q)] = & \sum_{k=0}^{Q-I^t} \Big(- Qv + p_1 k + \mathbb{E}[\Pi^2(Q-k,\xi_2)|p_2^L] \Big) f_{\xi_1}(k) \\
& + \sum_{k=Q-I^t+1}^{Q} \Big(- Qv + p_1 k + \mathbb{E}[\Pi^2(Q-k,\xi_2)|p_2^H] \Big) f_{\xi_1}(k) \\
& + \sum_{k=Q+1}^{\infty} \Big(- Qv + p_1 Q - B(k-Q) \Big) f_{\xi_1}(k)
\end{aligned}
$$

After some algebraic operations, which are shown in Appendix F.2, the expected total profit function with discrete demand is

$$
\begin{aligned}
\mathbb{E}[\Pi(Q,\xi_1)] = & Q(p_1 - v) - B(\mu - Q) + \Big[p_2^L \mu^L - p_2^H \mu^H - q(\mu^L - \mu^H) \Big] \sum_{k=0}^{Q-I^t} f_{\xi_1}(k) \\
& + (p_2^H - q)\mu^H \sum_{k=0}^{Q} f_{\xi_1}(k) \\
& - (p_2^H - q) \sum_{k=0}^{Q} \sum_{l=Q-k+1}^{\infty} (l - (Q-k)) f_{\xi_2|p_2^H}(l) f_{\xi_1}(k) \\
& - (p_2^L - q) \sum_{k=0}^{Q-I^t} \sum_{l=Q-k+1}^{\infty} (l - (Q-k)) f_{\xi_2|p_2^L}(l) f_{\xi_1}(k) \\
& + (p_2^H - q) \sum_{k=0}^{Q-I^t} \sum_{l=Q-k+1}^{\infty} (l - (Q-k)) f_{\xi_2|p_2^H}(l) f_{\xi_1}(k) \\
& - (p_1 + B - q) \sum_{k=0}^{Q} (Q-k) f_{\xi_1}(k)
\end{aligned}
\tag{2.25}
$$

For continuous case, the equation (2.25) turns out to be

$$
\begin{aligned}
\mathbb{E}[\Pi(Q,\xi_1)] =& Q(p_1-v) - B(\mu - Q) + \left[p_2^L\mu^L - p_2^H\mu^H - q(\mu^L-\mu^H)\right]\int_0^{Q-I^t} f_{\xi_1}(x)dx \\
&+(p_2^H-q)\mu^H\int_0^Q f_{\xi_1}(x)dx \\
&-(p_2^H-q)\int_0^Q\int_{Q-x}^{\infty}(y-(Q-x))f_{\xi_2|p_2^H}(y)dyf_{\xi_1}(x)dx \\
&-(p_2^L-q)\int_0^{Q-I^t}\int_{Q-x}^{\infty}(y-(Q-x))f_{\xi_2|p_2^L}(y)f_{\xi_1}(x)dx \\
&+(p_2^H-q)\int_0^{Q-I^t}\int_{Q-x}^{\infty}(y-(Q-x))f_{\xi_2|p_2^H}(y)dyf_{\xi_1}(x)dx \\
&-(p_1+B-q)\int_0^Q(Q-x)f_{\xi_1}(x)dx
\end{aligned} \tag{2.26}
$$

The first and second derivative of the expected total profit function are shown in Appendix F. The second derivative

$$
\frac{\partial^2\mathbb{E}[\Pi^1(Q_0,\xi_1)]}{\partial Q^2}
$$

is negative under the following condition:

$$
\begin{aligned}
&f_{\xi_1}(Q-I_1^t)\left[(p_2^H-q)F_{\xi_2|p_2^H}(I_1^t)-(p_2^L-q)F_{\xi_2|p_2^L}(I_1^t)\right] \le f_{\xi_1}(Q)\left[(p_1+B)-p_2^H\right] \\
&+ f_{\xi_1}(Q-I_1^t)\left[p_2^H-p_2^L\right] + (p_2^L-q)\int_0^{Q-I_1^t} f_{\xi_2|p_2^L}(Q-x)f_{\xi_1}(x)dx \\
&+ (p_2^H-q)\int_{Q-I_1^t}^{Q} f_{\xi_2|p_2^H}(Q-x)f_{\xi_1}(x)dx
\end{aligned} \tag{2.27}
$$

To conclude, if demand distributions are known, this chapter provides optimal threshold value and optimal order quantity when distribution of the first period and distribution of the second period for given price selected are Poisson&Poisson, Uniform&Uniform, as well as Triangular&Triangular. When we use Triangular distribution, we consider Symmetric, Right-Skewed and Left-Skewed Triangular distributions. The optimal solutions are also obtained when the first period demand is Symmetric Triangular distributed and the second period demand distribution is Uniform distributed. The conditions 2.10 and 2.11 shows that we achieve the results we want and they will be used later on. Also, we will use these exact results to compare the performance of data-driven models provided in Chapter 3.

Chapter 3

Model Formulation

In the previous section, we solve the problem under the assumption that the demand distribution is known prior to the selling period. In the real world, however, the demand distribution is not available. Generally, data-driven approaches using the past data are developed in such cases.

In this section, we assume that the only information available to the decision maker is a set of historical demand data and the demand distributions are not known explicitly. For the first period, we have a set of independent samples of the demand ξ_1, denoted by $\xi_1^1, \ldots, \xi_1^n$, instead of using the demand distribution of ξ_1, which is not available. We consider two possible marked-down prices p_2^H and p_2^L for the second period problem. Likewise, we have N_j iid samples of the demand $\xi_2(p_2^j)$, denoted by $\xi_2^1(p_2^j) \ldots \xi_2^{N_j}(p_2^j)$ where $j = H, L$ for the second period.

There are two decision variables to be determined in the data-driven newsvendor problem with resource action, the threshold and the order quantity. In this study, we proposed eight data-driven approaches to determine these decision variables. The outline of this section is scheduled as follows. Section 3.1 provides three approaches for the second period problem and Section 3.2 provides two approaches to solve the first period problem. Both of approaches provided for the first period problem can work with any approach provided for the second period.

Therefore, an explanation is given about how we combine proposed data-driven models in Section 3.3. Moreover, we propose another data-driven approach, which is a mixed integer linear programming model to make the order quantity and the marked-down price decisions at the same time in Section 3.4. In addition to that, we also develop a simple heuristic where the marked-down decision has to be made prior to the first period to quantify the value of information in Section 3.5. After identifying these eight approaches to solve the data-driven newsvendor problem with a recourse decision, we also define the pre-processing and post-processing that are important steps for models.

3.1 Data-Driven Models for The Second Stage

To begin with, we develop an algorithm based on sample average approximation which is well-known data-driven approach. Our second way to solve the second stage is a linear regression model where a linear equation for each marked-down price is derived by using the historical demand observation. Quadratic regression model is the third way of solving the second period problem. Similar to the linear regression model, we obtain a quadratic equation for each marked-down price and use them to estimate the second period profit.

Sample Average Approximation for The Second Period Problem

Motivated by Ban and Rudin [9] and Beutel et al. [11], we propose sample average approximation type algorithm as a sensible non-parametric heuristic approach to find the optimal policy for the second period problem (for more detail about sample average approximation see Kleywegt et al. [25]). In this approach, the realizations of the demand which are obtained from past data are used to substitute true expectation.

For given leftover I, the sample average approximation (SAA) method to solve the second period SAA problem is as follows.

$$\max_{p_2^j}\{\frac{1}{N_j}\sum_{i=1}^{N_j}\Pi^2(I,\xi_2^i(p_2^j))\} \tag{3.1}$$

where the true problem is

$$\max_{p_2}\mathbb{E}_{\xi_2(p_2)}[\Pi^2(I,\xi_2(p_2))] \qquad \forall I=1,2,\ldots,n \tag{3.2}$$

For the two price case, the second period SAA problem is expressed by:

$$\max\{\frac{1}{N_H}\sum_{i=1}^{N_H}\Pi^2(I,\xi_2^i(p_2^H)),\frac{1}{N_L}\sum_{i=1}^{N_L}\Pi^2(I,\xi_2^i(p_2^L))\} \tag{3.3}$$

We decide which price to choose for each possible inventory level I in this problem. To apply SAA algorithm, we first discretize the continuous variable I. Before introducing the algorithm, we discuss the possible data problems. We know the values of the profit function for a limited number of inventory values, and hence there are gaps in the data which include the results of SAA method. One of the possible problem is that we do not know the profit of the value which is outside the range of leftover I. In this case, we perform extrapolation to estimate the new profit value. The missing value between two known values is the other possible problem. Likewise, we perform interpolation to predict the new value within the range of a discrete set of known profit values. For two possible marked-down price, the SAA algorithm is as follows.

Figure 3.1 Sample average algorithm for the second period problem

Given $p_2 = (p_2^H, p_2^L)$
for each value of I **do**
  for $p_2^j = H, L$ **do**
    for i=1... N_j **do**
      Process/Read $\xi_2^i(p_2^j)$
      if $I \geq \xi_2^i(p_2^j)$ **then**
        $\Pi^2(I, p_2^j) \leftarrow \Pi^2(I, p_2^j) + p_2^j \xi_2^i(p_2^j) + q(I - \xi_2^i(p_2^j))$
      else
        $\Pi^2(I, p_2^j) \leftarrow \Pi^2(I, p_2^j) + p_2^j I$
      end if
    end for
    $\Pi^2(I, p_2^j) \leftarrow \frac{\Pi^2(I, p_2^j)}{N_j}$
  end for
  $\bar{\Pi}^2(I) \leftarrow \max\{\Pi^2(I, p_2^H), \Pi^2(I, p_2^L)\}$
  if $\bar{\Pi}^2(I) == \Pi^2(I, \xi_2^i(p_2^H))$ **then**
    $p_2^\star(I) \leftarrow p_2^H$
  else
    $p_2^\star(I) \leftarrow p_2^L$
  end if
end for

The code gives the optimal average profit $\bar{\Pi}^2$ and the optimal price for each inventory level $I = 1, \ldots, N$.

Linear Regression Model for The Second Period Problem

Beutel and Minner [11] determine the optimal order quantity as an linear function of exogenous variables whose coefficients are determined by a linear programming model. Motivated by [11], we propose an approach where the linear model is used to find the threshold. Given the marked-down price p_2, the second period profit can be partially explained by independent variable I. The main idea is to fit a linear model to the sample data that derives the optimal second period profit depending on the values of the variable I. Linear regression model is built by using leftover I and profit $\bar{\Pi}_2$ variables. The second period profit is treated as a dependent variable and leftover inventory is treated as an independent variable. The relation between profit Π^2 and leftover I may be modelled with,

$$\Pi^2 = \beta_0 + \beta_1 I + \varepsilon. \tag{3.4}$$

The unknown parameters β_0, and β_1 in the equation may be estimated from data with least-squares method. We can force the fitted regression lines to go through the origin. This is reasonable since the profit of zero inventory level should be zero. By fitting linear regressions without intercept, we obtain the slope β_1 which is used as a salvage value in the model. As the second period problem has no meaning with this constrained regression models, we do not apply the constrained regression method. Although we look at r-squared value to see how well the regression model fits the data, we do not record them. Generally, a higher r-squared indicates a better fit for the model and we can say that linear regression model fits well statistically.

After fitting linear regression model for both prices, we have two linear models: one for the high price and the other for the low price. The intersection of two lines can be found as the parameters of equations, β_0, and β_1, are known. The intersection point is expressed as follows.

$$I^t = \frac{\beta_0^H - \beta_0^L}{\beta_1^L - \beta_1^H}$$

The β_0 coefficients of the linear model can be positive. This means that the seller makes profit even if she does not have inventory on hand, which is counter-intuitive. Linear regression line, however, can be used from point ϵ close to zero and the line segment between zero and that point can be created manually. Thus, we ensure that the profit is zero when the remaining inventory is zero.

There might be cases where there is only past data for the low price or for the high price. In such cases, it is assumed that the coefficients of linear equation of the price whose past data is not available are zero.

This intersection is positive when

$$\beta_1^L > \beta_1^H \text{ and } \beta_0^H > \beta_0^L$$

and it becomes the threshold value I^t where the linear model of high price will be used to estimate the second period profit. If the inventory level at the end of the first period is equal or more than the threshold value I^t, the linear model of the low price will be used to estimate the second period profit of the leftover.

If the coefficients of the linear regression models are one of the following cases

$$\begin{gathered}
\beta_0^L = 0 \text{ and } \beta_1^L = 0 \\
\beta_1^L < \beta_1^H \text{ and } \beta_0^H > \beta_0^L \\
\beta_1^L = \beta_1^H \text{ and } \beta_0^H = \beta_0^L \\
\beta_0^H = \beta_0^L \text{ and } \beta_1^H > \beta_1^L \\
\beta_0^H > \beta_0^L \text{ and } \beta_1^H = \beta_1^L
\end{gathered}$$

then the linear model of the high price is always used after the point ϵ close to zero.

There may be cases where the linear model of the low price is always used after the point ϵ. These cases are shown below.

$$\beta_0^H = 0 \text{ and } \beta_1^H = 0$$
$$\beta_1^L > \beta_1^H \text{ and } \beta_0^H < \beta_0^L$$
$$\beta_1^L > \beta_1^H \text{ and } \beta_0^H = \beta_0^L$$
$$\beta_1^L = \beta_1^H \text{ and } \beta_0^H < \beta_0^L$$

In such cases, threshold value is set to ϵ. Moreover, if the following case occurs

$$\beta_1^L < \beta_1^H \text{ and } \beta_0^H < \beta_0^L$$

it is assumed that the threshold value equals to the ϵ and the linear equation of the low price is always used for the remaining inventory up to the intersection unless the intersection is smaller than the ϵ.

Otherwise, the linear model of the high price is used between the point ϵ and I^t. After the threshold value I^t, the linear model of low price is applied (see Figure 3.2).

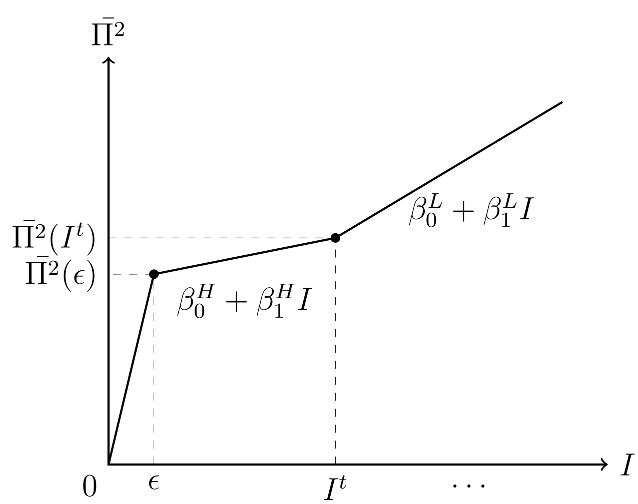

Figure 3.2: The intersection point of two linear regression models

Quadratic Regression Model for The Second Period Problem

In addition to the linear regression model, the relation between profit Π^2 and leftover I variables may be modelled with

$$P = \beta_0 + \beta_1 I + \beta_2 I^2 + \varepsilon. \tag{3.5}$$

In papers [11] and [9], it is assumed that the order quantity is the linear function of the explanatory features. We would like to propose a model that also works with quadratic function as an alternative to the linear function. Therefore, quadratic regression model can be used instead of linear model to obtain the second period profits. The unknown parameters β_0, β_1, and β_2 in the equation may be estimated from data with least-squares method. The coefficients of the quadratic function is used in the mixed integer programming model to determine the order quantity. Also, a constrained regression which has an equation of the form $\Pi_2 = \beta_0 + \beta_1 I + \beta_2 (I)^2$ can be fit to data with constraints $\beta_0 = 0$ and $\beta_1 < 0$. By setting β_0 to zero, the fitted quadratic model can be forced to pass the origin as well. Although we look at r-squared value to see how well the quadratic regression model fits the data, we do not record them. Generally, a higher r-squared indicates a better fit for the model and we can say that regression model fits well statistically. After fitting quadratic model to data for both prices, we can approximate this nonlinear quadratic regression model by a piece-wise linear curve in order to make the objective function of the model linear.

There might be cases where there is only past data for the low price or for the high price. In such cases, it is assumed that the coefficients of quadratic equation of the price whose past data is not available are zero. We calculate the intersection of two quadratic equations since it derives the threshold value. The intersection point is found by the following equation.

$$I^t = \frac{\beta_1^H - \beta_1^L}{\beta_2^L - \beta_2^H}$$

When

$$\beta_2^L > \beta_2^H \text{ and } \beta_1^H > \beta_1^L$$

the quadratic equation of high price is used in order to estimate the second period profit if the remaining inventory is less than the threshold value. For leftovers which are greater than or equal to the threshold, the quadratic regression model of low price is used.

If the high price is chosen for all leftovers, the threshold value is set to the optimal order quantity. The high price is always selected under one of the following conditions.

$$\begin{aligned}
&\beta_1^L = 0 \text{ and } \beta_2^L = 0\\
&\beta_2^L < \beta_2^H \text{ and } \beta_1^H > \beta_1^L\\
&\beta_2^H = \beta_2^L \text{ and } \beta_1^H = \beta_1^L\\
&\beta_1^H = \beta_1^L \text{ and } \beta_2^H > \beta_2^L\\
&\beta_1^H > \beta_1^L \text{ and } \beta_2^H = \beta_2^L
\end{aligned}$$

Moreover, if one of the following conditions hold, then we set the threshold value to ϵ which is introduced in the explanation of the linear regression model. Therefore, the low price is always determined as the marked-down price.

$$\begin{aligned}
&\beta_1^H = 0 \text{ and } \beta_2^H = 0\\
&\beta_2^L > \beta_2^H \text{ and } \beta_1^H < \beta_1\\
&\beta_2^L > \beta_2^H \text{ and } \beta_1^H = \beta_1^L\\
&\beta_2^L = \beta_2^H \text{ and } \beta_1^H < \beta_1^L
\end{aligned}$$

3.2 Data-Driven Models for The First Stage

With the solution of the second period problem, we provide two different data-driven models to find the optimal order quantity. We solve the two-stage problem considered hierarchically and thereby we first obtain the threshold as well as the second period average profit to use in the order quantity decision. The second period average profit, denoted by $\bar{\Pi}^2$, is obtained from one of the methods used for Period 2 and it is used as an input in the approaches proposed for the first period problem: One of them is an algorithm based on sample average approximation whereas the second one is the mixed integer linear programming model.

Sample Average Approximation for The First Period Problem

Let us define the first period SAA problem:

$$\max_{Q \geq 0} \frac{1}{N} \sum_{i=1}^{N} \Pi(Q, \xi_1^i) \tag{3.6}$$

where the true problem is

$$\max_{Q \geq 0} \mathbb{E}_{\xi_1}[\Pi(Q, \xi_1)] \tag{3.7}$$

The algorithm based on SAA takes the expected second period profit function as an input to find the optimal order quantity. Thus, the expected second period function is calculated by one of the models defined above. The SAA algorithm of the first problem is given below.

Figure 3.3 Sample average algorithm for the first period problem

```
for each value of Q do
    for i=1...N do
        Process ξ_1^i
        if Q ≥ ξ_1^i then
            Compute E[Π^2(Q − ξ_1^i, ξ_2)|p_2] with one of data-driven models of the second period problem
            Π(Q) ← Π(Q) − Qv + p_1 ξ_1^i + E[Π^2(Q − ξ_1^i, ξ_2)|p_2]
        else if Q > ξ_1^i then
            Π(Q) ← Π(Q) − Qv + p_1 Q − B(ξ_1^i − Q)
        end if
    end for
    Π(Q, ξ_1) ← Π(Q)/N
end for
Q̄ ← max_Q Π(Q, ξ_1)
```

Then, we have average total profit values for each order quantity Q. Among these profits, the order quantity that yields the maximum average total profit value determines the optimum order quantity.

Mixed Integer Linear Programming Model for The First Period Problem

Motivated by [36] and [11], we formulate the mixed integer programming problem with some linearization constraints to use the second period average profit function which is obtained from one of data-driven models in 3.1. In this approach, the newsvendor aims to find the quantity Q such that total revenues comprising of first period and second period are maximized.

The objective function consists of selling price p_1, the second period average profit (salvage value) $\bar{\Pi}^2$, the cost of purchasing v, and the underage or shortage cost B. $\bar{\Pi}^2$ is a piece-wise linear function made by $m-1$ line segments and m breakpoints. In Figure 3.4, the points where the slope of piece-wise function changes are called the breakpoints. We use the piece-wise linear function to estimate the average profit value of the remaining inventory that is not one of the inventory level whose profit function is calculated. Thus, we present this piece-wise linear function in a linear form to solve our optimization problem. Moreover, we still consider a set of historical demand observations $\xi_1^1, \ldots, \xi_1^n$ in this approach. To describe the problem, the following notation is used:

$s_i =$ satisfied demand when the demand is ξ_1^i

$u_i =$ the amount of underage or shortage when the demand is ξ_1^i

$r_i =$ the revenue received when the leftover is $Q - \xi_1$

$b_k =$ breakpoints of the piecewise linear function $\bar{\Pi}^2$

$z_{ik} =$ the convex combination multipliers for breakpoints of the piece-wise linear function $\bar{\Pi}^2$

$y_{ik} =$ a binary variable to represent the piece-wise linear function $\bar{\Pi}^2$ in a linear form

where

$$y_{ik} = \begin{cases} 1, & \text{if } Q - \xi_1^i \text{ is in interval } k = 1, \ldots, m-1 \\ 0, & otherwise \end{cases}$$

For some k ($k = 1, \ldots, m$), we know that $b_k \leq Q - s_i \leq b_{k+1} \quad (\forall i = 1, \ldots, n)$. Hence, for some z_{ik} ($0 \leq z_{ik} \leq 1 \quad \forall i$), $r_i = \bar{\Pi}^2(Q - s_i)$ and $Q_0 - s_i$ can be replaced by

$$r_i = \bar{\Pi}^2(Q - s_i) = z_{ik}\bar{\Pi}^2(b_k) + z_{i(k+1)}\bar{\Pi}^2(b_{k+1})$$
$$Q - s_i = z_{ik}b_k + z_{i(k+1)}b_{k+1}$$

respectively.

Given the second period average total profit, a non-parametric formulation of the data-driven newsvendor problem is to find Q is shown below:

$$\max_{Q} \quad \frac{1}{n}\sum_{i=1}^{n}(-vQ + p_1 s_i + r_i - Bu_i) \tag{3.8}$$

$$s.t \quad s_i \leq Q \qquad \forall i = 1,\dots,n \tag{3.9}$$

$$s_i \leq \xi_1^i \qquad \forall i = 1,\dots,n \tag{3.10}$$

$$u_i \geq \xi_1^i - Q \qquad \forall i = 1,\dots,n \tag{3.11}$$

$$Q - s_i = \sum_{k=1}^{m} z_{ik} b_k \qquad \forall i = 1,\dots,n \tag{3.12}$$

$$r_i = \sum_{k=1}^{m} z_{ik} \bar{\Pi}^2(b_k) \qquad \forall i = 1,\dots,n \tag{3.13}$$

$$\sum_{k=1}^{m} z_{ik} = 1 \qquad \forall i = 1,\dots,n \tag{3.14}$$

$$\sum_{k=1}^{m-1} y_{ik} = 1 \qquad \forall i = 1,\dots,n \tag{3.15}$$

$$z_{i1} \leq y_{i1} \qquad \forall i = 1,\dots,n \tag{3.16}$$

$$z_{ik} \leq y_{ik-1} + y_{ik} \qquad \forall i = 1,\dots,n \quad \forall k = 2,\dots,m-1 \tag{3.17}$$

$$z_{im} \leq y_{im-1} \qquad \forall i = 1,\dots,n \tag{3.18}$$

$$Q \geq 0 \tag{3.19}$$

$$y_{ik} = 0 \text{ or } 1 \qquad \forall i = 1,\dots,n \quad \forall k = 1,\dots,m-1 \tag{3.20}$$

$$z_{ik}, s_i, u_i, r_i \geq 0 \qquad \forall i = 1,\dots,n \quad \forall k = 1,\dots,m \tag{3.21}$$

The objective function includes the purchasing cost v, selling price p_1, the second period average profit r_i, as well as the underage cost B. The constraints (3.9) and (3.10) are for linearization of the expression $s_i = min(q, \xi_1^i)$. By these constraints, it is ensured that satisfied demand is equal to the minimum of the order quantity Q and the demand ξ_1^i. Similarly, the constraints (3.11) and $u_i \geq 0$ in (3.21) are equivalent to the term $u_i = max(0, \xi_1^i - Q)$. These constraints with the maximization objective determine the unmet demand or underage u_i corresponding to each demand observation ξ_1^i. We express the second period average profit r_i, which is a piece-wise linear function, by adding linear constraints (3.12) and (3.13) into the model. The constraints (3.14), (3.15), (3.16), (3.17) and (3.18) are also added to the model in order to represent the piece-wise linear function in linear form.

3.3 Combination of Data-Driven Models

In this section, we explain how to deal with both two periods of a fashionable product by using a data-driven approaches provided for each period. We have three data-driven models for the second period and two data-driven models for the first period. They are defined in detail in the previous sections 3.1 and 3.2. The combination of these models are all used to find the optimal policy. In total, we have six models to solve the data-driven two-stage newsvendor problem (see Table 3.1 for notations of models).

Table 3.1: Notation for models

#	Proposed Models	Description
1	SAA-P	The algorithm based on SAA is used to solve both stage
2	MIP-SAA	The algorithm based on SAA is used for the second stage and results are used in mixed integer program
3	SAA-LR	Linear regression is used for the second stage and results are used in the algorithm based on SAA
4	MIP-LR	Linear regression is used for the second stage and results are used in mixed integer program
5	SAA-QR	Quadratic regression is used for the second stage and results are used in the algorithm based on SAA
6	MIP-QR	Quadratic regression is used for the second stage and results are used in mixed integer program

SAA-P

We first implement the algorithm based on SAA for the second period problem and use the obtained results as an input in the algorithm based on SAA for first period problem. With the calculated expected second period function, we choose the order quantity that maximizes the expected average profit. Before writing the algorithm, we transform continuous variable Q into a discrete form. For the unknown second profit values of the remaining inventory that fall between known inventory level values, interpolation method may be performed. Extrapolation method might be used to estimate unknown values which are outside the range of the inventory level.

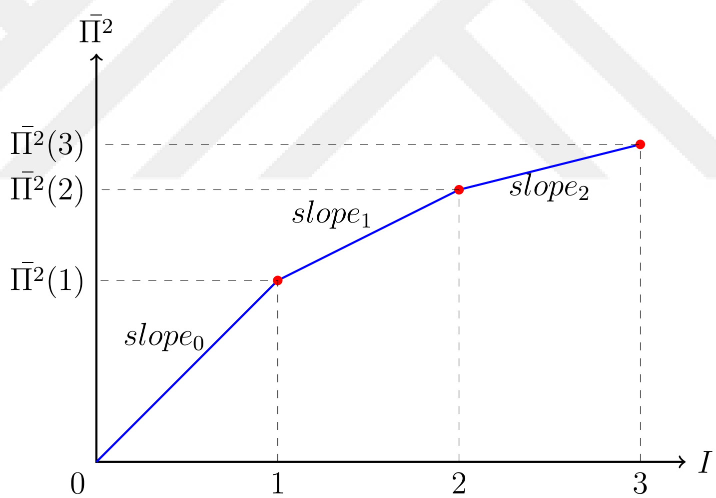

Figure 3.4: Multiple breakpoint function

The piece-wise function depicted in Figure 3.4 can be represented as follows.

$$\Pi_2(I) = \begin{cases} slope_0(I-0) + \bar{\Pi}^2(0) & \text{if } 0 \le I < 1 \\ slope_1(I-1) + \bar{\Pi}^2(1) & \text{if } 1 \le I < 2 \\ slope_2(I-2) + \bar{\Pi}^2(2) & \text{if } 2 \le I < 3 \\ slope_2(I-3) + \bar{\Pi}^2(3) & \text{if } 3 \le I \end{cases}$$

The SAA algorithm of the first problem is given below.

Figure 3.5 Algorithm for SAA-P

for $i \in \{0 \dots I\}$ **do**

    $slope_i \leftarrow \bar{\Pi}^2(i+1) - \bar{\Pi}^2(i)$

end for

for each value of Q **do**

    for i=1…N **do**

        Process ξ_1^i

        if $Q_0 \geq \xi_1^i$ **then**

            for $j \in \{0, \dots, I\}$ **do**

                if $Q - \xi_1^i \geq j$ and $Q - \xi_1^i < j+1$ **then**

                    $\Pi(Q) \leftarrow \Pi(Q) - Qv + p_1\xi_1^i + [slope_j(Q - \xi_1^i - j) + \bar{\Pi}^2(j)]$

                end if

            end for

            if $Q - \xi_1^i \geq I$ **then**

                $\Pi(Q) \leftarrow \Pi(Q) - Qv + p_1\xi_1^i + [slope_{(I-1)}(Q - \xi_1^i - I) + \bar{\Pi}^2(I)]$

            end if

        else

            $\Pi(Q) \leftarrow \Pi(Q) - Qv + p_1Q - B(\xi_1^i - Q)$

        end if

    end for

    $\Pi(Q, \xi_1) \leftarrow \frac{\Pi(Q)}{N}$

end for

$\bar{Q} \leftarrow \max_Q \Pi(Q, \xi_1)$

The order quantity that gives the maximum average total profit determines the optimum order quantity.

In the algorithm based on SAA for the first period, we determine an upper bound M for the order quantity. To do so, We run the same algorithm for the first period, assuming that the remaining inventory is salvaged at the high price, regardless of the second period demand. The output of this algorithm is used as the upper limit for the order quantity, as the retailer tends to order more when the salvage value is set at the high price.

MIP-SAA

While we use the algorithm based on SAA to estimate the second period average profit $\bar{\Pi}^2$, we solve the first period problem with the mixed integer programming model. The second period average profit $\bar{\Pi}^2$ observed from the algorithm based on SAA is used as an input in the model. The algorithm based on SAA for the second period is identified in Section 3.1. Also we define the mixed integer linear programming model in Section 3.2. To illustrate, a possible outcome of the algorithm based on SAA is shown in Table 3.6.

Inventory level I	p_2 value	Second period average profit value
$I = 1$	$p_2 = p_2^H$	$\Pi^2(1)$
$I = 2$	$p_2 = p_2^H$	$\Pi^2(2)$
$\vdots$	$\vdots$	$\vdots$
$I = n$	$p_2 = p_2^L$	$\Pi^2(n)$

Figure 3.6: The output of the algorithm based on SAA for period 2

and we use the piece-wise linear function to integrate the observed average profit values $\bar{\Pi}^2(I), \quad \forall I$. We express this piece-wise linear function by adding linear constraints (3.12) and (3.13) into the model.

The optimal order quantity $Q^\star$ found by the model is a continuous value, yet we may would like to find an integer value approximately. To do so, instead of continuous decision variable Q, we can also define an integer variable Q with an upper bound $BigM$ so that the proposed model can give an integer order quantity. One should be careful in choosing an appropriate number for the upper bound $BigM$ since it must be large enough so that the optimal order quantity that maximizes the objective value and satisfies all constraints can be found. To find an appropriate positive number for the upper bound of the order quantity variable, we solve the newsvendor model where the remaining inventory is salvaged at the high price, regardless of what the demand of the second period will be. Setting salvage value to the high price increases the tendency of the retailer to order more in the first period. Therefore, the optimal solution determined by this model will be used as an upper bound $BigM$ for the order quantity variable.

SAA-LR

We combine the algorithm based on SAA with the linear model to find the order quantity and the threshold value which is used for the marked-down price decision. After fitting a linear model to the sample data and obtaining coefficients of linear equations, the second period profit for the inventory level I can be estimated.

We can always use the high price linear equation or we can always use the low price linear equation to estimate the second period profit based on the values of the coefficients $\beta_0^H, \beta_1^H, \beta_0^L, \beta_1^L$. Otherwise, we use the linear equation of the high price for the leftovers which are smaller than the threshold whereas we use the linear equation of the low price for the leftovers above the threshold. For the first period, the following algorithm based on SAA working with linear regression model is written considering each possible case.

Figure 3.7 Algorithm for SAA-LR when the linear model of the high price is used for all $I = Q - \xi_1 = 1, \ldots, N$

$slope_H \leftarrow (\epsilon\beta_1^H + \beta_0^H)/(\epsilon)$
for each value of Q **do**
  for i=1...N **do**
    Process ξ_1^i
    $\bar{\Pi}^2(\epsilon) \leftarrow \epsilon\beta_1^H + \beta_0^H$
    if $Q \geq \xi_1^i$ **then**
      if $Q - \xi_1^i \geq 0$ and $Q - \xi_1^i < \epsilon$ **then**
        $\Pi(Q) \leftarrow \Pi(Q) - Qv + p_1\xi_1^i + [slope_H(Q - \xi_1^i)]$
      else if $Q - \xi_1^i \geq \epsilon$ **then**
        $\Pi(Q) \leftarrow \Pi(Q) - Qv + p_1\xi_1^i + [\beta_0^H(Q - \xi_1^i - \epsilon) + \bar{\Pi}^2(\epsilon)]$
      end if
    else if $Q \leq \xi_1^i$ **then**
      $\Pi(Q) \leftarrow \Pi(Q) - Qv + p_1 Q - B(\xi_1^i - Q)$
    end if
  end for
  $\Pi(Q, \xi_1) \leftarrow \frac{\Pi(Q)}{N}$
end for
$\bar{Q} \leftarrow \max_Q \Pi(Q, \xi_1)$
$I^t \leftarrow \bar{Q}$

This algorithm is also used when the linear model of the low price is used for all $I = Q - \xi_1 = 1, \ldots, N$. The threshold value is set to ϵ when the low price is selected for all leftovers (i.e, $I^t \leftarrow \epsilon$).

Figure 3.8 Algorithm for SAA-LR when both linear models are used

$I^t \leftarrow \frac{\beta_0^H - \beta_0^L}{\beta_1^L - \beta_1^H}$

$slope_H \leftarrow (\epsilon\beta_1^H + \beta_0^H)/(\epsilon)$

$slope_L \leftarrow (\epsilon\beta_1^L + \beta_0^L)/(\epsilon)$

for each value of Q **do**

  for i=1…N **do**

    Process ξ_1^i

    if $Q \geq \xi_1^i$ **then**

      if $I^t > \epsilon$ **then**

        $\bar{\Pi}^2(\epsilon) \leftarrow \epsilon\beta_1^H + \beta_0^H$

        $\bar{\Pi}^2(I^t) \leftarrow I^t\beta_1^L + \beta_0^L$

        if $Q - \xi_1^i \geq 0$ and $Q - \xi_1^i < \epsilon$ **then**

          $\Pi(Q) \leftarrow \Pi(Q) - Qv + p_1\xi_1^i + [slope_H(Q - \xi_1^i)]$

        else if $Q - \xi_1^i \geq \epsilon$ and $Q - \xi_1^i < I^t$ **then**

          $\Pi(Q) \leftarrow \Pi(Q) - Qv + p_1\xi_1^i + [\beta_0^H(Q - \xi_1^i - \epsilon) + \bar{\Pi}^2(\epsilon)]$

        else if $Q - \xi_1^i \geq I^t$ **then**

          $\Pi(Q) \leftarrow \Pi(Q) - Qv + p_1\xi_1^i + [\beta_0^L(Q - \xi_1^i - I^t) + \bar{\Pi}^2(I^t)]$

        end if

      else

        $\bar{\Pi}^2(\epsilon) \leftarrow \epsilon\beta_1^L + \beta_0^L$

        $\bar{\Pi}^2(I^t) \leftarrow I^t\beta_1^L + \beta_0^L$

        if $Q - \xi_1^i \geq 0$ and $Q - \xi_1^i < \epsilon$ **then**

          $\Pi(Q) \leftarrow \Pi(Q) - Qv + p_1\xi_1^i + [slope_L(Q - \xi_1^i)]$

        else if $Q - \xi_1^i \geq \epsilon$ **then**

          $\Pi(Q) \leftarrow \Pi(Q) - Qv + p_1\xi_1^i + [\beta_0^L(Q - \xi_1^i - I^t) + \bar{\Pi}^2(I^t)]$

        end if

      end if

    else

      $\Pi(Q) \leftarrow \Pi(Q) - Qv + p_1 Q - B(\xi_1^i - Q)$

    end if

  end for

  $\Pi(Q, \xi_1) \leftarrow \frac{\Pi(Q)}{N}$

end for

$\bar{Q} \leftarrow \max_Q \Pi(Q, \xi_1)$

Similarly, the order quantity that gives the maximum average total value determines the optimum order quantity.

As we stated before in Section 3.1, we determine an upper bound for the order quantity.

MIP-LR

Another data-driven approach where the results of the linear regression are integrated into the mixed integer programming model is proposed. After fitting linear regression model for both prices, a piece-wise linear function whose shape is depending on the value of the threshold I^t is obtained (See Figure 3.2). As a result of linear regression models, we have the threshold value and the piece-wise linear function to use in the model. As we stated in SAA-LR model, the value of the intersection point determines which linear model we will use at which interval. One should be careful in choosing the last break point of the piece-wise linear function. It can be determined according to the sample data. To illustrate, it can be simply twice the largest data point. Moreover, a pre-processing step is needed to use linear regression models in the mixed integer programming model. The preprocess for the model is as follows.

Figure 3.9 Pre-process for MIP-LR when the linear model of the high price is used for all $I = Q - \xi_1 = 1, \dots, N$

$\bar{\Pi}^2(0) \leftarrow 0$
$\bar{\Pi}^2(\epsilon) \leftarrow \epsilon\beta_1^H + \beta_0^H$
$\bar{\Pi}^2(I^{max}) \leftarrow (I^{max})\beta_1^H + \beta_0^H$

This pre-process algorithm is also used when the linear model of the low price is used for all $I = Q - \xi_1 = 1, \dots, N$.

Figure 3.10 Pre-process for MIP-LR when both linear models are used

$\bar{\Pi}^2(0) \leftarrow 0$

$\bar{\Pi}^2(\epsilon) \leftarrow \epsilon\beta_1^H + \beta_0^H$

if $I^t > \epsilon$ **then**

  $\bar{\Pi}^2(I^t) \leftarrow (I^t)\beta_1^L + \beta_0^L$

  $\bar{\Pi}^2(I^{max}) \leftarrow (I^{max})\beta_1^L + \beta_0^L$

else

  $\bar{\Pi}^2(\epsilon) \leftarrow \epsilon\beta_1^L + \beta_0^L$

  $\bar{\Pi}^2(I^{max}) \leftarrow (I^{max})\beta_1^L + \beta_0^L$

end if

After the process for observing the average profit values for each possible remaining inventory by using the linear regression models, the mixed integer programming model introduced in Section 3.2 determines the order quantity that maximizes the average total profit.

SAA-QR

As an alternative to the SAA-LR approach defined in the previous subsection, we introduce the approach where the marked down pricing policy is found by using quadratic regression whereas the algorithm based on SAA which works with the quadratic regression results derives the optimal order quantity. For the first period, the following algorithm based on SAA working with quadratic regression model is written considering each possible case.

Figure 3.11 Algorithm for SAA-QR when the quadratic equation of the high price is used for all $I = Q - \xi_1 = 1, \ldots, N$

$slope_H \leftarrow (\epsilon\beta_1^H + \beta_0^H)/(\epsilon)$
for each value of Q **do**
  for i=1…N **do**
    Process ξ_1^i
    $\bar{\Pi}^2(\epsilon) \leftarrow \epsilon\beta_1^H + \beta_0^H$
    if $Q \geq \xi_1^i$ **then**
      $\Pi(Q) \leftarrow \Pi(Q) - Qv + p_1\xi_1^i + [\beta_1^H(Q - \xi_1^i) + \beta_2^H(Q - \xi_1^i)^2]$
    else if $Q \leq \xi_1^i$ **then**
      $\Pi(Q) \leftarrow \Pi(Q) - Qv + p_1 Q - B(\xi_1^i - Q)$
    end if
  end for
  $\Pi(Q, \xi_1) \leftarrow \frac{\Pi(Q)}{N}$
end for
$\bar{Q} \leftarrow \max_Q \Pi(Q, \xi_1)$
$I^t \leftarrow \bar{Q}$

This algorithm is also used when the quadratic model of the low price is used for all $I = Q - \xi_1 = 1, \ldots, N$. The threshold value is set to ϵ when the low price is selected for all leftovers (i.e, $I^t \leftarrow \epsilon$).

Figure 3.12 Algorithm for SAA-QR when both quadratic equations are used

$I^t \leftarrow \frac{\beta_0^H - \beta_0^L}{\beta_1^L - \beta_1^H}$
for each value of Q **do**
  for i=1…N **do**
    Process ξ_1^i
    if $Q \geq \xi_1^i$ **then**
      if $Q - \xi_1^i < I^t$ **then**
        $\Pi(Q) \leftarrow \Pi(Q) - Qv + p_1\xi_1^i + [\beta_1^H(Q - \xi_1^i) + \beta_2^H(Q - \xi_1^i)^2]$
      else if $Q - \xi_1^i \geq I^t$ **then**
        $\Pi(Q) \leftarrow \Pi(Q) - Qv + p_1\xi_1^i + [\beta_1^L(Q - \xi_1^i) + \beta_2^L(Q - \xi_1^i)^2]$
      end if
    else if $Q > \xi_1^i$ **then**
      $\Pi(Q) \leftarrow \Pi(Q) - Qv + p_1 Q - B(\xi_1^i - Q)$
    end if
  end for
  $\Pi(Q, \xi_1) \leftarrow \frac{\Pi(Q)}{N}$
end for
$\bar{Q} \leftarrow \max_Q \Pi(Q, \xi_1)$

Likewise, we select the order quantity that gives the maximum average total profit value.

MIP-QR

When we use quadratic regression model to obtain the second period average profits, the mixed-integer linear programming model is used for the first period problem. After estimating the coefficient of the quadratic equations, we used them in the mixed integer programming model to determine the order quantity. We approximate this nonlinear quadratic regression model by a piece-wise linear curve in order to make the objective function of the model linear.

According to the estimated coefficients and thereby threshold value, the high price might be always selected for all leftovers I or the low price might be always determined as the marked-down price. Otherwise, the quadratic regression of the high price might be used for the remaining inventory which are less than the threshold value and the quadratic regression equation of the low marked-down price is used for the remaining inventory greater than or equal to the threshold.

Moreover, a pre-processing step is needed to use quadratic regression models in the mixed integer programming model. The preprocess for the model is as follows.

Figure 3.13 Pre-process for MIP-QR when the quadratic equation of the high price is used for all $I = Q - \xi_1 = 1, \ldots, N$

$\bar{\Pi}^2(0) \leftarrow 0$
for $i \in \{1 \ldots I^{max}\}$ **do**
  $\bar{\Pi}^2(i) \leftarrow \beta_1^H i + \beta_2^H i^2$
end for

This pre-process algorithm is also used when the quadratic equation of the low price is used for all $I = Q - \xi_1 = 1, \ldots, N$.

Figure 3.14 Pre-process for MIP-QR when both quadratic equations are used

$\bar{\Pi}^2(0) \leftarrow 0$
for $i \in \{1 \ldots I^t\}$ **do**
  $\bar{\Pi}^2(i) \leftarrow \beta_1^H i + \beta_2^H i^2$
end for
for $i \in \{I^t \ldots I^{max}\}$ **do**
  $\bar{\Pi}^2(i) \leftarrow \beta_1^L i + \beta_2^L i^2$
end for

3.4 Mixed Integer Linear Model Formulation

We also solve the two-stage data-driven inventory problem as a mixed integer linear program which finds the order quantity and the threshold value simultaneously. This model is denoted by MIP-P. We motivated by the model in [36] and [11] when we proposed this approach. The parameters and decision variables are shown in Table 3.2.

Table 3.2: Notation for the MIP approach that finds two decisions simultaneously

Parameters	Definition
v	Unit purchasing cost
q	Unit salvage value for the leftover at the end of period 2
p_1	Unit selling price in period 1
p_2^j	The value of the marked-down price-j in period 2 $\forall j = H, L$
B	Goodwill/Shortage cost
ξ_1^i	The demand data point in period 1 $\forall i = 1, \ldots, n$
$\xi_{2\|p_2^j}^k$	The demand data point of the marked-down price-j in period 2 $\forall j = H, L \quad \forall k = 1, \ldots, n_j$
Decision Variables	**Definition**
Q	Order quantity
s_i^1	Satisfied demand in period 1 when the demand is $\xi_1^i \quad \forall i = 1, \ldots, n$
x_{ij}	Binary variable which takes the value 1 if the p_2^j is chosen for the inventory level $Q - s_i^1$, 0 otherwise $\forall i = 1, \ldots, n \quad \forall j = H, L$
u_i	Amount of underage or shortage when the demand is $\xi_1^i \quad \forall i = 1, \ldots, n$
s_{ik}^j	Satisfied demand in period 2 when the demand is $\xi_{2\|p_2^j}^k \quad \forall j = H, L \quad \forall k = 1, \ldots, n_j$

Notice that

- $s_i^1 = \min\{Q, \xi_1^i\}$
- $u_i = \max\{\xi_1^i - Q, 0\}$
- $s_{ik}^j = \min\{Q - s_i^1, \xi_{2|p_2^j}^k\} \quad \forall i = 1, \dots, n \quad \forall j = H, L \quad \forall k = 1, \dots, n_j$

The MIP model to find the solutions simultaneously, denoted by MIP-P, is as follows.

$$\max_{Q} \quad \frac{1}{n} \sum_{i=1}^{n} \Bigg(- vQ + p_1 s_i^1 - Bu_i + \sum_{j=1}^{2} \Bigg(x_{ij} \sum_{k=1}^{n_j} \frac{1}{n_j} (p_2^j s_{ik}^j + q[Q - s_i^1 - s_{ik}^j]) \Bigg) \Bigg) \tag{3.22}$$

$$\begin{aligned}
\text{subject to} \quad & s_i^1 \leq Q && \forall i = 1, \dots, n && (3.23)\\
& s_i^1 \leq \xi_1^i && \forall i = 1, \dots, n && (3.24)\\
& u_i \geq \xi_1^i - Q && \forall i = 1, \dots, n && (3.25)\\
& s_{ik}^j \leq Q - s_i^1 && \forall i = 1, \dots, n \quad \forall j = H, L \quad \forall k = 1, \dots, n_j && (3.26)\\
& s_{ik}^j \leq \xi_{2|p_2^j}^k && \forall i = 1, \dots, n \quad \forall j = H, L \quad \forall k = 1, \dots, n_j && (3.27)\\
& \sum_{j=1}^{2} x_{ij} = 1 && \forall i = 1, \dots, n && (3.28)\\
& Q, s_i^1, s_{ik}^j, u_i \geq 0 && \forall i = 1, \dots, n \quad \forall j = H, L \quad \forall k = 1, \dots, n_j && (3.29)\\
& x_{ij} \in \{0, 1\} && \forall i = 1, \dots, n \quad \forall j = H, L && (3.30)
\end{aligned}$$

New decision variables: Let's define the continuous variables for linearization of the objective function

$$
\begin{aligned}
y_{ij} &= x_{ij}Q \quad \forall i=1,\dots,n \quad \forall j=H,L \\
z_{ik}^{j} &= x_{ij}s_{ik}^{j} \quad \forall i=1,\dots,n \quad \forall j=H,L \quad \forall k=1,\dots,n_j \\
t_{ij} &= x_{ij}s_{i}^{1} \quad \forall i=1,\dots,n \quad \forall j=H,L
\end{aligned}
$$

Let's use these additional variables in the objective (3.22) instead of the products

$$
\max_{Q} \quad \frac{1}{n}\sum_{i=1}^{n}\left(-vQ+p_1 s_i^1-Bu_i+\sum_{j=1}^{2}\sum_{k=1}^{n_j}\frac{1}{n_j}(p_2^j z_{ik}^j+q[y_{ij}-t_{ij}-z_{ik}^j])\right) \tag{3.31}
$$

and add the following constraints

$$
\begin{aligned}
& y_{ij} \le BigMx_{ij} && \forall i=1,\dots,n \quad \forall j=H,L && (3.32)\\
& y_{ij} \le Q && \forall i=1,\dots,n \quad \forall j=H,L,\dots,m && (3.33)\\
& y_{ij} \ge Q-(1-x_{ij})BigM && \forall i=1,\dots,n \quad \forall j=H,L && (3.34)\\
& z_{ik}^j \le \xi_{2|p_2^j}^k s_{ik}^j && \forall i=1,\dots,n \quad \forall j=H,L \quad \forall k=1,\dots,n_j && (3.35)\\
& z_{ik}^j \le s_{ik}^j && \forall i=1,\dots,n \quad \forall j=H,L \quad \forall k=1,\dots,n_j && (3.36)\\
& z_{ik}^j \ge s_{ik}^j-(1-x_{ij})\xi_{2|p_2^j}^k && \forall i=1,\dots,n \quad \forall j=H,L \quad \forall k=1,\dots,n_j && (3.37)\\
& t_{ij} \le \xi_1^i x_{ij} && \forall i=1,\dots,n \quad \forall j=H,L && (3.38)\\
& t_{ij} \le s_i^1 && \forall i=1,\dots,n \quad \forall j=H,L && (3.39)\\
& t_{ij} \ge s_i^1-(1-x_{ij})\xi_1^i && \forall i=1,\dots,n \quad \forall j=H,L && (3.40)\\
& y_{ij}, z_{ik}^j, t_{ij} \ge 0 && \forall i=1,\dots,n \quad \forall j=H,L \quad \forall k=1,\dots,n_j && (3.41)
\end{aligned}
$$

In inequalities (3.32) and (3.34), $BigM$ must be large enough to ensure that y_{ij} will be less than or equal to BigM. To find an appropriate large number for $BigM$, the same process mentioned in MIP-SAA model explanation is used.

As a result, we have the following formulation

$$\max_{Q} \quad \frac{1}{n}\sum_{i=1}^{n}\left(-vQ + p_1 s_i^1 - Bu_i \quad + \sum_{j=1}^{2}\sum_{k=1}^{n_j}\frac{1}{n_j}(p_2^j z_{ik}^j + q[y_{ij} - t_{ij} - z_{ik}^j])\right) \tag{3.42}$$

$$\text{subject to} \quad s_i^1 \leq Q \qquad \forall i = 1,\ldots,n \tag{3.43}$$

$$s_i^1 \leq \xi_1^i \qquad \forall i = 1,\ldots,n \tag{3.44}$$

$$u_i \geq \xi_1^i - Q \qquad \forall i = 1,\ldots,n \tag{3.45}$$

$$s_{ik}^j \leq Q - s_i^1 \qquad \forall i = 1,\ldots,n \quad \forall j = H,L \quad \forall k = 1,\ldots,n_j \tag{3.46}$$

$$s_{ik}^j \leq \xi_{2|p_2^j}^k \qquad \forall i = 1,\ldots,n \quad \forall j = H,L \quad \forall k = 1,\ldots,n_j \tag{3.47}$$

$$\sum_{j=1}^{2} x_{ij} = 1 \qquad \forall i = 1,\ldots,n \tag{3.48}$$

$$y_{ij} \leq BigM x_{ij} \qquad \forall i = 1,\ldots,n \quad \forall j = H,L \tag{3.49}$$

$$y_{ij} \leq Q \qquad \forall i = 1,\ldots,n \quad \forall j = H,L \tag{3.50}$$

$$y_{ij} \geq Q - (1 - x_{ij})BigM \qquad \forall i = 1,\ldots,n \quad \forall j = H,L \tag{3.51}$$

$$z_{ik}^{j} \leq \xi_{2|p_2^j}^{k} s_{ik}^{j} \qquad \forall i = 1,\ldots,n \quad \forall j = H, L \quad \forall k = 1,\ldots,n_j \tag{3.52}$$

$$z_{ik}^{j} \leq s_{ik}^{j} \qquad \forall i = 1,\ldots,n \quad \forall j = H, L \quad \forall k = 1,\ldots,n_j \tag{3.53}$$

$$z_{ik}^{j} \geq s_{ik}^{j} - (1 - x_{ij})\xi_{2|p_2^j}^{k} \qquad \forall i = 1,\ldots,n \quad \forall j = H, L \quad \forall k = 1,\ldots,n_j \tag{3.54}$$

$$t_{ij} \leq \xi_1^i x_{ij} \qquad \forall i = 1,\ldots,n \quad \forall j = H, L \tag{3.55}$$

$$t_{ij} \leq s_i^1 \qquad \forall i = 1,\ldots,n \quad \forall j = H, L \tag{3.56}$$

$$t_{ij} \geq s_i^1 - (1 - x_{ij})\xi_1^i \qquad \forall i = 1,\ldots,n \quad \forall j = H, L \tag{3.57}$$

$$y_{ij}, z_{ik}^{j}, t_{ij}, Q, s_i^1, s_{ik}^{j}, u_i \geq 0 \qquad \forall i = 1,\ldots,n \quad \forall j = H, L \quad \forall k = 1,\ldots,n_j \tag{3.58}$$

$$x_{ij} \in \{0, 1\} \qquad \forall i = 1,\ldots,n \quad \forall j = H, L \tag{3.59}$$

The objective function (3.42) is the sum of the first period average profit and the second period average profit. The first two constraints (3.43) and (3.44) together convert the term $s_i^1 = min(Q, \xi_1^i)$ into a linear form. Similarly, the constraints (3.45) and $u_i \geq 0$ in (3.58) with the maximization objective describe the expression of unmet demand $u_i = max(0, \xi_1^i - Q)$. Likewise, conditions (3.46) and (3.47) ensures that the second period sales quantity is equal to the minimum of leftover $Q - s_i^1$ and second period demand observation (i.e, $s_{ik}^{j} = \min\{Q - s_i^1, \xi_{2|p_2^j}^{k}\}$). With the condition (3.48), we enforce that only one marked-down price is chosen for the inventory level $Q - s_i^1$ corresponding to the demand observation ξ_1^i. For linearization of the product $x_{ij}Q$ in the objective function, the constraints (3.49), (3.50), and (3.51) are added. In a similar way, the constraints (3.52), (3.53), and (3.54) are equal to the product $x_{ij}s_{ik}^{j}$. Lastly, the constraints (3.55), (3.56), and (3.57) are used to linearize the product of two decision variables $x_{ij}s_i^1$. The constraints (3.58) and (3.59) stand for the non-negative decision variables and the binary variable respectively.

This model determines one marked-down price for the inventory level $Q - s_i^1$ corresponding to the demand observation ξ_1^i. Although we do not restrict the marked-down price selection of model, it is possible to limit the model about the price decision. For the limited version of mixed integer linear programming formulation, the following constraint should be added.

$$x_{iL} \leq x_{(i+1)L} \qquad \forall i = 1, \ldots, n-1 \tag{3.60}$$

By knowing there exist a threshold for price decision for leftovers, the constraint (3.60) ensures that if the low marked-down price is not selected for inventory level $Q - s^1_{i+1}$, then it is not selected for inventory level $Q - s^1_i$ for all $i = 1, \ldots, n-1$.

The model defined to find the optimal order quantity and the optimal threshold gives a range for the low marked-down price and for the high marked-down price. If the largest inventory level $Q - \xi_1$ that the marked-down price is used is less than the smallest inventory level that the low marked-down price is used, then there exists a gap where there are inventory levels that we did not determine the marked-down price for. We use the following algorithm to find this gap. If $Gap \leq BigM$ and $Gap \geq 1$, then we use the sample average algorithm defined to find threshold and check if the threshold is in this gap.

Figure 3.15 Post-process for MIP that finds solutions simultaneously

$Max \leftarrow 0$, and $Min \leftarrow BigM$
for $i \in \{1 \ldots N\}$ **do**
  if $x_{iH} = 1$ for leftover $Q - \xi^i_1$ **then**
    if $Q - \xi^i_1 > Max$ **then**
      $Max \leftarrow Q - \xi^i_1$
    end if
  else if $x_{iL} = 1$ for leftover $Q - \xi^i_1$ **then**
    if $Q - \xi^i_1 < Min$ **then**
      $Min \leftarrow Q - \xi^i_1$
    end if
  end if
end for
$Gap \leftarrow Min - Max$

3.5 A Heuristic: One Period Linear Programming Formulation for Each Given Marked-down Price

In this section, the heuristic where the order quantity and the marked-down price are determined prior to the beginning of the first period is proposed. It is assumed that the possible marked-down prices are given at the start of the first period.

The following notation is used to describe the new model which is denoted by MIP-GP.

Q:	Order quantity
p_1:	Unit Selling price in period 1
p_2:	Unit marked-down price in period 2
B:	Goodwill/Shortage cost
ξ_i:	The demand in period i=1,2
s_i^1:	Satisfied demand in period 1 when the demand is ξ_1^i
u_i:	Amount of underage or shortage when the demand is ξ_1^i
s_{ij}^2:	Satisfied demand in period 2 when the demand ξ_1^i and ξ_2^j

Given the marked-down price, the linear programming formulation is as follows.

$$\max_{Q} \quad \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}(-vQ + p_1 s_i^1 + p_2 s_{ij}^2 + q(Q - s_i^1 - s_{ij}^2) - Bu_i) \tag{3.61}$$

$$
\begin{aligned}
\text{subject to} \quad & s_i \le Q && \forall i = 1,\dots,n && (3.62)\\
& s_i \le \xi_1^i && \forall i = 1,\dots,n && (3.63)\\
& s_{ij}^2 \le Q - s_i && \forall i = 1,\dots,n \quad \forall j = 1,2,\dots,m && (3.64)\\
& s_{ij}^2 \le \xi_2^j && \forall i = 1,\dots,n \quad \forall j = 1,2,\dots,m && (3.65)\\
& u_i \ge \xi_1^i - Q && \forall i = 1,\dots,n && (3.66)\\
& s_{ij}^2, s_i, u_i \ge 0 && \forall i = 1,\dots,n \quad \forall j = 1,2,\dots,m && (3.67)\\
& && && (3.68)
\end{aligned}
$$

The objective function (3.61) is the sum of the first period average profit and the second period average profit of given marked-down price. In the first period, the description of sales quantity $s_i = min(Q, \xi_1^i)$ is made by the constraints (3.62) and (3.63). Similarly, the satisfied demand for the second period is forced to be the minimum of leftover and the second period demand observation by the constraints (3.64) and (3.65). The unmet demand is determined by the constraints 3.66 and $u_i \ge 0$ in (3.67) with the objective. Lastly, the constraint (3.67) is about the non-negativity of decision variables.

To sum up, the heuristic is to run the model defined above for the high marked-down price as well as the low marked-down price and select the price and order quantity combination that yields to the larger average total profit. This heuristic differs substantially from the model we proposed before in terms of dynamic-decision making process. The threshold is set to the optimal order quantity if the model defined for the low price gives the larger average total profit. Otherwise, the threshold value is set to 1 which means that the high price is determined for all remaining inventory levels.

Chapter 4

Numerical Analysis and Comparison

In this section we provide the numerical study using generated data. We compare the performance of proposed data-driven models for different parameter settings. We describe our parameter setting in order to perform the numerical analysis in Section 4.1. A number of measurements are derived to make sure that the parameter settings are robust. In addition to that, how we generated the data is explained in Section 4.2. Section 4.3 introduces what we want to measure as well as how we conduct the performance test. In particular,

- The data is a function of decision made in the past. Thus, we define the relevancy level of the data according to the way the decision is made. This description is made in three ways. The relevancy level 1 is for the data in which the pricing decision is made as if demand distributions are known. The relevancy level 2 is for the data where the pricing decision is made as if the insights of demand distributions are available. The relevancy level 3 is for the data where the pricing decision is made randomly. Section 4.2 also gives a detail explanation about them. Then, we measure the effects of relevancy of the given data for a given data-driven model.

- The generated data is divided into an in-sample data, which is used to train data-driven models, and an out-of-sample data which is used to test their performance. Hence, we measure the effects of size of the training data for a given data-driven model.

- We carry out a comparison of data-driven models to see how important to select a data-driven model. As a part of comparison, we measure how much the percentage differences of their average profits from expected values and differences in terms of standard deviations. We measure how many times one method is the best among out-of-samples.

After testing proposed data-driven models, results are examined and more general insights of their performance are provided in Section 4.4.

4.1 Parameter Setting

Two different parameter categories are considered, namely low profit and high profit margin cases. The parameter selection is made to meet the conditions in (2.10) and in (2.11). The threshold value for price selection is either zero or infinite if these conditions are not satisfied with selected parameters. Condition in (2.10) ensures that the threshold can be any number except zero. Moreover, the threshold value is forced not to be infinite by condition in (2.11).

In addition to these conditions, we consider a number of measurements for parameter selection. To ensure the robustness, these performance measures are used to determine the value of the second period problem so that we make sure that solving the two period problem is significant in our setting.

Selection of Performance Measures:

The following notations are used to describe the performance measure.

Table 4.1: Notation for the performance measures

Notation	Definition
Q	Quantity ordered in the beginning of period 1
I	Ending inventory for period 1
$\mathbb{E}[\Pi^2(I)\vert p_2^j]$	Expected second period profit given p_2^j as the marked-down price for I $\forall j = H, L$
I_t	Threshold ending inventory
$\mathbb{E}[\Pi(Q, I_t)]$	Expected total profit for the two-period problem

First, we solve the newsvendor problem with a recourse action analytically. We provide a decision rule to determine the marked-down price for the second period by finding the threshold I_t. For continuous demand distributions, the threshold I_t corresponds to the point where the expected profit difference equation in (2.12) is zero. For the discrete demand distributions, it corresponds to the first point where the condition in (2.13) holds.

Note that the lower price p_2^L is determined if $I \geq I_t$, higher price p_2^H is determined if $I < I_t$ according to the decision rule for the second period problem.

Then, we find the order quantity that maximizes the total expected profit by solving the following problem.

$$\max_{Q\geq 0} \mathbb{E}[\Pi(Q, I_t)] \text{ where} \tag{4.1}$$

$$\mathbb{E}[\Pi(Q, I_t)] = -vQ + p_1\mathbb{E}[\min(Q, \xi_1)] - B\mathbb{E}[\max(\xi_1 - Q, 0)]$$
$$+ \mathbb{1}_{\{0\leq \max(Q-\xi_1,0)<I_t\}}\mathbb{E}[\Pi^2(I)|p_2^H] + \mathbb{1}_{\{\max(Q-\xi_1,0)\geq I_t\}}\mathbb{E}[\Pi^2(I)|p_2^L]$$
$$\mathbb{E}[\Pi(Q, I_t)] = -vQ + p_1\mathbb{E}[\min(Q, \xi_1)] - B\mathbb{E}[\max(\xi_1 - Q, 0)]$$
$$+ \mathbb{1}_{\{0\leq \max(Q-\xi_1,0)<I_t\}}\Big(p_2^H\mathbb{E}[\min(Q-\xi_1, \xi_2(p_2^H))] + q\mathbb{E}[\max((Q-\xi_1-\xi_2(p_2^H), 0)]\Big)$$
$$+ \mathbb{1}_{\{\max(Q-\xi_1,0)\geq I_t\}}\Big(p_2^L\mathbb{E}[\min(Q-\xi_1, \xi_2(p_2^L))] + q\mathbb{E}[\max((Q-\xi_1-\xi_2(p_2^L), 0)]\Big)$$

Let Q^* be the optimal solution of problem (4.1).

We consider the following remarks to show some properties as well as how important the second period problem can be.

Remark 1: (Classical Newsvendor-the marked-down prices are disregarded)

We solve the classical newsvendor problem where we do not consider any decision for the second period. In this case, the marked-down prices are disregarded and the remaining inventory at the end of the period is sold at the price we set for the salvage value. There is no knowledge of the second period prices and distributions. Hence, the decision maker does not use those opportunities and salvages leftovers at the end of the period 1 directly. The solution of the following problem gives the lower bound for the newsvendor problem with a recourse action.

$$\begin{aligned}&\max_{Q\geq 0} \mathbb{E}[\Pi(Q, None)] \text{ where} \\ &\mathbb{E}[\Pi(Q, None)] = -vQ + p_1\mathbb{E}[\min(Q, \xi_1)] - B\mathbb{E}[\max(\xi_1 - Q, 0)] \\ &+ q\mathbb{E}[\max((Q - \xi_1, 0)]\end{aligned} \tag{4.2}$$

Let Q^N be the optimal newsvendor solution of the problem (4.2) where the ending inventory is only salvaged.

For given order quantity, Q^N, we solve the newsvendor problem with a recourse action. Then, the difference between this realization of the given order quantity and the optimal order quantity is used to show how it is important not to disregard the marked-down prices. The following formulation that gives the percentage increase in the expected profit when the second period operations are considered.

$$\left(\frac{\mathbb{E}[\Pi(Q^*, I_t)] - \mathbb{E}[\Pi(Q^N, None)]}{\mathbb{E}[\Pi(Q^N, None)]}\right) \times 100 \tag{4.3}$$

Remark 2: (Lower price + demand for higher price)

In addition, we find the optimal solution of the problem where where the decision maker has the information but no threshold value is used for decisions. The retailer sells the remaining inventory at a lower marked-down price p_2^L and observes the demand of higher price and solves the following problem:

$$\begin{aligned}\max_{Q\geq 0} & -vQ + p_1\mathbb{E}[\min(Q,\xi_1)] - B\mathbb{E}[\max(\xi_1 - Q, 0)] \\ & + \left(p_2^L\mathbb{E}[\min(Q-\xi_1, \xi_2(p_2^H))] + q\mathbb{E}[\max((Q-\xi_1-\xi_2(p_2^H), 0)]\right)\end{aligned} \tag{4.4}$$

Let Q^L be the optimal solution of problem (4.4). Q^L is a lower bound for the optimal quantity ordered in the problem (4.1). I.e, $Q^L \leq Q_0^*$. Furthermore,

$$\mathbb{E}[\Pi(Q^L, I_t)] \leq \mathbb{E}[\Pi(Q^*, I_t)] \tag{4.5}$$

The following formulation is used to calculate the expected profit loss if the second period problem is not explicitly considered to solve for quantity.

$$\left(\frac{\mathbb{E}[\Pi(Q^*, I_t)] - \mathbb{E}[\Pi(Q^L, I_t)]}{\mathbb{E}[\Pi(Q^*, I_t)]}\right) \times 100 \tag{4.6}$$

Also note that

$$\mathbb{E}[\Pi(Q^L, I_t)] \geq \mathbb{E}[\Pi(Q^N, None)] \tag{4.7}$$

since Q^L is the tighter lower bound for the optimal order quantity Q^*.

Remark 3: (Classical Newsvendor-salvage value=higher price)

We find the order quantity by solving the classical newsvendor problem where the salvage value is equal to higher price. In this case, the decision maker requires only the knowledge of higher price p_2^H.

Let us define

$$\mathbb{E}[\Pi^1(Q)] = -vQ + p_1\mathbb{E}[\min(Q,\xi_1)] - B\mathbb{E}[\max(\xi_1 - Q, 0)] + p_2^H\mathbb{E}[\max((Q - \xi_1, 0)]$$

and let Q^{U_1} be the optimal solution for the following problem

$$\max_{Q\geq 0} \mathbb{E}[\Pi^1(Q)] \tag{4.8}$$

Q^{U_1} is an upper bound for the solution of the newsvendor problem with a recourse action. It is realistic upper bound if p_2^H is realistic. In other words, $Q^{U_1} \geq Q^*$. Furthermore,

$$\mathbb{E}[\Pi^1(Q^{U_1})] \geq \mathbb{E}[\Pi(Q^*, I_t)] \tag{4.9}$$

Remark 4: (Higher price + demand for lower price)

We assume that the marked-down price is given as the higher price p_2^H with the demand of lower price, and then solve the problem where the remaining inventory at the end of the first period is offered at higher price regardless of the second period decision rule. The decision maker requires the full knowledge of the second period parameters. No threshold value is used for decisions but this is optimistic.

Let us define

$$\mathbb{E}[\Pi^2(Q)] = -vQ + p_1\mathbb{E}[\min(Q,\xi_1)] - B\mathbb{E}[\max(\xi_1 - Q, 0)]$$
$$+ \left(p_2^H \mathbb{E}[\min(Q-\xi_1, \xi_2(p_2^L))] + q\mathbb{E}[\max((Q-\xi_1-\xi_2(p_2^L), 0)] \right)$$

and let Q^{U_2} be the optimal solution of the problem

$$\max_{Q\geq 0} \mathbb{E}[\Pi^2(Q)] \tag{4.10}$$

Q^{U_2} is an upper bound for the solution of the newsvendor problem with a recourse action. I.e, $Q^{U_2} \geq Q^*$. Furthermore,

$$\mathbb{E}[\Pi^2(Q^{U_2})] \geq \mathbb{E}[\Pi(Q^*, I_t)] \tag{4.11}$$

Also note that

$$\mathbb{E}[\Pi^2(Q^{U_2})] \leq \mathbb{E}[\Pi^1(Q^{U_1})] \tag{4.12}$$

and

$$\mathbb{E}[\Pi(Q^{U_2}, I_t)] \geq \mathbb{E}[\Pi(Q^{U_1}, I_t)]$$

since Q^{U_2} is a tighter upper bound than Q^{U_1} for the optimal order quantity Q^*.

Remark 5: (Potential of solving 2^{nd} period problem)

Both Remark 3 and Remark 4 consider the potential of gains of we solve the problem considering the second period decisions in an implicit way.

(1) Bounds on the order quantity. (Importance for computing the optimal order quantity)

(2) The longer the difference in the upper bound and the lower bound, the longer potential we have for learning about the second period parameters, as well as finding the threshold for decision making.

If parameters of the second period problem other than p_2 is not known, the potential of investigating/estimating those can be calculated using the less tighter bounds:

$$\left(\frac{\mathbb{E}[\Pi^1(Q^{U_1})] - \mathbb{E}[\Pi(Q^N, None)]}{\mathbb{E}[\Pi(Q^N, None)]}\right) \times 100 \tag{4.13}$$

The expected profit percentage difference as given above states the importance of knowing the second stage best solution to decide on the remaining quantity.

On the other hand, the following formulation is another indication of the value of solving the second period problem. Note that if this quantity is small, the importance of the second period problem is small.

$$\left(\frac{\mathbb{E}[\Pi^2(Q^{U_2})] - \mathbb{E}[\Pi(Q^L, I_t)]}{\mathbb{E}[\Pi(Q^L, I_t)]}\right) \times 100 \tag{4.14}$$

To summarize, we propose four measures to test robustness of the parameters we used in the numerical analysis. The formula (4.3) is used for the percentage benefit of solving two stage problem. Other formula in (4.6) is used for the percentage expected profit loss of incorrect order quantity. the potential of learning more about Period 2 calculated by the equation (4.13), and (4.14) is used to find the potential of solving Period 2.

Implementation of the Selected Measures for the Parameter Setting Used

Considering the performance measures under the constraints (2.10)-(2.11), parameters are selected as follows. We think that the key parameters are μ_1, μ_2^H, and μ_2^L describing the mean of the demand distributions. So, we first set the remaining parameters p_1, q, v, p_2^H, p_2^L and B to a value and find μ_1, μ_2^H, and μ_2^L, accordingly.

Table 4.2: Parameter values used in numerical experiments

parameter	p_1	q	v	p_2^H	p_2^L	B
Low Profit Margin	1	0.1	0.7	0.6	0.4	$\{0, 0.2\}$

parameter	p_1	q	c	p_2^H	p_2^L	B
High Profit Margin	1	0.1	0.5	0.4	0.25	$\{0, 0.5\}$

Parameter values are given in Table 4.2. We determine the effects of no unmet demand penalty cost. Thus, we have two alternatives for penalty.

Table 4.3: Mean of demand distributions for generation

	μ_1	μ_2^H	μ_2^L
Option 1	80	10	60
Option 2	40	5	30

We consider two options for the mean of each demand distribution (see Table 4.3). By combining the values in Tables 4.2 and 4.3, we define the parameter setting shown in Table 4.4 to use in the numerical experiments.

Table 4.4: Parameter Setting used in numerical experiments

parameter	p_1	q	v	p_2^H	p_2^L	B	μ_1	μ_2^H	μ_2^L
1	1	0.1	0.7	0.6	0.4	0	80	10	60
2	1	0.1	0.7	0.6	0.4	0	40	5	30
3	1	0.1	0.7	0.6	0.4	0.2	80	10	60
4	1	0.1	0.7	0.6	0.4	0.2	40	5	30
5	1	0.1	0.5	0.4	0.25	0	80	10	60
6	1	0.1	0.5	0.4	0.25	0	40	5	30
7	1	0.1	0.5	0.4	0.25	0.5	80	10	60
8	1	0.1	0.5	0.4	0.25	0.5	40	5	30

The data is generated from different demand distributions. Namely, we want to use Poisson from deterministic demand distributions, uniform and triangular from continuous demand distributions. Symmetric, right-skewed and left-skewed triangular distributions are used by varying the parameters when we generate data from triangular distribution. In addition to this, we also consider a case where the first period demand is symmetric triangular distributed and the second period demand is uniform distributed. As a result, we are using 12 different demand models:

Table 4.5: Demand Options

#	Expected Demand Options	ξ_1
1	Poisson(μ)	Pois(80)
2	Poisson(μ)	Pois(40)
3	U$(0, 2\mu)$	U(0,160)
4	U $(0, 2\mu)$	U(0,80)
5	Trian$(0, \mu, 2\mu)$	Trian(0,80,160)
6	Trian$(0, \mu, 2\mu)$	Trian(0,40,80)
7	Trian$(0, 3/5\mu, 12/5\mu)$	Trian(0,48,192)
8	Trian$(0, 3/5\mu, 12/5\mu)$	Trian(0,24,96)
9	Trian$(0, 9/7\mu, 12/7\mu)$	Trian(0,102.85,137.14)
10	Trian$(0, 9/7\mu, 12/7\mu)$	Trian(0,51.42,68.57)
11	Trian$(0, \mu, 2\mu)$-U$(0, 2\mu)$	Trian(0,80,160)
12	Trian$(0, \mu, 2\mu)$-U$(0, 2\mu)$	Trian(0,40,80)

Table 4.6: Demand Options

#	Expected Demand Options	ξ_2^H	ξ_2^L
1	Poisson(μ)	Pois(10)	Pois(60)
2	Poisson(μ)	Pois(5)	Pois(30)
3	U$(0, 2\mu)$	U(0,20)	U(0,120)
4	U $(0, 2\mu)$	U(0,10)	U(0,60)
5	Trian$(0, \mu, 2\mu)$	Trian(0,10,20)	Trian(0,60,120)
6	Trian$(0, \mu, 2\mu)$	Trian(0,5,10)	Trian(0,30,60)
7	Trian$(0, 3/5\mu, 12/5\mu)$	Trian(0,6,24)	Trian(0,36,144)
8	Trian$(0, 3/5\mu, 12/5\mu)$	Trian(0,3,12)	Trian(0,18,72)
9	Trian$(0, 9/7\mu, 12/7\mu)$	Trian(0,12.85,17.14)	Trian(0,77.14,102.85)
10	Trian$(0, 9/7\mu, 12/7\mu)$	Trian(0,6.42,8.57)	Trian(0,38.57,51.42)
11	Trian$(0, \mu, 2\mu)$-U$(0, 2\mu)$	U(0,20)	U(0,120)
12	Trian$(0, \mu, 2\mu)$-U$(0, 2\mu)$	U(0,10)	U(0,60)

For each demand model in Tables 4.5 and 4.6, the historical data is generated by using parameters in Table 4.2.

Given the parameters of the problem and the demand distributions, we find two different ranges, one of which is tighter, not only for the optimal order quantity but also for the optimal objective function. Ranges which may help to find the optimal order quantity are given in Appendix G. All observed bounds for the solution to the newsvendor problem with recourse action follows inequalities in (4.5), (4.7), (4.9), and (4.11) which are mentioned in remarks. Moreover, the results of our four performance measures for determined parameters are given in Tables 4.7-4.14. We use MATLAB to apply these performance measures for parameters that we select. The formula (4.3) is used for the percentage benefit of solving the two-stage problem. Other formulas (4.6), (4.13), and (4.14) are used for the percentage expected profit loss of incorrect order quantity, the potential of learning more about Period 2 and the potential of solving Period 2 respectively.

Table 4.7: Value of performance measures when p_1=1,q=0.1,v=0.7, p_2^H=0.6,p_2^L=0.4,B=0,μ_1=80,μ_2^H=10,μ_2^L=60

	Demand Option					
	Poisson	Uniform	Symmetric Triangular	Right-Skewed Triangular	Left-Skewed Triangular	Combination of Symmetric Triangular and Uniform
% Benefit of Two Stage Problem	5.59	37.32	21.03	22.72	21.00	18.72
% Expected Profit Loss of Incorrect Q	0.27	3.58	2.30	2.33	2.02	1.75
% Potential of Learning More of Period 2	8.29	125.01	51.34	65.16	50.01	51.34
% Potential of Solving Period 2	2.84	30.36	18.91	21.80	18.48	16.14

Table 4.8: Value of performance measures when p_1=1,q=0.1,v=0.7, p_2^H=0.6,p_2^L=0.4,B=0,μ_1=40,μ_2^H=5,μ_2^L=30

	Demand Option					
	Poisson	Uniform	Symmetric Triangular	Right-Skewed Triangular	Left-Skewed Triangular	Combination of Symmetric Triangular and Uniform
% Benefit of Two Stage Problem	6.83	37.33	21.03	22.69	20.99	18.73
% Expected Profit Loss of Incorrect Q	0.19	4.16	2.29	2.33	2.02	1.76
% Potential of Learning More of Period 2	12.25	125.03	51.36	65.16	50.00	51.36
% Potential of Solving Period 2	5.27	31.14	18.91	21.83	18.47	16.14

Table 4.9: Value of performance measures when p_1=1,q=0.1,v=0.7, p_2^H=0.6,p_2^L=0.4,B=0.2,μ_1=80,μ_2^H=10,μ_2^L=60

	Demand Option					
	Poisson	Uniform	Symmetric Triangular	Right-Skewed Triangular	Left-Skewed Triangular	Combination of Symmetric Triangular and Uniform
% Benefit of Two Stage Problem	8.30	217.71	42.29	68.12	38.95	37.85
% Expected Profit Loss of Incorrect Q	0.51	7.59	2.80	5.10	2.51	2.18
% Potential of Learning More of Period 2	12.58	694.51	102.43	192.60	88.49	102.43
% Potential of Solving Period 2	4.48	70.37	29.42	44.35	26.62	25.92

From the Tables 4.7-4.14, it is observed that the highest benefit of the solving two-stage problem is always when the demand are uniform distributed. It varies between 14.03 % and 217.94 % for uniform demand. When the demand distribution is Poisson, the benefit of solving two-stage problem is always lowest, in contrast to uniform demand, and it's range for discrete demand is between 3.23 % and 10.69 %. According to the calculations for the expected profit loss of incorrect order quantity, the biggest loss occurs when the demand is right-skewed triangular distributed or uniform distributed. It's values are between 0.05 % and 7.59 % across all demand distributions. Consistently, we can say that the case where the demand is uniform distributed always gives the biggest percentage value for the potential of investigating more about the parameters of the second period problem and also for the value of solving the second period problem. Given the high marked-down price p_2^H, the potential for other second period parameter estimation among all demand models changes between 4.96 % and 695.06 % which means that there is a benefit of knowing the solution to the second period problem. Moreover, the results of the formula 4.14, which is another demonstration for the benefit of solving Period 2, change from 2.84 % to 70.37 %. Lastly, it can be seen that the results obtained from each formula increase when the penalty cost B is increased.

To conclude, the parameter setting is defined and verified according to our four performance measures. We are satisfied with the parameters selected and think that we have reasonable system to fit the two-period problem framework considered.

Table 4.10: Value of performance measures when p_1=1,q=0.1,v=0.7, p_2^H=0.6,p_2^L=0.4,B=0.2,μ_1=40,μ_2^H=5,μ_2^L=30

	Demand Option					
	Poisson	Uniform	Symmetric Triangular	Right-Skewed Triangular	Left-Skewed Triangular	Combination of Symmetric Triangular and Uniform
% Benefit of Two Stage Problem	10.69	217.94	42.24	68.14	38.89	37.85
% Expected Profit Loss of Incorrect Q	1.07	6.47	3.50	5.09	3.13	2.83
% Potential of Learning More of Period 2	19.11	695.06	102.43	192.75	88.46	102.43
% Potential of Solving Period 2	8.78	68.29	30.40	44.37	27.47	26.76

Table 4.11: Value of performance measures when p_1=1,q=0.1,v=0.5, p_2^H=0.4,p_2^L=0.25,B=0,μ_1=80,μ_2^H=10,μ_2^L=60

	Demand Option					
	Poisson	Uniform	Symmetric Triangular	Right-Skewed Triangular	Left-Skewed Triangular	Combination of Symmetric Triangular and Uniform
% Benefit of Two Stage Problem	3.23	14.03	9.05	10.92	8.85	8.20
% Expected Profit Loss of Incorrect Q	0.11	0.61	0.57	0.72	0.48	0.46
% Potential of Learning More of Period 2	4.94	50.00	24.78	34.59	22.52	24.78
% Potential of Solving Period 2	1.76	14.56	10.18	12.57	9.41	9.00

Table 4.12: Value of performance measures when p_1=1,q=0.1,v=0.5, p_2^H=0.4,p_2^L=0.25,B=0,μ_1=40,μ_2^H=5,μ_2^L=30

	Demand Option					
	Poisson	Uniform	Symmetric Triangular	Right-Skewed Triangular	Left-Skewed Triangular	Combination of Symmetric Triangular and Uniform
% Benefit of Two Stage Problem	3.83	14.03	9.05	10.92	8.84	8.19
% Expected Profit Loss of Incorrect Q	0.05	0.78	0.57	0.91	0.47	0.46
% Potential of Learning More of Period 2	7.21	50.01	24.79	34.61	22.51	24.79
% Potential of Solving Period 2	3.31	14.76	10.18	12.80	9.41	9.00

Table 4.13: Value of performance measures when p_1=1,q=0.1,v=0.5, p_2^H=0.4,p_2^L=0.25,B=0.5,μ_1=80,μ_2^H=10,μ_2^L=60

	Demand Option					
	Poisson	Uniform	Symmetric Triangular	Right-Skewed Triangular	Left-Skewed Triangular	Combination of Symmetric Triangular and Uniform
% Benefit of Two Stage Problem	4.65	25.86	15.23	21.79	13.00	13.88
% Expected Profit Loss of Incorrect Q	0.19	0.66	0.56	0.88	0.49	0.45
% Potential of Learning More of Period 2	7.42	90.91	42.07	70.06	32.49	42.07
% Potential of Solving Period 2	2.84	22.35	15.15	20.65	12.38	13.75

Table 4.14: Value of performance measures when p_1=1,q=0.1,v=0.5, p_2^H=0.4,p_2^L=0.25,B=0.5,μ_1=40,μ_2^H=5,μ_2^L=30

	Demand Option					
	Poisson	Uniform	Symmetric Triangular	Right-Skewed Triangular	Left-Skewed Triangular	Combination of Symmetric Triangular and Uniform
% Benefit of Two Stage Problem	5.72	25.85	15.22	21.79	12.96	13.87
% Expected Profit Loss of Incorrect Q	0.22	0.44	0.80	0.87	0.24	0.67
% Potential of Learning More of Period 2	11.03	90.91	42.07	70.09	32.49	42.07
% Potential of Solving Period 2	5.26	22.08	15.44	20.64	12.14	14.00

4.2 Data Generation

In this subsection, we explain three different data sets and what is required to generate these data sets. The data contains how much the product was sold during the first period, marked-down price determined for the remaining inventory, as well as how much the product was sold during the second period. Each row of the data looks like $\xi_1, p_2, \xi_2(p_2)$.

Data Description:

We define the relevancy of data using three levels. Relevancy levels differ depending on how second period decision is handled. We have three different data sets based on data relevance.

We consider a company which keeps the data so that it can be used for decision making. The first data set is created as if it were stored by a company that uses the analytical results to make better decisions. In other words, the company has an accurate, sufficient and perfect data. Therefore, the data with relevancy level 1 and the second period optimization we carry out are in full accordance.

For relevancy level 1, the data is generated by implementing the optimal pricing decision.

The second data set created as if it were recorded by a company that chooses an accurate ordering amount according to its insights for the long run by knowing demand distribution. Hence, the data with relevancy level 2 and second period optimization we carry out are probabilistically in accordance. Price decisions are independent of a realization, but probability of taking a correct action reflects "optimal" solution. In other words, the company has the long run probabilities that help to handle the second period decision. Probabilities that the company calculated are correct yet they do not have to be correct at the right time. To generate the past data, therefore, we calculated the probability of being used for each marked down price. As a result, the data with relevancy level 2 is generated by implementing the partially reasonable pricing decision.

The last data set is generated randomly. Actually, we generate the data such that we take either price equally likely. In this case, we assumed that the company does not know how to handle the second period decision and wants to choose the high marked down price with probability 0.5 and the low marked down price with probability 0.5. Thus, the data with relevancy level 3 and second period optimization we carry out are not related. For relevancy level 3, the data is generated by implementing the random pricing decision.

To summarize, the relevance of data is specified with respect to how it was obtained with what kind of a decision making process. These three data sets defined above are noticed by $Data_1, Data_2$, and $Data_3$, respectively.

To generate the data sets defined above, exact solutions to the newsvendor problem with a recourse action are needed. Therefore, we obtain the optimal order quantity and the threshold value that is used to decide marked down price for each demand distribution assumption by using MATLAB. There are eight different results to be input for data generation of each desired distribution. The results are shown in Appendix H, Table H.1 and H.2. Moreover, we generate a bunch of U(0,1) random numbers and store them to use in data generation. We applied the basic known tests in [37] to the random number generator before we start to generate our data sets. Chi square test is one of the empirical test that we implemented to test uniformly distributed random variates U_i's. However, we also implement different kinds of tests such as serial, runs-up, and computation of correlation so as to make sure that this random number generator is "good" enough. These tests for random number generator are applied in MATLAB. The test statistics results related to the random number generator are given in Appendix I.1, Tables I.1, I.2, and I.3. Since the results are reasonable, we can trust the random number generator that is used.

Use of Random Numbers:

Let us say that J is the size of the sample and M is the number of replication. We need $2JM + 1$ random number streams for $Data_2, Data_3$ and $2JM$ random number streams for $Data_1$. Each stream which we want to control actually generates an $U \sim U(0, 1)$ value. By doing this, we try to reduce variance of an output random variable without disturbing its expectation and use the correlated variates. This is known as common random number technique which is one of the variance-reduction techniques (VRT). To make sure that we generate problems with positive correlation among the generated values, for each data set we use the same U value for the same purpose. Specifically, let u_{jk} be a vector that holds the U(0,1) random numbers used to generate the j^{th} sample for the k^{th} purpose where $j = 1, \dots, J$ and $k = \xi_1, p_2, \xi_2$. The following table shows how we are going to use the U(0,1) random numbers in data generation:

Table 4.15: U(0,1) random numbers that are used in generation

Sample Number	ξ_1	p_2	ξ_2
1	$u_{1\xi_1}$	u_{1p_2}	$u_{1\xi_2}$
2	$u_{2\xi_1}$	u_{2p_2}	$u_{2\xi_2}$
3	$u_{3\xi_1}$	u_{3p_2}	$u_{3\xi_2}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
J	$u_{J\xi_1}$	u_{Jp_2}	$u_{J\xi_2}$

Note that the above generation can also be used if we have different alternatives for parameters in Table 4.2 and demand rates in Table 4.3. We first generate all the U(0,1) random numbers for 500 samples for each of the 50 replications and store them. Then for ξ_1 and ξ_2, we transform these U_i's to obtain a realization of demand. We use the same U(0,1) random numbers for all approach we want to test so that we reduce the variance for comparison. Also, they can be used if we change any of the parameters. For generation of Poisson, Uniform and Triangular random variates, the algorithms in [37] are applied.

In the following algorithms, we outline the steps in Figures 4.1, 4.2, and 4.3 to generate the data sets in details. We use MATLAB to generate the data for each relevancy level.

Figure 4.1 $Data_1$: Data with the relevancy level 1

1: Compute the optimal order quantity $Q_0^\star$ and threshold I_1^t inventory level using the equations (2.25), (2.26), (2.12) and (2.13) under the assumption that demand distributions are known.

2: Generate ξ_1

3: I=$(Q^\star - \xi_1)^+$. Price follows using the comparison of Threshold I^t with Inventory I

4: Generate ξ_2 using the distribution for the determined price.

Figure 4.2 *Data*$_2$: Data with the relevancy level 2

Compute the optimal order quantity $Q_0^\star$ and threshold I_1^t inventory level using the equations (2.25), (2.26), (2.12) and (2.13) under the assumption that demand distributions are known.

Compute P_1 = Probability of using p_2^H using the cumulative distribution function F_{ξ_1} in equations (2.25), (2.26).

Similarly, compute P_2 = Probability of using p_2^L. Note that there will be cases where ending inventory value will be zero and hence no price is needed. Hence $P_1 + P_2 < 1$.

Compute $q_1 = P_1/(P_1 + P_2)$ and $q_2 = 1 - q_1$

Generate ξ_1

Generate U

if $U < q_1$ **then**

    $p_2^\star \leftarrow p_2^H$

else

    $p_2^\star \leftarrow p_2^L$

end if

Generate ξ_2 using the distribution for the determined price.

Figure 4.3 *Data*$_3$: Data with the relevancy level 3

Generate ξ_1

Generate U

if $U < \frac{1}{2}$ **then**

    $p_2^\star \leftarrow p_2^H$

else

    $p_2^\star \leftarrow p_2^L$

end if

Generate ξ_2 using the distribution for the determined price.

4.3 Testing Procedure/Experimental Design

This section provides the outline of the test conducted for the performance of proposed data-driven models. We first analyze the effect of the relevancy level of data which the approaches work with and then the the effect of the sample size. When we do the statistical comparison of models, we perform three different tests. The first one is for the importance of selecting a data-driven model. Another way to do comparison is considering the difference from the expected values. Lastly, we examine the number of times each method is best in samples. In our tests, we take the advantage of generating random varieties utilizing common random number method (CRN) so to decrease the overall variability. In the following part, we describe how we accumulate the data to perform the analysis.

The Effect of Relevancy of the Given Data

For a given demand, parameter setting and train data size, we define the following notations.

- $P(1, j, r)$ is the profit obtained by relevancy level 1 for the j^th test period ($j = 1, \ldots, 50$) for replication $r = 1, \ldots, 50$.
- $P(2, j, r)$ is the profit obtained by relevancy level 2 for the j^th test period ($j = 1, \ldots, 50$) for replication $r = 1, \ldots, 50$.
- $P(3, j, r)$ is the profit obtained by relevancy level 3 for the j^th test period ($j = 1, \ldots, 50$) for replication $r = 1, \ldots, 50$.

We compute the average profit obtained by relevancy level 1,2 and 3 for replication $r = 1, \ldots, 50$ over all 50 test periods, respectively. In other words,

$$AP(1,r) = \frac{1}{50}\sum_{j=1}^{50} P(1,j,r) \forall r = 1,\ldots,50$$
$$AP(2,r) = \frac{1}{50}\sum_{j=1}^{50} P(2,j,r) \forall r = 1,\ldots,50$$
$$AP(3,r) = \frac{1}{50}\sum_{j=1}^{50} P(3,j,r) \forall r = 1,\ldots,50$$

Next, we calculate the difference between the average profit obtained by relevancy level 1 and 2. Likewise, the difference between the average profit obtained from relevancy level 1 and 3 is calculated:

$$D(1-2,r) = AP(1,r) - AP(2,r)$$
$$D(1-3,r) = AP(1,r) - AP(3,r)$$

Lastly, we conduct two different one sample t test for the data set obtained for $r = 1,\ldots,50$. We have 50 data points for each test. The purpose of each test is to statistically check whether the defined difference is significantly different than zero for the 95 % confidence interval. The two tests are: "Relevancy Level 1 compared to Relevancy Level 2" and "Relevancy Level 1 compared to Relevancy Level 3". We record the significance level where the result is significantly different than 0.

We again use the significance testing approach and report the p-values in this test. The significance level is 0.05 and p-value smaller than 0.05 is a significant test result. We interpret small p-value as meaning that the defined difference is significantly different than zero for the 95 % confidence interval. As a result, the test is said to be statistically significant at the significance level $\alpha = 0.05$ if the p-value is smaller than $\alpha = 0.05$.

We only use the data with relevancy level 1, $Data_1$, to conduct other tests.

The Effect of the Sample Size

After generation of all data sets, we divide this generated data set into a training set and test set. We increase the size of the learning data set while keeping the test data constant to see the effect of the sample size on the results. More specifically:

(a) **SET1-Small learning data:** We use $j = 1, \ldots, 50$ to determine Q and rule for selecting second period price and $j = 451, \ldots, 500$ to test.

(b) **SET2-Medium learning data:** We use $j = 1, \ldots, 250$ to determine Q and rule for selecting second period price and $j = 451, \ldots, 500$ to test.

(c) **SET3-Large learning data:** We use $j = 1, \ldots, 450$ to determine Q and rule for selecting second period price and use $j = 451, \ldots, 500$ to test.

We compute the profit associated with each sample in each set and keep it for testing. Effects of size of the training data is analyzed in this section. The test is conducted for only $Data_1$ (relevancy level 1). To do so, we define the following three profit function for a given data-driven model, and for each parameter setting.

- $P(1, j, r)$ is the profit obtained for training data size $= 50$ for the j^{th} test period ($J = 1, \ldots, 50$) for replication $r = 1, \ldots, 50$
- $P(2, j, r)$ is the profit obtained for training data size $= 250$ for the j^{th} test period ($J = 1, \ldots, 50$) for replication $r = 1, \ldots, 50$
- $P(3, j, r)$ is the profit obtained for training data size $= 450$ for the j^{th} test period ($J = 1, \ldots, 50$) for replication $r = 1, \ldots, 50$

After obtaining profits, we compute the average profit for each training data, and for each replication $r = 1, \ldots, 50$ over all 50 test periods. In other words, we calculate the following equations

- $$AP(1,r) = \frac{1}{50}\sum_{j=1}^{50} P(1,j,r) \quad \forall r = 1,\dots,50$$
  is the average profit obtained for training data size = 50 for replication $r = 1,\dots,50$ over all 50 test periods.

- $$AP(2,r) = \frac{1}{50}\sum_{j=1}^{50} P(2,j,r) \quad \forall r = 1,\dots,50$$
  is the average profit obtained for training data size = 250 for replication $r = 1,\dots,50$ over all 50 test periods.

- $$AP(3,r) = \frac{1}{50}\sum_{j=1}^{50} P(3,j,r) \quad \forall r = 1,\dots,50$$
  is the average profit obtained for training data size = 450 for replication $r = 1,\dots,50$ over all 50 test periods.

Then, we look at the difference between the average profit for training size 50 and the average profit for training size 250 and the average profit for training size 450, respectively. The differences are denoted by (1,2), (1,3) and (2,3) respectively. Thus, we compute

$$D(1-2,r) = AP(1,r) - AP(2,r) \forall r = 1,\dots,50$$
$$D(1-3,r) = AP(1,r) - AP(3,r) \forall r = 1,\dots,50$$
$$D(2-3,r) = AP(2,r) - AP(3,r) \forall r = 1,\dots,50$$

We conduct three different one sample t test for the data set obtained for $r = 1,\dots,50$. We have 50 data points for each test. The purpose of each test is to statistically check whether the defined difference is significantly different than zero. After performing one sample t test, we record the significance level where the result is significantly different than 0.

In this test, we use the significance testing approach and report the p-value which is the level of statistical significance. The significance level is 0.05 and p-value smaller than 0.05 is a significant test result. We interpret small p-value as meaning that the defined difference is significantly different than zero for the 95 % confidence interval. As a result, the test is said to be statistically significant at the significance level $\alpha = 0.05$ if the p-value is smaller than $\alpha = 0.05$.

Profit comparison of data-driven models

We compare data-driven models in three different ways. First of all, we compare the best data-driven model with the worst one. To do so, we define the following notations for a given demand and for a parameter setting and for each train data size.

- $P(B, j, r)$ is the largest profit obtained by any method for the j^th test period ($J = 1, \ldots, 50$) for replication $r = 1, \ldots, 50$.
- $P(W, j, r)$ is the lowest profit obtained by any method for the j^th test period ($J = 1, \ldots, 50$) for replication $r = 1, \ldots, 50$.

Then, we compute

$$D(B - W, j, r) = P(B, j, r) - P(W, j, r) \forall j, r = 1, \ldots, 50$$

We have 50 data points for each test representing one replication. We conduct one sample t test for the data set obtained for $j = 1, \ldots, 50$. The purpose of each test is to statistically check whether the defined difference is significantly different than zero. We record for each case: if the result is within 90% CI set 1, otherwise 0, and if the result is within 99% CI set 1, otherwise 0.

We perform one sample t test for each replication and for each train data set size at significance level 0.1 and 0.01, respectively. The tests are conducted for only the data with relevancy level 1. 0 means that the defined difference is significantly different than zero. Since we have 50 replications and conduct this test for each replication, the largest number that we can obtain is 50.

Comparison with respective expected values

As a second way to compare data-driven models, we compute the percentage difference between the expected profit under the assumption that the demand is known beforehand and the profit achieved by data-driven models when the demand is unknown. Also we calculated the inverse of coefficient of Variation for the difference for comparison of data-driven models.

We perform this test for a given demand model, parameter set and train data size. The test is conducted for only the data with relevancy level 1.

For a given train data size, the following notations are defined.

- $P(k, j, r)$ is the profit obtained by method k for the j^th test period ($J = 1, \ldots, 50$) for replication $r = 1, \ldots, 50$.
- $P(k, r)$ is the average profit obtained by method k for the for replication r (Average over all $J = 1, \ldots, 50$).
- $P(k)$ is the average profit obtained by method k (Average over all $J = 1, \ldots, 50$ and $r = 1, \ldots, 50$).
- $StD(k)$ is the standard deviation estimate for the average profit obtained by method k overall replications.

and let E be the theoretically computed expected profit.

We define the inverse of coefficient of variation for the difference whose formulation is as follows.

$$CoVd(k) = \frac{E - P(k)}{StD(k)} \tag{4.15}$$

The average profits from expected value and the formula in (4.15) are observed for each train data set size, demand model and parameter setting. The small value of inverse CovD for a model is small means that the variance of the average profit obtained by the model is low. Thus, it can be said that a model is stable if its inverse CovD is small.

Number of times one method is the best

For a given demand model, parameter setting and train data size, the following steps are repeated.

We record the count the method that gives the maximum for the j^th test period ($J = 1, \ldots, 50$) and for replication $r = 1, \ldots, 50$. Let $Max(k)$ be the number of times method k is selected as maximum. If we have more than one method yielding the same result, then we count $Max(k)$ for all the methods yielding the result.

After finishing counting, the percentage success is calculated for a model by considering how many instances the method is the best out of 2500 samples. The test is conducted for only the data with relevancy level 1. Based on the percentage success, the comparison of models is made.

4.4 Discussion of Results and Main Findings

In the previous subsection 4.3, we implement data-driven models proposed by using Eclipse IDE for Java and RStudio. We select eight different parameter settings, and six different demand distributions for the data. Since we define three different relevancy levels, the data is generated for each of them. The data includes 500 samples for each of 50 replications. Therefore, we run data-driven models with 7200 (3x6x8x50) data sets for each training data size. There are total of 21600 instances. We train models with the training data of size 50, 250 and 450 respectively. During the training stage, The time of data-driven models to solve samples is restricted to 20 minutes as there is a bunch of samples. All data-driven models except MIP-SAA and MIP-P can solve samples within 20 minutes. We observe that MIP-SAA and MIP-P can solve samples in 20 minutes if the training data size is 50. For the samples when the training data size is 250 or 450, MIP-SAA and MIP-P has to find solutions within 20 minutes although they need more time. For MIP-SAA, 46 % of instances are solved to optimality in our termination limit of 20 minutes. For MIP-P, 35 % of instances are solved to optimality in our termination limit of 20 minutes is 7467. We record minimum gap, average gap, maximum gap and standard deviation of gap values for other instances which are shown in Table 4.16.

Table 4.16: Average gap of MIP-SAA and MIP-P models over instances not solved to optimality in 20 minutes

Models	Min Gap	Average Gap	Max Gap	Standard Deviation
MIP-SAA	0.01 %	1.19 %	74.98 %	0.02
MIP-P	0.01 %	25.94 %	329.37 %	0.27

After training the models and obtaining the order quantity and the threshold value estimated by each model, we test models with each sample in the testing data. Each testing data is composed of 50 samples for each of 50 replications. Hence, we use the estimated solutions obtained from each training data with each size in testing stage and we test model for each solution with 50 samples of testing data and obtain the average profit value. We conduct the tests explained in 4.3 by using RStudio.

The following case can arise when we are testing the models: Considering the sample in the test data set and the estimated decision variables, the model may want to choose a price different than the marked-down price in testing sample depending on the leftover. In such cases, one should change the marked-down price in testing sample according to the threshold value estimated by model. Furthermore, the one should change not only the marked-down price but also the second period demand accordingly. This is because second period demand would be less if the high price is selected, or the second period demand would be higher when the low price is selected. To deal with this case and to obtain right average profit value, we calculate a ratio by using the demand of the high price and the lower price in the training data. Thus, we use the estimated ratio to forecast the new demand of the second period in the testing sample according to the marked-down price decision. The algorithm used to change the price and the second period demand when needed is given in the Appendix I.2.

In this section, an overall discussion is made by using the obtained numerical results. The performance of data-driven models is evaluated from different perspectives. We make the following observations:

The Effect of Relevancy of the Given Data

To investigate the effect of the relevancy level of the data, another test is conducted. We obtain p-value between zero and one, which expresses the level of statistical significance and give the results in Appendix I.4, in Tables I.10-I.57. We test if the difference between the relevancy level 1 and 2 is zero or not. Similarly, we conduct a test to check whether the difference between the relevancy level 1 and 3 is zero. A p-value less than 0.05 is statistically significant. At significance level of 0.05, significant test results (or p-value less than 0.05) are counted to calculate the percentage importance about the difference between the average profits obtained by relevancy levels. The number of observations of averages is 384. The number of significant results and percentage values are given in Table 4.17. We observe that there is a significant difference between the relevancy level 1 and 2 in 9 % of all results. In 28 % of all results, it is observed that performance of proposed data-driven models that work with the relevancy level 1, is different than the relevancy level 3 at significance level of 0.05.

Table 4.17: At significance level of 0.05, percentage of significant test results for the effect of relevancy level(percentage of p-values which is smaller than $\alpha = 0.05$ is given for each training data size)

Comparison	The relevancy level 1 and 2	The relevancy level 1 and 3
Training Data Size 50	7.29 %	24.74 %
Training Data Size 250	8.59 %	30.47%
Training Data Size 450	10.16 %	28.13%

When this comparison is made considering the training data size, we make the following observations. At significance level of 0.05, we can reject that the difference defined for the relevancy level 1 and 2 is zero in 7 % of the results when the training data size 50. When the training data size is 250, this difference is more noticeable as the number of p-values smaller than 0.05 is increased and 9 % of results are statistically significant. Furthermore, it is observed that there is a difference between the relevancy level 1 and 2 in 10 % of the results for the training data size 450, statistically. Although the number of significant p-values for the comparison of the relevancy level 1 and 2 increase as the training data size is increased, it may not be large enough to conclude that there is a difference between working with the relevancy level 1 and 2.

We make the following observations when we compare the relevancy level 1 and 3 based on the training data size. For the training data size 50, it is observed that our approaches perform different when the relevancy level is 1 instead of 3 in 25 % of the results. On the other hand, the difference between the relevancy level 1 and 3 is observed statistically in 30 % of the results if the training data set is 250. Likewise, we also observe that the difference between the relevancy level 1 and 3 is significant in 28 % of results when the training data size is 450. Therefore, we can say that the performance of data-driven models for the relevancy level 1 increases if we use training data size 250 or 450 instead of 50.

Then, we look at the differences defined for the relevancy levels according to data-driven models in Table 4.18. The number of observations of averages is 144. We make the following observations. For SAA-P, working with the relevancy level 1 and 2 is different significantly in 6 % of the results. For only MIP-SAA, this number of p-values smaller than significance level 0.05 indicating the difference for the relevancy level 1 and 2 decreases and we observe significant p-values in 2 % of results. We observe this difference statistically in 10 % of results for SAA-QR and for SAA-LR. There is a difference between the relevancy level 1 and 2 significantly in 10 % of the results for MIP-QR and in 9 % of the results for MIP-LR, respectively. Although this difference is observed for MIP-GP in 4 % of the results, it is observed in 17 % for MIP-P at significance level 0.05. We can say that selecting the relevancy level that MIP-P works with as 1 or 2 is more important when we compare it with the others.

Table 4.18: At significance level of 0.05, percentage of significant test results for the effect of relevancy level. (percentage of p-values which is smaller than $\alpha = 0.05$ is given for each model)

Data-driven Models	The relevancy level 1 and 2	The relevancy level 1 and 3
SAA-P	6.25 %	30.56 %
MIP-SAA	2.08 %	32.64 %
SAA-QR	10.42 %	23.61 %
MIP-QR	9.72 %	23.61 %
SAA-LR	10.42 %	22.92 %
MIP-LR	9.03 %	22.22 %
MIP-GP	4.17 %	30.56 %
MIP-P	17.36 %	36.11 %

In contrast to the comparison for the relevancy level 1 and 2, the number of significant test results at significance level 0.05 which we can see difference defined for the relevancy level 1 and 3 is larger for data-driven models. To begin with, working with the relevancy level 1 and 3 for SAA-P is significantly different in 31 % of the results for SAA-P and for MIP-GP whereas it is different for MIP-SAA in 33 % of the results. This different is observed statistically in 24 % of the results for SAA-QR and for MIP-QR. We observe that there is a difference between the relevancy level 1 and 3 significantly in 23 % of the results for SAA-LR and in 22 % of the results for MIP-LR. Lastly, we observe the difference between the relevancy level 1 and 3 statistically for MIP-P in 36 % of the results. The largest percentage of significant test results for the difference between relevancy level 1 and 3 belongs to MIP-P.

As a result, we observe that the difference between the relevancy level 1 and 3 is more significant than the difference defined for the relevancy level 1 and 2. The training data size affects these two differences. The differences are more noticeable when the training data size increases. However, the difference between the relevancy level 1 and 3 is decreased negligibly when the training data size is increased from 250 to 450. Moreover, the differences defined for relevancy level 1 and 2 as well as for the relevancy level 1 and 3 are more obvious for MIP-P, statistically. This may be due to that it is more difficult to work with the relevancy level 2 or 3 for MIP-P. Also, MIP-P consider each sample of data as an united to find the decisions simultaneously. This can be another reason why the largest percentage of significant test results for the difference between relevancy levels belongs to MIP-P. On the other hand, SAA-P and MIP-SAA handle the relevancy level 1 and 2 of data better than others whereas SAA-LR and MIP-LR handle the relevancy level 1 and 3 better. Lastly, we can say that if the data is not inaccurate, there is not a significant effect of the relevancy level of data, especially for data-driven models that solve the problem considered hierarchically. This is because they look at the data separately as the first period and the second period data.

The Effect of the Sample Size

First of all, we change the training data size and examine the performance of data-driven models. To show the difference between the training size 50 and

250, we apply one sample t test to the differences in average profits obtained for training data size 50 and 250. A similar one sample t test is applied to the difference between average profits calculated for the training data size 50 and 450. We conduct a test for the difference defined for the training data size 250 and 450 in a similar way. We obtain p-value between zero and one, which expresses the level of statistical significance and give the results in Appendix I.3. A p-value less than 0.05 is statistically significant. At significance level of 0.05, we record how many significant test results (# of p-values which is smaller than $\alpha = 0.05$) there are from all results presented in Tables I.4-I.9. The percentage of significant test results are given in Table 4.19. The number of observations of averages is 64. In 32 % of the results, when we examine the difference between the training data size 50 and 250, it can be said that our data-driven models working with the training data size 450 and working working with the training data size 250 give significantly different average profits. Similarly, we count the significant test results and determine the difference between the training data size 50 and 450 as a percentage. At significance level of 0.05, we observe that there is a difference between training data size 450 and the training data size 50 in of 44 % all results. When we look at the difference between the average profit for training size 250 and the average profit for training size 450, it is observed that the increase in the training data size from 250 to 450 statistically affects the average profit of the model in 9 % of the results.

When we consider the demand distributions selected, we can infer that the percentage of significant test results is the largest when the demand distributions are Poisson for each comparison for the training data size. In contrast, the number of significant test results is the smallest when the demand distributions are assumed to be Uniform.

As a result, it can be said that the effect of the training data size on the performance of data-driven models may be dependent on the amount of increase. The percentage of significant test results or p-values smaller than 0.05 indicates that there is a difference between the training data size 50 and 250 as well as between the training data size 50 and 450 significantly. This difference is increased as the training data size is increased from 50 to 450. However, the percentage of the significant test results are small for the difference between the training data size 250 and 450. There might be a bound on the size of the training data size required to provide good decisions and the training data size exceeding this bound may affect performance less. Moreover, we observe that there is a relation between sample size and coefficient of variation. The difference between training data sizes is more precise for demand with low variation. For discrete demand, the training data size has significant effect. We make general insight about the effect of the training data size in Table 4.19, and we elaborate our discussion in the following tests.

Table 4.19: At significance level of 0.05, percentage of significant test results for the effect of training data size (# of p-values which is smaller than $\alpha = 0.05$ is given for each demand distributions selected)

Demand Option	Training Data Size 50&250	Training Data Size 50&450	Training Data Size 250&450
Poisson	64%	70%	13%
Uniform	14%	27%	2%
Symmetric Triangular and Uniform	34%	36%	8%
Right-Skewed Triangular	25%	41%	5%
Symmetric Triangular	31%	38%	13%
Left-Skewed Triangular	25%	53%	13%

Profit Comparison of Data-driven Models

First, we compare the profit of the best model with the profit of the worst one by performing one sample t test. The aim of the test is to check if the difference between the best and the worst model is zero or not. We set 1 if the test result is within 90 % CI, otherwise we set 0. Similarly, we set 1 if the result is within 99 % CI, otherwise 0. At significance level 0.1 and 0.01, 0 means that defined difference is significantly different than zero. We count the number of zeros significantly and give the results in Appendix I.5, Tables I.58-I.63. Since we have 50 replications and conduct this test for each replication, the largest number that we can obtain is 50. In Table 4.20, we observe a lot of zeroes for all training sizes and parameter settings. Based on this statistical results, we can say that it is important to make a good selection of data-driven models when the demand distribution is Poisson. For this table, it can be said that if a method is the best, the difference with the worst one is affected slightly as the training data size increases.

Table 4.20: Comparison of the best model with the worst one for Poisson demand

	sum of 0's					
	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	# with 90% CI	# with 99% CI	# with 90% CI	# with 99% CI	# with 90% CI	# with 99% CI
1	49	49	50	50	50	50
2	49	49	50	50	50	50
3	49	49	50	50	50	50
4	50	50	50	50	50	50
5	49	49	50	50	50	50
6	50	49	50	50	50	50
7	50	50	50	50	50	50
8	50	50	50	50	50	50

For other tables in Appendix I.5, we can say that results remains similar. When the data is generated from Uniform, Symmetric Triangular, Right-skewed Triangular and Left-Skewed Triangular as well as the combination of Symmetric Triangular and Uniform, we observe that the number of zeros is 50 for each parameter setting and training data size, statistically. According to these results, we can say that selecting a data-driven model is significant regardless of which distribution is used to generate the data.

Comparison with respective expected values

Second, we do comparison of the profit of data-driven models with respective expected values. We investigate how close the average profit a data-driven method to the expected profit obtained under the demand distribution assumption. The percentage difference is calculated and observed results are given in Appendix I.6, Tables I.64-I.69. According to the results, proposed data-driven models perform well generally. To illustrate, the percentage difference for SAA-P is maximum 7 % when we consider all demand model options. For MIP-SAA, the percentage difference from expected total profit is between 0 % and 7 %. For SAA-QR, the percentage difference from expected total profit is between 1 % and 8 %. The ranges of the percentage difference from expected total profit for other data-driven models are shown in Table 4.21.

Table 4.21: Summary table for the percentage difference between average profits and expected value

	SAA-P	MIP-SAA	SAA-QR	MIP-QR
The Range of % Difference with Expected Value	0%-6.60%	0%-6.72%	0.87%-8.01%	0.87%-8.01%
	SAA-LR	MIP-LR	MIP-GP	MIP-P
The Range of % Difference with Expected Value	0.14%-10.49%	0.24%-10.02%	0.08%-7.85%	0%-22.37%

We can assess the value of training data size for the performance of data-driven models by analyzing the percentage differences. To summarize, the mean values of the percentage difference with the expected value is calculated over the parameter settings and given for each data-driven models in the Figures 4.4-4.11. From Figure 4.4, it can be said that the mean percentage difference from expected profit for SAA-P decreases when the training data size is increased for each demand distribution selected. When we look at the mean percentage difference of SAA-P, the demand distributions that we select to generate data are ordered from the smallest to largest as follows: Poisson, Left-skewed Triangular, Symmetric Triangular, the combination of Symmetric Triangular and Uniform, Right-skewed Triangular and Uniform distributions, respectively. This means that SAA-P gives the average profit closest to the expected value if we generate the data from Poisson distribution.

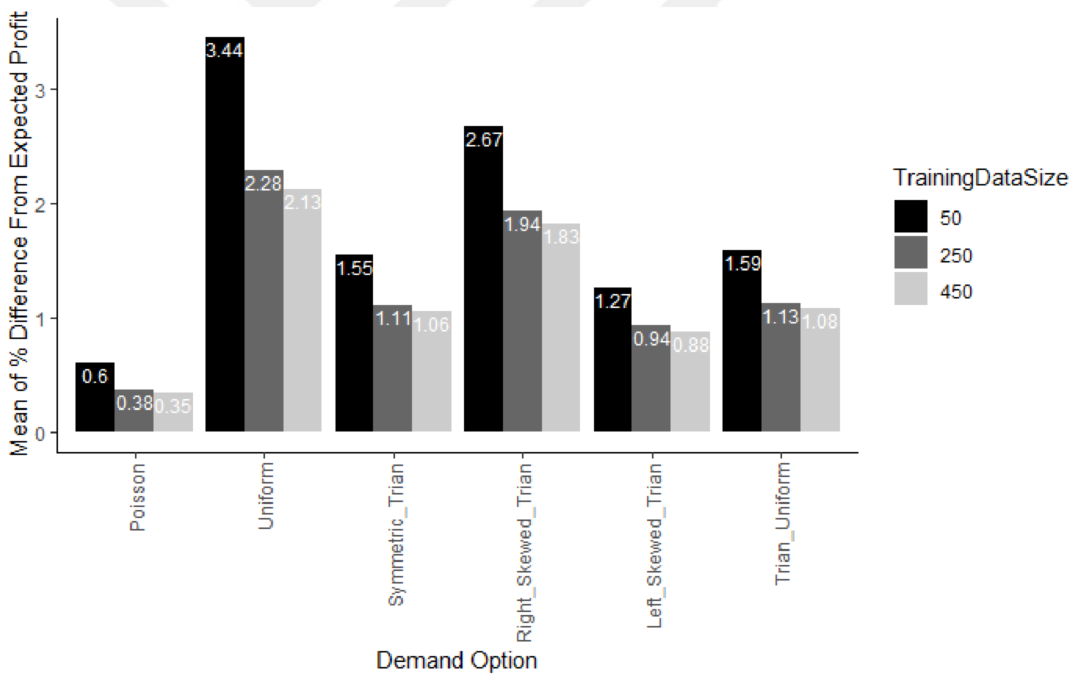

Figure 4.4: Mean of the percentage difference between average profit of SAA-P and expected value over parameter settings

For each training data size, we observe that the mean percentage difference of MIP-SAA is close to the mean percentage difference of SAA-P when the the data is generated from Poisson and Uniform distributions. For other demand distributions, same evaluation can be made when the training data size is set to

50 and 250, respectively. Furthermore, this percentage value for MIP-SAA decreases as the training data size is increased for the data generated from Poisson and Uniform distributions (see Figure 4.5). When the training data size is set from 50 or 250 to 450, however, it is increased for the data generated from Symmetric, Right-skewed, Left-skewed Triangular distributions and the combination of Symmetric Triangular and Uniform. This may due to that MIP-SAA cannot handle the largest size of the training data for most of the demand distributions selected.

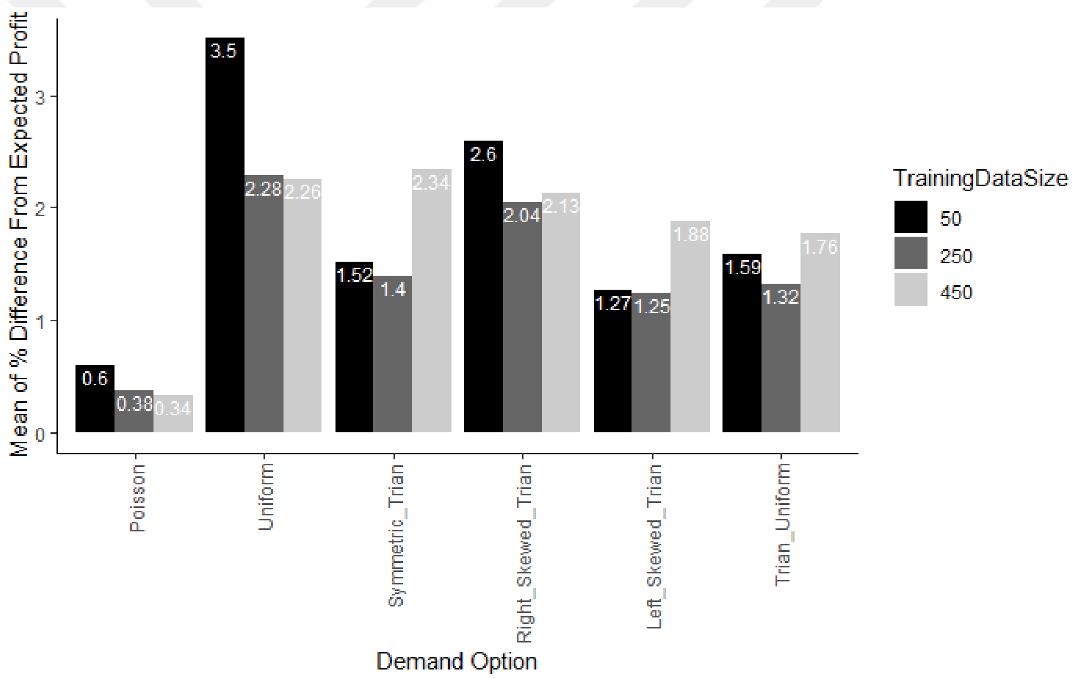

Figure 4.5: Mean of the percentage difference between average profit of MIP-SAA and expected value over parameter settings

From Figures 4.6 and 4.7, we make the following observations. For both SAA-LR and MIP-LR, the largest mean value of the percentage difference occurs when the data is generated from Uniform distribution and the smallest value is obtained for the Poisson distribution. Also, we observe that their mean percentage difference decreases when the training data size increases. From the smallest mean percentage difference of MIP-LR to its largest mean percentage difference, the order of the distributions selected for generating data is as follows: Poisson, Left-skewed Triangular, the combination of Symmetric Triangular and Uniform, then Symmetric Triangular, Right-skewed Triangular and Uniform distributions. This order is also valid for SAA-LR.

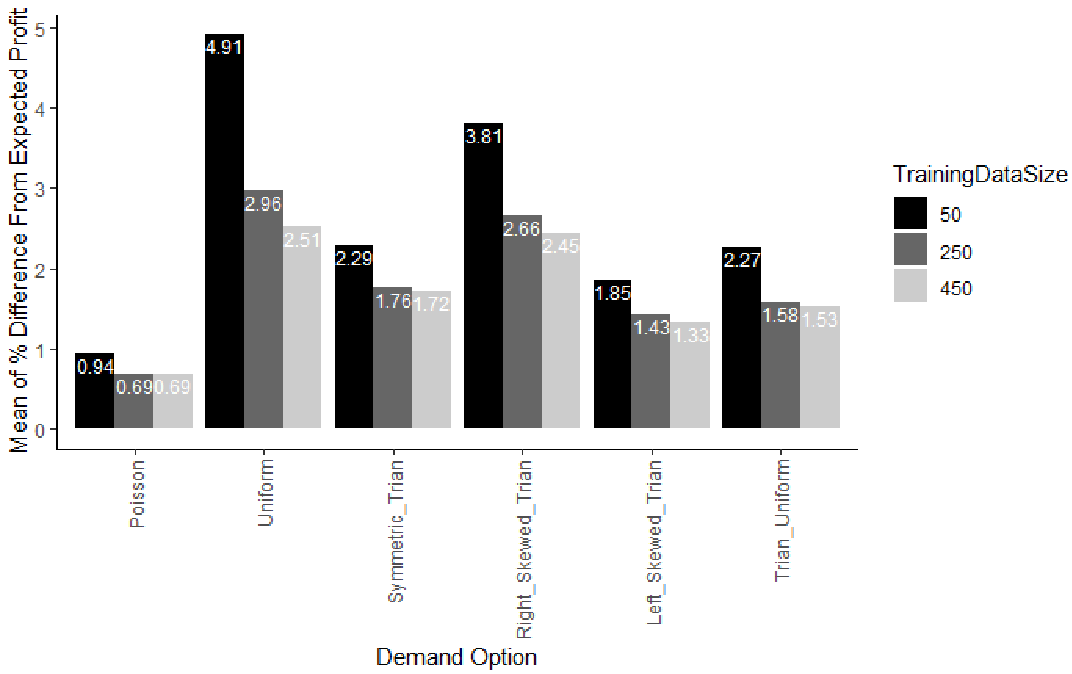

Figure 4.6: Mean of the percentage difference between average profit of SAA-LR and expected value over parameter settings

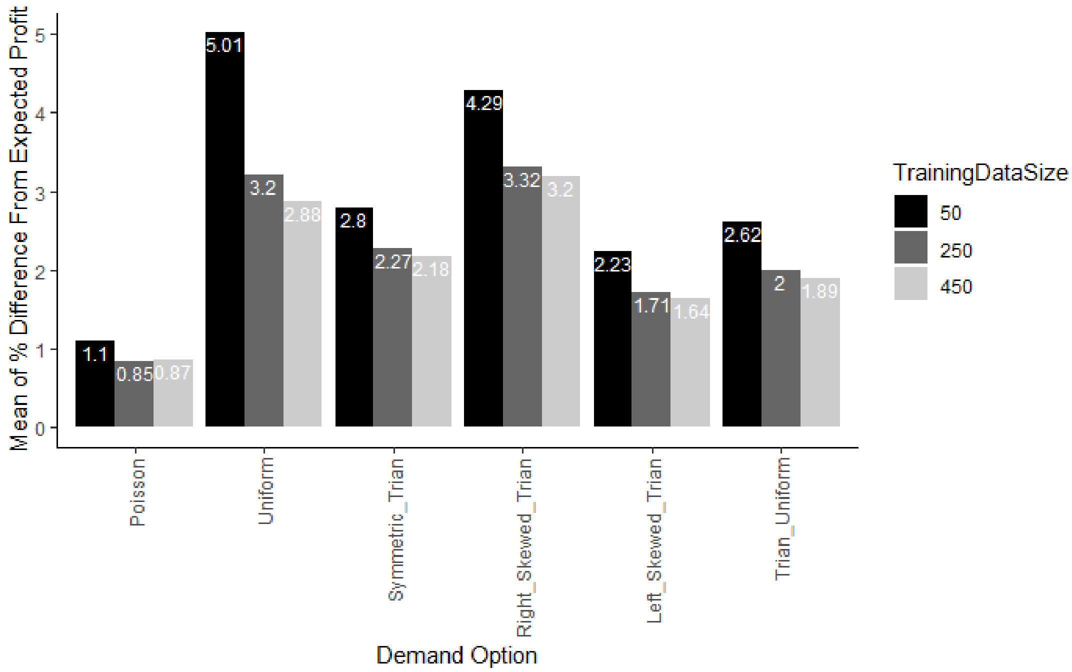

Figure 4.7: Mean of the percentage difference between average profit of MIP-LR and expected value over parameter settings

For SAA-QR and MIP-QR, the mean percentage difference is almost similar for each demand distribution chosen and for each training data size. In Figures 4.8 and 4.9, the largest value is observed when the data is generated by using Uniform distribution for both SAA-QR and MIP-QR. In addition to, the mean values of percentage difference are decreased as the training data size is increased for most of the demand distributions. It is observed that there is a negligible difference between these values obtained for each training data size when the data is generated from Poisson distribution.

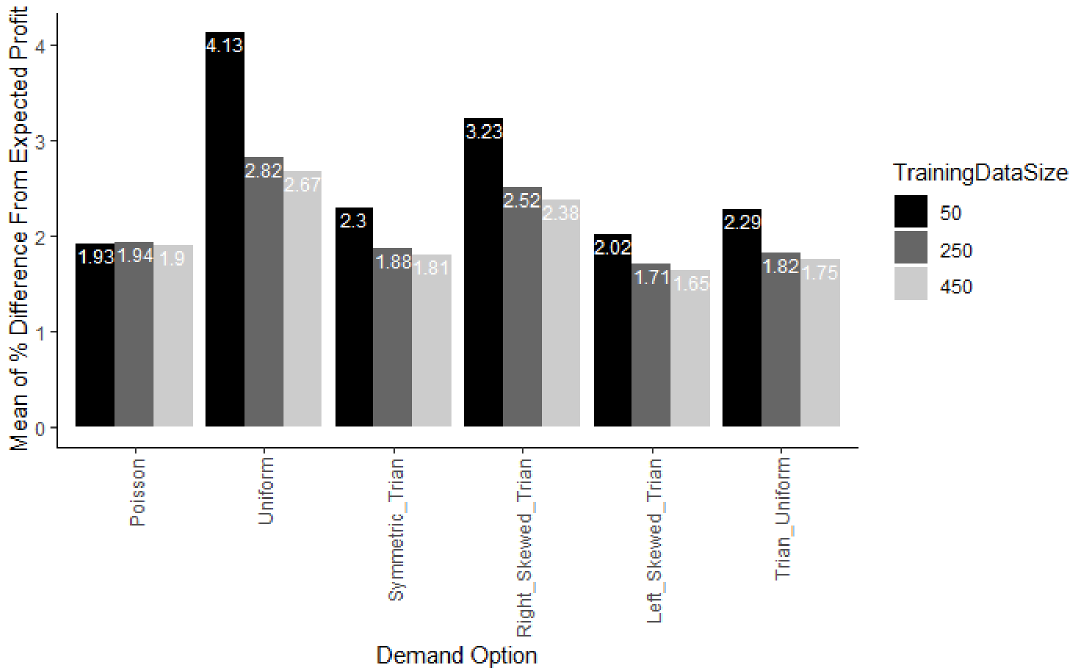

Figure 4.8: Mean of the percentage difference between average profit of SAA-QR and expected value over parameter settings

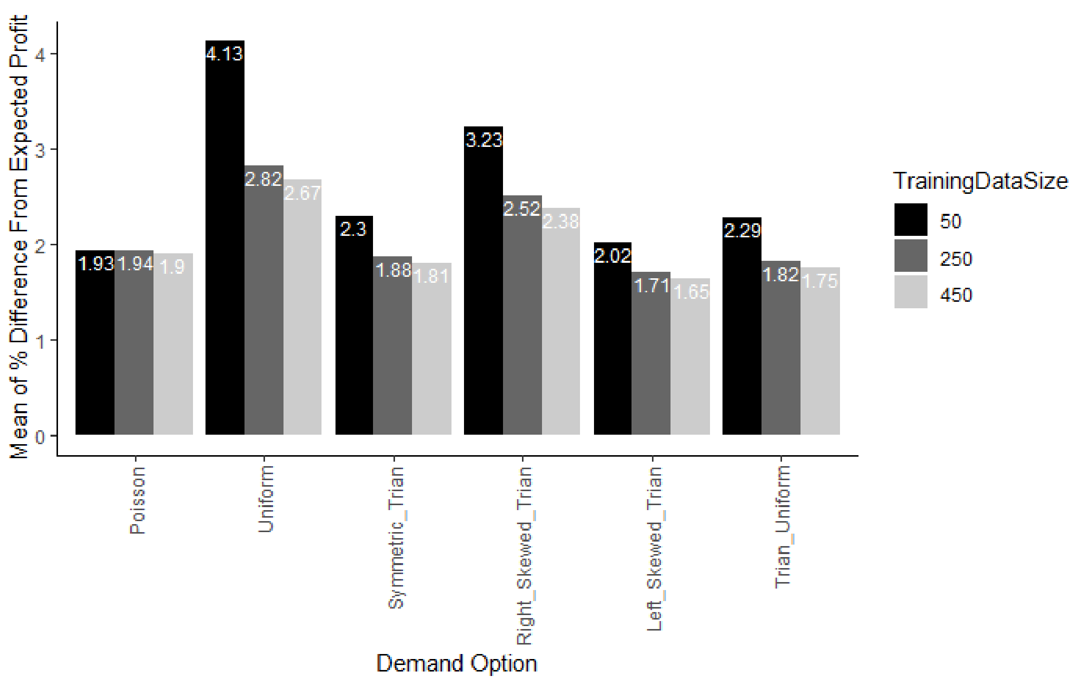

Figure 4.9: Mean of the percentage difference between average profit of MIP-QR and expected value over parameter settings

From Figures 4.10, the following observations are made. For MIP-GP, the largest mean percentage difference occurs when the data is generated from Uniform distribution and the smallest mean percentage difference occurs when the Poisson distribution is selected. Also, we observe that mean percentage difference is decreased as the training data size increases for MIP-GP. In ascending order, the distributions selected for generating data are as follows: Poisson, Left-skewed Triangular, the combination of Symmetric Triangular and Uniform, then Symmetric Triangular, Right-skewed Triangular and Uniform distributions. Based on the values of mean percentage difference, its performance compared to other data-driven models can be categorized as intermediate level.

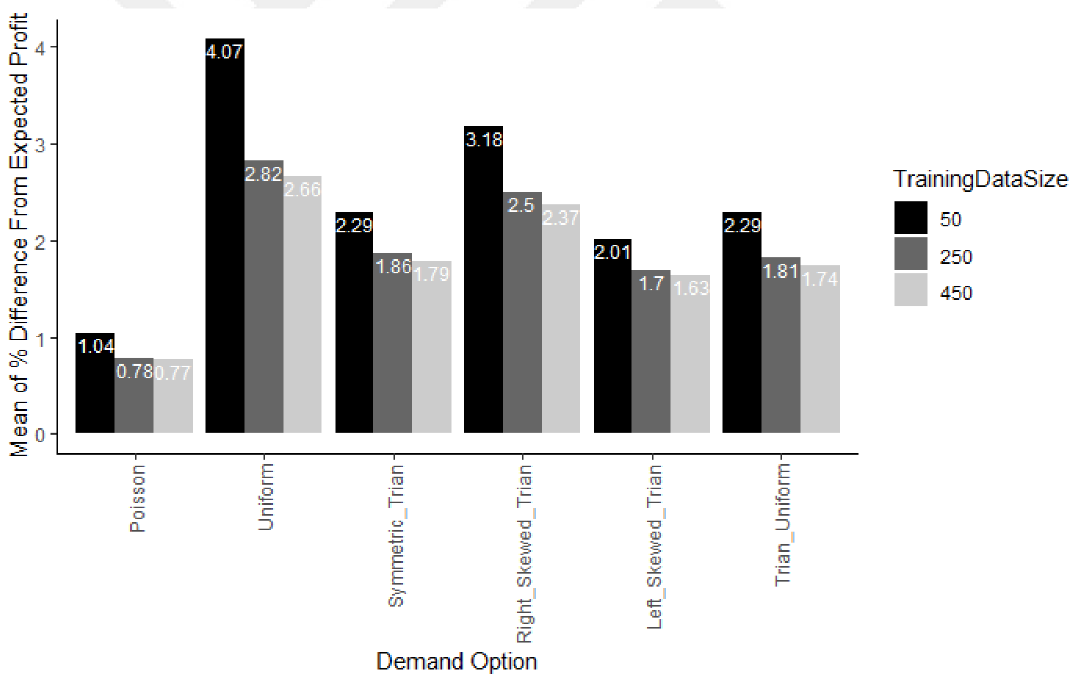

Figure 4.10: Mean of the percentage difference between average profit of MIP-GP and expected value over parameter settings

Lastly, we observe that the smallest mean percentage difference defined for MIP-P occur when the Poisson distribution is selected to generate the data in Figure 4.11. When the training data size is increased, the mean percentage difference is decreased for Poisson distribution. However, the mean percentage difference between average profit of MIP-P and expected value increase as the training size is increased from 50 or 250 to 450. We can infer that it is difficult for MIP-P to deal with large size of the training data for most of the demand distributions.

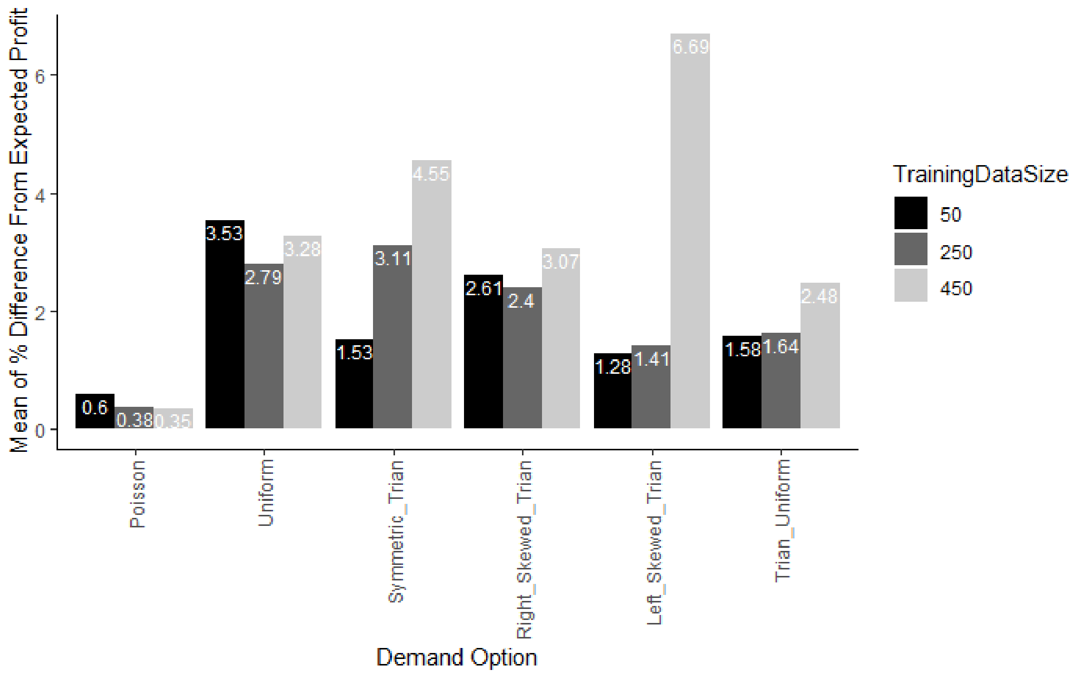

Figure 4.11: Mean of the percentage difference between average profit of MIP-P and expected value over parameter settings

As a result, we can make the following observations when we look at the Figures shown above. For each data-driven model, the smallest mean percentage difference from expected value is observed for Poisson distribution. In most of the results, SAA-P gives smaller percentage difference than others and it can be used for each size of the training data as it performance is good. Furthermore, which model is chosen for larger training data size is significant. Generally, the mean values of percentage difference defined for MIP-SAA and MIP-P are also closer to the expected values when they are used for small size of the training data. When the training data size is increased, however, time to solve the problem is increased exponentially for MIP-SAA and MIP-P. To illustrate, MIP-P can give the largest percentage difference when the training data size is set to 450. For other sizes of the training data, larger mean percentage difference from expected value belong to MIP-LR although its performance against the expected value is better as the training data size increases generally. For MIP-LR, it is easier to solve the problem because of the small number of inputs. Lastly, it can be said that the performances of SAA-QR and MIP-QR are similar based on the percentage difference with the expected value. The percentage difference decreases as the size of training data is increased for data driven models except MIP-SAA and MIP-P.

Moreover, the inverse of coefficient of variation for the difference is also calculated to test data-driven models. The results are shown in Appendix I.7, Tables I.70-I.75. The coefficient of variation can be defined as the ratio of the standard deviation to the mean. This means that when the the ratio is decreased, the variance also decreases. In our formula for Covd, we divide the difference $E - P(k)$ by the standard deviation. Therefore, it can be said that the variance of the average profit obtained is low when the result of CovD formula is low. If difference $E - P(k)$ is large, it turns out to be large in every problem. According to the obtained ratios, variance is lower when the training size is increased for all of data-driven models generally. We observe that the variance is small for comparison in most of the results. To illustrate, for SAA-P, the inverse of the coefficient of variation is between 0 and 0.35 whereas it is between 0 and 0.43 for MIP-SAA. The range of the inverse CovD for models are given in Table 4.22. Based on the ranges, it can be said that model SAA-P is more stable than others. Since the largest range belongs to MIP-P, we can say that it is less stable than others. To give general insight for data-driven models, we take the average of the inverse CovD over parameter settings for each model and give the results in Figures 4.12-4.19.

Table 4.22: Summary table for the inverse of coefficient of variation for the difference

	SAA-P	MIP-SAA	SAA-QR	MIP-QR
The Range of the inverse of CovD	0.00-0.35	0.00-0.43	0.07-1.10	0.07-1.10
	SAA-LR	MIP-LR	MIP-GP	MIP-P
The Range of the inverse of CovD	0.06-0.53	0.06-0.67	0.04-0.63	0.00-3.65

In Figure 4.12, we observe that the mean values of inverse CovD of SAA-P are decreased as the training data size increases. This means that SAA-P is more stable when the training data size is increased. Also, it can be said that it is more stable when the data is generated from Uniform, Symmetric Triangular, Left-Skewed Triangular, and the combination of Symmetric Triangular and Uniform distributions. The largest values of the inverse of CovD are observed when the training data size is set to 50 and the data is generated from Poisson and Right-Skewed Triangular distributions.

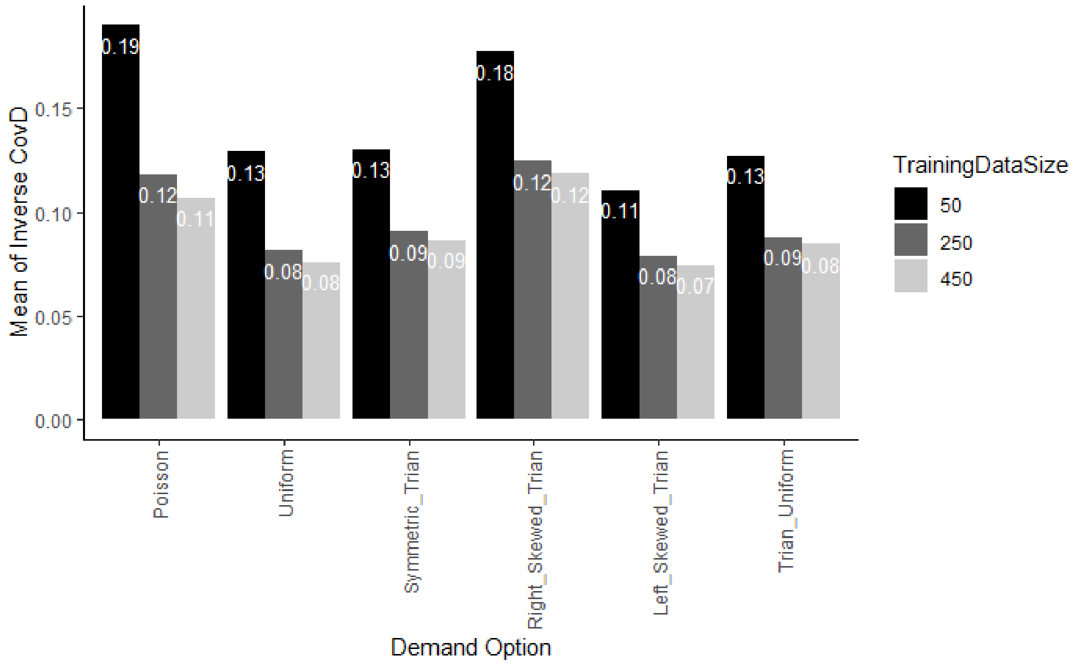

Figure 4.12: Mean of Inverse CovD over parameter settings for SAA-P

In Figure 4.13, we make the following observations. The stability of MIP-SAA changes based on the training data size. When the data is generated from Poisson and Uniform distributions, the mean values of inverse CovD are decreased as the training data size increases. For other distributions, however, these values are increased when the training data size is increased from 50 or 250 to 450. The largest mean value of inverse CovD is obtained when the training data size is 50 and the data is generated from Symmetric Triangular distribution.

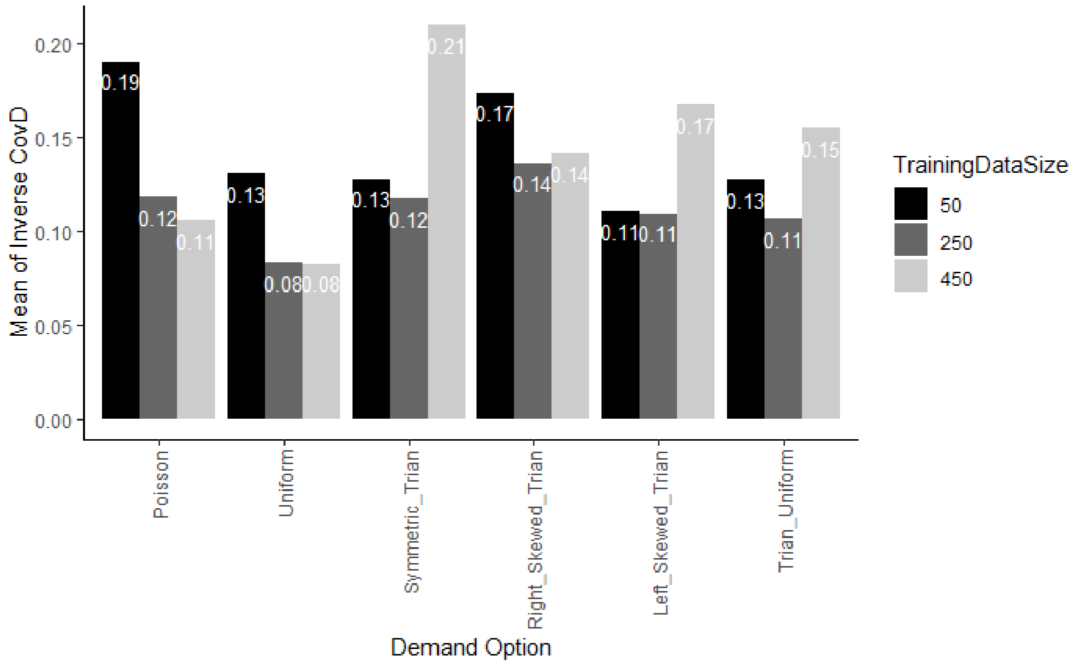

Figure 4.13: Mean of Inverse CovD over parameter settings for MIP-SAA

From Figure 4.14 and 4.15, we can say that mean values of the inverse CovD for both SAA-LR and MIP-LR are decreased as we increase the training data size for all distribution selected. When the data is generated from Poisson, however, a negligible difference of the mean values of inverse CovD when the training data size is 250 and 450 is observed for MIP-LR. For both SAA-LR and MIP-LR, the largest two mean values of inverse CovD are observed when the distribution is Poisson and Right-skewed Triangular, respectively.

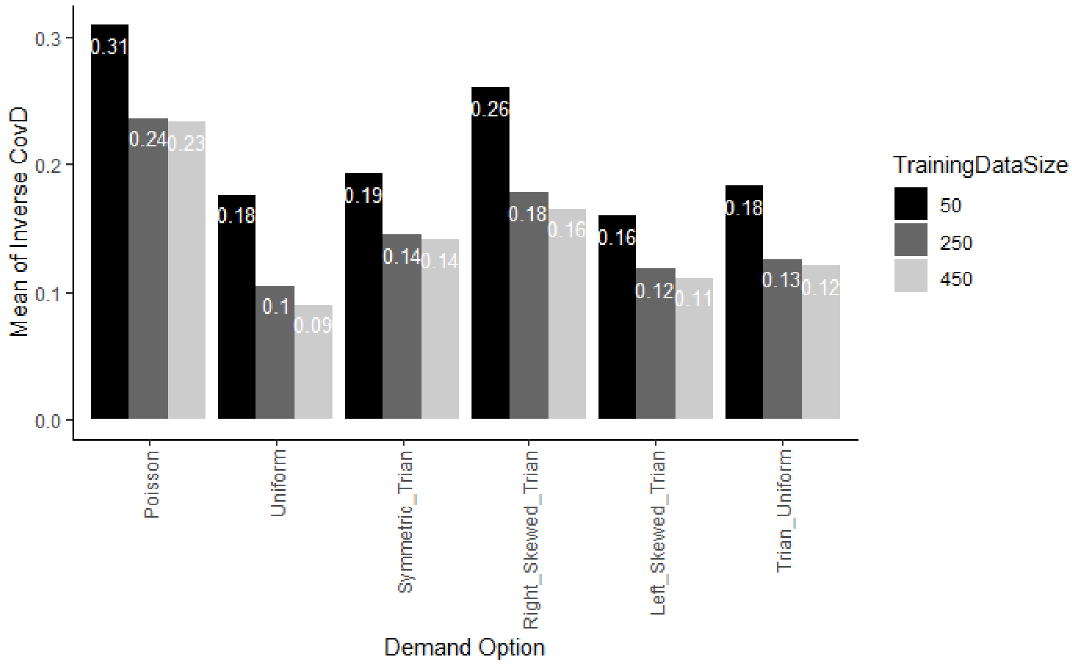

Figure 4.14: Mean of Inverse CovD over parameter settings for SAA-LR

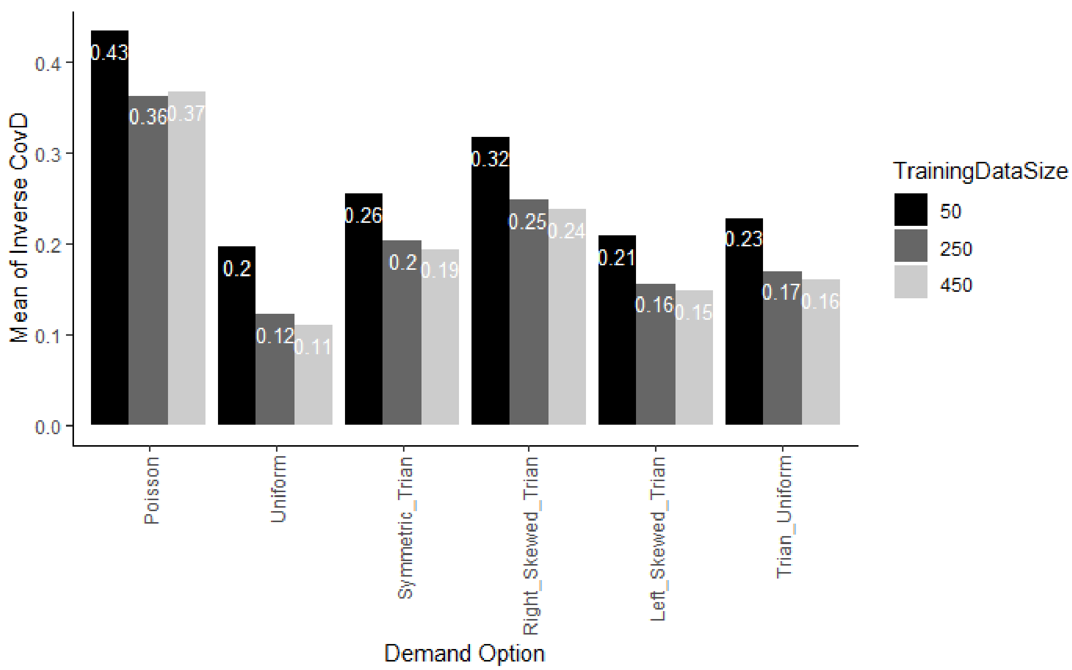

Figure 4.15: Mean of Inverse CovD over parameter settings for MIP-LR

The Figures 4.16 and 4.17 shows that the mean values of inverse CovD of SAA-QR and MIP-QR are similar. Similar with SAA-P, for both SAA-QR and MIP-QR, the mean value of the inverse CovD is diminished when the training data size increases in most of the results. For the data is generated from Poisson or Uniform distribution, however, this value is increased when the training data size is 450. Among all distributions selected, they are less stable for Poisson distribution. We can say that when the data is generated from Poisson, an arbitrary result can be obtained if the right regression model is not used. The mean values of inverse CovD are close to each other and small for other distributions chosen to generate the data.

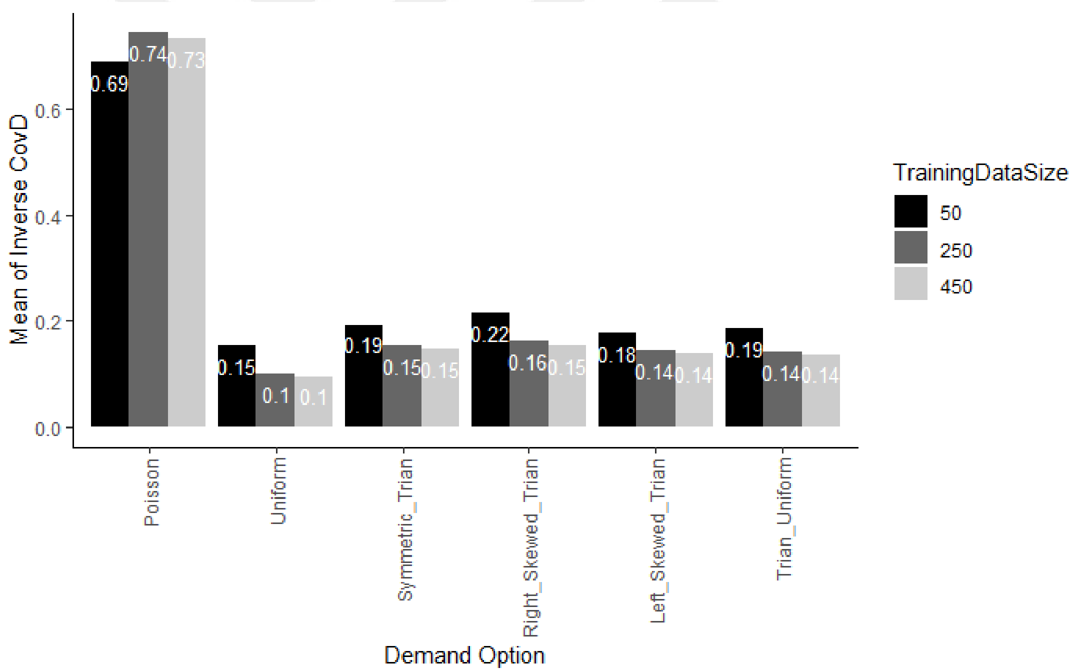

Figure 4.16: Mean of Inverse CovD over parameter settings for SAA-QR

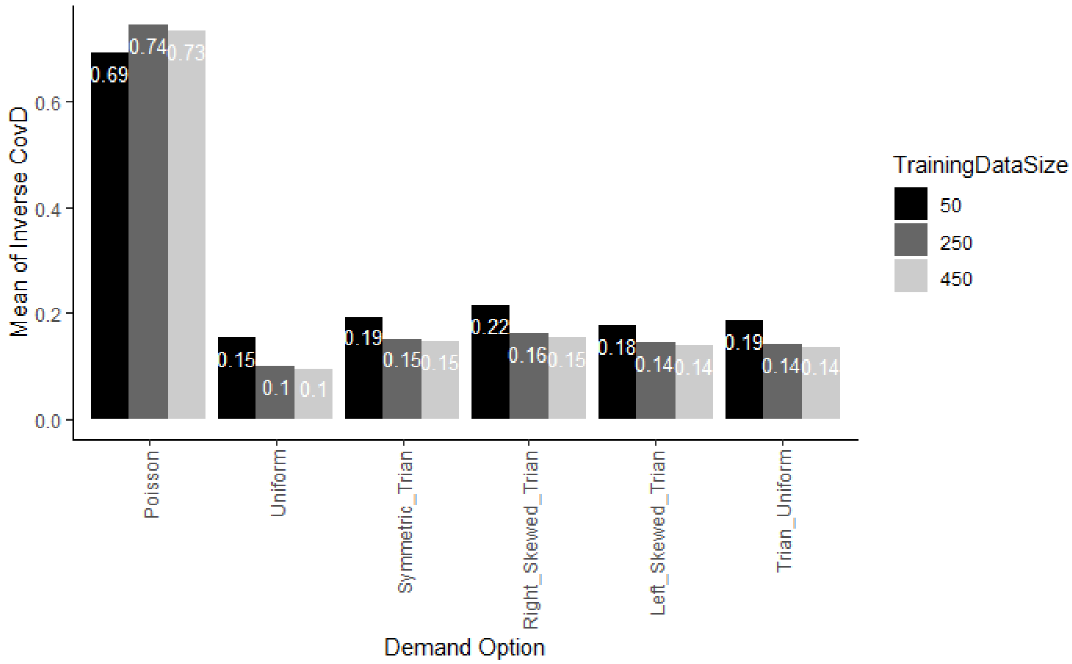

Figure 4.17: Mean of Inverse CovD over parameter settings for MIP-QR

We can make the similar observations for MIP-GP based on the results given in Figure 4.18. For MIP-GP, mean values of inverse CovD are diminished as the training data size increases and the largest mean value of inverse CovD occurs when the data is generated by using Poisson distribution. The smallest value is obtained for the data generated from Uniform distribution.

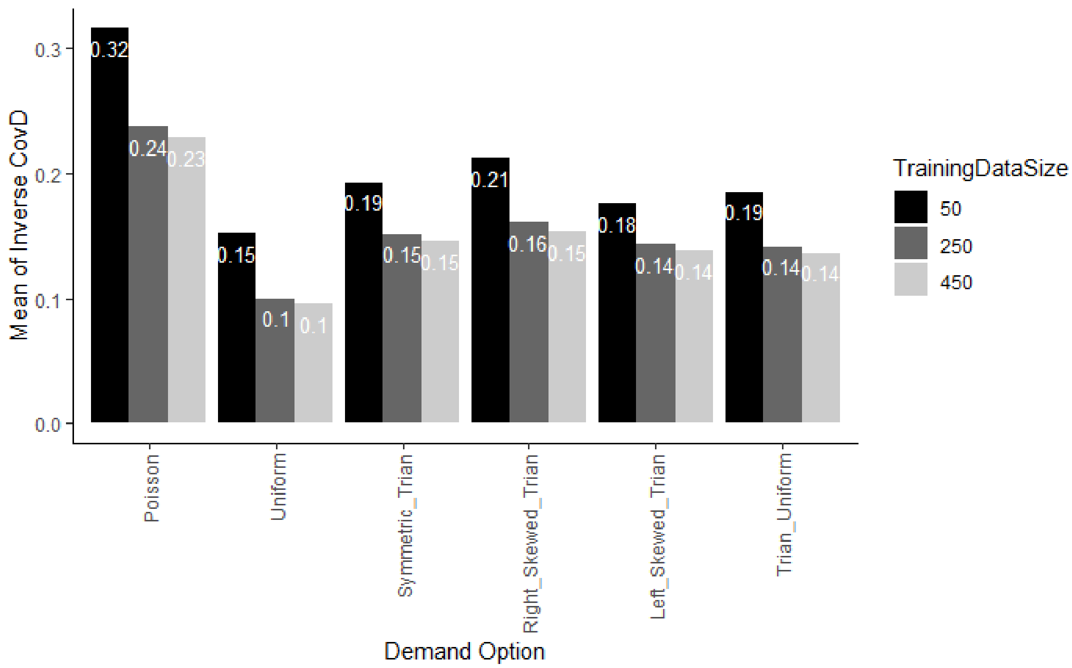

Figure 4.18: Mean of Inverse CovD over parameter settings for MIP-GP

The Figure 4.19 shows that MIP-P is more stable when we set the training data size to 50 for all distributions. The mean values of inverse CovD for MIP-P are increased for large size of the training data size for all distributions except Poisson. Although MIP-P can handle the larger size of the training data when the data is generated from Poisson distribution, it has difficulty to handle for continuous distributions selected. The largest value is obtained for Left-skewed Triangular distribution due to an extreme outlier.

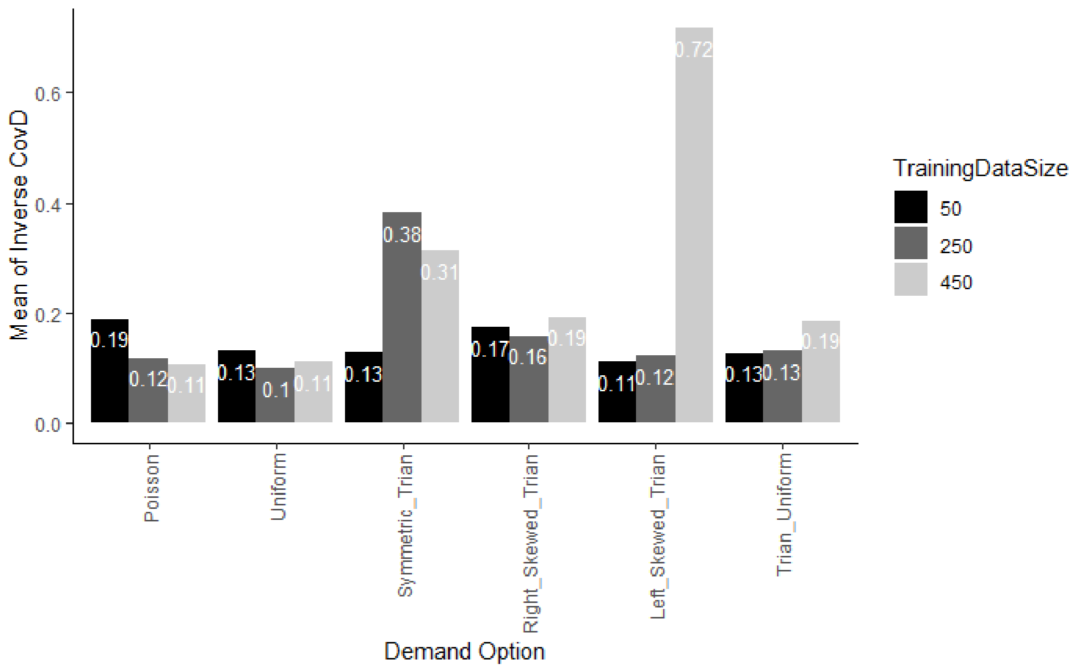

Figure 4.19: Mean of Inverse CovD over parameter settings for MIP-P

To summarize, we can again conclude that MIP-P and MIP-SAA should be used when the training data size is small whereas other data-driven models can be used for all training data sizes. For most of proposed models, variation is smaller when the training data size is increased. For MIP-SAA and MIP-P, variation increases when the training data size is set to 450. Similar with the general insight about the percentage difference from the expected value, the mean values of inverse CovD for SAA-QR and MIP-QR are close to each other. In such cases, the optimal policy found by models are generally similar. In addition, SAA-P is more stable than others when all of the results for inverse CovD are considered. For continuous distributions, MIP-LR can be less stable than other when the training data size is 50 or 250 yet it is observed that the mean values of inverse CovD of MIP-P can be larger for the training data size 450. When we use Poisson distribution to generate the data, SAA-QR and MIP-QR are less stable than others.

Number of times one method is the best

Lastly, we observe the number of times one method is the best in Appendix I.8, Tables I.76-I.81. We compute the percentage success by considering how many instances the method is the best out of 2500 instances. We give the ranges of the results for each model in Table 4.23. Then we take average of the percentage success defined over the parameter settings an give the results in Figures 4.20-4.25.

Table 4.23: Summary table for the percentage success of data-driven models

	SAA-P	MIP-SAA	SAA-QR	MIP-QR
The Range of % Success	1.68%-63.6%	3.6%-63.6%	0.44%-46.48%	0.08%-47.04%
	SAA-LR	MIP-LR	MIP-GP	MIP-P
The Range of % Success	1.28%-66.08%	8.72%-72.4%	0.6%-66.72%	1.16%-63.6%

According to the results in Figures 4.20-4.25, MIP-LR yields the larger percentage success than other in most of the time. For MIP-LR, the percentage success changes from 9 % to 72 % which are given in Table 4.23. The performance of MIP-LR is better than others for most of the distributions selected. Therefore, it can be said that MIP-LR can be used for each training data size. This comparison test is conducted for each instances in testing data. There might be instances or samples where the estimations of MIP-LR are good. On the other hand, there might be also instances in which MIP-LR found a solution which is not good. By considering our observations about other tests, we can infer that instances where MIP-LR derives good estimations dominates other instances where MIP-LR finds incorrect solutions. However, these incorrect solutions to instances might be enough to yield that the average profit of MIP-LR is further from the expected values.

In addition to that, Figures 4.21-4.25 show that the mean of percentage success for MIP-LR is diminished as the training data size increases. This may due to that the number of instances where incorrect solutions are found may increase when the training data size increases. However, the number of instances where good solutions are found dominates them as its performance is still better than others for most of the distributions selected.

These cases mentioned above should be considered for each data-driven models. There are some models whose the percentage success is decreased when the training data size is increased and there are some models whose percentage success is increased when the training data size increases. Therefore, the ranking of models from the largest percentage success to smallest changes. In general, it is observed that the percentage success of models SAA-P and MIP-SAA increases when the training data size is increased whereas the percentage success of models SAA-QR, MIP-QR, SAA-LR, MIP-LR and MIP-P decreases as the training data size is increased. Figures 4.20-4.22 gives the order for each distributions selected.

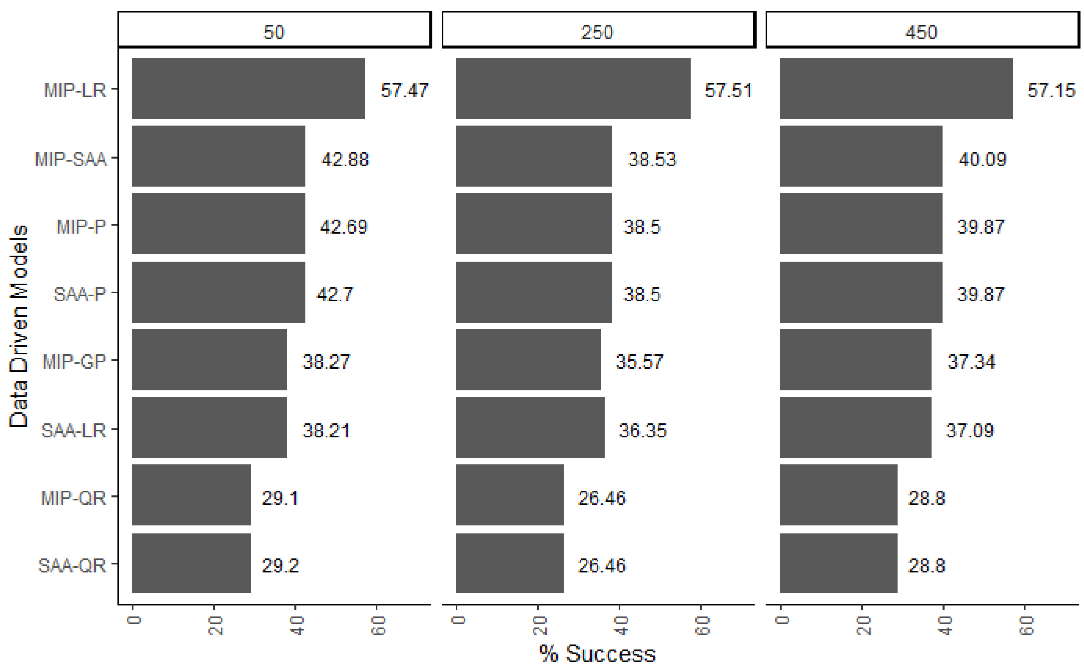

Figure 4.20: Mean of the percentage success over parameter settings when demand is Poisson distributed

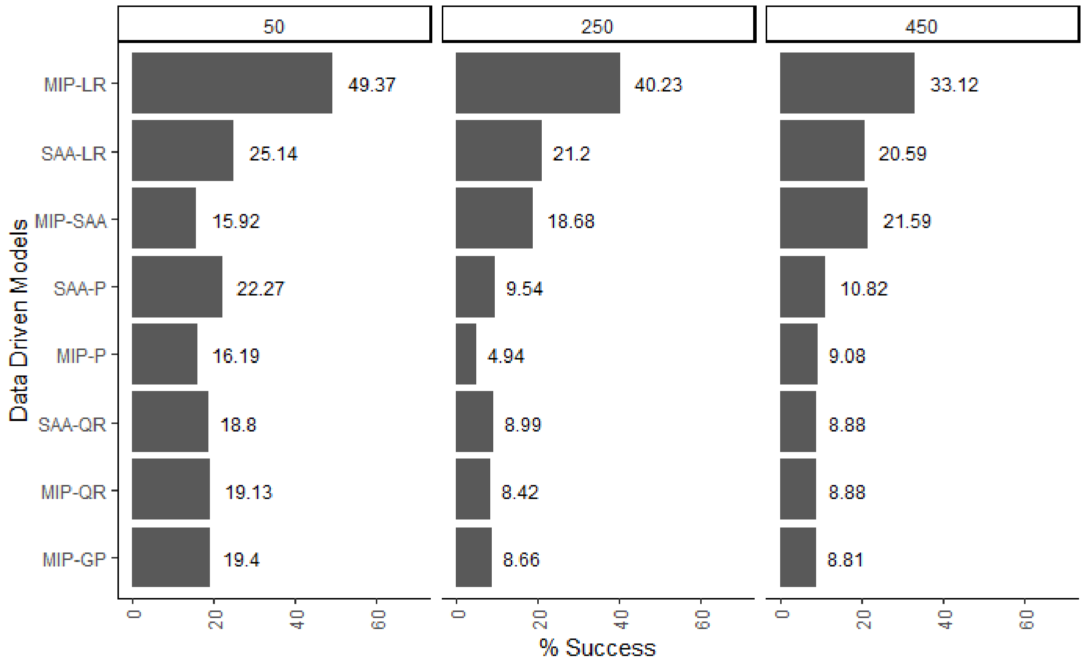

Figure 4.21: Mean of the percentage success over parameter settings when demand is Uniform distributed

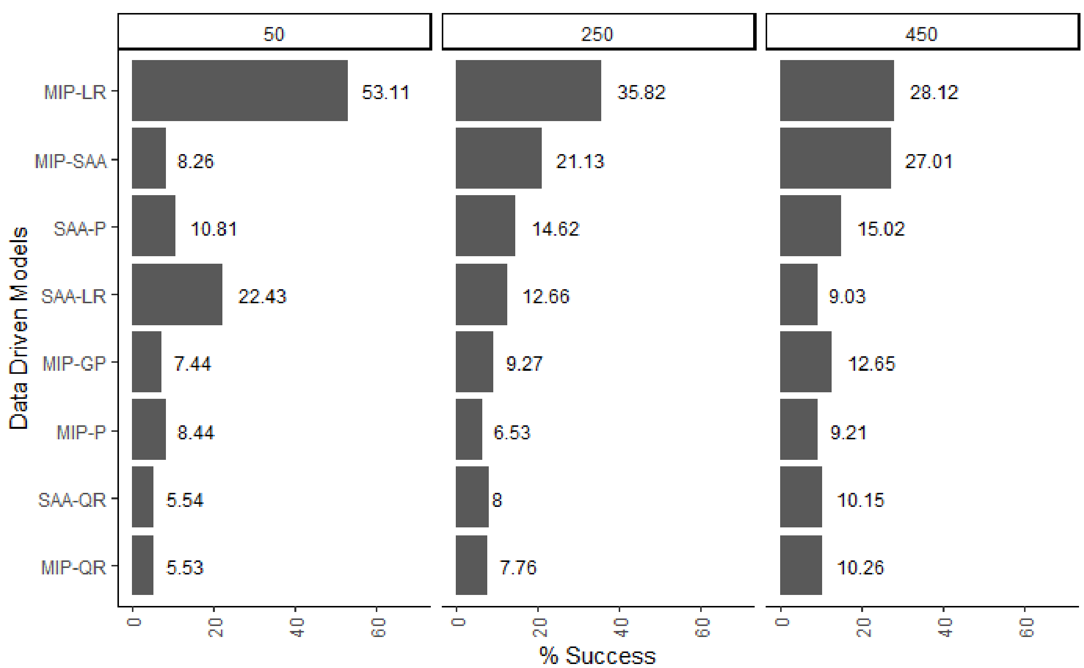

Figure 4.22: Mean of the percentage success over parameter settings when demand is Symmetric Triangular distributed

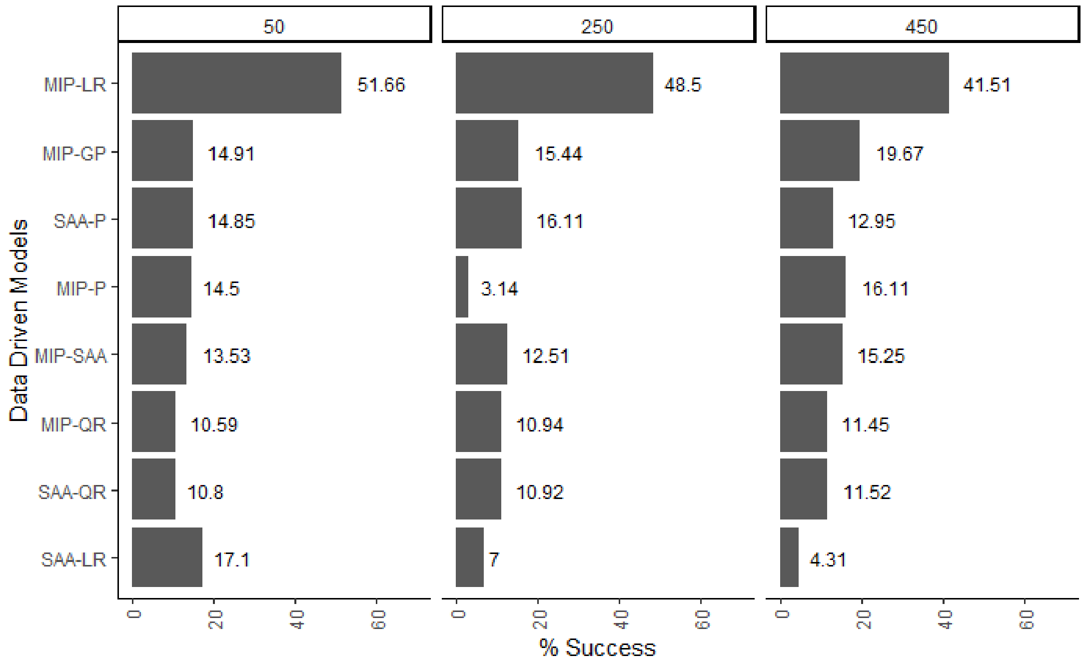

Figure 4.23: Mean of the percentage success over parameter settings when demand is Right-Skewed Triangular distributed

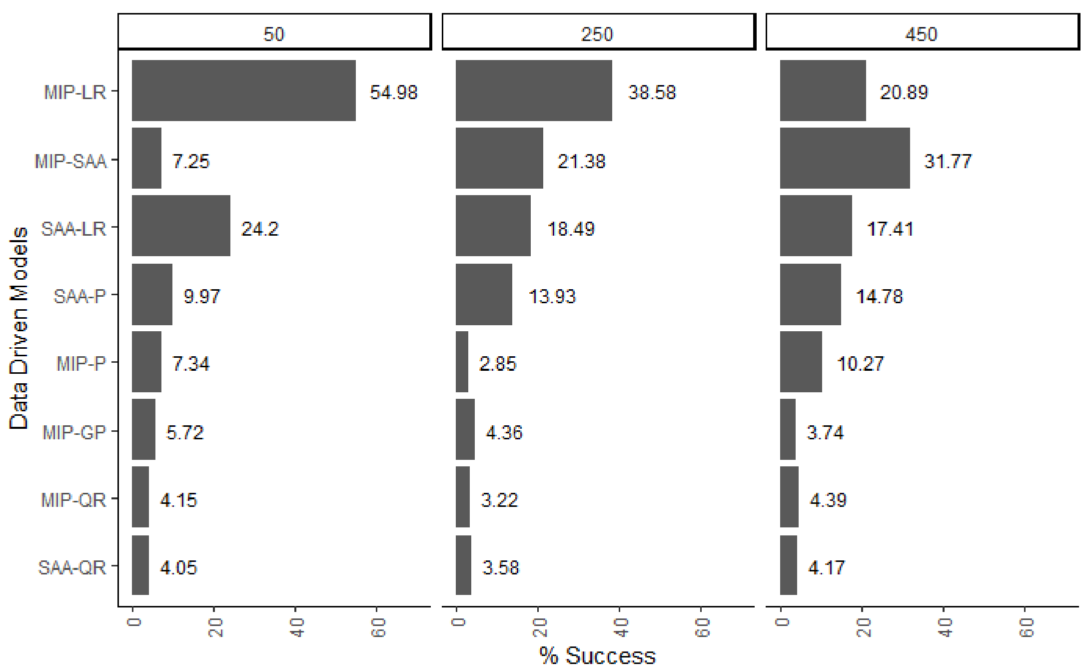

Figure 4.24: Mean of the percentage success over parameter settings when demand is Left-Skewed Triangular distributed

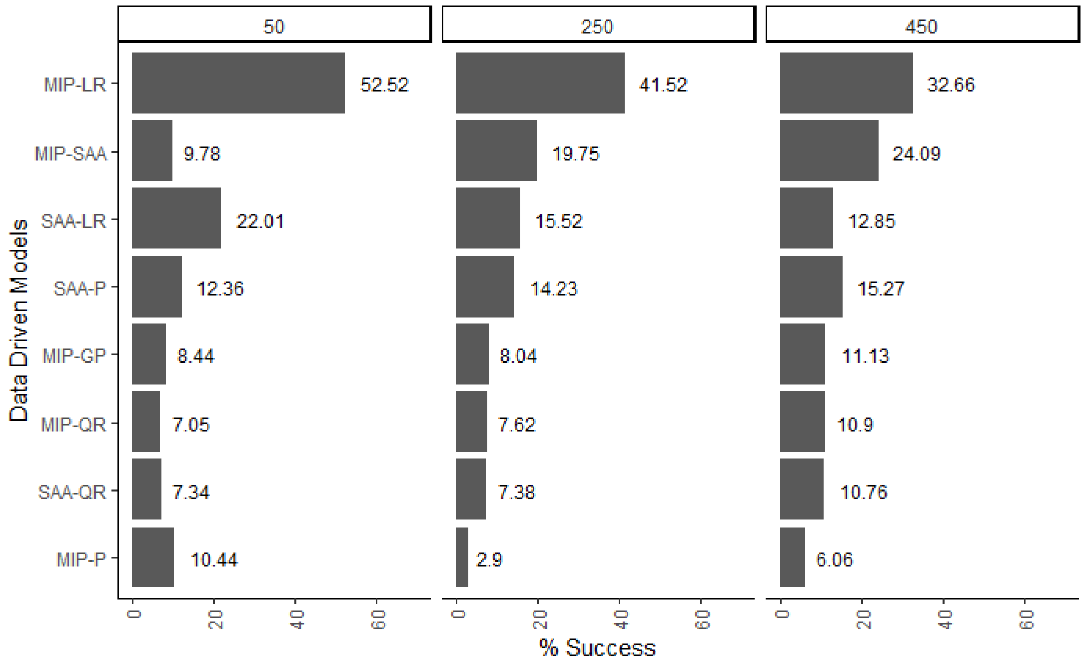

Figure 4.25: Mean of the percentage success over parameter settings when demands are Symmetric Triangular and Uniform distributed

In a nutshell, we conclude that if the data is not inaccurate, there is not a significant effect of what decision was made in the data. It is more important to consider the way how decisions were made in the past data for MIP-P model as it considers each sample of the data as united to find the decisions simultaneously. At significance level 0.05, we infer that there is a difference between the training data size 50 and 250 as well as between the training data size 50 and 450 for each data-driven models. The difference between training data sizes is more precise for demand with low variation. Also, we show the significance of making a good selection among data-driven models statistically by performing a test for the difference between profit of best model and profit of the worst model. From the numerical experiments for profit comparison with respective expected values, we can conclude that the optimal order quantity and the optimal threshold found by our data-driven models are consistently good estimates and they are close to the exact results obtained. Also, which model is chosen for larger training data size is significant. According to the results for inverse CovD, which is calculated by the ratio of the difference between average profit of a model and expected profit computed theoretically to the standard deviation (Formula in 4.15), we can say that the variance of the average profit obtained by each method overall replications is small for comparison and data-driven models are stable in most of the results. Lastly, we make the following observation by looking the number of times one method is the best. In general, it is observed that the percentage success of models SAA-P and MIP-SAA increases when the training data size is increased. In contrast, the percentage success of models SAA-QR, MIP-QR, SAA-LR, MIP-LR and MIP-P decreases as the training data size is increased. Based on the mean values of percentage success, MIP-LR outperforms other data-driven models consistently. When the training data size is 50, SAA-LR is the second best model in general. The ranking of models from the largest percentage success to smallest changes for each training data size and for each demand distribution selected. Data-driven models show their best performance when the data is generated from Poisson distribution. Another observation made is that time to solve the problem is increased exponentially when the training data size is increased for MIP-SAA and MIP-P models and this may affect their performances. Therefore, implementing them with smaller training data size can be better.

Chapter 5

Conclusion

In this thesis, we investigate the newsvendor problem with recourse action. It is known that optimal solution to the traditional newsvendor problem is found analytically unless the demand distribution is unknown. As demand distributions are likely to be unknown in real life, we are interested in proposing data-driven methods that do not require demand distribution and work with the historical data. We are also interested in offering the remaining unsold products at a marked-down price one more period as a second period problem as the salvage value is quite low in a single period newsvendor problem. Therefore, we consider a retailer who has a single fashion product is sold through two stages. To the extent of our knowledge, this is the first work combines the newsvendor problem with a recourse action considering non-parametric data-driven models. There are two decisions, which are the order quantity and the marked-down price, in our setting. We decide on the order quantity for the first stage as well as we set the marked-down price for the second stage based on the leftover inventory. For simplicity, the second stage or recourse decision is made between two marked-down prices: the low price and the high price. Under the conditions (2.10) and (2.11), there exist a threshold value which describes the range of the selection for both marked-down prices. Hence, the second period problem becomes determining the threshold to use in marked-down price decision for the remaining inventory at the end of first period.

First, we analyze the problem defined under the assumption that the demand distributions of the periods are known. The equations for the expected total profit of the first stage and the second stage are derived. The optimal order quantity and the optimal threshold are obtained when the demand in each stage follow Poisson distribution, Uniform distribution and Triangular distribution. Right-skewed, Left-skewed and Symmetric Triangular distributions are all used by varying the parameters. The exact results are also found if the distribution of the first stage is Symmetric Triangular and the distribution of the second stage is Uniform. These exact solutions are used in the data generation process as well as in the comparison of data-driven models.

Then, we propose data-driven models for the two-period newsvendor problem that we consider by assuming the demand distributions are not known. When the only information available to the decision maker is a set of historical demand observations, data-driven models first find the threshold to solve the second period problem and consider this threshold with the average total profit of the second period to derive the optimal order quantity. There are six data-driven models that we proposed to solve the newsvendor problem with recourse action hierarchically. They are comprised of the combination of models which are proposed for the second period and proposed for the first period. In particular, we provide three data-driven model to find the optimal threshold based on the remaining inventory at the end of the first period. An algorithm based on SAA, linear regression (LR) model and quadratic regression (QR) model are presented to solve the second period problem. By using the results obtained from one of these models as inputs, two data-driven models are introduced to find the optimal order quantity in the first period, which are an algorithm based on SAA and a mixed integer linear program (MIP).

To summarize, we propose an approach where the algorithm based on SAA is used to solve both stage and call is SAA-P. In the second approach proposed, the algorithm based on SAA is used for the second stage and its results are used as inputs in mixed integer program to solve the first stage. It is called MIP-SAA. In another approach denoted by SAA-LR, we combine the algorithm based on SAA with the linear regression model to find the optimal order quantity and the optimal threshold which is used for the marked-down price decision. Likewise, we propose MIP-LR approach where the results of the linear regression are used as inputs in the mixed integer programming model. we propose another approach where the algorithm based on SAA is used to find the optimal order quantity based on the results obtained by the quadratic regression model in the second stage. It is denoted by SAA-QR. Lastly, we provide MIP-QR approach where the optimal threshold is found by using quadratic regression model whereas the optimal order quantity is found by the mixed-integer programming model which works with the quadratic regression results.

This thesis apart from data-driven models solving the problem defined hierarchically presents an approach where the optimal order quantity and the optimal threshold are found simultaneously and a simple heuristic. We call them MIP-P and MIP-GP, respectively. In heuristic MIP-GP, we run the model defined for the high marked-down price as well as the low marked-down price. Then, we select the price and order quantity combination that yields to the larger average total profit prior to the beginning of the first stage.

Our data-driven models do not require any assumption about the demand distribution. Instead, they use the historical demand data which is generated from selected distributions. We first describe the parameter settings and test their robustness by four performance measures that we defined. These performance measures determine the value of the second period problem and verify the parameters selected. To generate the data, we use the exact solutions to the newsvendor problem with a recourse action obtained under the demand distribution assumption. The data contains how much the product was sold during the first period, marked-down price determined for the remaining inventory, as well as how much the product was sold during the second period. Thus, it is a function of the way that decision made historically. For each demand distribution and for each parameter setting, we generate three different data sets. Namely, they are the data with relevancy level 1, the data with relevancy level 2, and the data with relevancy level 3. The relevancy level is defined based on what decisions the data depends on. For relevancy level 1, the data is generated by implementing the optimal decision. The data with relevancy level 2 is generated by implementing the partially reasonable decision. For relevancy level 3, the data is generated by implementing the random decision.

After generating data sets, they are divided into training data and testing data. We increase the size of the learning data set from 50 to 250 and to 450 while keeping the test data constant at 50. After training data-driven models and obtaining the average profit value from the testing data, we examine the effect of the relevancy levels of data and the effect of the training data size numerically. We carry out a comparison of their performance in three different ways. First of all, we compare the best data-driven model with the worst one. Second of all, we do a comparison with respective expected values obtained by assuming demand distributions are known. Lastly, we measure how many times one method is the best among out-of-samples and calculate its percentage success over all testing samples.

Based on the test conducted for the relevancy level of data, we observe that the difference between the relevancy level 1 and 3 is more significant than the difference defined for the relevancy level 1 and 2. Also, the training data size affects these two differences. In general, the differences defined about the relevancy levels are more noticeable when the training data size increases. However, if the data is not inaccurate, there is not a significant effect of what decision was made in the data for data-driven models. Also, we make the following observation based on the test related to the effect of the training data size. The percentage value calculated by counting the significant test results indicates that there is a difference between the training data size 50 and 250 as well as between the training data size 50 and 450. This difference is increased as the training data size is increased from 50 to 450.

In addition to, we infer that it is important to make a good selection of data-driven models regardless of which distribution is used to generate the data when we compare the best model with the worst one. From the numerical experiments for profit comparison with respective expected values, we can conclude our data-driven models provide a consistently good estimate of the optimal order quantity and the optimal threshold against the exact results obtained in most of the results.

According to the results for inverse CovD, which is calculated by the ratio of the difference between average profit of a model and expected profit computed theoretically to the standard deviation (Formula in 4.15), we can say that the variance of the average profit obtained by each method overall replications is small for comparison and data-driven models are stable in most of the results. To illustrate, the values of inverse CovD of SAA-P is between 0 and 0.35 whereas it is between 0 and 0.43 for MIP-SAA. Lastly, we make the following observation when we look the number of times one method is the best. In general, it is observed that the percentage success of models SAA-P and MIP-SAA increases when the training data size is increased whereas the percentage success of models SAA-QR, MIP-QR, SAA-LR, MIP-LR and MIP-P decreases as the training data size is increased.

Based on the mean values of percentage success, the performance of MIP-LR is better than others for most of the distributions selected. When the training data size is 50, SAA-LR is the second best model in general. the ranking of models from the largest results to smallest changes for each training data size and for each demand distribution selected. Data-driven models show their best performance when the data is generated from Poisson distribution. Another observation made is that MIP-SAA and MIP-P models can have difficulty about dealing with larger training data size and this may affect their performances. Therefore, implementing them with smaller training data size can be better.

There are three directions for follow-up work. First, the feature data can be used with the historical demand observations for data-driven models. We used the historical data in our data-driven models yet our approaches can be extended such that feature data such as weather condition and seasonality is also considered to estimate the decision variables. The second future work is that more than two possible marked down prices might be considered for the second stage. Although, we solve the second period problem when the recourse decision is made between two prices, a generalization for several prices can be made. Similarly, our approaches can be extended in a way that we can find the optimal solutions if more than two marked-down prices are considered. Lastly, we may consider two fashionable products that substitute each other in two-stage newsvendor problem and modify our data-driven models to solve this problem.

Bibliography

[1] E. L. Porteus, "Chapter 12 stochastic inventory theory," in *Stochastic Models*, vol. 2 of *Handbooks in Operations Research and Management Science*, pp. 605 – 652, Elsevier, 1990.

[2] T. M. Whitin, "Inventory control and price theory," *Management Science*, vol. 2, no. 1, pp. 61–68, 1955.

[3] E. L.Porteus, "The newsvendor problem," in *Building intuition insights from basic operations management models and principles* (D. Chhajed and T. J. Lowe, eds.), ch. 7, pp. 115–134, Springer, 2008.

[4] H. Wagner, *Principles of Operations Research: With Applications to Managerial Decisions*. Prentice-Hall international series in management, Prentice-Hall, 1969.

[5] E. A. Silver, D. F. Pyke, and D. J. Thomas, *Inventory and production management in supply chains*. CRC Press, 2016.

[6] T. Nguyen, L. ZHOU, V. Spiegler, P. Ieromonachou, and Y. Lin, "Big data analytics in supply chain management: A state-of-the-art literature review," *Computers and Operations Research*, vol. 98, pp. 254–264, 2018.

[7] S. Saghafian and B. Tomlin, "The newsvendor under demand ambiguity: Combining data with moment and tail information," *Operations Research*, vol. 64, no. 1, pp. 167–185, 2016.

[8] J. Gallien, A. J. Mersereau, A. Garro, A. D. Mora, and M. N. Vidal, "Initial shipment decisions for new products at zara," *Operations Research*, vol. 63, no. 2, pp. 269–286, 2015.

[9] G.-Y. Ban and C. Rudin, "The big data newsvendor: Practical insights from machine learning," *Oper. Res.*, vol. 67, pp. 90–108, 2019.

[10] D. Bertsimas and N. Kallus, "From predictive to prescriptive analytics," *Management Science*, vol. 66, no. 3, pp. 1025–1044, 2020.

[11] A. L. Beutel and S. Minner, "Safety stock planning under causal demand forecasting," in *International Journal of Production Economics*, 2012.

[12] R. Levi, G. Perakis, and J. Uichanco, "The data-driven newsvendor problem: New bounds and insights," *Operations Research*, vol. 63, no. 6, pp. 1294–1306, 2015.

[13] J. Huber, S. Müller, M. Fleischmann, and H. Stuckenschmidt, "A data-driven newsvendor problem: From data to decision," *European Journal of Operational Research*, 2019.

[14] S. Mitra, "Newsvendor problem with clearance pricing," *European Journal of Operational Research*, vol. 268, no. 1, pp. 193–202, 2018.

[15] F. Caro and J. Gallien, "Clearance pricing optimization for a fast-fashion retailer," *Operations Research*, vol. 60, no. 6, pp. 1404–1422, 2012.

[16] Y. Aviv, M. M. Wei, and F. Zhang, "Responsive pricing of fashion products: The effects of demand learning and strategic consumer behavior," *Management Science*, vol. 65, no. 7, pp. 2982–3000, 2019.

[17] A. Aflaki and R. Swinney, "Inventory Integration with Rational Consumers," *Operations Research*, vol. 69, no. 4, pp. 1025–1043, 2021.

[18] J. Rubio-Herrero, M. Baykal-Gursoy, and A. Jaśkiewicz, "A price-setting newsvendor problem under mean-variance criteria," *European Journal of Operational Research*, vol. 247, pp. 575–587, 2015.

[19] M. A. Lariviere and E. L. Porteus, “Stalking information: Bayesian inventory management with unobserved lost sales,” *Management Science*, vol. 45, no. 3, pp. 346–363, 1999.

[20] H. Scarf, “Bayes solutions of the statistical inventory problem,” *Ann Math Stat*, vol. 30, 06 1959.

[21] D. Barrow and N. Kourentzes, “The impact of special days in call arrivals forecasting: A neural network approach to modelling special days,” *European Journal of Operational Research*, vol. 264, no. 3, pp. 967 – 977, 2018.

[22] S. F. Crone, M. Hibon, and K. Nikolopoulos, “Advances in forecasting with neural networks? empirical evidence from the nn3 competition on time series prediction,” *International Journal of Forecasting*, vol. 27, no. 3, pp. 635 – 660, 2011. Special Section 1: Forecasting with Artificial Neural Networks and Computational Intelligence Special Section 2: Tourism Forecasting.

[23] D. Bertsimas and A. Thiele, *Robust and Data-Driven Optimization: Modern Decision Making Under Uncertainty*, ch. Chapter 4, pp. 95–122.

[24] L. Liyanage and J. Shanthikumar, “A practical inventory control policy using operational statistics,” *Oper. Res. Lett.*, vol. 33, pp. 341–348, 07 2005.

[25] A. Kleywegt, A. Shapiro, and T. Homem-De-Mello, “The sample average approximation method for stochastic discrete optimization,” *SIAM Journal on Optimization*, vol. 12, no. 2, pp. 479–502, 2002.

[26] S. Müller, J. Huber, M. Fleischmann, and H. Stuckenschmidt, “Data-driven inventory management under customer substitution,” *Available at SSRN 3624026*, 2020.

[27] A. Oroojlooyjadid, L. V. Snyder, and M. Takáč, “Applying deep learning to the newsvendor problem,” *IISE Transactions*, pp. 1–37, 2019.

[28] H. Scarf, “A min-max solution of an inventory problem,” *Studies in the Mathematical Theory of Inventory and Production*, pp. 201–209, 1958.

[29] G. Gallego and I. Moon, "The Distribution Free Newsboy Problem: Review and Extensions," *Journal of the Operational Research Society*, vol. 44, no. 8, pp. 825–834, 1993.

[30] G. Perakis and G. Roels, "Regret in the newsvendor model with partial information," *Operations Research*, vol. 56, no. 1, pp. 188–203, 2008.

[31] H. S. Lau and A. H. L. Lau, "Estimating the demand distributions of single-period items having frequent stockouts," *European Journal of Operational Research*, vol. 92, no. 2, pp. 254–265, 1996.

[32] W. T. Huh, R. Levi, P. Rusmevichientong, and J. B. Orlin, "Adaptive data-driven inventory control with censored demand based on kaplan-meier estimator," *Operations Research*, vol. 59, no. 4, pp. 929–941, 2011.

[33] D. Bertsimas and A. Thiele, "A data-driven approach to newsvendor problems," *Optimization Online*, no. May, pp. 1–31, 2006.

[34] D. Bertsimas and A. Thiele, "A robust optimization approach to inventory theory," *Operations Research*, vol. 54, no. 1, pp. 150–168, 2006.

[35] R. Levi, R. O. Roundy, and D. B. Shmoys, "Provably near-optimal sampling-based policies for stochastic inventory control models," *Mathematics of Operations Research*, vol. 32, no. 4, pp. 821–839, 2007.

[36] A. L. Sachs and S. Minner, "The data-driven newsvendor with censored demand observations," *International Journal of Production Economics*, vol. 149, pp. 28–36, 2014.

[37] A. M. Law, *Simulation Modeling & Analysis*. New York, NY, USA: McGraw-Hill, 5 ed., 2015.

Appendix A

An Alternative Expression for The Expected Second Period Profit Function

The total expected profit of the second period when the retailer has I inventory level is written as below.

$$
\begin{aligned}
\mathbb{E}[\Pi^2(I)|p_2^i] &= \int_0^I \big[p_2^i y + q(I-y)\big] f_{\xi_2|p_2^i}(y)dy + \int_I^\infty p_2^i I f_{\xi_2|p_2^i}(y)dy \quad \forall i,j = H,L \\
&= p_2^i \int_0^I y f_{\xi_2|p_2^i}(y)dy + qI \int_0^I f_{\xi_2|p_2^i}(y)dy - q \int_0^I y f_{\xi_2|p_2^i}(y)dy \\
&+ p_2^i I\big(1 - \int_0^I f_{\xi_2|p_2^i}(y)dy\big) \quad \forall i,j = H,L \\
&= p_2^i I + \int_0^I ((p_2^i - q)(y-I)) f_{\xi_2|p_2^i}(y)dy \quad \forall i,j = H,L \qquad \text{(A.1)}
\end{aligned}
$$

Appendix B

An Alternative Expression for The Total Profit Function

Recall that the total profit function is

$$\Pi(Q,\xi_1) = \begin{cases} -Qv + p_1\xi_1 + \mathbb{E}[\Pi^2(I)|p_2^H], & \text{if } 0 \le I \le I^t \\ -Qv + p_1\xi_1 + \mathbb{E}[\Pi^2(I)|p_2^L], & \text{if } I \ge I^t \\ -Qv + p_1 Q - B(\xi_1 - Q), & \text{if } I \le 0 \end{cases} \tag{B.1}$$

where

$$\Pi(I|p_2^i) = \begin{cases} p_2^i\xi_2 + q(I-\xi_2), & \text{if } I \ge \xi_2 \\ p_2^i I, & \text{if } I \le \xi_2 \end{cases} \quad \forall i = H, L \tag{B.2}$$

Since $I = Q-\xi_1$, the first conditions of the equation in (B.1) and the equation in (B.2) can be written as $\xi_2 \le Q - \xi_1 \le I^t$. We can rewrite the inequality $Q - \xi_1 \le I^t$ from the equation (B.1) and the inequality $Q - \xi_1 \le \xi_2$ from the equation (B.2) by using $\min(I^t, \xi_2)$. Similarly, we can write the condition $Q-\xi_1 \ge \max(I^t, \xi_2)$ for the case where $Q - \xi_1 \ge I^t$ and $Q - \xi_1 \ge \xi_2$. By combining these inequalities, the total profit turns out to be

$$\Pi(Q,\xi_1) = \begin{cases} -Qv + p_1\xi_1 + (p_2^1\xi_2 + q(Q-\xi_1-\xi_2)), & \text{if } \xi_2 \leq Q-\xi_1 \leq I^t \\ -Qv + p_1\xi_1 + p_2^1(Q-\xi_1), & \text{if } Q-\xi_1 \leq \min(I^t,\xi_2) \\ -Qv + p_1\xi_1 + (p_2^2\xi_2 + q(Q-\xi_1-\xi_2)), & \text{if } Q-\xi_1 \geq \max(I^t,\xi_2) \\ -Qv + p_1\xi_1 + p_2^1(Q-\xi_1), & \text{if } I^t \leq Q-\xi_1 \leq \xi_2 \\ -Qv + p_1 Q - B(\xi_1 - Q), & \text{if } Q \leq \xi_1 \end{cases} \tag{B.3}$$

To sum up, the first part of the function in (B.3) is used for the case where $0 \leq I \leq I^t$ and $I \geq \xi_2$. The second part of the function is used for all values of I, which is $Q-\xi_1$, less than or equal to I^t and ξ_2. Similarly, the condition $Q-\xi_1 \geq \max(I^t,\xi_2)$ should be satisfied for the third sub-function. Moreover, for all values of $Q-\xi_1$ satisfying $Q-\xi_1 \leq \xi_2$ and $Q-\xi_1 \geq I^t$, the next sub-function is used. The last part is applied for all values of $Q-\xi_1$ less than or equal to zero.

Appendix C

Algorithm for Finding Optimal Solution with Poisson Distributed Demand

C.1 Algorithm for The Second Period Problem

The threshold value as a solution of the second period and the optimal order quantity are found by using the following algorithms:

Pdf of Poisson Distribution $(f_{\xi_1}(k), f_{\xi_2|p_2}(k))$

Input: λ, k

Figure C.1 Algorithm to find pdf of Poisson distribution

result $\leftarrow e^{-\lambda}$
if $k == 0$ **then**
    result $\leftarrow e^{-\lambda}$
else
    for $i \in \{1 \ldots k\}$ **do**
        result $\leftarrow$ result $\times \frac{\lambda}{i}$
    end for
end if
return result

The Second Period Expected Profit Function: $(\mathbb{E}[\Pi^2(I, \xi_2)|p_2])$

Input: I_1, q, p_2

Figure C.2 Algorithm to find the second period expected profit when demand is Poisson

$SUM_1 \leftarrow 0$
$SUM_2 \leftarrow 0$
for $i \in \{0 \ldots I\}$ **do**
    $SUM_1 \leftarrow SUM_1 + [p_2 i + q(I - i)] \times f_{\xi_2|p_2}(i)$
    $SUM_2 \leftarrow SUM_2 + [p_2 I] \times f_{\xi_2|p_2}(i)$
end for
return SUM_1+$p_2 I$-SUM_2

Figure C.3 Algorithm for the second period problem when demand is Poisson

Given $p_2 = (p_2^H, p_2^L)$
$I^t \leftarrow \infty$
for $i \in \{0 \dots I^{max}\}$ **do**
  if $\mathbb{E}[\Pi^2(i, \xi_2)|p_2^H] - \mathbb{E}[\Pi^2(i, \xi_2)|p_2^L] \geq 0$ and $\mathbb{E}[\Pi^2(i+1, \xi_2)|p_2^H] - \mathbb{E}[\Pi^2(i+1, \xi_2)|p_2^L] < 0$ **then**
    $I^t \leftarrow i + 1$
  end if
end for

C.2 Algorithm for The First Period Problem

The First Period Expected Total Profit Function ($\mathbb{E}[\Pi(Q, \xi_1)]$)

Input: $Q, v, p_1, B, q, p_2 = (p_2^H, p_2^L), \lambda_1, \lambda_2 = (\lambda_2^H, \lambda_2^L)$

Figure C.4 Algorithm to calculate the first period expected total profit when demand is Poisson

$SUM_1, SUM_2, SUM_3, SUM_4 \leftarrow 0$
for $i \in \{0 \dots Q - I^t\}$ **do**
  $SUM_1 \leftarrow SUM_1 + [-Qv + p_1 i + \mathbb{E}[\Pi^2(Q - i, \xi_2)|p_2^L)] f_{\xi_1}(i)$
end for
for $i \in \{(Q - I^t + 1) \dots Q\}$ **do**
  $SUM_2 \leftarrow SUM_2 + [-Qv + p_1 i + \mathbb{E}[\Pi^2(Q - i, \xi_2)|p_2^H)] f_{\xi_1}(i)$
end for
for $i \in \{0 \dots Q\}$ **do**
  $SUM_3 \leftarrow SUM_3 + [-Qv + p_1 Q + BQ] f_{\xi_1}(i)$
  $SUM_4 \leftarrow SUM_4 + [-Bi] f_{\xi_1}(i)$
end for
return $SUM_1 + SUM_2 + [-Qv + p_1 Q + BQ] - SUM_3 - B\lambda_1 - SUM_4$

Figure C.5 Algorithm for the newsvendor problem with recourse action when demand is Poisson

$Q^{\star} \leftarrow 0$
$\mathbb{E}[\Pi^1(Q^{\star}, \xi_1)] \leftarrow -\infty$
for $j \in \{0 \dots Q^{max}\}$ **do**
    if $\mathbb{E}[\Pi(j, \xi_1)] > \mathbb{E}[\Pi(Q^{\star}, \xi_1)]$ **then**
        $Q^{\star} \leftarrow j$
        $\mathbb{E}[\Pi(Q^{\star}, \xi_1)] \leftarrow \mathbb{E}[\Pi(j, \xi_1)]$
    end if
end for
return $Q^{\star}, \mathbb{E}[\Pi(Q^{\star}, \xi_1)]$

Appendix D

The Second Period Problem for Uniform Distributed Demand

Let's assume that $\xi_{2|p_2^H} \sim U(a^H, b^H)$ and $\xi_{2|p_2^L} \sim U(a^L, b^L)$.

Suppose that $I \geq 0$, $a^H = a^L = 0$ and $\mu^H = \frac{a^H+b^H}{2} < \mu^L = \frac{a^L+b^L}{2}$.

The uniform probability density function

$$f(y) = \begin{cases} \frac{1}{b-a} & \text{for } y \in [a,b] \\ 0 & \text{otherwise} \end{cases} \tag{D.1}$$

and the uniform cumulative distribution function

$$F_{\xi_{2|p_2^i}}(I) = \int_0^I f_{\xi_{2|p_2^i}}(y)dy = \begin{cases} 0 & \text{if } I < a^i \\ \frac{I-b^i}{b^i-a^i} & \text{if } a^i \leq I \leq b^i \\ 1 & \text{if } b^i < I \end{cases} \quad \forall i = H, L \tag{D.2}$$

$$\int_{I}^{b^i} f_{\xi_2|p_2^i}(y)dy = \begin{cases} 1 & \text{if } I < a^i \\ \frac{b^i - I}{b^i - a^i} & \text{if } a^i \leq I \leq b^i \ \forall i = H, L \\ 0 & \text{if } b^i < I \end{cases} \tag{D.3}$$

$$\int_{0}^{I} y f_{\xi_2|p_2^i}(y)dy = \begin{cases} 0 & \text{if } I < a^i \\ \frac{I^2 - (b^i)^2}{2(b^i - a^i)} & \text{if } a^i \leq I \leq b^i \ \forall i = H, L \\ \mu^i & \text{if } b^i < I \end{cases} \tag{D.4}$$

$$\int_{I}^{b^i} y f_{\xi_2|p_2^i}(y)dy = \begin{cases} \mu^i & \text{if } I < a^i \\ \frac{(b^i)^2 - I^2}{2(b^i - a^i)} & \text{if } a^i \leq I \leq b^i \ \forall i = H, L \\ 0 & \text{if } b^i < I \end{cases} \tag{D.5}$$

$$\mu^i = \frac{b^i}{2} \quad \forall i = H, L$$

1. For $a^i \leq I \leq b^i \quad \forall i = H, L$

   -
     $$\int_{I}^{b^i} y f_{\xi_2|p_2^i}(y)dy = \frac{1}{2(b^i - a^i)} y^2 \bigg|_{I}^{b^i} = \frac{(b^i)^2 - I^2}{2(b^i - a^i)} \quad \forall i = H, L$$

   -
     $$\int_{I}^{b^i} f_{\xi_2|p_2^i}(y)dy = \frac{1}{2(b^i - a^i)} y \bigg|_{I}^{b^i} = \frac{b^i - I}{b^i - a^i} \quad \forall i = H, L$$

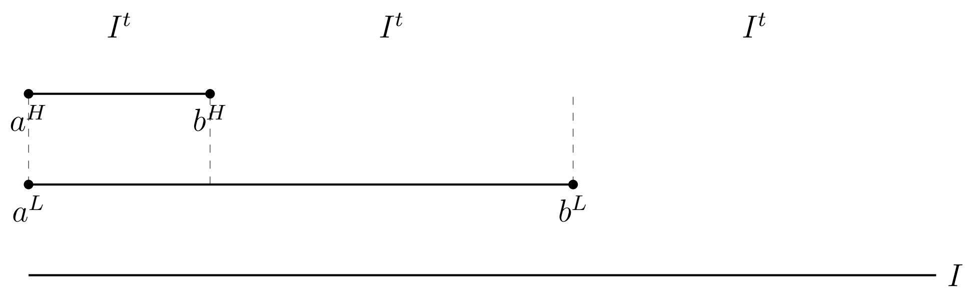

Figure D.1: Intervals for Uniform demand distribution

1. $0 \leq I \leq b^i \quad \forall i = H, L$

$$\begin{aligned}
\mathbb{E}[\Pi^2(I)] &= qI + (p_2^i - q)\Big(\mu^i + \int_I^{b^i} (I - y) f_{\xi_2|p_2^i}(y) dy\Big) \\
&= qI + (p_2^i - q)\Big(\frac{b^i}{2} + \frac{2(b^i - I)I}{2b^i} - \frac{(b^i)^2 - I^2}{2b^i}\Big) \\
&= -I^2\Big(\frac{p_2^i - q}{2b^H}\Big) + I\Big(p_2^i - q + q\Big) + (p_2^i - q)\Big(\frac{b^i}{2} - \frac{b^i}{2}\Big) \\
\mathbb{E}[\Pi^2(I)] &= -I^2\Big(\frac{p_2^i - q}{2b^i}\Big) + Ip_2^i
\end{aligned} \tag{D.6}$$

2. $b^i \leq I \quad \forall i = H, L$

$$\begin{aligned}
\mathbb{E}[\Pi^2(I)] &= qI + (p_2^i - q)\Big(\mu^i + \cancelto{0}{\int_I^{b^i} (I - y) f_{\xi_2|p_2^i}(y) dy}\Big) \\
\mathbb{E}[\Pi^2(I)] &= qI + (p_2^i - q)\frac{b^i}{2}
\end{aligned} \tag{D.7}$$

Given the marked down price, the expected second period profit of the inventory level I and is

$$\mathbb{E}[\Pi^2(I)|p_2^i] = \begin{cases} -I^2\Big(\frac{p_2^i - q}{2b^i}\Big) + Ip_2^i, & \text{if } I < I^t \text{and } 0 \leq I \leq b^i \\ Iq + (p_2^i - q)\frac{b^i}{2}, & \text{if } I < I^t \text{and } b^i < I \end{cases} \quad \forall i = H, L \tag{D.8}$$

D.1 Expected 2^{nd} Period Profit when Demand is Uniform

If $0 \leq I^t \leq b^H$, the solution of the second period can be shown as follows.

$$\mathbb{E}[\Pi^2(I)] = \begin{cases} -I^2\left(\frac{p_2^H - q}{2b^H}\right) + Ip_2^H, & \text{if } I < I^t \text{and } 0 \leq I \leq b^H \\ -I^2\left(\frac{p_2^L - q}{2b^L}\right) + Ip_2^L, & \text{if } I \geq I^t \text{and } 0 \leq I \leq b^L \\ Iq + (p_2^L - q)\frac{b^L}{2}, & \text{if } I \geq I^t \text{and } b^L < I \end{cases} \tag{D.9}$$

If $b^H \leq I^t \leq b^L$, the solution of the second period can be shown as follows.

$$\mathbb{E}[\Pi^2(I)] = \begin{cases} -I^2\left(\frac{p_2^H - q}{2b^H}\right) + Ip_2^H, & \text{if } I < I^t \text{and } 0 \leq I \leq b^H \\ Iq + (p_2^H - q)\frac{b^H}{2}, & \text{if } I < I^t \text{and } b^H < I \\ -I^2\left(\frac{p_2^L - q}{2b^L}\right) + Ip_2^L, & \text{if } I \geq I_1^{(t)} \text{and } 0 \leq I \leq b^L \\ Iq + (p_2^L - q)\frac{b^L}{2}, & \text{if } I \geq I^t \text{and } b^L < I \end{cases} \tag{D.10}$$

If $b^L \leq I^t$, the solution of the second period can be shown as follows.

$$\mathbb{E}[\Pi^2(I)] = \begin{cases} -I^2\left(\frac{p_2^H - q}{2b^H}\right) + Ip_2^H, & \text{if } I < I^t \text{and } 0 \leq I \leq b^H \\ Iq + (p_2^H - q)\frac{b^H}{2}, & \text{if } I < I^t \text{and } b^H < I \\ Iq + (p_2^L - q)\frac{b^L}{2}, & \text{if } I \geq I^t \text{and } b^L < I \end{cases} \tag{D.11}$$

Appendix E

The Second Period Problem for Triangular Distributed Demand

For the Triangular distributed demand, let's assume that

$$\xi_{2|p_2^H} \sim Trian(a^H, c^H, b^H)$$

and

$$\xi_{2|p_2^L} \sim Trian(a^L, c^L, b^L)$$

It is assumed that $a^H = a^L = 0$ and $\mu^H = \frac{a^H+b^H+c^H}{2} < \mu^L = \frac{a^L+b^L+c^H}{2}$. However, there is no assumption about the relation between the parameters b^H and c^L. Thus, there can be a case where the upper limit of Triangular distribution for high price b^H can be less than the shape parameter of Triangular distribution for low price c^L. Other possible case is that the parameter b^H can be greater than the parameter c^L. As a result, The solution to the expected second period problem is derived for two possible cases. Namely case 1 which is $a^H = a^L = 0 < c^H < b^H < c^L < b^L$ and case 2 which is $a^H = a^L = 0 < c^H < c^L < b^H < b^L$.

Figures E.1, E.2, E.3, and E.4 show these two cases.

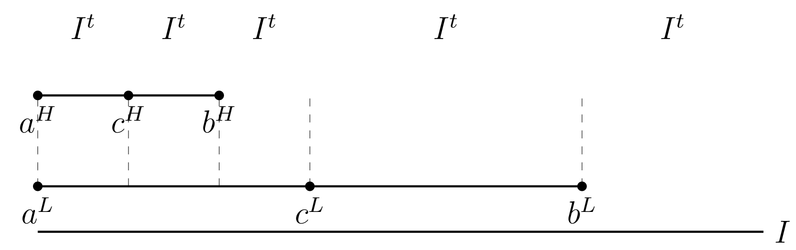

Figure E.1: Intervals for Symmetric Triangular: Case 1

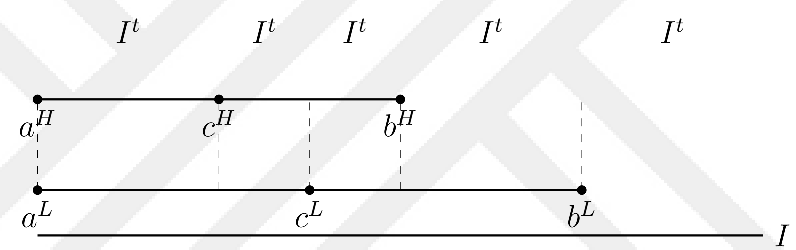

Figure E.2: Intervals for Symmetric Triangular: Case 2

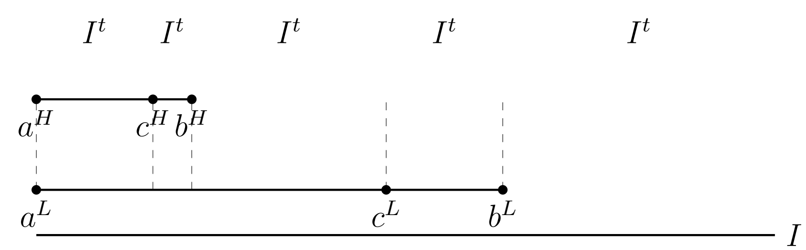

Figure E.3: Intervals for Nonsymmetric Triangular: Case 1

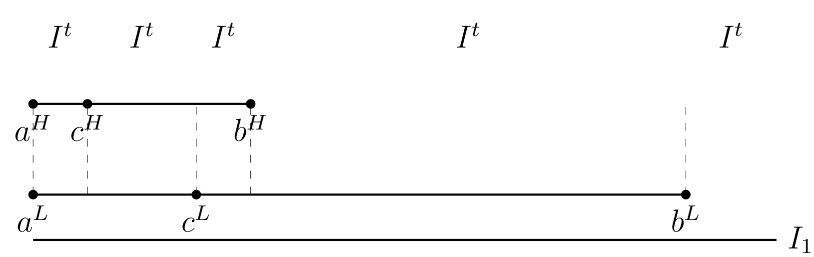

Figure E.4: Intervals for Nonsymmetric Triangular: Case 2

The triangular probabilty density function

$$f_{\xi_2|p_2^i}(y) = \begin{cases} 0 & \text{if } y < a^i \\ \frac{2(y-a^i)}{(c^i-a^i)(b^i-a^i)} & \text{if } a^i \leq y \leq c^i \\ \frac{2(b^i-y)}{(b^i-c^i)(b^i-a^i)} & \text{if } c^i \leq y \leq b^i \\ 0 & \text{if } b^i < y \end{cases} \quad \forall i = H, L \tag{E.1}$$

and the triangular cumulative distribution function

$$F_{\xi_2|p_2^i}(I) = \int_0^I f_{\xi_2|p_2^i}(y)dy = \begin{cases} 0 & \text{if } I < a^i \\ \frac{(I-a^i)^2}{(c^i-a^i)(b^i-a^i)} & \text{if } a^i \leq I \leq c^i \\ 1 - \frac{(b^i-I)^2}{(b^i-c^i)(b^i-a^i)} & \text{if } c^i \leq I \leq b^i \\ 1 & \text{if } b^i < I \end{cases} \quad \forall i = H, L \tag{E.2}$$

$$\int_I^\infty f_{\xi_2|p_2^i}(y)dy = \begin{cases} 1 & \text{if } I < a^i \\ 1 - \frac{(I-a^i)^2}{(c^i-a^i)(b^i-a^i)} & \text{if } a^i \leq I \leq c^i \\ \frac{(b^i-I)^2}{(b^i-c^i)(b^i-a^i)} & \text{if } c^i \leq I \leq b^i \\ 0 & \text{if } b^i < I \end{cases} \quad \forall i = H, L \tag{E.3}$$

$$\int_0^I y f_{\xi_2|p_2^i}(y)dy = \begin{cases} 0 & \text{if } I < a^i \\ \frac{(I-a^i)^2(2I+a^i)}{3(b^i-a^i)(c^i-a^i)} & \text{if } a^i \leq I \leq c^i \\ \mu^i - \frac{(b^i-I)^2(2I+b)}{3(b^i-c^i)(b^i-a^i)} & \text{if } c^i \leq I \leq b^i \\ \mu^i & \text{if } b^i < I \end{cases} \quad \forall i = H, L \tag{E.4}$$

$$\int_I^\infty y f_{\xi_2|p_2^i}(y)dy = \begin{cases} \mu_i & \text{if } I < a^i \\ \mu^i - \frac{(I-a^i)^2(2I+a^i)}{3(b^i-a^i)(c^i-a^i)} & \text{if } a^i \leq I \leq c^i \\ \frac{(b^i-I)^2(2I+b)}{3(b^i-c^i)(b^i-a^i)} & \text{if } c^i \leq I \leq b^i \\ 0 & \text{if } b^i < I \end{cases} \quad \forall i = H, L \tag{E.5}$$

$$\mu^i = \frac{a^i + b^i + c^i}{3} \quad \forall i = H, L$$

1. For $0 \leq I \leq c^i \quad \forall i = H, L$

   -
   $$\begin{aligned} \int_I^\infty y f_{\xi_2|p_2^i}(y)dy &= \int_0^\infty y f_{\xi_2|p_2^i}(y)dy - \int_0^I y f_{\xi_2|p_2^i}(y)dy \\ &= \mu^i - \underbrace{\int_0^I \frac{y2(y-a^i)}{(b^i-a^i)(c^i-a^i)}dy}_{\frac{2y^3-3y^2a^i}{3(b^i-c^i)(b^i-a^i)}\Big|_0^I} \\ &= \mu^i - \frac{(I-a^i)^2(2I+a^i)}{3(b^i-a^i)(c^i-a^i)} \end{aligned}$$

   -
   $$\begin{aligned} \int_I^\infty f_{\xi_2|p_2^i}(y)dy &= \int_0^\infty f_{\xi_2|p_2^i}(y)dy - \int_0^I f_{\xi_2|p_2^i}(y)dy \\ &= 1 - \underbrace{\int_0^I \frac{2(y-a^i)}{(b^i-a^i)(c^i-a^i)}dy}_{\frac{y^2-2a^iy}{(b^i-a^i)(c^i-a^i)}\Big|_0^I} \\ &= 1 - \frac{(I-a^i)^2}{(c^i-a^i)(b^i-a^i)} \end{aligned}$$

2. For $c^i \leq I \leq b^i \quad \forall i = H, L$

- $$\int_I^\infty y f_{\xi_2|p_2^i}(y)dy = \underbrace{\int_I^{b^i} \frac{y2(b^i-y)}{(b^i-c^i)(b^i-a^i)}dy}_{\frac{3y^2b^i-2y^3}{3(b^i-c^i)(b^i-a^i)}\Big|_I^{b^i}} + \underbrace{\int_{b^i}^\infty y f_{\xi_2|p_2^i}(y)dy}_{0}$$
$$= \frac{(b^i-I)^2(2I+b)}{3(b^i-c^i)(b^i-a^i)}$$

- $$\int_I^\infty f_{\xi_2|p_2^i}(y)dy = \underbrace{\int_I^{b^i} \frac{2(b^i-y)}{(b^i-c^i)(b^i-a^i)}dy}_{\frac{2b^iy-y^2}{(b^i-c^i)(b^i-a^i)}\Big|_I^{b^i}} = \frac{(b^i-I)^2}{(b^i-c^i)(b^i-a^i)}$$

E.1 Expected Profit Function when Demand is Triangular

Given the marked down price, the second period profit function can be written by using equations in (E.3) and (E.5) as follows.

$$\mathbb{E}[\Pi^2(I)|p_2^i] = \begin{cases} \left[-\frac{(p_2^i-q)}{3b^ic^i}\right]I^3 + p_2^iI & \text{if } 0 \le I \le c^i \\ \left[\frac{p_2^i-q}{3(b^i-c^i)b^i}\right]I^3 - \frac{p_2^i-q}{b^i-c^i}I^2 & \text{if } c^i < I < b^i \\ +\left[\frac{(p_2^i-q)b^i}{b^i-c^i}+q\right]I - \frac{(p_2^i-q)(c^i)^2}{3(b^i-c^i)} & \\ qI + (p_2^i-q)\mu^i & \text{if } b^i < I \end{cases} \quad \forall i = H, L \tag{E.6}$$

The threshold value I^t is defined as the point where the expected profit difference is zero. Hence,

$$
\begin{aligned}
&\mathbb{E}[\Pi^2(I)|p_2^H] - \mathbb{E}[\Pi^2(I)|p_2^L] = 0 \\
&\left[qI + (p_2^H - q)\left(\mu^H - \int_I^\infty (y-I) f_{\xi_2|p_2^H}(y)dy\right)\right] \\
-&\left[qI + (p_2^L - q)\left(\mu^L - \int_I^\infty (y-I) f_{\xi_2|p_2^L}(y)dy\right)\right] = 0
\end{aligned} \tag{E.7}
$$

E.2 Expected 2^{nd} Period Profit for Case 1

If $0 \leq I^t \leq c^H$, the solution of the second period can be shown as follows.

$$
\mathbb{E}[\Pi^2(I)] = \begin{cases}
\left[-\frac{(p_2^H - q)}{3b^H c^H}\right] I^3 + p_2^H I, & \text{if } I < I^t \text{and } 0 \leq I \leq c^H \\
\left[-\frac{(p_2^L - q)}{3b^L c^L}\right] I^3 + p_2^L I, & \text{if } I \geq I^t \text{and } 0 \leq I \leq c^L \\
\left[\frac{p_2^L - q}{3(b^L - c^L)b^L}\right] I^3 - \frac{p_2^L - q}{b^L - c^L} I^2, & \text{if } I \geq I^t \text{and } c^L < I \leq b^L \\
qI + (p_2^L - q)\mu^L & \text{if } I \geq I^t \text{and } b^L < I
\end{cases} \tag{E.8}
$$

If $c^H \leq I^t \leq b^H$, the solution of the second period can be shown as follows.

$$
\mathbb{E}[\Pi^2(I)] = \begin{cases}
\left[-\frac{(p_2^H - q)}{3b^H c^H}\right] I^3 + p_2^H I, & \text{if } I < I^t \text{and } 0 \leq I \leq c^H \\
\left[\frac{p_2^H - q}{3(b^H - c^H)b^H}\right] I^3 - \frac{p_2^H - q}{b^H - c^H} I^2, & \text{if } I < I^t \text{and } c^H \leq I \leq b^H \\
\left[-\frac{(p_2^L - q)}{3b^L c^L}\right] I^3 + p_2^L I, & \text{if } I \geq I^t \text{and } 0 \leq I \leq c^L \\
\left[\frac{p_2^L - q}{3(b^L - c^L)b^L}\right] I^3 - \frac{p_2^L - q}{b^L - c^L} I^2, & \text{if } I \geq I^t \text{and } c^L < I \leq b^L \\
qI + (p_2^L - q)\mu^L & \text{if } I \geq I^t \text{and } b^L < I
\end{cases} \tag{E.9}
$$

If $b^H \leq I^t \leq c^L$, the solution of the second period can be shown as follows.

$$\mathbb{E}[\Pi^2(I)] = \begin{cases} \left[-\frac{(p_2^H-q)}{3b^Hc^H}\right]I^3 + p_2^HI, & \text{if } I < I^t\text{and } 0 \leq I \leq c^H \\ \left[\frac{p_2^H-q}{3(b^H-c^H)b^H}\right]I^3 - \frac{p_2^H-q}{b^H-c^H}I^2, & \text{if } I < I^t\text{and } c^H \leq I \leq b^H \\ qI + (p_2^H - q)\mu^H & \text{if } I < I^t\text{and } b^H < I \\ \left[-\frac{(p_2^L-q)}{3b^Lc^L}\right]I^3 + p_2^LI, & \text{if } I \geq I^t\text{and } 0 \leq I \leq c^L \\ \left[\frac{p_2^L-q}{3(b^L-c^L)b^L}\right]I^3 - \frac{p_2^L-q}{b^L-c^L}I^2, & \text{if } I \geq I^t\text{and } c^L < I \leq b^L \\ qI + (p_2^L - q)\mu^L & \text{if } I \geq I^t\text{and } b^L < I \end{cases} \tag{E.10}$$

If $c^L \leq I^t \leq b^L$, the solution of the second period can be shown as follows.

$$\mathbb{E}[\Pi^2(I)] = \begin{cases} \left[-\frac{(p_2^H-q)}{3b^Hc^H}\right]I^3 + p_2^HI, & \text{if } I < I^t\text{and } 0 \leq I \leq c^H \\ \left[\frac{p_2^H-q}{3(b^H-c^H)b^H}\right]I^3 - \frac{p_2^H-q}{b^H-c^H}I^2, & \text{if } I < I^t\text{and } c^H \leq I \leq b^H \\ qI + (p_2^H - q)\mu^H & \text{if } I < I^t\text{and } b^H < I \\ \left[\frac{p_2^L-q}{3(b^L-c^L)b^L}\right]I^3 - \frac{p_2^L-q}{b^L-c^L}I^2, & \text{if } I \geq I^t\text{and } c^L < I \leq b^L \\ qI + (p_2^L - q)\mu^L & \text{if } I \geq I^t\text{and } b^L < I \end{cases} \tag{E.11}$$

If $b^L \leq I^t$, the solution of the second period can be shown as follows.

$$\mathbb{E}[\Pi^2(I)] = \begin{cases} \left[-\frac{(p_2^H-q)}{3b^Hc^H}\right]I^3 + p_2^HI, & \text{if } I < I^t\text{and } 0 \leq I \leq c^H \\ \left[\frac{p_2^H-q}{3(b^H-c^H)b^H}\right]I^3 - \frac{p_2^H-q}{b^H-c^H}I^2, & \text{if } I < I^t\text{and } c^H \leq I \leq b^H \\ qI + (p_2^H - q)\mu^H & \text{if } I < I^t\text{and } b^H < I \\ qI + (p_2^L - q)\mu^L & \text{if } I \geq I^t\text{and } b^L < I \end{cases} \tag{E.12}$$

E.3 Expected 2^{nd} Period Profit for Case 2

If $a^H \leq I^t \leq c^H$, the solution of the second period is shown in equation (E.8). For $c^H \leq I^t \leq c^L$, then it is expressed in equation (E.9).

For $c^L \leq I^t \leq b^H$, the solution of the second period can be shown as follows.

$$\mathbb{E}[\Pi^2(I)] = \begin{cases} \left[-\frac{(p_2^H - q)}{3b^H c^H}\right] I^3 + p_2^H I, & \text{if } I < I^t \text{and } 0 \leq I \leq c^H \\ \left[\frac{p_2^H - q}{3(b^H - c^H)b^H}\right] I^3 - \frac{p_2^H - q}{b^H - c^H} I^2, & \text{if } I < I^t \text{and } c^H \leq I \leq b^H \\ \left[\frac{p_2^L - q}{3(b^L - c^L)b^L}\right] I^3 - \frac{p_2^L - q}{b^L - c^L} I^2, & \text{if } I \geq I^t \text{and } c^L < I \leq b^L \\ qI + (p_2^L - q)\mu^L & \text{if } I \geq I^t \text{and } b^L < I \end{cases} \tag{E.13}$$

If $b^H \leq I^t \leq b^L$, the solution of the second period is shown in equation (E.11).

If $b^L \leq I^t$, the solution of the second period is shown in equation (E.12).

$$\mathbb{E}[\Pi^2(I)] = \begin{cases} \left[-\frac{(p_2^H - q)}{3b^H c^H}\right] I^3 + p_2^H I, & \text{if } I < I^t \text{and } 0 \leq I \leq c^H \\ \left[\frac{p_2^H - q}{3(b^H - c^H)b^H}\right] I^3 - \frac{p_2^H - q}{b^H - c^H} I^2, & \text{if } I < I^t \text{and } c^H \leq I \leq b^H \\ \left[\frac{p_2^L - q}{3(b^L - c^L)b^L}\right] I^3 - \frac{p_2^L - q}{b^L - c^L} I^2, & \text{if } I \geq I^t \text{and } c^L < I \leq b^L \\ qI + (p_2^L - q)\mu^L & \text{if } I \geq I^t \text{and } b^L < I \end{cases} \tag{E.14}$$

E.4 Expected Profit Difference Function for Case 1

Using equations in (E.3) and (E.5), $\mathbb{E}[\Pi^2(I)|p_2^H] - \mathbb{E}[\Pi^2(I)|p_2^L]$ is

$$\begin{cases} \underbrace{\left[\frac{(p_2^L-q)}{3b^Lc^L} - \frac{(p_2^H-q)}{3b^Hc^H}\right]}_{A\leq 0} I^3 + \underbrace{\left[p_2^H - p_2^L\right]}_{B\geq 0} I & \text{if } 0 \leq I \leq c^H < c^L \\ \underbrace{\left[\frac{p_2^H-q}{3(b^H-c^H)b^H} + \frac{p_2^L-q}{3b^Lc^L}\right]}_{A\geq 0} I^3 \underbrace{-\frac{p_2^H-q}{b^H-c^H}}_{B\leq 0} I^2 & \text{if } c^H < I < b^H \\ + \underbrace{\left[\frac{(p_2^H-q)b^H}{b^H-c^H} - (p_2^L-q)\right]}_{C} I \underbrace{-\frac{(p_2^H-q)(c^H)^2}{3(b^H-c^H)}}_{D\leq 0} & \text{and } 0 \leq I < c^L \\ \underbrace{(p_2^L-q)}_{A\geq 0} I^3 \underbrace{-3(p_2^L-q)b^Lc^L}_{C\leq 0} I & \text{if } b^H < I \\ \underbrace{+(p_2^H-q)(b^H+c^H)b^Lc^L}_{D\geq 0} & \text{and } a^L \leq I \leq c^L \\ \underbrace{\left[\frac{-(p_2^L-q)}{3(b^L-c^L)b^L}\right]}_{A\leq 0} I^3 \underbrace{+\frac{p_2^L-q}{b^L-c^L}}_{B\geq 0} I^2 \underbrace{-\frac{(p_2^L-q)b^L}{b^L-c^L}}_{C\leq 0} I & \text{if } b^H < I \\ \underbrace{+(p_2^H-q)\frac{(b^H+c^H)}{3} + \frac{(p_2^L-q)(c^L)^2}{3(b^L-c^L)}}_{D\geq 0} & \text{and } c^L \leq I \leq b^L \\ (p_2^H-q)\mu^H - (p_2^L-q)\mu^L & \text{if } b^H < b^L < I \end{cases} \tag{E.15}$$

1. $a^i \leq I \leq c^i \quad \forall i = H, L$

   (Note that $a^i = 0 \quad \forall i = H, L$ and $a^H = a^L < c^H < b^H < c^L < b^L$)

$$
\begin{aligned}
&\left[qI + (p_2^H - q)\left(\mu^H - \int_I^{b^H} (y-I) f_{\xi_2|p_2^H}(y)dy\right)\right] \\
&-\left[qI + (p_2^L - q)\left(\mu^L - \int_I^{b^L} (y-I) f_{\xi_2|p_2^L}(y)dy\right)\right] = 0 \\
&\left[\cancel{qI} + (p_2^H - q)\left(\cancel{\frac{b^H + c^H}{3}} - \left(\cancel{\frac{b^H + c^H}{3}} - \frac{2I^3}{3b^H c^H}\right) + I\left(1 - \frac{3I^2}{3b^H c^H}\right)\right)\right] \\
&-\left[\cancel{qI} + \cancel{(p_2^L - q)\left(\frac{b^L + c^L}{3}\right)} - \cancel{(p_2^L - q)\left(\frac{b^L + c^L}{3}\right)} + (p_2^L - q)\left(\frac{2I^3}{3b^L c^L}\right)\right. \\
&\left. +(p_2^L - q)I\left(1 - \frac{3I^2}{3b^L c^L}\right)\right] = 0 \\
&\left[\frac{(p_2^L - q)}{3b^L c^L} - \frac{(p_2^H - q)}{3b^H c^H}\right] I^3 + \left[(p_2^H - \cancel{q}) - (p_2^L - \cancel{q})\right] I = 0 \\
&\underbrace{\left[\frac{(p_2^L - q)}{3b^L c^L} - \frac{(p_2^H - q)}{3b^H c^H}\right]}_{A \leq 0} I^3 + \underbrace{\left[p_2^H - p_2^L\right]}_{B \geq 0} I = 0 \\
&I_1^t = \frac{B}{A} = \frac{9b^L c^L b^H c^H (p_2^H - p_2^L)}{3b^H c^H (p_2^L - q) - 3b^L c^L (p_2^H - q)}
\end{aligned}
\tag{E.16}
$$

2. $c^H \leq I \leq b^H$ and $a^L \leq I \leq c^L$

   (Note that $a^i = 0 \quad \forall i = H, L$ and $a^H = a^L < c^H < b^H < c^L < b^L$)

$$
\begin{aligned}
&\left[qI + (p_2^H - q)\left(\mu^H - \int_I^{b^H} (y - I) f_{\xi_2|p_2^H}(y)dy\right)\right] \\
&-\left[qI + (p_2^L - q)\left(\mu^L - \int_I^{b^L} (y - I) f_{\xi_2|p_2^L}(y)dy\right)\right] = 0 \\
&\left[\cancel{qI} + (p_2^H - q)\left(\frac{b^H + c^H}{3} - \frac{(b^H - I)^2(2I + b^H)}{3(b^H - c^H)b^H} + \frac{3I(b^H - I)^2}{3(b^H - c^H)b^H}\right)\right] \\
&-\left[\cancel{qI} + \cancel{(p_2^L - q)\left(\frac{b^L + c^L}{3}\right)} - \cancel{(p_2^L - q)\left(\frac{b^L + c^L}{3}\right)} + (p_2^L - q)\left(\frac{2I^3}{3b^L c^L}\right)\right. \\
&\left.+(p_2^L - q)I\left(1 - \frac{3I^2}{3b^L c^L}\right)\right] = 0 \\
&\left[(p_2^H - q)\left(\frac{b^H + c^H}{3} + \frac{(b^H - I)^2(I - b^H)}{3(b^H - c^H)b^H}\right)\right] \\
&+\frac{(p_2^L - q)}{3b^L c^L} I^3 - (p_2^L - q)I = 0 \\
&\left[(p_2^H - q)\left(\frac{b^H + c^H}{3} + \frac{I^3 + 3(b^H)^2 I - 3b^H I^2 - (b^H)^3}{3(b^H - c^H)b^H}\right)\right] \\
&+\frac{(p_2^L - q)}{3b^L c^L} I^3 - (p_2^L - q)I = 0 \\
&\underbrace{\left[\frac{p_2^H - q}{3(b^H - c^H)b^H} + \frac{p_2^L - q}{3b^L c^L}\right]}_{A \geq 0} I^3 \underbrace{- \frac{p_2^H - q}{b^H - c^H}}_{B \leq 0} I^2 \\
&+ \underbrace{\left[\frac{(p_2^H - q)b^H}{b^H - c^H} - (p_2^L - q)\right]}_{C} I \underbrace{- \frac{(p_2^H - q)(c^H)^2}{3(b^H - c^H)}}_{D \leq 0} = 0
\end{aligned}
\tag{E.17}
$$

3. $b^H < I$ and $a^L \leq I \leq c^L$

   (Note that $a^i = 0 \quad \forall i = H, L$ and $a^H = a^L < c^H < b^H < c^L < b^L$)

$$
\begin{aligned}
&\left[qI + (p_2^H - q)\left(\mu^H - \int_{b^H}^{\infty}(y-I)f_{\xi_2|p_2^H}(y)dy\right)\right] \\
&-\left[qI + (p_2^L - q)\left(\mu^L - \int_{I}^{b^L}(y-I)f_{\xi_2|p_2^L}(y)dy\right)\right] = 0 \\
&\left[\cancel{qI} + (p_2^H - q)\left(\frac{b^H + c^H}{3} - \cancelto{0}{\int_{b^H}^{\infty}(y-I)f_{\xi_2|p_2^H}(y)dy}\right)\right] \\
&-\left[\cancel{qI} + \cancel{(p_2^L - q)\left(\frac{b^L + c^L}{3}\right)} - \cancel{(p_2^L - q)\left(\frac{b^L + c^L}{3}\right)} + (p_2^L - q)\left(\frac{2I^3}{3b^L c^L}\right)\right. \\
&\left. + (p_2^L - q)I\left(1 - \frac{3I^2}{3b^L c^L}\right)\right] = 0 \\
&(p_2^H - q)\frac{(b^H + c^H)}{3} + \frac{(p_2^L - q)}{3b^L c^L}I^3 - (p_2^L - q)I = 0 \\
&\underbrace{(p_2^L - q)}_{A \geq 0} I^3 \underbrace{- 3(p_2^L - q)b^L c^L}_{C \leq 0} I \underbrace{+ (p_2^H - q)(b^H + c^H)b^L c^L}_{D \geq 0} = 0
\end{aligned}
\tag{E.18}
$$

4. $b^H < I$ and $c^L \leq I \leq b^L$

   (Note that $a^i = 0 \quad \forall i = H, L$ and $a^H = a^L < c^H < b^H < c^L < b^L$)

$$
\begin{aligned}
&\left[qI + (p_2^H - q)\left(\mu^H - \int_{b^H}^{\infty}(y-I)f_{\xi_2|p_2^H}(y)dy\right)\right] \\
-&\left[qI + (p_2^L - q)\left(\mu^L - \int_{I}^{b^L}(y-I)f_{\xi_2|p_2^L}(y)dy\right)\right] = 0 \\
&\left[\cancel{qI} + (p_2^H - q)\left(\frac{b^H + c^H}{3} - \cancelto{0}{\int_{b^H}^{\infty}(y-I)f_{\xi_2|p_2^H}(y)dy}\right)\right] \\
-&\left[\cancel{qI} + (p_2^L - q)\left(\frac{b^L + c^L}{3} - \frac{(b^L - I)^2(2I + b^L)}{3(b^L - c^L)b^L} + \frac{3I(b^L - I)^2}{3(b^L - c^L)b^L}\right)\right] = 0 \\
&(p_2^H - q)\frac{(b^H + c^H)}{3} - \left[(p_2^L - q)\left(\frac{b^L + c^L}{3} + \frac{I^3 + 3(b^L)^2 I - 3b^L I^2 - (b^L)^3}{3(b^L - c^L)b^L}\right)\right] = 0 \\
&\underbrace{\left[\frac{-(p_2^L - q)}{3(b^L - c^L)b^L}\right]}_{A\leq 0} I^3 + \underbrace{\frac{p_2^L - q}{b^L - c^L}}_{B\geq 0} I^2 - \underbrace{\frac{(p_2^L - q)b^L}{b^L - c^L}}_{C\leq 0} I \\
&\underbrace{+(p_2^H - q)\frac{(b^H + c^H)}{3} + \frac{(p_2^L - q)(c^L)^2}{3(b^L - c^L)}}_{D\geq 0} = 0
\end{aligned}
\tag{E.19}
$$

E.5 Expected Profit Difference Function for Case 2

Using equations in (E.3) and (E.5), $\mathbb{E}[\Pi^2(I)|p_2^H] - \mathbb{E}[\Pi^2(I)|p_2^L]$ is

$$
\begin{cases}
I(p_2^H - p_2^L) & \text{if } I < 0 \\
\underbrace{\left[\dfrac{(p_2^L - q)}{3b^L c^L} - \dfrac{(p_2^H - q)}{3b^H c^H}\right]}_{A \leq 0} I^3 + \underbrace{\left[p_2^H - p_2^L\right]}_{B \geq 0} I & \text{if } 0 \leq I \leq c^H < c^L \\
\underbrace{\left[\dfrac{p_2^H - q}{3(b^H - c^H)b^H} + \dfrac{p_2^L - q}{3b^L c^L}\right]}_{A \geq 0} I^3 - \underbrace{\dfrac{p_2^H - q}{b^H - c^H}}_{B \leq 0} I^2 & \text{if } c^H < I < b^H \\
+ \underbrace{\left[\dfrac{(p_2^H - q)b^H}{b^H - c^H} - (p_2^L - q)\right]}_{C} I - \underbrace{\dfrac{(p_2^H - q)(c^H)^2}{3(b^H - c^H)}}_{D \leq 0} & \text{and } 0 \leq I < c^L \\
\underbrace{\left[\dfrac{p_2^H - q}{3(b^H - c^H)b^H} - \dfrac{p_2^L - q}{3(b^L - c^L)b^L}\right]}_{A} I^3 + \underbrace{\left[\dfrac{p_2^L - q}{b^L - c^L} - \dfrac{p_2^H - q}{b^H - c^H}\right]}_{B} I^2 & \text{if } c^H < I < b^H \\
+ \underbrace{\left[\dfrac{(p_2^H - q)b^H}{b^H - c^H} - \dfrac{(p_2^L - q)b^L}{b^L - c^L}\right]}_{C} I & \text{and } c^L \leq I \leq b^L \\
\underbrace{+(p_2^H - q)\dfrac{(b^H)^2 - 2(c^H)^2}{3(b^H - c^H)} - (p_2^L - q)\dfrac{(b^L)^2 - 2(c^L)^2}{3(b^L - c^L)}}_{D} & \\
\underbrace{\left[\dfrac{-(p_2^L - q)}{3(b^L - c^L)b^L}\right]}_{A \leq 0} I^3 + \underbrace{\dfrac{p_2^L - q}{b^L - c^L}}_{B \geq 0} I^2 - \underbrace{\dfrac{(p_2^L - q)b^L}{b^L - c^L}}_{C \leq 0} I & \text{if } b^H < I \\
\underbrace{+(p_2^H - q)\dfrac{(b^H + c^H)}{3} + \dfrac{(p_2^L - q)(c^L)^2}{3(b^L - c^L)}}_{D \geq 0} & \text{and } c^L \leq I \leq b^L \\
(p_2^H - q)\mu^H - (p_2^L - q)\mu^L & \text{if } b^H < b^L < I
\end{cases}
\tag{E.20}
$$

1. $a^i \leq I \leq c^i \quad \forall i = H, L$

   (Note that $a^i = 0 \quad \forall i = H, L$ and $a^H = a^L < c^H < c^L < b^H < b^L$)

   It is given in the previous section.

2. $c^H \leq I \leq b^H$ and $a^L \leq I \leq c^L$

   It is given in the previous section.

3. $c^H \leq I \leq b^H$ and $c^L \leq I \leq b^L$

(Note that $a^i = 0 \quad \forall i = H, L$ and $a^H = a^L < c^H < c^L < b^H < b^L$)

$$\begin{aligned}
&\left[qI + (p_2^H - q)\left(\mu^H - \int_I^{b^H} (y-I) f_{\xi_2|p_2^H}(y)dy\right)\right] \\
-&\left[qI + (p_2^L - q)\left(\mu^L - \int_I^{b^L} (y-I) f_{\xi_2|p_2^L}(y)dy\right)\right] = 0 \\
&\left[(p_2^H - q)\left(\frac{b^H + c^H}{3} + \frac{(b^H - I)^2(I - b^H)}{3(b^H - c^H)b^H}\right)\right] \\
-&\left[(p_2^L - q)\left(\frac{b^L + c^L}{3} + \frac{I^3 + 3(b^L)^2 I - 3b^L I^2 - (b^L)^3}{3(b^L - c^L)b^L}\right)\right] = 0 \\
&\left[(p_2^H - q)\left(\frac{b^H + c^H}{3} + \frac{I^3 + 3(b^H)^2 I - 3b^H I^2 - (b^H)^3}{3(b^H - c^H)b^H}\right)\right] \\
-&\left[(p_2^L - q)\left(\frac{b^L + c^L}{3} + \frac{I^3 + 3(b^L)^2 I - 3b^L I^2 - (b^L)^3}{3(b^L - c^L)b^L}\right)\right] = 0 \\
&\left[\frac{p_2^H - q}{3(b^H - c^H)b^H} - \frac{p_2^L - q}{3(b^L - c^L)b^L}\right] I^3 + \left[\frac{p_2^L - q}{b^L - c^L} - \frac{p_2^H - q}{b^H - c^H}\right] I^2 \\
+&\left[\frac{(p_2^H - q)b^H}{b^H - c^H} - \frac{(p_2^L - q)b^L}{b^L - c^L}\right] I + (p_2^H - q)\frac{(b^H + c^H)}{3} \\
&- \frac{(p_2^H - q)(c^H)^2}{3(b^H - c^H)} - (p_2^L - q)\frac{b^L + c^L}{3} + \frac{(p_2^L - q)(c^L)^2}{3(b^L - c^L)} = 0 \\
&\underbrace{\left[\frac{p_2^H - q}{3(b^H - c^H)b^H} - \frac{p_2^L - q}{3(b^L - c^L)b^L}\right]}_{A} I^3 + \underbrace{\left[\frac{p_2^L - q}{b^L - c^L} - \frac{p_2^H - q}{b^H - c^H}\right]}_{B} I^2 \\
&\underbrace{+\left[\frac{(p_2^H - q)b^H}{b^H - c^H} - \frac{(p_2^L - q)b^L}{b^L - c^L}\right]}_{C} I \\
&\underbrace{+(p_2^H - q)\frac{(b^H)^2 - 2(c^H)^2}{3(b^H - c^H)} - (p_2^L - q)\frac{(b^L)^2 - 2(c^L)^2}{3(b^L - c^L)}}_{D} = 0
\end{aligned} \tag{E.21}$$

4. $b^H < I$ and $c^L \leq I \leq b^L$

   (Note that $a^i = 0 \quad \forall i = H, L$ and $a^H = a^L < c^H < c^L < b^H < b^L$)

   It is given in the previous section.

Appendix F

The Newsvendor Problem with Recourse Action

The expected total profit difference for continuous distributions is as follows:

$$
\begin{aligned}
\mathbb{E}[\Pi(Q,\xi_1)] = &\int_0^{Q-I^t} \Bigg(- Qv + xp_1 \\
&+ \Bigg[q(Q-x) + (p_2^L - q)\Bigg(\mu^L - \int_{Q-x}^{\infty} (y-(Q-x)) f_{\xi_2|p_2^L}(y)dy\Bigg)\Bigg]\Bigg) f_{\xi_1}(x)dx \\
&+ \int_{Q-I^t}^{Q} \Bigg(- Qv + xp_1 + \Bigg[q(Q-x) \\
&+ (p_2^H - q)\Bigg(\mu^H - \int_{Q-x}^{\infty} (y-(Q-x)) f_{\xi_2|p_2^H}(y)dy\Bigg)\Bigg) f_{\xi_1}(x)dx \\
&+ \int_{Q}^{\infty} \Bigg(- Qv + Qp_1 - B(x-Q)\Bigg) f_{\xi_1}(x)dx
\end{aligned}
$$

$$
\frac{\partial \mathbb{E}[\Pi^1(Q_0,\xi_1)]}{\partial Q} = 0
$$

By using **Leibniz's Rule:**

1. $\frac{\partial}{\partial Q}\int_0^{Q-I_1^t}\Big(-Qv + xp_1 + q(Q-x) + (p_2^L - q)\mu^L\Big)f_{\xi_1}(x)dx$

$$= 1[-Qv+(Q-I_1^t)p_1+q(Q-Q+I_1^t)+(p_2^L-q)\mu^L]f_{\xi_1}(Q-I_1^t)-0+\int_0^{Q-I_1^t} -v+qf_{\xi_1}(x)dx$$

2. $\frac{\partial}{\partial Q}\int_{Q-I_1^t}^{Q}\Big(-Qv + xp_1 + q(Q-x) + (p_2^H - q)\mu^H\Big)f_{\xi_1}(x)dx$

$$\begin{aligned}
&= 1[-Qv + Qp_1 + \cancelto{0}{q(Q-Q)} + (p_2^H - q)\mu^H]f_{\xi_1}(Q) \\
&- 1[-Qv + (Q-I_1^t)p_1 + q(Q-Q+I_1^t) + (p_2^H - q)\mu^H]f_{\xi_1}(Q-I_1^t) \\
&+ \int_{Q-I_1^t}^{Q} -v + qf_{\xi_1}(x)dx
\end{aligned}$$

3. $\frac{\partial}{\partial Q}\int_0^{Q-I_1^t}\int_{Q-x}^{\infty}(y-(Q-x))f_{\xi_2|p_2^L}(y)dyf_{\xi_1}(x)dx$

$$\begin{aligned}
&= 1\int_{Q-Q+I_1^t}^{\infty}(y-Q+Q-I_1^t)f_{\xi_2|p_2^L}(y)dyf_{\xi_1}(Q-I_1^t) - 0 \\
&+ \int_0^{Q-I_1^t}\frac{\partial}{\partial Q}\int_{Q-x}^{\infty}(y-Q+x)f_{\xi_2|p_2^L}(y)dyf_{\xi_1}(x)dx \\
&= \int_{I_1^t}^{\infty}(y-I_1^t)f_{\xi_2|p_2^L}(y)dyf_{\xi_1}(Q-I_1^t) - 0 \\
&+ \int_0^{Q-I_1^t}\Big[0 - 1\cancelto{0}{[Q-x-Q+x]}f_{\xi_2|p_2^L}(Q-x)f_{\xi_1}(x) \\
&+ \int_{Q-x}^{\infty}\frac{\partial}{\partial Q}(y-Q+x)f_{\xi_2|p_2^L}(y)dyf_{\xi_1}(x)\Big]dx \\
&= \int_{I_1^t}^{\infty}(y-I_1^t)f_{\xi_2|p_2^L}(y)dyf_{\xi_1}(Q-I_1^t) + \int_0^{Q-I_1^t}\int_{Q-x}^{\infty} -f_{\xi_2|p_2^L}(y)dyf_{\xi_1}(x)dx
\end{aligned}$$

4. $\frac{\partial}{\partial Q}\int_{Q-I_1^t}^{Q}\int_{Q-x}^{\infty}(y-(Q-x))f_{\xi_2|p_2^H}(y)dyf_{\xi_1}(x)dx$

$$= 1\underbrace{\int_0^{\infty} yf_{\xi_2|p_2^H}(y)dy}_{\mu^H} f_{\xi_1}(Q) - 1\int_{I_1^t}^{\infty}(y-I_1^t)f_{\xi_2|p_2^H}(y)dyf_{\xi_1}(Q-I_1^t)$$

$$+ \int_{Q-I_1^t}^{Q}\frac{\partial}{\partial Q}\int_{Q-x}^{\infty}(y-Q+x)f_{\xi_2|p_2^H}(y)dyf_{\xi_1}(x)dx$$

$$= \mu^H f_{\xi_1}(Q) - \int_{I_1^t}^{\infty}(y-I_1^t)f_{\xi_2|p_2^H}(y)dyf_{\xi_1}(Q-I_1^t)$$

$$+ \int_{Q-I_1^t}^{Q}\Big[0 - 1[\underbrace{Q-x-Q+x}_{0}]f_{\xi_2|p_2^H}(Q-x)f_{\xi_1}(x)$$

$$+ \int_{Q-x}^{\infty}\frac{\partial}{\partial Q}(y-Q+x)f_{\xi_2|p_2^H}(y)dyf_{\xi_1}(x)\Big]dx$$

$$= \mu^H f_{\xi_1}(Q) - \int_{I_1^t}^{\infty}(y-I_1^t)f_{\xi_2|p_2^H}(y)dyf_{\xi_1}(Q-I_1^t)$$

$$+ \int_{Q-I_1^t}^{Q}\int_{Q-x}^{\infty} -f_{\xi_2|p_2^H}(y)dyf_{\xi_1}(x)dx$$

$$\frac{\partial\mathbb{E}[\Pi^1(Q_0,\xi_1)]}{\partial Q} = [\cancel{-Qv} + \cancel{(Q-I_1^t)p_1} + \cancel{qI_1^t} + (p_2^L-q)\mu^L]f_{\xi_1}(Q-I_1^t)$$

$$+ \int_0^{Q-I_1^t}(-v+q)f_{\xi_1}(x)dx - \int_{I_1^t}^{\infty}(p_2^L-q)(y-I_1^t)f_{\xi_2|p_2^L}(y)dyf_{\xi_1}(Q-I_1^t)$$

$$- \int_0^{Q-I_1^t}\int_{Q-x}^{\infty} -(p_2^L-q)f_{\xi_2|p_2^L}(y)dyf_{\xi_1}(x)dx + [\cancel{-Qv} + \cancel{Qp_1} + (p_2^H-q)\mu^H]f_{\xi_1}(Q)$$

$$- 1[\cancel{-Qv} + \cancel{(Q-I_1^t)p_1} + \cancel{qI_1^t} + (p_2^H-q)\mu^H]f_{\xi_1}(Q-I_1^t) + \int_{Q-I_1^t}^{Q}(-v+q)f_{\xi_1}(x)dx$$

$$- (p_2^H-q)\mu^H f_{\xi_1}(Q) - \int_{I_1^t}^{\infty}(p_2^H-q)(y-I_1^t)f_{\xi_2|p_2^H}(y)dyf_{\xi_1}(Q-I_1^t)$$

$$- \int_{Q-I_1^t}^{Q}\int_{Q-x}^{\infty} -(p_2^H-q)f_{\xi_2|p_2^H}(y)dyf_{\xi_1}(x)dx$$

$$+ 0 - 1[\cancel{-Qv} + \cancel{Qp_1} - \underbrace{B(Q-Q)}_{0}]f_{\xi_1}(Q)\int_Q^{\infty}(-v+p_1+B)f_{\xi_1}dx$$

$$
\begin{aligned}
\frac{\partial \mathbb{E}[\Pi^1(Q_0,\xi_1)]}{\partial Q} &= (p_2^L-q)\mu^L f_{\xi_1}(Q-I_1^t) - (p_2^H-q)\mu^H f_{\xi_1}(Q-I_1^t) \\
&+ \underbrace{\int_0^{Q-I_1^t} -v f_{\xi_1}(x)dx + \int_{Q-I_1^t}^{Q} -v f_{\xi_1}(x)dx + \int_Q^{\infty} -v f_{\xi_1}dx}_{-v} \\
&+ \int_0^{Q-I_1^t} q f_{\xi_1}(x)dx + \int_{Q-I_1^t}^{Q} q f_{\xi_1}(x)dx \\
&- \Bigg[\int_{I_1^t}^{\infty}(p_2^L-q)(y-I_1^t)f_{\xi_2|p_2^L}(y)dy - \int_{I_1^t}^{\infty}(p_2^H-q)(y-I_1^t)f_{\xi_2|p_2^H}(y)dy\Bigg] f_{\xi_1}(Q-I_1^t) \\
&+ \cancel{(p_2^H-q)\mu^H f_{\xi_1}(Q)} - \cancel{(p_2^H-q)\mu^H f_{\xi_1}(Q)} + \int_0^{Q-I_1^t}\int_{Q-x}^{\infty}(p_2^L-q)f_{\xi_2|p_2^L}(y)dy f_{\xi_1}(x)dx \\
&+ \int_{Q-I_1^t}^{Q}\int_{Q-x}^{\infty}(p_2^H-q)f_{\xi_2|p_2^H}(y)dy f_{\xi_1}(x)dx + (p_1+B)\Bigg[1-\int_0^Q f_{\xi_1}dx\Bigg]
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \mathbb{E}[\Pi^1(Q_0,\xi_1)]}{\partial Q} &= (p_2^L-q)\mu^L f_{\xi_1}(Q-I_1^t) - (p_2^H-q)\mu^H f_{\xi_1}(Q-I_1^t) - v \\
&+ (p_1+B)\Big[1-F_{\xi_1}(Q)\Big] + qF_{\xi_1}(Q) - \Bigg[-I_1^t(p_2^L-q)\big[1-F_{\xi_2|p_2^L}(I_1^t)\big] \\
&+ (p_2^L-q)\big[\mu^L - \int_0^{I_1^t} y f_{\xi_2|p_2^L}(y)dy\big] + I_1^t(p_2^H-q)\big[1-F_{\xi_2|p_2^H}(I_1^t)\big] \\
&- (p_2^H-q)\big[\mu^H - \int_0^{I_1^t} y f_{\xi_2|p_2^H}(y)dy\big]\Bigg] f_{\xi_1}(Q-I_1^t) \\
&+ (p_2^L-q)\Bigg(\int_0^{Q-I_1^t}[1-F_{\xi_2|p_2^L}(Q-x)]f_{\xi_1}(x)dx\Bigg) \\
&+ (p_2^H-q)\Bigg(\int_{Q-I_1^t}^{Q}[1-F_{\xi_2|p_2^H}(Q-x)]f_{\xi_1}(x)dx\Bigg)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \mathbb{E}[\Pi^1(Q_0,\xi_1)]}{\partial Q} &= p_2^L F_{\xi_1}(Q-I_1^t) - \cancel{qF_{\xi_1}(Q-I_1^t)} + p_2^H\Big[F_{\xi_1}(Q) - F_{\xi_1}(Q-I_1^t)\Big] \\
&- q\cancel{\Big[F_{\xi_1}(Q) - F_{\xi_1}(Q-I_1^t)\Big]} + \cancel{qF_{\xi_1}(Q)} - (p_2^L - q)\int_0^{Q-I_1^t} F_{\xi_2|p_2^L}(Q-x) f_{\xi_1}(x)dx \\
&- (p_2^H - q)\int_{Q-I_1^t}^{Q} F_{\xi_2|p_2^H}(Q-x) f_{\xi_1}(x)dx + \cancel{(p_2^L - q)\mu^L f_{\xi_1}(Q-I_1^t)} \\
&- \cancel{(p_2^L - q)\mu^L f_{\xi_1}(Q-I_1^t)} - \cancel{(p_2^H - q)\mu^H f_{\xi_1}(Q-I_1^t)} + \cancel{(p_2^H - q)\mu^H f_{\xi_1}(Q-I_1^t)} \\
&- v + (p_1 + B)\Big[1 - F_{\xi_1}(Q)\Big] + (p_2^L - q) f_{\xi_1}(Q-I_1^t)\int_0^{I_1^t} y f_{\xi_2|p_2^L}(y)dy \\
&- (p_2^H - q) f_{\xi_1}(Q-I_1^t)\int_0^{I_1^t} y f_{\xi_2|p_2^H}(y)dy \\
&+ I_1^t f_{\xi_1}(Q-I_1^t)\Big[(p_2^H - q)F_{\xi_2|p_2^H}(I_1^t) - (p_2^L - q)F_{\xi_2|p_2^L}(I_1^t)\Big] \\
&+ I_1^t f_{\xi_1}(Q-I_1^t)\Big[p_2^L - \cancel{q} - (p_2^H - \cancel{q})\Big] = 0
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \mathbb{E}[\Pi^1(Q_0,\xi_1)]}{\partial Q} &= p_2^L F_{\xi_1}(Q-I_1^t) + p_2^H\Big[F_{\xi_1}(Q) - F_{\xi_1}(Q-I_1^t)\Big] \\
&- (p_2^L - q)\int_0^{Q-I_1^t} F_{\xi_2|p_2^L}(Q-x) f_{\xi_1}(x)dx - (p_2^H - q)\int_{Q-I_1^t}^{Q} F_{\xi_2|p_2^H}(Q-x) f_{\xi_1}(x)dx \\
&- v + (p_1 + B)\Big[1 - F_{\xi_1}(Q)\Big] + (p_2^L - q) f_{\xi_1}(Q-I_1^t)\int_0^{I_1^t} y f_{\xi_2|p_2^L}(y)dy \\
&- (p_2^H - q) f_{\xi_1}(Q-I_1^t)\int_0^{I_1^t} y f_{\xi_2|p_2^H}(y)dy \\
&+ I_1^t f_{\xi_1}(Q-I_1^t)\Big[(p_2^H - q)F_{\xi_2|p_2^H}(I_1^t) - (p_2^L - q)F_{\xi_2|p_2^L}(I_1^t)\Big] \\
&+ I_1^t f_{\xi_1}(Q-I_1^t)\Big[p_2^L - \cancel{q} - (p_2^H - \cancel{q})\Big] = 0
\end{aligned}
$$

$$\frac{\partial \mathbb{E}[\Pi^1(Q_0,\xi_1)]}{\partial Q} = -v + p_1 + B + F_{\xi_1}(Q)\Big[p_2^H - (p_1+B)\Big] + F_{\xi_1}(Q-I_1^t)\Big[p_2^L - p_2^H\Big]$$
$$-(p_2^L - q)\int_0^{Q-I_1^t} F_{\xi_2|p_2^L}(Q-x) f_{\xi_1}(x)dx - (p_2^H - q)\int_{Q-I_1^t}^{Q} F_{\xi_2|p_2^H}(Q-x) f_{\xi_1}(x)dx$$
$$+(p_2^L - q) f_{\xi_1}(Q-I_1^t)\int_0^{I_1^t} y f_{\xi_2|p_2^L}(y)dy - (p_2^H - q) f_{\xi_1}(Q-I_1^t)\int_0^{I_1^t} y f_{\xi_2|p_2^H}(y)dy$$
$$+ I_1^t f_{\xi_1}(Q-I_1^t)\Big[p_2^L - p_2^H + (p_2^H - q)F_{\xi_2|p_2^H}(I_1^t) - (p_2^L - q)F_{\xi_2|p_2^L}(I_1^t)\Big] = 0$$

$$\frac{\partial \mathbb{E}[\Pi^1(Q_0,\xi_1)]}{\partial Q} = -v + p_1 + B + F_{\xi_1}(Q)\Big[p_2^H - (p_1+B)\Big] + F_{\xi_1}(Q-I_1^t)\Big[p_2^L - p_2^H\Big]$$
$$+ f_{\xi_1}(Q-I_1^t)\Big[(p_2^L - q)\Big(\int_0^{I_1^t} y f_{\xi_2|p_2^L}(y)dy - I_1^t F_{\xi_2|p_2^L}(I_1^t)\Big)$$
$$-(p_2^H - q)\Big(\int_0^{I_1^t} y f_{\xi_2|p_2^H}(y)dy - I_1^t F_{\xi_2|p_2^H}(I_1^t)\Big) + I_1^t(p_2^L - p_2^H)\Big]$$
$$-(p_2^L - q)\int_0^{Q-I_1^t} F_{\xi_2|p_2^L}(Q-x) f_{\xi_1}(x)dx - (p_2^H - q)\int_{Q-I_1^t}^{Q} F_{\xi_2|p_2^H}(Q-x) f_{\xi_1}(x)dx = 0$$

F.1 2^{nd} Derivative of The Expected Total Profit

$$\frac{\partial^2 \mathbb{E}[\Pi^1(Q_0,\xi_1)]}{\partial Q^2} = f_{\xi_1}(Q)\underbrace{\Big[p_2^H - (p_1+B)\Big]}_{\le 0} + f_{\xi_1}(Q-I_1^t)\underbrace{\Big[p_2^L - p_2^H\Big]}_{\le 0}$$
$$+\frac{\partial}{\partial Q} f_{\xi_1}(Q-I_1^t)\Big[(p_2^L - q)\Big(\int_0^{I_1^t} y f_{\xi_2|p_2^L}(y)dy - I_1^t F_{\xi_2|p_2^L}(I_1^t)\Big)$$
$$-(p_2^H - q)\Big(\int_0^{I_1^t} y f_{\xi_2|p_2^H}(y)dy - I_1^t F_{\xi_2|p_2^H}(I_1^t)\Big) + I_1^t(p_2^L - p_2^H)\Big]$$
$$-(p_2^L - q)\frac{\partial}{\partial Q}\int_0^{Q-I_1^t} F_{\xi_2|p_2^L}(Q-x) f_{\xi_1}(x)dx - (p_2^H - q)\frac{\partial}{\partial Q}\int_{Q-I_1^t}^{Q} F_{\xi_2|p_2^H}(Q-x) f_{\xi_1}(x)dx$$

$$
\begin{aligned}
\frac{\partial^2 \mathbb{E}[\Pi^1(Q_0,\xi_1)]}{\partial Q^2} &= f_{\xi_1}(Q)\underbrace{\Big[p_2^H-(p_1+B)\Big]}_{\leq 0}+f_{\xi_1}(Q-I_1^t)\underbrace{\Big[p_2^L-p_2^H\Big]}_{\leq 0}\\
&+\frac{\partial}{\partial Q}f_{\xi_1}(Q-I_1^t)\Bigg[(p_2^L-q)\Bigg(\int_0^{I_1^t} y f_{\xi_2|p_2^L}(y)dy - I_1^t F_{\xi_2|p_2^L}(I_1^t)\Bigg)\\
&-(p_2^H-q)\Bigg(\int_0^{I_1^t} y f_{\xi_2|p_2^H}(y)dy - I_1^t F_{\xi_2|p_2^H}(I_1^t)\Bigg)+I_1^t(p_2^L-p_2^H)\Bigg]\\
&-(p_2^L-q)\Bigg[F_{\xi_2|p_2^L}(I_1^t)f_{\xi_1}(Q-I_1^t)+\int_0^{Q-I_1^t}\frac{\partial}{\partial Q}F_{\xi_2|p_2^L}(Q-x)f_{\xi_1}(x)dx\Bigg]\\
&-(p_2^H-q)\Bigg[0-F_{\xi_2|p_2^H}(I_1^t)f_{\xi_1}(Q-I_1^t)+\int_{Q-I_1^t}^{Q}\frac{\partial}{\partial Q}F_{\xi_2|p_2^H}(Q-x)f_{\xi_1}(x)dx\Bigg]
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 \mathbb{E}[\Pi^1(Q_0,\xi_1)]}{\partial Q^2} &= f_{\xi_1}(Q)\underbrace{\Big[p_2^H-(p_1+B)\Big]}_{\leq 0}+f_{\xi_1}(Q-I_1^t)\underbrace{\Big[p_2^L-p_2^H\Big]}_{\leq 0}\\
&+\frac{\partial}{\partial Q}f_{\xi_1}(Q-I_1^t)\Bigg[(p_2^L-q)\Bigg(\int_0^{I_1^t} y f_{\xi_2|p_2^L}(y)dy - I_1^t F_{\xi_2|p_2^L}(I_1^t)\Bigg)\\
&-(p_2^H-q)\Bigg(\int_0^{I_1^t} y f_{\xi_2|p_2^H}(y)dy - I_1^t F_{\xi_2|p_2^H}(I_1^t)\Bigg)+I_1^t(p_2^L-p_2^H)\Bigg]\\
&+f_{\xi_1}(Q-I_1^t)\Bigg[(p_2^H-q)F_{\xi_2|p_2^H}(I_1^t)-(p_2^L-q)F_{\xi_2|p_2^L}(I_1^t)\Bigg]\\
&-(p_2^L-q)\int_0^{Q-I_1^t}\frac{\partial}{\partial Q}F_{\xi_2|p_2^L}(Q-x)f_{\xi_1}(x)dx-(p_2^H-q)\int_{Q-I_1^t}^{Q}\frac{\partial}{\partial Q}F_{\xi_2|p_2^H}(Q-x)f_{\xi_1}(x)dx
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 \mathbb{E}[\Pi^1(Q_0,\xi_1)]}{\partial Q^2} &= f_{\xi_1}(Q)\underbrace{\Big[p_2^H-(p_1+B)\Big]}_{\leq 0}+f_{\xi_1}(Q-I_1^t)\underbrace{\Big[p_2^L-p_2^H\Big]}_{\leq 0}\\
&+\frac{\partial}{\partial Q}f_{\xi_1}(Q-I_1^t)\Bigg[(p_2^L-q)\int_0^{I_1^t} y f_{\xi_2|p_2^L}(y)dy-(p_2^H-q)\int_0^{I_1^t} y f_{\xi_2|p_2^H}(y)dy+I_1^t(p_2^L-p_2^H)\Bigg]\\
&+\Bigg[\frac{\partial}{\partial Q}f_{\xi_1}(Q-I_1^t)I_1^t+f_{\xi_1}(Q-I_1^t)\Bigg]\Bigg[(p_2^H-q)F_{\xi_2|p_2^H}(I_1^t)-(p_2^L-q)F_{\xi_2|p_2^L}(I_1^t)\Bigg]\\
&-(p_2^L-q)\int_0^{Q-I_1^t}\frac{\partial}{\partial Q}F_{\xi_2|p_2^L}(Q-x)f_{\xi_1}(x)dx-(p_2^H-q)\int_{Q-I_1^t}^{Q}\frac{\partial}{\partial Q}F_{\xi_2|p_2^H}(Q-x)f_{\xi_1}(x)dx
\end{aligned}
$$

$$\frac{\partial^2 \mathbb{E}[\Pi^1(Q_0,\xi_1)]}{\partial Q^2} = \underbrace{\left[p_2^H - (p_1+B)\right]}_{\leq 0} f_{\xi_1}(Q) + \underbrace{\left[p_2^L - p_2^H\right]}_{\leq 0} f_{\xi_1}(Q - I_1^t)$$

$$\underbrace{+\left[(p_2^L - q)\int_0^{I_1^t} y f_{\xi_2|p_2^L}(y)dy - (p_2^H - q)\int_0^{I_1^t} y f_{\xi_2|p_2^H}(y)dy + I_1^t(p_2^L - p_2^H)\right]}_{\leq 0} \frac{\partial}{\partial Q} f_{\xi_1}(Q - I_1^t)$$

$$\underbrace{+\left[(p_2^H - q)F_{\xi_2|p_2^H}(I_1^t) - (p_2^L - q)F_{\xi_2|p_2^L}(I_1^t)\right]}_{\geq 0} \left[\frac{\partial}{\partial Q} f_{\xi_1}(Q - I_1^t) I_1^t + f_{\xi_1}(Q - I_1^t)\right]$$

$$\underbrace{-(p_2^L - q)\int_0^{Q-I_1^t} f_{\xi_2|p_2^L}(Q-x) f_{\xi_1}(x)dx - (p_2^H - q)\int_{Q-I_1^t}^{Q} f_{\xi_2|p_2^H}(Q-x) f_{\xi_1}(x)dx}_{\leq 0}$$

1. $p_2^H - (p_1+B) \leq 0$ since $p_2^H \leq p_1 + B$
2. $p_2^L - p_2^H \leq 0$ since $p_2^L \leq p_2^H$
3. $$(p_2^L - q)\int_0^{I_1^t} y f_{\xi_2|p_2^L}(y)dy - (p_2^H - q)\int_0^{I_1^t} y f_{\xi_2|p_2^H}(y)dy + I_1^t(p_2^L - p_2^H) \leq 0$$
   since
   $$\int_0^{I_1^t} y f_{\xi_2|p_2^L}(y)dy \leq \int_0^{I_1^t} y f_{\xi_2|p_2^H}(y)dy$$
   and $p_2^L \leq p_2^H$
4. $(p_2^H - q)F_{\xi_2|p_2^H}(I_1^t) - (p_2^L - q)F_{\xi_2|p_2^L}(I_1^t) \geq 0$ since $F_{\xi_2|p_2^H}(I_1^t) \geq F_{\xi_2|p_2^L}(I_1^t)$
5. $$(p_2^L - q)\int_0^{Q-I_1^t} f_{\xi_2|p_2^L}(Q-x) f_{\xi_1}(x)dx \geq 0$$
6. $$(p_2^H - q)\int_{Q-I_1^t}^{Q} f_{\xi_2|p_2^H}(Q-x) f_{\xi_1}(x)dx \geq 0$$

The second derivative $\frac{\partial^2 \mathbb{E}[\Pi^1(Q_0,\xi_1)]}{\partial Q^2}$ is negative under the following condition:

$$
\begin{aligned}
&f_{\xi_1}(Q-I_1^t)\Big[(p_2^H-q)F_{\xi_2|p_2^H}(I_1^t)-(p_2^L-q)F_{\xi_2|p_2^L}(I_1^t)\Big] \leq f_{\xi_1}(Q)\Big[(p_1+B)-p_2^H\Big] \\
&+ f_{\xi_1}(Q-I_1^t)\Big[p_2^H-p_2^L\Big] + (p_2^L-q)\int_0^{Q-I_1^t} f_{\xi_2|p_2^L}(Q-x)f_{\xi_1}(x)dx \\
&+ (p_2^H-q)\int_{Q-I_1^t}^{Q} f_{\xi_2|p_2^H}(Q-x)f_{\xi_1}(x)dx
\end{aligned}
\tag{F.1}
$$

F.2 Poisson Demand

The probability density function of Poisson distribution is

$$
f(k) = \begin{cases} \frac{e^{-\lambda}\lambda^k}{k!} & \text{if } k \geq 0 \\ 0 & \text{otherwise} \end{cases}
$$

When the demand is Poisson distributed, the total expected profit function is

$$
\begin{aligned}
\mathbb{E}[\Pi^1(Q,\xi_1)] = &\sum_{k=0}^{Q-I^t}\Big(-Qv+kp_1+\mathbb{E}[\Pi^2(Q-k,\xi_2)|p_2^L)]\Big)f_{\xi_1}(k) \\
&+ \sum_{k=Q_0-I^t+1}^{Q}\Big(-Qv+kp_1+\mathbb{E}[\Pi^2(Q-k,\xi_2)|p_2^H)]\Big)f_{\xi_1}(k) \\
&+ \sum_{k=Q+1}^{\infty}\Big(-Qv+p_1Q-B(k-Q)\Big)f_{\xi_1}(k)
\end{aligned}
\tag{F.2}
$$

$$
\begin{aligned}
\mathbb{E}[\Pi^1(Q,\xi_1)] = \sum_{k=0}^{Q-I^t} \Bigg(&- Qv + kp_1 + \Bigg[q(Q-k) \\
&+(p_2^L - q)\Bigg(\mu^L - \sum_{l=Q-k+1}^{\infty}(l-(Q-k))f_{\xi_2|p_2^L}(l)\Bigg)\Bigg]\Bigg) f_{\xi_1}(k) \\
&+ \sum_{k=Q-I^t+1}^{Q} \Bigg(- Qv + kp_1 + \Bigg[q(Q-k) \\
&+(p_2^H - q)\Bigg(\mu^H - \sum_{l=Q-k+1}^{\infty}(l-(Q-k))f_{\xi_2|p_2^H}(l)\Bigg)\Bigg]\Bigg) f_{\xi_1}(k) \\
&+Q(p_1 - v + B)\big(1 - \sum_{k=0}^{Q} f_{\xi_1}(k)\big) - B\big(\mu - \sum_{k=0}^{Q} kf_{\xi_1}(k)\big) \qquad \text{(F.3)}
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}[\Pi^1(Q,\xi_1)] = -Qv + \sum_{k=0}^{Q} kp_1 f_{\xi_1}(k) \\
+ \sum_{k=0}^{Q-I^t} \Bigg[q(Q-k) + (p_2^L - q)\Bigg(\mu^L - \sum_{l=Q-k+1}^{\infty}(l-(Q-k))f_{\xi_2|p_2^L}(l)\Bigg)\Bigg] f_{\xi_1}(k) \\
+ \sum_{k=Q-I^t+1}^{Q} \Bigg[q(Q-k) + (p_2^H - q)\Bigg(\mu^H - \sum_{l=Q-k+1}^{\infty}(l-(Q-k))f_{\xi_2|p_2^H}(l)\Bigg)\Bigg] f_{\xi_1}(k) \\
+Q(p_1 + B)\big(1 - \sum_{k=0}^{Q} f_{\xi_1}(k)\big) - B\big(\mu - \sum_{k=0}^{Q} kf_{\xi_1}(k)\big)
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}[\Pi^1(Q,\xi_1)] =& Q(p_1 - v) - B(\mu - Q) + \sum_{k=0}^{Q} kp_1 f_{\xi_1}(k) + q\sum_{k=0}^{Q}(Q-k)f_{\xi_1}(k) \\
&+ \sum_{k=0}^{Q-I^t} \Bigg[(p_2^L - q)\Bigg(\mu^L - \sum_{l=Q-k+1}^{\infty}(l-(Q-k))f_{\xi_2|p_2^L}(l)\Bigg)\Bigg] f_{\xi_1}(k) \\
&+ \sum_{k=Q-I^t+1}^{Q} \Bigg[(p_2^H - q)\Bigg(\mu^H - \sum_{l=Q-k+1}^{\infty}(l-(Q-k))f_{\xi_2|p_2^H}(l)\Bigg)\Bigg] f_{\xi_1}(k) \\
&-Qp_1\sum_{k=0}^{Q} f_{\xi_1}(k) - B\sum_{k=0}^{Q}(Q-k)f_{\xi_1}(k)
\end{aligned}
$$

$$\begin{aligned}\mathbb{E}[\Pi^1(Q,\xi_1)] =& Q(p_1-v) - B(\mu - Q) + \sum_{k=0}^{Q} kp_1 f_{\xi_1}(k)\\
&+ \sum_{k=0}^{Q-I^t}\left[(p_2^L - q)\left(\mu^L - \sum_{l=Q-k+1}^{\infty}(l-(Q-k))f_{\xi_2|p_2^L}(l)\right)\right]f_{\xi_1}(k)\\
&+ \sum_{k=Q-I^t+1}^{Q}\left[(p_2^H - q)\left(\mu^H - \sum_{l=Q-k+1}^{\infty}(l-(Q-k))f_{\xi_2|p_2^H}(l)\right)\right]f_{\xi_1}(k)\\
&- Qp_1\sum_{k=0}^{Q} f_{\xi_1}(k) - (B-q)\sum_{k=0}^{Q}(Q-k)f_{\xi_1}(k)\end{aligned}$$

$$\begin{aligned}&\mathbb{E}[\Pi^1(Q,\xi_1)] =\\
&Q(p_1-v) - B(\mu - Q) + \sum_{k=0}^{Q-I^t}\left[(p_2^L - q)\left(\mu^L - \sum_{l=Q-k+1}^{\infty}(l-(Q-k))f_{\xi_2|p_2^L}(l)\right)\right]f_{\xi_1}(k)\\
&+ \sum_{k=Q-I^t+1}^{Q}\left[(p_2^H - q)\left(\mu^H - \sum_{l=Q-k+1}^{\infty}(l-(Q-k))f_{\xi_2|p_2^H}(l)\right)\right]f_{\xi_1}(k)\\
&- (p_1 + B - q)\sum_{k=0}^{Q}(Q-k)f_{\xi_1}(k)\end{aligned} \tag{F.4}$$

$$\begin{aligned}&\mathbb{E}[\Pi^1(Q,\xi_1)] =\\
&Q(p_1-v) - B(\mu - Q) + \sum_{k=0}^{Q}\left[(p_2^H - q)\left(\mu^H - \sum_{l=Q-k+1}^{\infty}(l-(Q-k))f_{\xi_2|p_2^H}(l)\right)\right]f_{\xi_1}(k)\\
&+ \sum_{k=0}^{Q-I^t}\left[(p_2^L - q)\left(\mu^L - \sum_{l=Q-k+1}^{\infty}(l-(Q-k))f_{\xi_2|p_2^L}(l)\right)\right]f_{\xi_1}(k)\\
&- \sum_{k=0}^{Q-I^t}\left[(p_2^H - q)\left(\mu^H - \sum_{l=Q-k+1}^{\infty}(l-(Q-k))f_{\xi_2|p_2^H}(l)\right)\right]f_{\xi_1}(k)\\
&- (p_1 + B - q)\sum_{k=0}^{Q}(Q-k)f_{\xi_1}(k)\end{aligned} \tag{F.5}$$

$$
\begin{aligned}
\mathbb{E}[\Pi^1(Q,\xi_1)] =& Q(p_1-v) - B(\mu-Q) + \Big[p_2^L\mu^L - p_2^H\mu^H - q(\mu^L-\mu^H)\Big]\sum_{k=0}^{Q-I^t} f_{\xi_1}(k) \\
&+(p_2^H-q)\mu^H\sum_{k=0}^{Q} f_{\xi_1}(k) - (p_2^H-q)\sum_{k=0}^{Q}\sum_{l=Q-k+1}^{\infty}(l-(Q-k))f_{\xi_2|p_2^H}(l)f_{\xi_1}(k) \\
&-(p_2^L-q)\sum_{k=0}^{Q-I^t}\sum_{l=Q-k+1}^{\infty}(l-(Q-k))f_{\xi_2|p_2^L}(l)f_{\xi_1}(k) \\
&+(p_2^H-q)\sum_{k=0}^{Q-I^t}\sum_{l=Q-k+1}^{\infty}(l-(Q-k))f_{\xi_2|p_2^H}(l)f_{\xi_1}(k) \\
&-(p_1+B-q)\sum_{k=0}^{Q}(Q-k)f_{\xi_1}(k) \qquad \text{(F.6)}
\end{aligned}
$$

Alternatively, the total expected profit function can be expressed as follows.

$$
\begin{aligned}
&\mathbb{E}[\Pi^1(Q,\xi_1)] = \\
&\sum_{k=0}^{Q-I^t}\Bigg(-Qv+kp_1+\Big[p_2^L(Q-k)+\sum_{l=0}^{Q-k}((p_2^L-q)(l-(Q-k)))f_{\xi_2|p_2^L}(l)\Big]\Bigg)f_{\xi_1}(k) \\
&+\sum_{k=Q-I^t+1}^{Q}\Bigg(-Qv+kp_1+\Big[p_2^H(Q-k)+\sum_{l=0}^{Q-k}((p_2^H-q)(l-(Q-k)))f_{\xi_2|p_2^H}(l)\Big]\Bigg)f_{\xi_1}(k) \\
&+Q(p_1-v+B)\big(1-\sum_{k=0}^{Q} f_{\xi_1}(k)\big) - B\big(\mu-\sum_{k=0}^{Q} kf_{\xi_1}(k)\big) \qquad \text{(F.7)}
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}[\Pi^1(Q,\xi_1)] &= -Qv + \sum_{k=0}^{Q} kp_1 f_{\xi_1}(k) \\
&+ \sum_{k=0}^{Q-I^t} \left[p_2^L(Q-k) + \sum_{l=0}^{Q-k} ((p_2^L - q)(l-(Q-k))) f_{\xi_2|p_2^L}(l) \right] f_{\xi_1}(k) \\
&+ \sum_{k=Q-I^t+1}^{Q} \left[p_2^H(Q-k) + \sum_{l=0}^{Q-k} ((p_2^H - q)(l-(Q-k))) f_{\xi_2|p_2^H}(l) \right] f_{\xi_1}(k) \\
&+ Q(p_1+B)\big(1 - \sum_{k=0}^{Q} f_{\xi_1}(k)\big) - B\big(\mu - \sum_{k=0}^{Q} k f_{\xi_1}(k)\big)
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}[\Pi^1(Q,\xi_1)] &= Q(p_1-v) - B(\mu - Q) \\
&+ \sum_{k=0}^{Q} kp_1 f_{\xi_1}(k) + \sum_{k=0}^{Q-I^t} \left[p_2^L(Q-k) + \sum_{l=0}^{Q-k} ((p_2^L - q)(l-(Q-k))) f_{\xi_2|p_2^L}(l) \right] f_{\xi_1}(k) \\
&+ \sum_{k=Q-I^t+1}^{Q} \left[p_2^H(Q-k) + \sum_{l=0}^{Q-k} ((p_2^H - q)(l-(Q-k))) f_{\xi_2|p_2^H}(l) \right] f_{\xi_1}(k) \\
&- Qp_1 \sum_{k=0}^{Q} f_{\xi_1}(k) - B \sum_{k=0}^{Q} (Q-k) f_{\xi_1}(k)
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}[\Pi^1(Q,\xi_1)] &= Q(p_1-v) - B(\mu - Q) \\
&+ \sum_{k=0}^{Q-I^t} \left[p_2^L(Q-k) + \sum_{l=0}^{Q-k} ((p_2^L - q)(l-(Q-k))) f_{\xi_2|p_2^L}(l) \right] f_{\xi_1}(k) \\
&+ \sum_{k=Q-I^t+1}^{Q} \left[p_2^H(Q-k) + \sum_{l=0}^{Q-k} ((p_2^H - q)(l-(Q-k))) f_{\xi_2|p_2^H}(l) \right] f_{\xi_1}(k) \\
&- (p_1+B) \sum_{k=0}^{Q} (Q-k) f_{\xi_1}(k) \qquad \text{(F.8)}
\end{aligned}
$$

Appendix G

Results of Performance Measures

Numerically, we solve the problems (4.1), (4.2), (4.4), (4.8), and (4.10) for selected demand distributions and the exact results are observed. Q^* is the optimal solution of the newsvendor problem with recourse action (4.1). Q^N is the optimal newsvendor solution of the problem (4.2) where the ending inventory is only salvaged. Q^L is the optimal solution of problem 4.4 in which the decision maker has the information but no threshold value is used for decisions. In this problem it is assumed that the retailer sells the remaining inventory at a lower marked-down price p_2^L and observes the demand of higher price. Q^{U_1} is the optimal solution for the problem (4.8) where the salvage value is equal to higher price. Lastly, Q^{U_2} is the optimal solution of the problem (4.10) where the remaining inventory at the end of the first period is offered at higher price regardless of the second period decision rule. In this problem, we assume that the marked-down price is given as the higher price p_2^H with the demand of lower price. All of the solutions are found as integer and all terms are expected in Tables. By using the solutions of these problems, we derive two intervals for the optimal order quantity and for its objective function, one of which is tighter. The intervals of the order quantity, which are $[Q^N, Q^{U_1}]$ and $[Q^L, Q^{U_2}]$, as well as the intervals of the optimal objective function, which are $[\mathbb{E}[\Pi(Q^N, None)], \mathbb{E}[\Pi^1(Q^{U_1})]]$ and $[\mathbb{E}[\Pi(Q^L, I_t)], \mathbb{E}[\Pi^2(Q^{U_2})]]$, are shown below.

Table G.1: Range of order quantity when p_1=1,q=0.1,v=0.7,p_2^H=0.6, p_2^L=0.4,B=0,μ_1=80,μ_2^H=10,μ_2^L=60

Demand Option	Q^*	$[Q^N, Q^{U_1}]$	$[Q^L, Q^{U_2}]$
Poisson	81	[76,86]	[79,86]
Uniform	70	[53,120]	[57,84]
Symmetric Triangular	78	[65,103]	[68,89]
Right-Skewed Triangular	70	[56,109]	[60,84]
Left-Skewed Triangular	81	[69,103]	[72,93]
Combination of Symmetric Triangular and Uniform	76	[65,103]	[68,86]

Table G.2: Range of expected total profit when p_1=1,q=0.1,v=0.7,p_2^H=0.6, p_2^L=0.4,B=0,μ_1=80,μ_2^H=10,μ_2^L=60

Demand Option	$\mathbb{E}[\Pi(Q^*, I_t)]$	$[\mathbb{E}[\Pi(Q^N, None)], \mathbb{E}[\Pi^1(Q^{U_1})]]$	$[\mathbb{E}[\Pi(Q^L, I_t)], \mathbb{E}[\Pi^2(Q^{U_2})]]$
Poisson	22.28	[21.10,22.85]	[22.22,22.85]
Uniform	10.99	[8.00,18.00]	[10.59,13.81]
Symmetric Triangular	15.81	[13.06,19.77]	[15.45,18.37]
Right-Skewed Triangular	13.63	[11.11,18.34]	[13.31,16.21]
Left-Skewed Triangular	16.59	[13.71,20.57]	[16.26,19.26]
Combination of Symmetric Triangular and Uniform	15.51	[13.06,19.77]	[15.24,17.70]

Table G.3: Range of order quantity when p_1=1,q=0.1,v=0.7,p_2^H=0.6, p_2^L=0.4,B=0,μ_1=40,μ_2^H=5,μ_2^L=30

Demand Option	Q^*	$[Q^N, Q^{U_1}]$	$[Q^L, Q^{U_2}]$
Poisson	40	[37,44]	[39,44]
Uniform	35	[27,60]	[28,42]
Symmetric Triangular	39	[33,52]	[34,45]
Right-Skewed Triangular	35	[28,54]	[30,42]
Left-Skewed Triangular	41	[34,51]	[36,46]
Combination of Symmetric Triangular and Uniform	38	[33,52]	[34,43]

Table G.4: Range of expected total profit when p_1=1,q=0.1,v=0.7,p_2^H=0.6, p_2^L=0.4,B=0,μ_1=40,μ_2^H=5,μ_2^L=30

Demand Option	$\mathbb{E}[\Pi(Q^*, I_t)]$	$[\mathbb{E}[\Pi(Q^N, None)], \mathbb{E}[\Pi^1(Q^{U_1})]]$	$[\mathbb{E}[\Pi(Q^L, I_t)], \mathbb{E}[\Pi^2(Q^{U_2})]]$
Poisson	10.64	[9.96,11.18]	[10.62,11.18]
Uniform	5.49	[4.00,9.00]	[5.26,6.90]
Symmetric Triangular	7.90	[6.53,9.89]	[7.72,9.18]
Right-Skewed Triangular	6.81	[5.55,9.17]	[6.65,8.11]
Left-Skewed Triangular	8.30	[6.86,10.28]	[8.13,9.63]
Combination of Symmetric Triangular and Uniform	7.75	[6.53,9.89]	[7.62,8.85]

Table G.5: Range of order quantity when p_1=1,q=0.1,v=0.7,p_2^H=0.6, p_2^L=0.4,B=0.2,μ_1=80,μ_2^H=10,μ_2^L=60

Demand Option	Q^*	$[Q^N, Q^{U_1}]$	$[Q^L, Q^{U_2}]$
Poisson	84	[79,89]	[81,89]
Uniform	88	[73,133]	[75,99]
Symmetric Triangular	87	[76,114]	[79,98]
Right-Skewed Triangular	83	[69,124]	[72,95]
Left-Skewed Triangular	91	[80,109]	[83,100]
Combination of Symmetric Triangular and Uniform	86	[76,114]	[79,96]

Table G.6: Range of expected total profit when p_1=1,q=0.1,v=0.7,p_2^H=0.6, p_2^L=0.4,B=0.2,μ_1=80,μ_2^H=10,μ_2^L=60

Demand Option	$\mathbb{E}[\Pi(Q^*, I_t)]$	$[\mathbb{E}[\Pi(Q^N, None)], \mathbb{E}[\Pi^1(Q^{U_1})]]$	$[\mathbb{E}[\Pi(Q^L, I_t)], \mathbb{E}[\Pi^2(Q^{U_2})]]$
Poisson	21.78	[20.11,22.64]	[21.67,22.64]
Uniform	6.93	[2.18,17.33]	[6.40,10.91]
Symmetric Triangular	13.41	[9.43,19.08]	[13.04,16.87]
Right-Skewed Triangular	9.95	[5.92,17.33]	[9.45,13.64]
Left-Skewed Triangular	14.86	[10.69,20.15]	[14.48,18.34]
Combination of Symmetric Triangular and Uniform	12.99	[9.43,19.08]	[12.71,16.00]

Table G.7: Range of order quantity when p_1=1,q=0.1,v=0.7,p_2^H=0.6, p_2^L=0.4,B=0.2,μ_1=40,μ_2^H=5,μ_2^L=30

Demand Option	Q^*	$[Q^N, Q^{U_1}]$	$[Q^L, Q^{U_2}]$
Poisson	42	[39,46]	[40,46]
Uniform	44	[36,67]	[38,50]
Symmetric Triangular	44	[38,57]	[39,49]
Right-Skewed Triangular	42	[35,62]	[36,48]
Left-Skewed Triangular	46	[40,55]	[41,50]
Combination of Symmetric Triangular and Uniform	43	[38,57]	[39,48]

Table G.8: Range of expected total profit when p_1=1,q=0.1,v=0.7,p_2^H=0.6, p_2^L=0.4,B=0.2,μ_1=40,μ_2^H=5,μ_2^L=30

Demand Option	$\mathbb{E}[\Pi(Q^*, I_t)]$	$[\mathbb{E}[\Pi(Q^N, None)], \mathbb{E}[\Pi^1(Q^{U_1})]]$	$[\mathbb{E}[\Pi(Q^L, I_t)], \mathbb{E}[\Pi^2(Q^{U_2})]]$
Poisson	10.25	[9.26,11.03]	[10.14,11.03]
Uniform	3.47	[1.09,8.67]	[3.24,5.45]
Symmetric Triangular	6.70	[4.71,9.54]	[6.47,8.44]
Right-Skewed Triangular	4.98	[2.96,8.66]	[4.72,6.82]
Left-Skewed Triangular	7.42	[5.35,10.07]	[7.19,9.17]
Combination of Symmetric Triangular and Uniform	6.50	[4.71,9.54]	[6.31,8.00]

Table G.9: Range of order quantity when p_1=1,q=0.1,v=0.5,p_2^H=0.4, p_2^L=0.25,B=0,μ_1=80,μ_2^H=10,μ_2^L=60

Demand Option	Q^*	$[Q^N, Q^{U_1}]$	$[Q^L, Q^{U_2}]$
Poisson	85	[81,89]	[83,89]
Uniform	99	[89,133]	[91,109]
Symmetric Triangular	92	[85,114]	[86,102]
Right-Skewed Triangular	91	[81,124]	[83,102]
Left-Skewed Triangular	95	[89,109]	[90,103]
Combination of Symmetric Triangular and Uniform	91	[85114]	[86,101]

Table G.10: Range of expected total profit when p_1=1,q=0.1,v=0.5,p_2^H=0.4, p_2^L=0.25,B=0,μ_1=80,μ_2^H=10,μ_2^L=60

Demand Option	$\mathbb{E}[\Pi(Q^*, I_t)]$	$[\mathbb{E}[\Pi(Q^N, None)], \mathbb{E}[\Pi^1(Q^{U_1})]]$	$[\mathbb{E}[\Pi(Q^L, I_t)], \mathbb{E}[\Pi^2(Q^{U_2})]]$
Poisson	38.01	[36.82,38.64]	[37.97,38.64]
Uniform	25.34	[22.22,33.33]	[25.18,28.85]
Symmetric Triangular	30.66	[28.11,35.08]	[30.48,33.58]
Right-Skewed Triangular	27.46	[24.76,33.33]	[27.27,30.69]
Left-Skewed Triangular	32.12	[29.51,36.15]	[31.97,34.98]
Combination of Symmetric Triangular and Uniform	30.42	[28.11,35.08]	[30.28,33.00]

Table G.11: Range of order quantity when p_1=1,q=0.1,v=0.5,p_2^H=0.4, p_2^L=0.25,B=0,μ_1=40,μ_2^H=5,μ_2^L=30

Demand Option	Q^*	$[Q^N, Q^{U_1}]$	$[Q^L, Q^{U_2}]$
Poisson	43	[41,46]	[42,46]
Uniform	49	[44,67]	[45,54]
Symmetric Triangular	46	[42,57]	[43,51]
Right-Skewed Triangular	45	[41,62]	[41,51]
Left-Skewed Triangular	48	[44,55]	[45,51]
Combination of Symmetric Triangular and Uniform	46	[42,57]	[43,50]

Table G.12: Range of expected total profit when p_1=1,q=0.1,v=0.5,p_2^H=0.4, p_2^L=0.25,B=0,μ_1=40,μ_2^H=5,μ_2^L=30

Demand Option	$\mathbb{E}[\Pi(Q^*, I_t)]$	$[\mathbb{E}[\Pi(Q^N, None)], \mathbb{E}[\Pi^1(Q^{U_1})]]$	$[\mathbb{E}[\Pi(Q^L, I_t)], \mathbb{E}[\Pi^2(Q^{U_2})]]$
Poisson	18.43	[17.75,19.03]	[18.42,19.03]
Uniform	12.67	[11.11,16.67]	[12.57,14.42]
Symmetric Triangular	15.33	[14.06,17.54]	[15.24,16.79]
Right-Skewed Triangular	13.73	[12.38,16.66]	[13.61,15.35]
Left-Skewed Triangular	16.06	[14.75,18.07]	[15.98,17.49]
Combination of Symmetric Triangular and Uniform	15.21	[14.06,17.54]	[15.14,16.50]

Table G.13: Range of order quantity when p_1=1,q=0.1,v=0.5,p_2^H=0.4, p_2^L=0.25,B=0.5,μ_1=80,μ_2^H=10,μ_2^L=60

Demand Option	Q^*	$[Q^N, Q^{U_1}]$	$[Q^L, Q^{U_2}]$
Poisson	88	[85,92]	[86,92]
Uniform	121	[114,145]	[115,127]
Symmetric Triangular	106	[100,126]	[101,114]
Right-Skewed Triangular	111	[103,142]	[104,119]
Left-Skewed Triangular	105	[100,116]	[101,111]
Combination of Symmetric Triangular and Uniform	106	[100,126]	[101,113]

Table G.14: Range of expected total profit when p_1=1,q=0.1,v=0.5,p_2^H=0.4, p_2^L=0.25,B=0.5,μ_1=80,μ_2^H=10,μ_2^L=60

Demand Option	$\mathbb{E}[\Pi(Q^*, I_t)]$	$[\mathbb{E}[\Pi(Q^N, None)], \mathbb{E}[\Pi^1(Q^{U_1})]]$	$[\mathbb{E}[\Pi(Q^L, I_t)], \mathbb{E}[\Pi^2(Q^{U_2})]]$
Poisson	37.36	[35.70,38.35]	[37.29,38.35]
Uniform	21.58	[17.14,32.73]	[21.43,26.22]
Symmetric Triangular	27.80	[24.13,34.27]	[27.65,31.83]
Right-Skewed Triangular	23.02	[18.90,32.14]	[22.82,27.53]
Left-Skewed Triangular	30.42	[26.92,35.66]	[30.27,34.02]
Combination of Symmetric Triangular and Uniform	27.47	[24.13,34.27]	[27.35,31.11]

Table G.15: Range of order quantity when p_1=1,q=0.1,v=0.5,p_2^H=0.4, p_2^L=0.25,B=0.5,μ_1=40,μ_2^H=5,μ_2^L=30

Demand Option	Q^*	$[Q^N, Q^{U_1}]$	$[Q^L, Q^{U_2}]$
Poisson	45	[43,49]	[44,49]
Uniform	60	[57,73]	[58,64]
Symmetric Triangular	53	[50,63]	[50,57]
Right-Skewed Triangular	55	[52,71]	[52,60]
Left-Skewed Triangular	52	[50,58]	[51,55]
Combination of Symmetric Triangular and Uniform	53	[50,63]	[50,57]

Table G.16: Range of expected total profit when p_1=1,q=0.1,v=0.5,p_2^H=0.4, p_2^L=0.25,B=0.5,μ_1=40,μ_2^H=5,μ_2^L=30

Demand Option	$\mathbb{E}[\Pi(Q^*, I_t)]$	$[\mathbb{E}[\Pi(Q^N, None)], \mathbb{E}[\Pi^1(Q^{U_1})]]$	$[\mathbb{E}[\Pi(Q^L, I_t)], \mathbb{E}[\Pi^2(Q^{U_2})]]$
Poisson	17.92	[16.95,18.82]	[17.88,18.82]
Uniform	10.79	[8.57,16.36]	[10.74,13.11]
Symmetric Triangular	13.90	[12.06,17.14]	[13.79,15.92]
Right-Skewed Triangular	11.51	[9.45,16.07]	[11.41,13.76]
Left-Skewed Triangular	15.20	[13.46,17.83]	[15.17,17.01]
Combination of Symmetric Triangular and Uniform	13.74	[12.06,17.14]	[13.64,15.55]

Appendix H

Exact Results for Different Demand Distributions

Table H.1: Threshold and order quantity when demand distribution is known

p_1	q	v	p_2^H	p_2^L	B	μ_1	μ_2^H	μ_2^L	I^t	Q^*	$\mathbb{E}[\Pi(Q^*)]$	$\mathbb{P}(p_2 = p_2^H)$	$\mathbb{P}(p_2 = p_2^L)$	q_1	q_2
Poisson															
1	0.1	0.7	0.6	0.4	0	80	10	60	17	81	22.28	0.54	0.04	0.93	0.07
1	0.1	0.7	0.6	0.4	0	40	5	30	9	40	10.64	0.46	0.09	0.84	0.16
1	0.1	0.7	0.6	0.4	0.2	80	10	60	17	84	21.78	0.62	0.08	0.89	0.11
1	0.1	0.7	0.6	0.4	0.2	40	5	30	9	42	10.25	0.51	0.15	0.77	0.23
1	0.1	0.5	0.4	0.25	0	80	10	60	20	85	38.01	0.69	0.05	0.93	0.07
1	0.1	0.5	0.4	0.25	0	40	5	30	10	43	18.43	0.56	0.15	0.79	0.21
1	0.1	0.5	0.4	0.25	0.5	80	10	60	20	88	37.36	0.73	0.10	0.88	0.12
1	0.1	0.5	0.4	0.25	0.5	40	5	30	10	45	17.92	0.57	0.24	0.70	0.30
Uniform															
1	0.1	0.7	0.6	0.4	0	80	10	60	17.78	70.00	10.99	0.11	0.33	0.25	0.75
1	0.1	0.7	0.6	0.4	0	40	5	30	8.89	35.00	5.49	0.11	0.33	0.25	0.75
1	0.1	0.7	0.6	0.4	0.2	80	10	60	17.78	88.00	6.93	0.11	0.44	0.20	0.80
1	0.1	0.7	0.6	0.4	0.2	40	5	30	8.89	44.00	3.47	0.11	0.44	0.20	0.80
1	0.1	0.5	0.4	0.25	0	80	10	60	22.03	99.00	25.34	0.14	0.48	0.22	0.78
1	0.1	0.5	0.4	0.25	0	40	5	30	11.02	49.28	12.67	0.14	0.48	0.22	0.78
1	0.1	0.5	0.4	0.25	0.5	80	10	60	22.03	121.00	21.58	0.14	0.62	0.18	0.82
1	0.1	0.5	0.4	0.25	0.5	40	5	30	11.02	60.36	10.79	0.14	0.62	0.18	0.82
Symmetric Triangular															
1	0.1	0.7	0.6	0.4	0	80	10	60	16.80	78.00	15.81	0.18	0.29	0.38	0.62
1	0.1	0.7	0.6	0.4	0	40	5	30	8.40	39.00	7.91	0.18	0.29	0.38	0.62
1	0.1	0.7	0.6	0.4	0.2	80	10	60	16.80	87.29	13.41	0.20	0.39	0.34	0.66
1	0.1	0.7	0.6	0.4	0.2	40	5	30	8.40	44.00	6.70	0.20	0.40	0.33	0.67
1	0.1	0.5	0.4	0.25	0	80	10	60	20.40	92.17	30.66	0.24	0.40	0.37	0.63
1	0.1	0.5	0.4	0.25	0	40	5	30	10.20	46.08	15.33	0.24	0.40	0.37	0.63
1	0.1	0.5	0.4	0.25	0.5	80	10	60	20.40	106.10	27.80	0.20	0.57	0.26	0.74
1	0.1	0.5	0.4	0.25	0.5	40	5	30	10.20	53.05	13.90	0.20	0.57	0.26	0.74

Table H.2: Threshold and order quantity when demand distribution is known

p_1	q	v	p_2^H	p_2^L	B	μ_1	μ_2^H	μ_2^L	I^t	Q^*	$\mathbb{E}[\Pi(Q^*)]$	$\mathbb{P}(p_2 = p_2^H)$	$\mathbb{P}(p_2 = p_2^L)$	q_1	q_2
Right-Skewed Triangular															
1	0.1	0.7	0.6	0.4	0	80	10	60	16.39	67.00	13.78	0.16	0.28	0.36	0.64
1	0.1	0.7	0.6	0.4	0	40	5	30	8.20	33.28	6.89	0.16	0.27	0.37	0.63
1	0.1	0.7	0.6	0.4	0.2	80	10	60	16.39	83.28	9.96	0.14	0.43	0.24	0.76
1	0.1	0.7	0.6	0.4	0.2	40	5	30	8.20	42.00	4.98	0.14	0.44	0.24	0.76
1	0.1	0.5	0.4	0.25	0	80	10	60	20.49	91.00	27.46	0.16	0.47	0.26	0.74
1	0.1	0.5	0.4	0.25	0	40	5	30	10.25	45.40	13.73	0.17	0.46	0.26	0.74
1	0.1	0.5	0.4	0.25	0.5	80	10	60	20.49	111.00	23.02	0.14	0.63	0.18	0.82
1	0.1	0.5	0.4	0.25	0.5	40	5	30	10.25	55.44	11.51	0.14	0.63	0.18	0.82
Left-Skewed Triangular															
1	0.1	0.7	0.6	0.4	0	80	10	60	16.87	81.46	16.59	0.17	0.30	0.37	0.63
1	0.1	0.7	0.6	0.4	0	40	5	30	8.44	41.00	8.30	0.18	0.30	0.37	0.63
1	0.1	0.7	0.6	0.4	0.2	80	10	60	16.87	91.15	14.86	0.20	0.39	0.34	0.66
1	0.1	0.7	0.6	0.4	0.2	40	5	30	8.44	46.00	7.43	0.20	0.40	0.33	0.67
1	0.1	0.5	0.4	0.25	0	80	10	60	20.36	95.36	32.12	0.25	0.40	0.38	0.62
1	0.1	0.5	0.4	0.25	0	40	5	30	10.18	48.00	16.06	0.25	0.41	0.38	0.62
1	0.1	0.5	0.4	0.25	0.5	80	10	60	20.36	105.00	30.42	0.27	0.51	0.35	0.65
1	0.1	0.5	0.4	0.25	0.5	40	5	30	10.18	52.49	15.21	0.27	0.51	0.35	0.65
Combination of Triangular and Uniform															
1	0.1	0.7	0.6	0.4	0	80	10	60	17.78	76.28	15.51	0.19	0.27	0.41	0.59
1	0.1	0.7	0.6	0.4	0	40	5	30	8.89	38.14	7.75	0.19	0.27	0.41	0.59
1	0.1	0.7	0.6	0.4	0.2	80	10	60	17.78	86.15	12.99	0.21	0.37	0.36	0.64
1	0.1	0.7	0.6	0.4	0.2	40	5	30	8.89	43.08	6.50	0.21	0.37	0.36	0.64
1	0.1	0.5	0.4	0.25	0	80	10	60	22.03	91.49	30.42	0.26	0.38	0.40	0.60
1	0.1	0.5	0.4	0.25	0	40	5	30	11.02	46.00	15.21	0.26	0.38	0.40	0.60
1	0.1	0.5	0.4	0.25	0.5	80	10	60	22.03	106.00	27.48	0.22	0.55	0.29	0.71
1	0.1	0.5	0.4	0.25	0.5	40	5	30	11.02	53.00	13.74	0.22	0.55	0.29	0.71

Appendix I

Numerical Results

We conduct a numerical study to evaluate the performance of each data-driven model to the sample size for all parameter settings given in Table 4.4 and for all demand distributions selected.

As mentioned before, there are six data-driven models that solve the problem defined hierarchically. SAA-P is the model where the algorithm based on SAA is used to solve both stage whereas MIP-SAA is the model where the algorithm based on SAA is used for the second stage and its results are used as inputs in mixed integer program to solve the first stage. SAA-LR is another data-driven model which use linear regression for the second stage and its results are used as inputs in the algorithm based on SAA for the first stage. In MIP-LR, linear regression is used for the second stage and its results are used as inputs in mixed integer program for the fist stage. We also propose SAA-QR model where quadratic regression is used for the second stage and its results are used as inputs in the algorithm based on SAA for the fist stage. MIP-QR is the model where quadratic regression is used for the second stage and its results are used as inputs in mixed integer program to solve the first stage. In addition to that, MIP-P is the model where mixed integer program is used to solve both stages to find two decisions simultaneously. Lastly, MIP-GP is a simple heuristic where we run a mixed integer program for the high marked-down price as well as the low marked-down price and select the price and order quantity combination that yields to the larger average total profit prior to the beginning of the first stage.

I.1 Testing Random Number Generator

Before the numerical study, we generate our data sets in Matlab by using the random number generators. In the beginning of the data generation process, we applied the basic known tests in [37] to the random number generator. Chi square test, one of the empirical test, is implemented to test uniformly distributed random variates U_i's. It is used in order to test if the U_i's appear to be uniformly distributed between 0 and 1. According to the chi square test result, random variables generated from uniform distribution resemble the values of real random variables and hence we can say that this random number generator can be used. However, we also implement different kinds of tests such as serial, runs-up, and computation of correlation so as to make sure that this random number generator is "good" enough. Serial test is a generalization of the chi square test and it is implemented to check the uniformity in higher dimension. We apply the serial test for 2-dimensional uniformity. Runs-up test is a test of independence assumption. The last test, namely correlation test, is to examine if the generated U(0,1) random numbers demonstrate discernible correlation. We calculate an estimate of correlation at lags $j = 1, 2, 3, 4$ for the test statistic. The test statistics results related to the random number generator are given below.

I.2 The Estimated Ratio used in The Testing Data

In the testing stage, the model may want to choose a price different than the marked-down price in testing sample depending on the leftover. In such cases, one should change the marked-down price in testing sample according to the

Table I.1: Results of Random Number Generator Test

Tests (At level 0.05)	Results for (500x100) U'i s	
	Accepted Samples	Rejected Sample
Chi Square Test:	95	5
Serial Test:	97	3
Runs-up Test:	92	8
	Insignificant A_j's	Significant A_j's
Correlation Test at lag 1 (A_1):	94	6
Correlation Test at lag 2 (A_2):	97	3
Correlation Test at lag 3 (A_3):	96	4
Correlation Test at lag 4 (A_4):	92	8

Table I.2: Results of Random Number Generator Test

Tests (At level 0.05)	Results for (500x300) U'i s (1)	
	Accepted Samples	Rejected Sample
Chi Square Test:	290	10
Serial Test:	285	15
Runs-up Test:	281	19
	Insignificant A_j's	Significant A_j's
Correlation Test at lag 1 (A_1):	280	20
Correlation Test at lag 2 (A_2):	285	15
Correlation Test at lag 3 (A_3):	286	14
Correlation Test at lag 4 (A_4):	281	19

Table I.3: Results of Random Number Generator Test

Tests (At level 0.05)	Results for (500x300) U'i s (2)	
	Accepted Samples	Rejected Sample
Chi Square Test:	284	16
Serial Test:	281	19
Runs-up Test:	273	27
	Insignificant A_j's	Significant A_j's
Correlation Test at lag 1 (A_1):	283	17
Correlation Test at lag 2 (A_2):	286	14
Correlation Test at lag 3 (A_3):	288	12
Correlation Test at lag 4 (A_4):	288	12

threshold value estimated by model. Furthermore, the one should change not only the marked-down price but also the second period demand accordingly.The algorithm used to change the price and the second period demand when needed is given below.

Figure I.1 Algorithm for estimated ratio

if There is at least one demand observation for each p_2^j where $i = H, L$ **then**
    for each p_2^j where $i = H, L$ **do**
        for i=1... N_j **do**
            Process/Read $\xi_2^i(p_2^j)$ in the training data
            $SUM_j \leftarrow SUM_j + \xi_2^i(p_2^j)$
        end for
        $AVERAGE_j \leftarrow \frac{SUM_j}{N_j}$
    end for
    $EstimatedRatio \leftarrow \frac{SUM_H}{SUM_L}$
else
    $EstimatedRatio \leftarrow \frac{p_2^L}{p_2^H}$
end if

Figure I.2 The algorithm for the second period demand estimation by using Ratio

Process/Read Q, I^t
Process/Read ξ_1, p_2, ξ_2 from the testing sample
if $Q - \xi_1 < I^t$ **then**
    if $p_2 \neq p_2^H$ **then**
        $p_2 \leftarrow p_2^H$
        $\xi_2 \leftarrow \xi_2 \times$Estimated Ratio
    end if
else if $Q - \xi_1 \geq I^t$ **then**
    if $p_2 \neq p_2^L$ **then**
        $p_2 \leftarrow p_2^L$
        $\xi_2 \leftarrow \xi_2/$Estimated Ratio
    end if
end if

I.3 The Effect of Sample Size

In this test, we use the significance testing approach and report the P value which is the weight of evidence in the data regarding H_0. P value is the smallest significance level at which the observed result would lead to rejection of H_0. We look at the difference between the average profit for training size 50 and the average profit for training size 250 and the average profit for training size 450, respectively. The differences are denoted by (1,2), (1,3) and (2,3) respectively. The purpose of the each test is to statistically check whether the defined difference is significantly different than zero. We interpret small P value as meaning that the data provide strong evidence against H_0 and it should be rejected. Also, we interpret large P value as meaning that there is not much evidence against H_0 and it should not be rejected.

$$\text{P-value} = \left(\text{sample mean would be as small as the observed value if } H_0 \text{ is true}\right)$$

When the observed value of the test statistic is T$= \frac{\bar{D}-0}{S_D/\sqrt{n}}$, P-value for two sided test is:

$$P = 2(1 - t(T))$$

The test is said to be statistically significant at the significance level α if the P-value is smaller than α.

Since the P value is the smallest significance level, we would reject H_0 for a significance level greater than or equal to this P value.

This test is conducted for all parameter settings given in Table 4.4 and for all demand distributions selected. For each model in Table 3.1, results are obtained below. The results for MIP-P and MIP-GP are also shown below. MIP-P is the model where mixed integer program is used to solve both stages to find two decisions simultaneously. Lastly, MIP-GP is a simple heuristic where we run a mixed integer program for the high marked-down price as well as the low marked-down price and select the price and order quantity combination that yields to the larger average total profit prior to the beginning of the first stage.

Table I.4: The effect of sample size for Poisson demand

Data/Parameter Setting	Difference of Train Data Set Size	Significance level							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	1,2	0.6506	0.6865	0.0073	0.0095	0.4115	0.7344	0.3621	0.6506
	1,3	0.4425	0.4587	0.0197	0.0249	0.7520	0.6672	0.3149	0.4425
	2,3	0.2139	0.1819	0.0042	0.0042	0.3447	0.6741	0.5391	0.2139
2	1,2	0.0010	0.0019	0.0707	0.0549	0.0002	0.0009	0.0010	0.0010
	1,3	<0.0001	0.0001	0.0178	0.0128	0.0006	0.0045	0.0003	<0.0001
	2,3	0.5508	0.5449	0.3258	0.3258	0.4884	0.0452	0.1969	0.5508
3	1,2	0.0011	0.0009	0.6132	0.6132	0.1042	0.1611	0.0019	0.0011
	1,3	0.0022	0.0017	0.4754	0.4754	0.0222	0.1954	0.0004	0.0022
	2,3	0.4016	0.3984	0.2495	0.2495	0.1595	0.9788	0.0081	0.4016
4	1,2	0.0047	0.0062	0.0118	0.0118	0.0057	0.0004	0.0034	0.0048
	1,3	0.0012	0.0013	0.0072	0.0072	0.0296	0.0035	0.0123	0.0012
	2,3	0.3949	0.3183	0.2133	0.2133	0.1949	0.0342	0.4722	0.3949
5	1,2	0.0070	0.0044	<0.0001	<0.0001	0.0925	0.2262	0.0100	0.0060
	1,3	0.0181	0.0101	<0.0001	<0.0001	0.0807	0.1831	0.0027	0.0181
	2,3	0.6548	0.4957	0.7106	0.7106	0.7517	0.9306	0.0523	0.7109
6	1,2	0.0176	0.0117	0.7362	0.7362	0.0083	0.0036	0.0118	0.0176
	1,3	0.0012	0.0010	0.4026	0.4026	0.0105	0.0062	0.0026	0.0011
	2,3	0.0383	0.0628	0.2231	0.2231	0.9796	0.9008	0.1722	0.0372
7	1,2	0.0059	0.0050	0.2145	0.2529	0.0157	0.0780	0.0314	0.0059
	1,3	0.0014	0.0009	0.3565	0.4150	0.0068	0.0652	0.0127	0.0014
	2,3	0.0877	0.0312	0.2878	0.2878	0.2936	0.3967	0.1905	0.0877
8	1,2	0.0279	0.0271	0.1339	0.1339	0.0708	0.0717	0.0100	0.0276
	1,3	0.0126	0.0101	0.0750	0.0750	0.0785	0.0306	0.0117	0.0126
	2,3	0.1915	0.1432	0.2662	0.2662	0.8364	0.5582	0.6924	0.1982

Table I.5: The effect of sample size for Uniform demand

Data/Parameter Setting	Difference of Train Data Set Size	Significance level							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	1,2	0.4446	0.3733	0.2714	0.2697	0.1112	0.1224	0.3192	0.9800
	1,3	0.1985	0.0618	0.1860	0.1843	0.0228	0.0625	0.2214	0.8005
	2,3	0.1779	0.0450	0.4844	0.4813	0.1228	0.1794	0.4822	0.6399
2	1,2	0.4200	0.3601	0.2720	0.2635	0.0853	0.1138	0.3142	0.8665
	1,3	0.1681	0.1604	0.1855	0.1778	0.0238	0.0609	0.2211	0.5896
	2,3	0.1410	0.2909	0.4799	0.4758	0.1923	0.2082	0.4988	0.4310
3	1,2	0.0709	0.0477	0.0512	0.0510	0.0046	0.0242	0.0679	0.2108
	1,3	0.0697	0.1477	0.0392	0.0391	0.0043	0.0142	0.0507	0.5313
	2,3	0.6961	0.3419	0.3988	0.3987	0.1544	0.1269	0.3746	0.0556
4	1,2	0.0807	0.0600	0.0511	0.0510	0.0055	0.0188	0.0684	0.3200
	1,3	0.0779	0.0303	0.0392	0.0391	0.0041	0.0093	0.0515	0.6942
	2,3	0.5946	0.3724	0.3981	0.4013	0.1338	0.0688	0.3779	0.5188
5	1,2	0.0724	0.0999	0.0557	0.0557	0.0430	0.0373	0.0595	0.2041
	1,3	0.0901	0.1523	0.0731	0.0731	0.0226	0.0252	0.0778	0.2386
	2,3	0.7780	0.6634	0.6783	0.6776	0.3178	0.3737	0.6807	0.9634
6	1,2	0.0678	0.0584	0.0552	0.0549	0.0431	0.0330	0.0578	0.1776
	1,3	0.0699	0.1449	0.0691	0.0693	0.0228	0.0214	0.0733	0.2165
	2,3	0.8733	0.4444	0.7647	0.7498	0.3152	0.3777	0.7326	0.7765
7	1,2	0.0756	0.1789	0.0956	0.0957	0.1095	0.0672	0.0805	0.1733
	1,3	0.0726	0.5021	0.0702	0.0703	0.1077	0.0470	0.0619	0.1699
	2,3	0.9472	0.5410	0.3720	0.3735	0.7927	0.4277	0.4067	0.7037
8	1,2	0.0748	0.1093	0.0950	0.0956	0.0748	0.0544	0.0806	0.1455
	1,3	0.0688	0.3439	0.0699	0.0701	0.0701	0.0428	0.0625	0.2098
	2,3	0.7191	0.9372	0.3639	0.3603	0.7407	0.5077	0.4035	0.5563

Table I.6: The effect of sample size for Triangular distributed demand

Data/Parameter Setting	Difference of Train Data Set Size	Significance level							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	1,2	0.7456	0.9843	0.7741	0.7676	0.4047	0.3662	0.8136	0.0031
	1,3	0.5859	0.6495	0.4224	0.4236	0.4100	0.2187	0.4864	<0.0001
	2,3	0.4906	0.4428	0.1684	0.1754	0.9474	0.2604	0.2187	<0.0001
2	1,2	0.8966	0.9324	0.7805	0.7709	0.3323	0.4421	0.7863	<0.0001
	1,3	0.7671	0.5716	0.4254	0.4269	0.2972	0.2896	0.4681	<0.0001
	2,3	0.2656	0.4902	0.1651	0.1750	0.7537	0.3146	0.2237	0.0008
3	1,2	0.0254	0.6706	0.0422	0.0422	0.0287	0.0762	0.0433	0.6210
	1,3	0.0238	0.1040	0.0272	0.0274	0.0205	0.0570	0.0259	<0.0001
	2,3	0.5381	0.0367	0.1728	0.1765	0.3136	0.3910	0.1473	<0.0001
4	1,2	0.0323	0.1743	0.0468	0.0465	0.0347	0.0869	0.0456	0.6883
	1,3	0.0218	0.1076	0.0302	0.0306	0.0302	0.0581	0.0277	<0.0001
	2,3	0.2410	0.0073	0.1731	0.1791	0.4209	0.3152	0.1512	<0.0001
5	1,2	0.0457	0.6166	0.0226	0.0227	0.0531	0.0730	0.0188	0.9329
	1,3	0.0528	0.5258	0.0372	0.0374	0.0672	0.0573	0.0321	0.9747
	2,3	0.9023	0.2374	0.4887	0.4821	0.8608	0.3916	0.4370	0.7897
6	1,2	0.0320	0.1457	0.0229	0.0226	0.0521	0.0593	0.0188	0.9014
	1,3	0.0522	0.7505	0.0369	0.0368	0.0569	0.0515	0.0314	0.9174
	2,3	0.4193	0.1970	0.5253	0.5168	0.9232	0.7821	0.4632	0.9600
7	1,2	0.1427	0.7160	0.0830	0.0831	0.1407	0.0512	0.0843	0.5345
	1,3	0.0990	0.0326	0.0599	0.0598	0.1457	0.0321	0.0591	0.7458
	2,3	0.5022	0.0437	0.4631	0.4622	0.9298	0.2874	0.5064	0.4457
8	1,2	0.1056	0.6041	0.0829	0.0830	0.1433	0.0602	0.0841	0.5766
	1,3	0.0760	0.0028	0.0595	0.0600	0.1473	0.0335	0.0586	0.9397
	2,3	0.6487	0.0082	0.4590	0.4696	0.9361	0.2418	0.5064	0.0987

Table I.7: The effect of sample size for Right-Skewed Triangular distributed demand

Data/Parameter Setting	Difference of Train Data Set Size	Significance level							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	1,2	0.3986	0.8642	0.4778	0.4734	0.1211	0.3349	0.5230	0.3624
	1,3	0.1819	0.4401	0.2248	0.2229	0.0419	0.2011	0.2613	0.0088
	2,3	0.2968	0.2450	0.1677	0.1678	0.2921	0.2323	0.2258	0.0333
2	1,2	0.5101	0.7054	0.4615	0.4527	0.1066	0.2391	0.5359	0.3894
	1,3	0.1896	0.3907	0.2053	0.2054	0.0361	0.1509	0.2771	0.7825
	2,3	0.0845	0.4533	0.1605	0.1701	0.2932	0.2959	0.2358	0.6614
3	1,2	0.0573	0.0726	0.0596	0.0603	0.0036	0.0753	0.0715	0.4059
	1,3	0.0450	0.3299	0.0301	0.0305	0.0032	0.0587	0.0363	0.1642
	2,3	0.3253	0.4604	0.0826	0.0822	0.2741	0.4184	0.0804	0.0155
4	1,2	0.0364	0.0415	0.0575	0.0576	0.0038	0.0747	0.0693	0.4230
	1,3	0.0295	0.0847	0.0270	0.0269	0.0033	0.0613	0.0365	0.1445
	2,3	0.4944	0.6066	0.0650	0.0610	0.2384	0.5162	0.1042	0.0021
5	1,2	0.0573	0.0514	0.0290	0.0291	0.0274	0.0353	0.0334	0.3195
	1,3	0.0730	0.0550	0.0359	0.0360	0.0161	0.0239	0.0418	0.8617
	2,3	0.9995	0.4799	0.9518	0.9499	0.5473	0.3487	0.9525	0.2466
6	1,2	0.0718	0.2201	0.0290	0.0288	0.0476	0.0424	0.0346	0.3316
	1,3	0.0880	0.0416	0.0359	0.0358	0.0289	0.0348	0.0425	0.9933
	2,3	0.7825	0.3211	0.9622	0.9604	0.5453	0.4723	0.9256	0.0840
7	1,2	0.1935	0.5847	0.1850	0.1845	0.0754	0.0383	0.1903	0.4855
	1,3	0.1523	0.7389	0.1351	0.1348	0.0581	0.0258	0.1405	0.4010
	2,3	0.6993	0.8248	0.5559	0.5565	0.5124	0.4980	0.4588	0.6231
8	1,2	0.1926	0.1901	0.1846	0.1840	0.0740	0.0395	0.1935	0.4078
	1,3	0.1626	0.9294	0.1344	0.1346	0.0551	0.0280	0.1429	0.4583
	2,3	0.7185	0.1593	0.5572	0.5630	0.4649	0.5084	0.4582	0.8761

Table I.8: The effect of sample size for Left-Skewed Triangular distributed demand

Data/Parameter Setting	Difference of Train Data Set Size	Significance level							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	1,2	0.9209	0.2859	0.7470	0.7507	0.5009	0.3914	0.8415	0.0077
	1,3	0.6883	0.9059	0.4067	0.4069	0.2733	0.2093	0.5179	<0.0001
	2,3	0.0972	0.0681	0.1701	0.1685	0.4979	0.2234	0.2185	<0.0001
2	1,2	0.9978	0.8131	0.7604	0.7712	0.5202	0.3895	0.8171	0.0187
	1,3	0.6478	0.9128	0.3962	0.4025	0.3008	0.2129	0.4619	<0.0001
	2,3	0.2818	0.8509	0.1507	0.1494	0.4361	0.2772	0.1791	<0.0001
3	1,2	0.0560	0.4855	0.0533	0.0531	0.0770	0.0309	0.0565	0.5792
	1,3	0.0498	0.1790	0.0379	0.0378	0.0202	0.0250	0.0350	<0.0001
	2,3	0.4323	0.3851	0.2109	0.2129	0.0296	0.4583	0.1448	<0.0001
4	1,2	0.0458	0.1383	0.0550	0.0543	0.0898	0.0310	0.0581	0.8714
	1,3	0.0306	0.0169	0.0377	0.0374	0.0238	0.0282	0.0364	<0.0001
	2,3	0.2520	0.0043	0.1917	0.1961	0.0384	0.5280	0.1523	<0.0001
5	1,2	0.0320	0.4299	0.0291	0.0293	0.0700	0.0563	0.0270	0.9736
	1,3	0.0443	0.1301	0.0448	0.0446	0.1027	0.0403	0.0421	0.8325
	2,3	0.7162	0.2119	0.4177	0.4310	0.3826	0.4375	0.3971	0.7295
6	1,2	0.0289	0.5147	0.0289	0.0294	0.0716	0.0475	0.0271	0.6092
	1,3	0.0371	0.0856	0.0446	0.0446	0.1041	0.0481	0.0420	0.7673
	2,3	0.7433	0.0542	0.4143	0.4372	0.4147	0.8624	0.4082	0.5768
7	1,2	0.0685	0.1357	0.1235	0.1213	0.0640	0.0325	0.1010	0.6553
	1,3	0.0370	0.0121	0.0637	0.0625	0.0450	0.0195	0.0523	0.1658
	2,3	0.4001	0.5725	0.3111	0.3131	0.8600	0.3811	0.3295	0.0649
8	1,2	0.0589	0.5595	0.1221	0.1164	0.0871	0.0366	0.1026	0.8488
	1,3	0.0362	0.0226	0.0633	0.0598	0.0550	0.0177	0.0538	0.1565
	2,3	0.6828	0.0018	0.3147	0.3102	0.7843	0.2427	0.3388	0.0570

Table I.9: The effect of sample size for Symmetric Triangular distributed and Uniform distributed demand

Data/Parameter Setting	Difference of Train Data Set Size	Significance level							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	1,2	0.7272	0.9556	0.3460	0.3463	0.1850	0.1538	0.2778	0.1179
	1,3	0.5756	0.8908	0.1885	0.1886	0.3086	0.1046	0.1367	<0.0001
	2,3	0.5439	0.8014	0.3079	0.3076	0.5227	0.3827	0.2622	<0.0001
2	1,2	0.4543	0.4096	0.3453	0.3478	0.1666	0.1658	0.2678	0.2072
	1,3	0.3664	0.4088	0.1893	0.1922	0.2581	0.0984	0.1259	0.0108
	2,3	0.5772	0.9756	0.3110	0.3147	0.6238	0.3198	0.2399	0.0677
3	1,2	0.0206	0.2473	0.0372	0.0372	0.0053	0.0860	0.0313	0.8360
	1,3	0.0171	0.0718	0.0190	0.0190	0.0039	0.0679	0.0154	0.0081
	2,3	0.4168	0.5283	0.0919	0.0924	0.3880	0.3516	0.0798	0.0028
4	1,2	0.0355	0.0628	0.0373	0.0370	0.0149	0.0795	0.0350	0.7083
	1,3	0.0250	0.7317	0.0190	0.0191	0.0065	0.0557	0.0165	0.0078
	2,3	0.2617	0.0650	0.0903	0.0952	0.5070	0.2603	0.0731	0.0008
5	1,2	0.0461	0.3721	0.0335	0.0337	0.0857	0.0626	0.0341	0.8771
	1,3	0.0420	0.0841	0.0532	0.0532	0.0423	0.0312	0.0524	0.8184
	2,3	0.6143	0.0406	0.4817	0.4859	0.1079	0.1086	0.5212	0.8600
6	1,2	0.0241	0.0482	0.0329	0.0327	0.1056	0.0320	0.0342	0.6076
	1,3	0.0421	0.3788	0.0536	0.0534	0.0471	0.0282	0.0530	0.7872
	2,3	0.5071	0.1172	0.4483	0.4419	0.0750	0.6592	0.5011	0.6385
7	1,2	0.1626	0.7168	0.0992	0.0991	0.0847	0.0342	0.0986	0.7359
	1,3	0.1495	0.1277	0.0771	0.0771	0.0705	0.0148	0.0801	0.9033
	2,3	0.7968	0.1839	0.6410	0.6381	0.7004	0.1881	0.6927	0.5094
8	1,2	0.2568	0.8017	0.0992	0.0985	0.0778	0.0425	0.0989	0.9099
	1,3	0.2075	0.0388	0.0766	0.0768	0.0685	0.0169	0.0797	0.8837
	2,3	0.7032	0.0489	0.6352	0.6523	0.7509	0.1375	0.6839	0.4568

I.4 The Effect of Relevancy of Data

Similarly, we report the P value which is the weight of evidence in the data regarding H_0 for this test as well. We calculate the difference between the average profit obtained by relevancy level 1 and 2. Likewise, the difference between the average profit obtained from relevancy level 1 and 3 is calculated. The purpose of the test is to statistically check whether the defined difference is significantly different than zero for the 95 % confidence interval. This test is conducted for all parameter settings given in Table 4.4 and for all demand distributions selected. For each model in Table 3.1, results are obtained below. The results for MIP-P and MIP-GP are also shown below. MIP-P is the model where mixed integer program is used to solve both stages to find two decisions simultaneously. Lastly, MIP-GP is a simple heuristic where we run a mixed integer program for the high marked-down price as well as the low marked-down price and select the price and order quantity combination that yields to the larger average total profit prior to the beginning of the first stage.

Table I.10: SAA-P results when demand is Poisson

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.1527	0.4092	0.9375	<0.0001	0.1752	<0.0001
2	0.4476	<0.0001	0.8308	<0.0001	0.9776	0.0002
3	0.9722	0.0010	0.1426	<0.0001	0.1618	<0.0001
4	0.3992	0.0009	0.9348	0.0001	0.8215	0.0007
5	0.4063	0.0100	0.8485	<0.0001	0.5477	0.0001
6	0.4927	0.0006	0.4447	0.0003	0.9853	0.0116
7	0.9809	0.0668	0.1791	0.0001	0.1045	0.0215
8	0.3599	0.0827	0.1055	0.0922	0.1690	0.1454

Table I.11: MIP-SAA results when demand is Poisson

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.1343	0.5010	0.8681	<0.0001	0.1752	<0.0001
2	0.7021	0.0003	0.5975	<0.0001	0.7518	0.0006
3	0.8438	0.0007	0.1004	<0.0001	0.1102	<0.0001
4	0.3985	0.0027	0.8896	0.0001	0.8960	0.0822
5	0.7557	0.0061	0.9843	<0.0001	0.5868	0.0001
6	0.5317	0.0007	0.5585	0.0012	0.7249	0.0065
7	0.8314	0.0502	0.1557	0.0001	0.3636	0.0626
8	0.4153	0.2737	0.1569	0.1142	0.1899	0.2821

Table I.12: SAA-QR results when demand is Poisson

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.0300	0.0012	0.5534	0.0666	0.1583	0.0269
2	0.0405	0.1503	0.0057	0.0368	0.0037	0.0498
3	0.0859	0.1038	0.0362	0.6254	0.2609	0.4506
4	0.0933	0.7820	0.0004	0.0335	0.0032	0.1082
5	0.3147	<0.0001	0.6685	0.0625	0.1010	0.4625
6	0.0126	0.0612	0.0001	0.0451	0.0061	0.4029
7	0.9707	0.0033	0.7871	0.6371	0.1452	0.3727
8	0.0560	0.2136	<0.0001	0.0003	<0.0001	0.0025

Table I.13: MIP-QR results when demand is Poisson

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.0381	0.0015	0.5534	0.0666	0.1583	0.0269
2	0.0765	0.2424	0.0057	0.0368	0.0037	0.0498
3	0.0859	0.1038	0.0362	0.6254	0.2609	0.4506
4	0.0933	0.7820	0.0004	0.0335	0.0032	0.1082
5	0.3403	<0.0001	0.6685	0.0625	0.1010	0.4625
6	0.0126	0.0612	0.0001	0.0451	0.0061	0.4029
7	0.8714	0.0039	0.7871	0.6371	0.1452	0.3727
8	0.0560	0.2136	<0.0001	0.0003	<0.0001	0.0025

Table I.14: SAA-LR results when demand is Poisson

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.0302	0.1659	0.7331	0.2106	0.1855	0.0091
2	0.9138	0.7968	0.0635	0.0007	0.0083	0.0006
3	0.8380	0.9305	0.6562	0.4487	0.1605	0.0757
4	0.0434	0.2670	0.0020	0.0001	0.0360	0.0175
5	0.2118	0.9609	0.4997	0.0400	0.9557	0.6424
6	0.2847	0.7643	0.1778	0.2506	0.0502	0.2641
7	0.2080	0.0603	0.2462	0.6697	0.8638	0.3680
8	0.7607	0.2917	0.0503	0.8897	0.0232	0.6722

Table I.15: MIP-LR results when demand is Poisson

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.0190	0.0380	0.2133	0.1923	0.6076	0.1165
2	0.5802	0.1343	0.0281	0.0000	0.0622	0.0074
3	0.3376	0.4118	0.5092	0.3388	0.8214	0.5671
4	0.3704	0.9995	0.0014	0.0005	0.2117	0.0009
5	0.3588	0.7126	0.1410	0.1374	0.9991	0.4217
6	0.9965	0.1221	0.2017	0.7490	0.1227	0.3546
7	0.4364	0.1315	0.1703	0.2159	0.9260	0.9483
8	0.3452	0.0868	0.0741	0.9814	0.0295	0.9202

Table I.16: MIP-GP results when demand is Poisson

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.1048	0.7288	0.7962	<0.0001	0.9646	<0.0001
2	0.7113	0.0445	0.0928	<0.0001	0.1105	0.0001
3	0.9239	0.0015	0.7945	0.0005	0.8367	0.0002
4	0.4957	0.0586	0.0578	0.0014	0.0796	0.0006
5	0.5765	0.0262	0.7804	<0.0001	0.7410	<0.0001
6	0.1452	0.0440	0.0298	<0.0001	0.0772	0.0002
7	0.5638	0.1548	0.4881	0.0025	0.9917	0.0036
8	0.5810	0.1016	0.0867	0.0009	0.0269	0.0006

Table I.17: MIP-P results when demand is Poisson

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.1527	0.4092	0.8988	<0.0001	0.1752	<0.0001
2	0.4476	<0.0001	0.8308	<0.0001	0.9776	0.0002
3	0.9957	0.0010	0.1426	<0.0001	0.1618	<0.0001
4	0.4079	0.0010	0.9348	0.0001	0.8215	0.0007
5	0.4063	0.0100	0.9554	<0.0001	0.5079	<0.0001
6	0.4595	0.0006	0.5255	0.0003	0.9771	0.0120
7	0.9809	0.0668	0.1791	0.0001	0.1045	0.0215
8	0.2122	0.0827	0.1223	0.0965	0.1993	0.1454

Table I.18: SAA-P results when demand is Uniform

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.3172	0.4850	0.0889	0.4999	0.9019	0.5906
2	0.1983	0.7073	0.2262	0.7126	0.6484	0.9408
3	0.9319	0.8511	0.0113	0.1967	0.9291	0.9847
4	0.4214	0.9933	0.0572	0.7412	0.2632	0.3076
5	0.4311	0.6268	0.1396	0.1879	0.7514	0.8655
6	0.8023	0.7303	0.0426	0.8941	0.0325	0.3149
7	0.7914	0.8070	0.3317	0.1587	0.7180	0.6662
8	0.2944	0.9608	0.6371	0.2254	0.9274	0.4564

Table I.19: MIP-SAA results when demand is Uniform

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.2191	0.4902	0.2987	0.7399	0.7649	0.4382
2	0.2815	0.6457	0.4714	0.5740	0.6419	0.8379
3	0.9339	0.5364	0.0045	0.6159	0.8653	0.1473
4	0.9065	0.5957	0.1123	0.0984	0.5164	0.3160
5	0.3152	0.9476	0.8284	0.4374	0.2845	0.6185
6	0.3840	0.9046	0.5062	0.1922	0.6674	0.4552
7	0.7059	0.8781	0.9673	0.2757	0.1875	0.7119
8	0.6135	0.6466	0.4464	0.3002	0.7710	0.2439

Table I.20: SAA-QR results when demand is Uniform

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.4430	0.7711	0.3575	0.6916	0.1536	0.7286
2	0.4442	0.7722	0.3514	0.6797	0.1513	0.7363
3	0.9441	0.9735	0.2677	0.6250	0.4593	0.3197
4	0.9347	0.9735	0.2822	0.6418	0.4865	0.3286
5	0.7161	0.9991	0.5215	0.5146	0.9351	0.7190
6	0.7819	0.9240	0.7232	0.6888	0.9398	0.6717
7	0.9211	0.7165	0.6128	0.9211	0.9784	0.7110
8	0.8977	0.7199	0.6656	0.9043	0.9984	0.7039

Table I.21: MIP-QR results when demand is Uniform

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.4553	0.7741	0.3511	0.6894	0.1547	0.7310
2	0.4762	0.8204	0.3630	0.6853	0.1510	0.7315
3	0.9385	0.9862	0.2797	0.6413	0.4607	0.3130
4	0.9193	0.9789	0.2912	0.6652	0.4855	0.3112
5	0.7016	0.9957	0.5181	0.5107	0.9146	0.7205
6	0.7665	0.9249	0.6296	0.6633	0.9751	0.6638
7	0.9249	0.7161	0.6269	0.9068	0.9769	0.7102
8	0.8986	0.7182	0.6457	0.8916	0.9467	0.6902

Table I.22: SAA-LR results when demand is Uniform

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.6180	0.7657	0.3301	0.2220	0.3523	0.6595
2	0.7101	0.6065	0.7146	0.9310	0.2569	0.3538
3	0.3503	0.5388	0.5038	0.5720	0.7905	0.8065
4	0.1882	0.4694	0.4315	0.6734	0.7144	0.7254
5	0.2535	0.9858	0.5225	0.6547	0.7494	0.9506
6	0.2489	0.9902	0.6199	0.5946	0.9512	0.8648
7	0.7091	0.4369	0.5805	0.9048	0.3827	0.1375
8	0.9930	0.6702	0.3826	0.6439	0.9925	0.1807

Table I.23: MIP-LR results when demand is Uniform

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.2304	0.7146	0.56	0.5358	0.9481	0.8099
2	0.2373	0.6275	0.9229	0.5384	0.7046	0.9026
3	0.9641	0.8293	0.7346	0.866	0.2624	0.2668
4	0.8644	0.6952	0.9366	0.789	0.6351	0.953
5	0.4541	0.373	0.9393	0.785	0.7134	0.6393
6	0.6267	0.1888	0.8354	0.8931	0.3046	0.2431
7	0.6169	0.1665	0.3133	0.8371	0.9619	0.7965
8	0.7369	0.1508	0.7877	0.5181	0.9998	0.6372

Table I.24: MIP-GP results when demand is Uniform

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.4600	0.6757	0.7234	0.9779	0.2118	0.6889
2	0.4558	0.6631	0.7300	0.9736	0.2287	0.7344
3	0.9217	0.9382	0.2929	0.5770	0.5031	0.3265
4	0.9371	0.9319	0.2976	0.5945	0.5906	0.3416
5	0.6292	0.8671	0.3784	0.5649	0.8654	0.7637
6	0.7316	0.9527	0.5178	0.6992	0.9862	0.6642
7	0.9188	0.7683	0.3137	0.9478	0.9486	0.9092
8	0.8873	0.7657	0.3553	0.9454	0.9644	0.9087

Table I.25: MIP-P results when demand is Uniform

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.2258	0.5250	0.1977	0.5678	0.2890	0.1938
2	0.2748	0.6253	0.7665	0.7629	0.9818	0.1785
3	0.7885	0.5731	0.5665	0.2827	0.0508	0.9769
4	0.7173	0.5567	0.7378	0.8750	0.2597	0.5271
5	0.3200	0.8267	0.0945	0.8626	0.1126	0.7015
6	0.3134	0.8701	0.1212	0.9753	0.2627	0.5949
7	0.3968	0.5932	0.6631	0.9858	0.0006	0.8565
8	0.2970	0.5313	0.3175	0.7072	0.0158	0.4192

Table I.26: SAA-P results when demand is Symmetric Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.0971	0.7929	0.6944	0.8914	0.4580	0.9164
2	0.2761	0.2454	0.2282	0.6525	0.5618	0.5047
3	0.8883	0.2465	0.9066	0.2943	0.0179	0.4726
4	0.8229	0.8282	0.4155	0.5493	0.0726	0.1243
5	0.4633	0.5686	0.5665	0.4326	0.6393	0.8605
6	0.9985	0.6390	0.3683	0.7772	0.1541	0.0531
7	0.8442	0.5333	0.7485	0.3978	0.4941	0.7054
8	0.4275	0.5229	0.4855	0.4052	0.3964	0.7822

Table I.27: MIP-SAA results when demand is Symmetric Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.1011	0.5528	0.4577	0.8862	0.7823	0.9774
2	0.4450	0.3415	0.6941	0.8681	0.9668	0.0965
3	0.8871	0.6037	0.2872	0.2825	0.4653	0.7947
4	0.6622	0.7615	0.4369	0.9552	0.1903	0.5425
5	0.3348	0.4937	0.7992	0.6820	0.9756	0.1618
6	0.4155	0.3134	0.3751	0.1072	0.8689	0.7615
7	0.9980	0.4099	0.5652	0.4825	0.2766	0.4468
8	0.6383	0.4177	0.9776	0.8561	0.1029	0.1734

Table I.28: SAA-QR results when demand is Symmetric Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.0679	0.5760	0.3166	0.7139	0.8079	0.8410
2	0.0647	0.5622	0.3103	0.6930	0.8143	0.8366
3	0.3947	0.1976	0.3280	0.5624	0.1665	0.9552
4	0.4600	0.2235	0.4055	0.6089	0.2741	0.9769
5	0.7663	0.4822	0.0538	0.3826	0.0929	0.7536
6	0.7690	0.4767	0.0385	0.3117	0.1015	0.7712
7	0.9119	0.6164	0.4474	0.8007	0.9538	0.7303
8	0.9105	0.6160	0.4517	0.8008	0.8841	0.6768

Table I.29: MIP-QR results when demand is Symmetric Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.0685	0.5886	0.3269	0.7310	0.7754	0.8640
2	0.0652	0.5966	0.3330	0.6859	0.7923	0.8757
3	0.3980	0.1940	0.3232	0.5523	0.1366	0.9975
4	0.4563	0.2280	0.4254	0.5860	0.2774	0.9603
5	0.7595	0.4786	0.0562	0.3919	0.0902	0.7721
6	0.7513	0.4847	0.0358	0.3189	0.0975	0.7832
7	0.9148	0.6237	0.4319	0.7920	0.9473	0.7271
8	0.9373	0.6309	0.4517	0.7631	0.8948	0.6895

Table I.30: SAA-LR results when demand is Symmetric Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.0009	0.0108	0.1915	0.2468	0.6534	0.9354
2	0.0062	0.0468	0.1710	0.1795	0.1272	0.1279
3	0.5788	0.6445	0.1603	0.3226	0.1638	0.3206
4	0.1446	0.8362	0.0926	0.2512	0.1222	0.2491
5	0.7831	0.3898	0.2370	0.6351	0.4206	0.6889
6	0.7417	0.3022	0.7976	0.5660	0.5946	0.6712
7	0.0440	0.5299	0.4116	0.3733	0.3418	0.5679
8	0.0424	0.6087	0.4264	0.3800	0.5091	0.5202

Table I.31: MIP-LR results when demand is Symmetric Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.0022	0.1048	0.4463	0.6047	0.0260	0.0197
2	0.0137	0.0764	0.4204	0.3061	0.1866	0.4201
3	0.7530	0.9065	0.9709	0.6684	0.6531	0.5067
4	0.5757	0.6800	0.6698	0.8571	0.4454	0.9697
5	0.7820	0.8758	0.9884	0.3249	0.5135	0.6350
6	0.9235	0.9065	0.8639	0.4463	0.7109	0.6363
7	0.6399	0.5183	0.5358	0.1173	0.7015	0.2723
8	0.5014	0.8519	0.1115	0.1162	0.7453	0.8557

Table I.32: MIP-GP results when demand is Symmetric Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.0230	0.1810	0.3644	0.8776	0.5508	0.8902
2	0.0240	0.2032	0.2823	0.8311	0.5234	0.9221
3	0.3294	0.2156	0.3645	0.7421	0.2477	0.9564
4	0.3289	0.2006	0.4904	0.8082	0.4209	0.9078
5	0.8255	0.2818	0.0780	0.6121	0.1810	0.6406
6	0.8246	0.2852	0.0499	0.4995	0.1925	0.6549
7	0.7169	0.5501	0.8523	0.4313	0.3495	0.5075
8	0.7231	0.5478	0.8556	0.4397	0.3010	0.4831

Table I.33: MIP-P results when demand is Symmetric Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.2063	0.3675	0.8671	0.3618	0.0131	<0.0001
2	0.1943	0.3855	<0.0001	<0.0001	<0.0001	0.6427
3	0.6306	0.5496	0.4157	0.0534	0.7534	0.0033
4	0.4282	0.6362	0.4977	0.1944	0.0010	0.0022
5	0.2907	0.9490	0.0582	0.1413	0.0009	0.1405
6	0.2796	0.9150	0.2231	0.0204	<0.0001	0.0778
7	0.5314	0.7587	0.0405	0.6469	0.0329	0.3080
8	0.6428	0.6390	0.0320	0.8977	0.4768	0.2183

Table I.34: SAA-P results when demand is Right-Skewed Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.7776	0.8138	0.7141	0.6766	0.2577	0.9637
2	0.9381	0.6940	0.0128	0.0119	0.3467	0.5395
3	0.6758	0.4347	0.9260	0.9158	0.4290	0.7363
4	0.2411	0.5538	0.7306	0.9096	0.1540	0.6175
5	0.6810	0.7083	0.4520	0.3552	0.5420	0.6757
6	0.7360	0.2677	0.0805	0.1560	0.2066	0.3575
7	0.4584	0.3779	0.8668	0.8006	0.6586	0.8808
8	0.2697	0.2833	0.4078	0.3998	0.7770	0.1954

Table I.35: MIP-SAA results when demand is Right-Skewed Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.7516	0.4242	0.8191	0.7342	0.7867	0.4992
2	0.6121	0.3975	0.3301	0.3774	0.7720	0.1383
3	0.5980	0.5779	0.4734	0.0142	0.7719	0.1636
4	0.7744	0.6042	0.4062	0.0022	0.5952	0.8501
5	0.8446	0.6185	0.1563	0.0077	0.2273	0.5036
6	0.5670	0.3935	0.4716	0.7089	0.0906	0.4459
7	0.1547	0.2408	0.0418	0.0396	0.8493	0.7810
8	0.2467	0.2504	0.6939	0.2358	0.6729	0.2642

Table I.36: SAA-QR results when demand is Right-Skewed Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.2245	0.5859	0.8332	0.9625	0.4313	0.8657
2	0.3128	0.6298	0.5897	0.9274	0.4199	0.8986
3	0.6856	0.7335	0.5274	0.9928	0.5968	0.5345
4	0.6962	0.6978	0.5022	0.9962	0.9674	0.3655
5	0.0551	0.9536	0.2836	0.9490	0.3862	0.3124
6	0.0546	0.9374	0.2899	0.9292	0.4047	0.3005
7	0.6313	0.2079	0.1607	0.9942	0.6469	0.4904
8	0.5915	0.2121	0.1673	0.9986	0.5955	0.4763

Table I.37: MIP-QR results when demand is Right-Skewed Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.2256	0.6017	0.8488	0.9431	0.4511	0.8794
2	0.3095	0.6220	0.6574	0.8771	0.4011	0.8819
3	0.6762	0.7304	0.5497	0.9964	0.6217	0.5315
4	0.6958	0.7134	0.5985	0.9703	0.8775	0.3330
5	0.0495	0.9533	0.2795	0.9637	0.3880	0.3068
6	0.0533	0.9324	0.2862	0.9490	0.4132	0.2960
7	0.6279	0.2084	0.1647	0.9885	0.6537	0.4911
8	0.5605	0.2080	0.1780	0.9646	0.5897	0.4771

Table I.38: SAA-LR results when demand is Right-Skewed Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.1660	0.9896	0.4143	0.5153	0.4016	0.2287
2	0.2216	0.4634	0.4375	0.4852	0.4532	0.1518
3	0.5421	0.3225	0.3765	0.7626	0.8597	0.3632
4	0.5438	0.3953	0.3689	0.9492	0.5421	0.2766
5	0.5797	0.6524	0.1225	0.5071	0.6234	0.3758
6	0.5028	0.6803	0.0276	0.2113	0.6076	0.3838
7	0.9547	0.7848	0.4310	0.3022	0.2595	0.9048
8	0.8650	0.8095	0.7257	0.7211	0.1719	0.9583

Table I.39: MIP-LR results when demand is Right-Skewed Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.0379	0.0353	0.6552	0.5857	0.9321	0.7201
2	0.0650	0.0586	0.3621	0.5361	0.5945	0.8854
3	0.0498	0.3124	0.4681	0.7012	0.7625	0.8236
4	0.0547	0.3217	0.4484	0.4857	0.9404	0.8499
5	0.5590	0.4090	0.2294	0.5408	0.1951	0.4367
6	0.9268	0.9832	0.3960	0.5156	0.7554	0.7317
7	0.4691	0.3932	0.1682	0.2787	0.4385	0.7866
8	0.3700	0.3727	0.6234	0.2322	0.4173	0.9502

Table I.40: MIP-GP results when demand is Right-Skewed Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.4152	0.5921	0.2359	0.6279	0.3438	0.8851
2	0.4539	0.5566	0.1451	0.5313	0.2167	0.7560
3	0.7222	0.9121	0.5244	0.9817	0.6002	0.6817
4	0.7192	0.8618	0.4092	0.8785	0.5674	0.7454
5	0.0927	0.4111	0.4076	0.9185	0.5563	0.4499
6	0.0656	0.3433	0.4889	0.8742	0.5850	0.4884
7	0.4989	0.1548	0.6068	0.8292	0.4157	0.2227
8	0.4937	0.1521	0.5363	0.8570	0.4193	0.2265

Table I.41: MIP-P results when demand is Right-Skewed Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.9125	0.6691	0.3088	0.8143	0.7288	0.0054
2	0.6972	0.5451	0.7334	0.5087	0.4154	0.0551
3	0.4494	0.5072	0.9898	0.8074	0.8321	0.1882
4	0.3088	0.4840	0.7986	0.2304	0.4280	0.0690
5	0.8877	0.5045	0.9645	0.6778	0.2801	0.1595
6	0.9232	0.4962	0.5771	0.4455	0.1394	0.5925
7	0.2830	0.2188	0.8986	0.2033	0.0062	0.3438
8	0.2791	0.1951	0.2339	0.2972	0.0685	0.1751

Table I.42: SAA-P results when demand is Left-Skewed Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.0019	<0.0001	0.1298	<0.0001	0.7715	<0.0001
2	0.0038	<0.0001	0.7626	<0.0001	0.9837	<0.0001
3	0.2979	<0.0001	0.6053	<0.0001	0.1794	<0.0001
4	0.2270	<0.0001	0.9015	<0.0001	0.2946	<0.0001
5	0.8310	<0.0001	0.4436	<0.0001	0.0459	<0.0001
6	0.8530	<0.0001	0.8880	<0.0001	0.3025	<0.0001
7	0.8337	<0.0001	0.5666	<0.0001	0.2146	<0.0001
8	0.7361	<0.0001	0.1547	<0.0001	0.8902	<0.0001

Table I.43: MIP-SAA results when demand is Left-Skewed Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.1397	<0.0001	0.1476	<0.0001	0.1373	<0.0001
2	0.0367	<0.0001	0.6074	<0.0001	0.3374	<0.0001
3	0.6396	<0.0001	0.3273	<0.0001	0.1278	<0.0001
4	0.4409	<0.0001	0.3536	<0.0001	0.0767	0.0038
5	0.2913	<0.0001	0.8285	<0.0001	0.5711	0.0003
6	0.5956	<0.0001	0.2477	<0.0001	0.2759	<0.0001
7	0.5930	<0.0001	0.6927	<0.0001	0.9762	<0.0001
8	0.4091	<0.0001	0.1089	<0.0001	0.2780	<0.0001

Table I.44: SAA-QR results when demand is Left-Skewed Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.1770	<0.0001	0.2725	<0.0001	0.8767	<0.0001
2	0.2099	<0.0001	0.1970	<0.0001	0.9810	<0.0001
3	0.5275	<0.0001	0.7353	<0.0001	0.1556	<0.0001
4	0.5041	<0.0001	0.7563	<0.0001	0.3451	<0.0001
5	0.6257	<0.0001	0.6121	<0.0001	0.6248	<0.0001
6	0.6695	<0.0001	0.6494	<0.0001	0.6701	<0.0001
7	0.4874	<0.0001	0.7201	<0.0001	0.6573	<0.0001
8	0.5247	<0.0001	0.8068	<0.0001	0.7399	<0.0001

Table I.45: MIP-QR results when demand is Left-Skewed Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.1751	<0.0001	0.2599	<0.0001	0.8697	<0.0001
2	0.2130	<0.0001	0.1823	<0.0001	0.8968	<0.0001
3	0.5266	<0.0001	0.7227	<0.0001	0.1424	<0.0001
4	0.5041	<0.0001	0.7344	<0.0001	0.3231	<0.0001
5	0.6209	<0.0001	0.5663	<0.0001	0.6345	<0.0001
6	0.6499	<0.0001	0.5312	<0.0001	0.6614	<0.0001
7	0.4993	<0.0001	0.7249	<0.0001	0.6627	<0.0001
8	0.5636	<0.0001	0.7674	<0.0001	0.7194	<0.0001

Table I.46: SAA-LR results when demand is Left-Skewed Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.1000	<0.0001	0.2765	<0.0001	0.8977	<0.0001
2	0.0050	<0.0001	0.0144	<0.0001	0.8113	<0.0001
3	0.1926	<0.0001	0.8959	<0.0001	0.8379	<0.0001
4	0.1828	<0.0001	0.9803	<0.0001	0.1382	<0.0001
5	0.7140	<0.0001	0.5713	<0.0001	0.3330	<0.0001
6	0.8123	<0.0001	0.5790	<0.0001	0.2701	<0.0001
7	0.5990	<0.0001	0.9015	<0.0001	0.6417	<0.0001
8	0.9517	<0.0001	0.8894	<0.0001	0.8394	<0.0001

Table I.47: MIP-LR results when demand is Left-Skewed Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.1916	<0.0001	0.5391	<0.0001	0.3527	<0.0001
2	0.0982	<0.0001	0.1910	<0.0001	0.1094	<0.0001
3	0.0129	<0.0001	0.4252	<0.0001	0.3754	<0.0001
4	0.0310	<0.0001	0.2961	<0.0001	0.3017	<0.0001
5	0.7717	<0.0001	0.4517	<0.0001	0.8306	<0.0001
6	0.8341	<0.0001	0.4340	<0.0001	0.6055	<0.0001
7	0.4790	<0.0001	0.5539	<0.0001	0.6354	<0.0001
8	0.6868	<0.0001	0.6659	<0.0001	0.1827	<0.0001

Table I.48: MIP-GP results when demand is Left-Skewed Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.0803	<0.0001	0.2420	<0.0001	0.5865	<0.0001
2	0.1372	<0.0001	0.1245	<0.0001	0.6822	<0.0001
3	0.6182	<0.0001	0.5885	<0.0001	0.2691	<0.0001
4	0.5748	<0.0001	0.7007	<0.0001	0.2279	<0.0001
5	0.7032	<0.0001	0.6931	<0.0001	0.3270	<0.0001
6	0.7226	<0.0001	0.6165	<0.0001	0.2378	<0.0001
7	0.8495	<0.0001	0.9251	<0.0001	0.3876	<0.0001
8	0.8719	<0.0001	0.9809	<0.0001	0.3969	<0.0001

Table I.49: MIP-P results when demand is Left-Skewed Triangular

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.0574	<0.0001	0.5519	<0.0001	0.0305	0.0513
2	0.0557	<0.0001	0.0555	<0.0001	<0.0001	<0.0001
3	0.7484	<0.0001	0.0442	<0.0001	<0.0001	0.1564
4	0.6867	<0.0001	0.9470	<0.0001	0.2131	0.0685
5	0.3938	<0.0001	0.0376	<0.0001	0.0002	<0.0001
6	0.2704	<0.0001	0.5795	<0.0001	<0.0001	<0.0001
7	0.9023	<0.0001	0.0003	<0.0001	0.9598	<0.0001
8	0.9776	<0.0001	0.6198	<0.0001	0.5082	<0.0001

Table I.50: SAA-P results when demand is the combination of Symmetric Triangular and Uniform

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.0479	0.2177	0.3432	0.0531	0.7888	0.1423
2	0.7020	0.9706	0.4492	0.2360	0.7304	0.1834
3	0.2360	0.4681	0.3807	0.2760	0.6880	0.5806
4	0.5346	0.8506	0.5276	0.3766	0.9655	0.7586
5	0.8298	0.2691	0.0628	0.1123	0.4693	0.0900
6	0.6725	0.8871	0.7873	0.6754	0.6028	0.8859
7	0.5326	0.5874	0.6002	0.7547	0.7971	0.9591
8	0.5225	0.0706	0.6522	0.8488	0.8718	0.6330

Table I.51: MIP-SAA results when demand is the combination of Symmetric Triangular and Uniform

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.6875	0.4630	0.3201	0.5456	0.7534	0.4965
2	0.7551	0.5254	0.0838	0.0877	0.5496	0.6424
3	0.1994	0.8550	0.8806	0.5202	0.0750	0.0217
4	0.4439	0.6431	0.5572	0.4729	0.7187	0.9455
5	0.1483	0.9878	0.6101	0.4314	0.1302	0.3759
6	0.1336	0.9492	0.3202	0.8708	0.9688	0.7716
7	0.1425	0.3570	0.5111	0.7346	0.1224	0.4733
8	0.0675	0.0367	0.3836	0.2797	0.9724	0.3478

Table I.52: SAA-QR results when demand is the combination of Symmetric Triangular and Uniform

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.9150	0.8338	0.9809	0.4354	0.4645	0.9530
2	0.9268	0.8230	0.9959	0.4494	0.4451	0.9856
3	0.8473	0.5008	0.3921	0.6866	0.8272	0.6680
4	0.8402	0.4997	0.3863	0.6750	0.8278	0.6834
5	0.9478	0.9888	0.5841	0.7016	0.0241	0.2755
6	0.9060	0.9789	0.8530	0.5127	0.0494	0.3575
7	0.0833	0.2689	0.6345	0.1425	0.8198	0.9219
8	0.0877	0.2680	0.6611	0.1475	0.8059	0.9212

Table I.53: MIP-QR results when demand is the combination of Symmetric Triangular and Uniform

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.9258	0.8378	0.9653	0.4318	0.4726	0.9559
2	0.9343	0.8223	0.9899	0.4755	0.4453	0.9698
3	0.8415	0.5021	0.3963	0.6772	0.8258	0.6797
4	0.8429	0.4901	0.4227	0.7035	0.8398	0.6979
5	0.9485	0.9955	0.5714	0.7073	0.0230	0.2762
6	0.9083	0.9728	0.8623	0.4873	0.0509	0.3669
7	0.0799	0.2721	0.6436	0.1524	0.8161	0.9234
8	0.0885	0.2668	0.5602	0.1213	0.7982	0.9242

Table I.54: SAA-LR results when demand is the combination of Symmetric Triangular and Uniform

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.0029	0.0265	0.8968	0.7684	0.2955	0.5420
2	0.0040	0.0645	0.8131	0.4921	0.8620	0.9692
3	0.5070	0.5214	0.9700	0.8163	0.3181	0.7591
4	0.9729	0.8930	0.8421	0.7616	0.2731	0.5922
5	0.9021	0.7645	0.2419	0.1931	0.2642	0.5491
6	0.5800	0.9331	0.2585	0.1825	0.4495	0.9309
7	0.4959	0.8478	0.2099	0.5318	0.6568	0.9404
8	0.5581	0.8520	0.3240	0.5825	0.6462	0.9773

Table I.55: MIP-LR results when demand is the combination of Symmetric Triangular and Uniform

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.0966	0.1760	0.7239	0.7044	0.7545	0.9993
2	0.0884	0.1488	0.9616	0.9504	0.6877	0.8083
3	0.3766	0.7619	0.4270	0.9665	0.4648	0.8689
4	0.4675	0.9166	0.0572	0.4227	0.3661	0.8953
5	0.2100	0.1929	0.0172	0.0290	0.0315	0.3217
6	0.0544	0.0544	0.8972	0.7450	0.0829	0.1094
7	0.9521	0.6595	0.3595	0.4485	0.3670	0.1368
8	0.8970	0.7036	0.3446	0.7867	0.1710	0.0970

Table I.56: MIP-GP results when demand is the combination of Symmetric Triangular and Uniform

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.7994	0.7594	0.9934	0.4941	0.6262	0.8546
2	0.8049	0.7696	0.9596	0.4821	0.5599	0.9449
3	0.9102	0.4773	0.3278	0.7602	0.7783	0.6792
4	0.8974	0.5015	0.2673	0.7115	0.8774	0.6720
5	0.8976	0.6248	0.6102	0.7115	0.0396	0.2822
6	0.9598	0.5942	0.7801	0.6668	0.0690	0.3230
7	0.0566	0.1956	0.3997	0.0666	0.8571	0.8923
8	0.0595	0.1992	0.4283	0.0737	0.8496	0.9186

Table I.57: MIP-P results when demand is the combination of Symmetric Triangular and Uniform

	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3	significance level for 1-2	significance level for 1-3
1	0.8788	0.7681	0.9730	0.9072	0.5152	0.0153
2	0.8178	0.7233	0.2475	0.9761	0.2329	0.8410
3	0.1147	0.6767	0.0993	0.0000	0.6883	0.7648
4	0.1342	0.6825	0.3821	0.0000	0.0043	0.8342
5	0.3454	0.9691	0.0037	0.0891	0.0152	0.0005
6	0.2097	0.8146	0.0009	0.0007	0.0060	0.0007
7	0.1794	0.3905	0.0640	0.5283	0.1152	0.3298
8	0.1964	0.3449	0.5593	0.0773	0.8341	0.8287

I.5 Comparison of Methods

We conduct one sample t test for the data set obtained for $j = 1, \ldots, 50$ for comparison of the best model with the worst one. The purpose of each test is to statistically check whether the defined difference is significantly different than zero. We record for each case: if the result is within 90% CI set 1, otherwise 0, and if the result is within 99% CI set 1, otherwise 0. This test is conducted for all parameter settings given in Table 4.4 and for all demand distributions selected.

Table I.58: Comparison of the best model with the worst one for Poisson demand

	sum of 0's					
	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	# with 90% CI	# with 99% CI	# with 90% CI	# with 99% CI	# with 90% CI	# with 99% CI
1	49	49	50	50	50	50
2	49	49	50	50	50	50
3	49	49	50	50	50	50
4	50	50	50	50	50	50
5	49	49	50	50	50	50
6	50	49	50	50	50	50
7	50	50	50	50	50	50
8	50	50	50	50	50	50

Table I.59: Comparison of the best model with the worst one for Uniform demand

	sum of 0's					
	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	# with 90% CI	# with 99% CI	# with 90% CI	# with 99% CI	# with 90% CI	# with 99% CI
1	50	49	50	50	50	50
2	50	50	50	50	50	50
3	50	50	50	50	50	50
4	50	50	50	50	50	50
5	50	49	50	50	50	50
6	50	49	50	50	50	50
7	49	45	50	50	50	50
8	50	46	50	50	50	50

Table I.60: Comparison of the best model with the worst one for Triangular distributed demand

	sum of 0's					
	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	# with 90% CI	# with 99% CI	# with 90% CI	# with 99% CI	# with 90% CI	# with 99% CI
1	50	50	50	50	50	50
2	50	50	50	50	50	50
3	50	50	50	50	50	50
4	50	50	50	50	50	50
5	50	50	50	50	50	50
6	50	50	50	50	50	50
7	49	49	50	50	50	50
8	50	49	50	50	50	50

Table I.61: Comparison of the best model with the worst one for Right-Skewed Triangular distributed demand

	sum of 0's					
	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	# with 90% CI	# with 99% CI	# with 90% CI	# with 99% CI	# with 90% CI	# with 99% CI
1	50	50	50	50	50	50
2	50	50	50	50	50	50
3	50	50	50	50	50	50
4	50	50	50	50	50	50
5	50	50	50	50	50	50
6	50	50	50	50	50	50
7	50	46	50	50	50	50
8	50	47	50	50	50	50

Table I.62: Comparison of the best model with the worst one for Left-Skewed Triangular distributed demand

	sum of 0's					
	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	# with 90% CI	# with 99% CI	# with 90% CI	# with 99% CI	# with 90% CI	# with 99% CI
1	50	50	50	50	50	50
2	50	50	50	50	50	50
3	50	50	50	50	50	50
4	50	50	50	50	50	50
5	50	50	50	50	50	50
6	50	50	50	50	50	50
7	50	50	50	50	50	50
8	50	50	50	50	50	50

Table I.63: Comparison of the best model with the worst one for the combination of Symmetric Triangular distributed demand and Uniform distributed demand

	sum of 0's					
	Train Data Size = 50		Train Data Size =250		Train Data Size = 450	
Data/ Parameter Setting	# with 90% CI	# with 99% CI	# with 90% CI	# with 99% CI	# with 90% CI	# with 99% CI
1	50	50	50	50	50	50
2	50	50	50	50	50	50
3	50	50	50	50	50	50
4	50	50	50	50	50	50
5	50	50	50	50	50	50
6	50	50	50	50	50	50
7	50	50	50	50	50	50
8	50	50	50	50	50	50

I.6 The Difference From The Expected Profit

We compute the percentage difference between the expected profit under the assumption that the demand is known beforehand and the profit achieved by data-driven models when the demand is unknown. This test is conducted for all parameter settings given in Table 4.4 and for all demand distributions selected. For each model in Table 3.1, results are obtained below. The results for MIP-P and MIP-GP are also shown below. MIP-P is the model where mixed integer program is used to solve both stages to find two decisions simultaneously. Lastly, MIP-GP is a simple heuristic where we run a mixed integer program for the high marked-down price as well as the low marked-down price and select the price and order quantity combination that yields to the larger average total profit prior to the beginning of the first stage.

Table I.64: The difference from the expected profit for Poisson demand

Data/Parameter Setting	Train Data Set Size	% difference from the expected value - 100%E-P(k)/E							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	50	0.07	0.07	1.52	1.53	0.24	0.88	0.18	0.07
	250	0.04	0.05	1.95	1.95	0.31	0.84	0.12	0.04
	450	0.03	0.03	1.89	1.89	0.27	0.82	0.11	0.03
2	50	1.02	1.00	2.45	2.47	1.68	1.61	1.76	1.02
	250	0.61	0.61	2.22	2.22	1.08	1.13	1.30	0.61
	450	0.59	0.59	2.18	2.18	1.12	1.23	1.26	0.59
3	50	0.23	0.23	2.51	2.51	0.62	1.13	0.51	0.23
	250	0.02	0.02	2.46	2.46	0.43	0.98	0.30	0.02
	450	0.00	0.00	2.43	2.43	0.36	0.98	0.24	0.00
4	50	1.36	1.35	3.05	3.05	2.07	2.18	2.62	1.36
	250	0.85	0.85	2.67	2.67	1.45	1.42	2.02	0.85
	450	0.80	0.79	2.62	2.62	1.57	1.60	2.10	0.80
5	50	0.16	0.17	0.87	0.87	0.23	0.32	0.20	0.16
	250	0.07	0.07	1.28	1.28	0.15	0.24	0.10	0.07
	450	0.06	0.06	1.29	1.29	0.14	0.24	0.08	0.06
6	50	0.76	0.76	1.60	1.60	0.94	0.94	1.04	0.76
	250	0.60	0.59	1.57	1.57	0.75	0.74	0.82	0.60
	450	0.52	0.53	1.54	1.54	0.75	0.74	0.80	0.52
7	50	0.26	0.27	1.45	1.46	0.49	0.50	0.37	0.26
	250	0.11	0.11	1.56	1.56	0.28	0.37	0.23	0.11
	450	0.08	0.07	1.53	1.53	0.26	0.33	0.21	0.08
8	50	0.97	0.97	1.98	1.98	1.26	1.28	1.63	0.97
	250	0.74	0.74	1.79	1.79	1.04	1.06	1.35	0.74
	450	0.69	0.68	1.75	1.75	1.03	1.03	1.33	0.69

Table I.65: The difference from the expected profit for Uniform demand

Data/Parameter Setting	Train Data Set Size	% difference from the expected value - 100%E-P(k)/E							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	50	2.62	2.63	3.25	3.25	4.06	4.59	3.21	2.62
	250	2.09	2.02	2.57	2.57	2.70	3.24	2.57	2.60
	450	1.71	1.42	2.40	2.40	2.12	2.82	2.39	2.82
2	50	2.71	2.65	3.25	3.26	4.07	4.61	3.20	2.66
	250	2.15	2.01	2.57	2.57	2.63	3.22	2.55	2.55
	450	1.71	1.68	2.40	2.40	2.15	2.80	2.38	2.30
3	50	6.60	6.72	8.01	8.01	10.23	9.86	7.85	6.73
	250	4.02	3.95	5.17	5.17	5.67	6.13	5.19	4.96
	450	3.87	4.48	4.84	4.84	4.62	5.55	4.84	8.16
4	50	6.40	6.70	8.01	8.01	10.49	10.02	7.83	6.76
	250	4.16	4.07	5.17	5.17	5.74	6.15	5.17	5.36
	450	3.95	3.67	4.84	4.84	4.63	5.40	4.82	6.06
5	50	1.72	1.72	2.01	2.01	2.14	2.22	2.00	1.74
	250	0.93	1.02	1.15	1.15	1.19	1.17	1.15	1.20
	450	0.96	1.09	1.19	1.19	1.05	1.06	1.19	1.21
6	50	1.72	1.73	2.03	2.03	2.15	2.24	2.01	1.74
	250	0.93	0.89	1.16	1.16	1.20	1.17	1.15	1.15
	450	0.92	1.04	1.19	1.19	1.06	1.05	1.18	1.18
7	50	2.88	2.92	3.22	3.22	3.01	3.27	3.25	2.98
	250	1.97	2.19	2.38	2.38	2.28	2.28	2.37	2.28
	450	1.96	2.50	2.26	2.26	2.23	2.18	2.26	2.22
8	50	2.89	2.93	3.22	3.22	3.11	3.29	3.25	3.01
	250	1.99	2.12	2.39	2.39	2.28	2.26	2.37	2.24
	450	1.93	2.17	2.26	2.26	2.23	2.17	2.26	2.32

Table I.66: The difference from the expected profit for Triangular distributed demand

Data/Parameter Setting	Train Data Set Size	% difference from the expected value - 100%E-P(k)/E							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	50	1.14	1.12	2.04	2.04	2.07	2.87	2.00	1.13
	250	1.06	1.11	1.97	1.97	1.80	2.54	1.95	1.90
	450	1.00	1.26	1.85	1.85	1.79	2.39	1.85	8.15
2	50	1.08	1.13	2.03	2.04	2.12	2.84	1.99	1.14
	250	1.12	1.16	1.97	1.97	1.82	2.55	1.92	13.71
	450	1.01	1.34	1.85	1.85	1.78	2.42	1.82	7.57
3	50	2.50	2.39	3.61	3.61	3.81	4.57	3.62	2.39
	250	1.54	2.19	2.83	2.83	2.74	3.68	2.84	2.20
	450	1.48	3.68	2.70	2.70	2.61	3.58	2.69	7.68
4	50	2.44	2.35	3.58	3.58	3.71	4.55	3.57	2.36
	250	1.56	1.87	2.82	2.82	2.72	3.69	2.80	2.21
	450	1.45	3.65	2.68	2.68	2.61	3.57	2.65	7.86
5	50	0.93	0.90	1.41	1.41	1.30	1.46	1.42	0.88
	250	0.55	0.79	1.05	1.05	0.87	1.05	1.04	0.87
	450	0.55	1.05	1.07	1.07	0.88	1.00	1.07	0.88
6	50	0.91	0.89	1.41	1.41	1.31	1.45	1.41	0.88
	250	0.55	0.62	1.05	1.05	0.88	1.04	1.03	0.90
	450	0.58	0.96	1.07	1.07	0.87	1.03	1.06	0.90
7	50	1.70	1.68	2.16	2.16	2.01	2.33	2.15	1.72
	250	1.28	1.57	1.67	1.67	1.62	1.78	1.66	1.55
	450	1.22	2.90	1.61	1.61	1.61	1.70	1.60	1.63
8	50	1.73	1.70	2.15	2.15	2.02	2.34	2.15	1.71
	250	1.26	1.89	1.66	1.66	1.63	1.80	1.65	1.56
	450	1.22	3.87	1.61	1.61	1.62	1.71	1.59	1.73

Table I.67: The difference from the expected profit for Right-Skewed Triangular distributed demand

Data/Parameter Setting	Train Data Set Size	% difference from the expected value - 100%E-P(k)/E							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	50	2.69	2.65	3.36	3.36	3.83	4.64	3.35	2.69
	250	2.36	2.71	3.13	3.12	3.19	4.17	3.13	3.06
	450	2.19	2.33	2.96	2.96	3.00	4.00	2.97	4.05
2	50	2.73	2.63	3.38	3.38	3.86	4.74	3.33	2.67
	250	2.46	2.48	3.13	3.13	3.19	4.18	3.12	2.96
	450	2.19	2.28	2.96	2.96	3.00	4.03	2.96	2.79
3	50	3.97	3.97	5.12	5.11	6.38	7.03	5.02	3.98
	250	2.67	2.74	3.65	3.65	3.98	5.40	3.62	3.37
	450	2.47	3.29	3.36	3.36	3.53	5.21	3.33	5.44
4	50	4.29	3.95	5.12	5.11	6.37	7.04	4.99	3.95
	250	2.61	2.54	3.63	3.63	3.98	5.39	3.57	3.33
	450	2.46	2.78	3.32	3.32	3.50	5.24	3.30	5.46
5	50	1.22	1.21	1.56	1.56	1.81	1.97	1.51	1.22
	250	0.73	0.69	1.02	1.02	1.05	1.22	1.02	0.98
	450	0.73	0.75	1.02	1.02	0.98	1.14	1.02	1.17
6	50	1.26	1.22	1.56	1.56	1.72	1.93	1.50	1.22
	250	0.71	0.85	1.03	1.03	1.05	1.21	1.02	0.98
	450	0.72	0.66	1.02	1.02	0.99	1.15	1.01	1.23
7	50	2.59	2.56	2.88	2.88	3.25	3.48	2.86	2.57
	250	1.98	2.33	2.28	2.28	2.39	2.47	2.28	2.27
	450	1.93	2.40	2.19	2.19	2.30	2.38	2.18	2.19
8	50	2.60	2.58	2.88	2.88	3.27	3.49	2.85	2.57
	250	1.97	1.99	2.28	2.28	2.41	2.49	2.28	2.21
	450	1.92	2.53	2.19	2.19	2.31	2.40	2.18	2.23

Table I.68: The difference from the expected profit for Left-Skewed Triangular distributed demand

Data/Parameter Setting	Train Data Set Size	% difference from the expected value - 100%E-P(k)/E							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	50	1.06	1.13	2.01	2.01	1.94	2.48	1.97	1.10
	250	1.09	0.88	1.94	1.94	1.68	2.15	1.92	1.77
	450	0.95	1.10	1.82	1.82	1.57	1.99	1.82	8.35
2	50	1.07	1.08	2.01	2.00	1.92	2.47	1.95	1.10
	250	1.08	1.02	1.94	1.94	1.69	2.15	1.90	1.63
	450	0.96	1.05	1.81	1.82	1.58	2.00	1.79	22.37
3	50	1.97	2.01	3.12	3.12	2.98	3.62	3.11	2.03
	250	1.33	2.30	2.46	2.46	2.27	2.70	2.47	1.84
	450	1.27	2.81	2.37	2.37	1.94	2.61	2.35	9.54
4	50	2.01	1.98	3.08	3.08	2.94	3.64	3.05	2.00
	250	1.30	1.45	2.44	2.44	2.25	2.70	2.42	2.04
	450	1.20	3.83	2.33	2.33	1.94	2.63	2.30	8.61
5	50	0.81	0.81	1.30	1.30	1.12	1.20	1.30	0.81
	250	0.52	0.89	1.01	1.01	0.72	0.80	1.01	0.81
	450	0.53	1.24	1.03	1.03	0.76	0.77	1.03	0.84
6	50	0.80	0.79	1.29	1.29	1.13	1.20	1.28	0.79
	250	0.50	0.69	1.00	1.01	0.73	0.78	0.99	0.86
	450	0.51	1.15	1.03	1.03	0.77	0.78	1.01	0.83
7	50	1.20	1.18	1.69	1.69	1.41	1.62	1.72	1.20
	250	0.86	1.71	1.43	1.43	1.04	1.20	1.43	1.12
	450	0.82	1.92	1.38	1.38	1.03	1.16	1.38	1.44
8	50	1.21	1.18	1.69	1.70	1.40	1.63	1.71	1.20
	250	0.85	1.04	1.43	1.43	1.06	1.23	1.42	1.24
	450	0.83	1.96	1.38	1.38	1.04	1.17	1.38	1.52

Table I.69: The difference from the expected profit for Symmetric Triangular distributed and Uniform distributed demand

Data/Parameter Setting	Train Data Set Size	% difference from the expected value - 100%E-P(k)/E							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	50	1.15	1.25	2.04	2.04	1.96	2.66	2.06	1.28
	250	1.05	1.27	1.80	1.80	1.51	2.13	1.80	1.72
	450	1.00	1.21	1.72	1.72	1.59	2.01	1.70	3.88
2	50	1.29	1.25	2.04	2.04	2.01	2.67	2.05	1.28
	250	1.09	1.03	1.80	1.80	1.53	2.16	1.78	1.60
	450	1.04	1.04	1.72	1.72	1.60	2.02	1.67	2.21
3	50	2.55	2.51	3.57	3.57	3.89	4.10	3.60	2.47
	250	1.59	1.92	2.78	2.78	2.36	3.19	2.77	2.55
	450	1.49	1.65	2.61	2.61	2.20	3.08	2.60	4.44
4	50	2.57	2.57	3.57	3.57	3.68	4.18	3.56	2.48
	250	1.65	1.70	2.77	2.77	2.35	3.25	2.75	2.33
	450	1.49	2.41	2.60	2.60	2.23	3.11	2.57	4.30
5	50	0.90	0.89	1.39	1.39	1.23	1.37	1.39	0.89
	250	0.54	0.70	1.03	1.03	0.87	0.95	1.03	0.87
	450	0.52	1.49	1.05	1.05	0.76	0.86	1.05	0.85
6	50	0.89	0.87	1.39	1.39	1.23	1.40	1.38	0.88
	250	0.50	0.54	1.02	1.02	0.88	0.91	1.01	0.80
	450	0.53	1.20	1.05	1.05	0.76	0.89	1.04	0.83
7	50	1.69	1.69	2.15	2.15	2.08	2.28	2.15	1.67
	250	1.29	1.82	1.68	1.68	1.58	1.68	1.68	1.58
	450	1.27	2.44	1.64	1.64	1.55	1.57	1.64	1.64
8	50	1.64	1.66	2.15	2.15	2.09	2.30	2.14	1.67
	250	1.32	1.59	1.68	1.68	1.59	1.70	1.67	1.64
	450	1.29	2.68	1.64	1.64	1.56	1.58	1.63	1.71

I.7 The Inverse of Coefficient of Variation for The Difference

We define the inverse of coefficient of variation for the difference whose formulation is as follows.

$$CoVd(k) = \frac{E - P(k)}{StD(k)} \quad \text{for all method } k$$

The results are obtained for all parameter settings given in Table 4.4 and for all demand distributions selected as well as for each model in Table 3.1. The results for MIP-P and MIP-GP are also shown below. MIP-P is the model where mixed integer program is used to solve both stages to find two decisions simultaneously. Lastly, MIP-GP is a simple heuristic where we run a mixed integer program for the high marked-down price as well as the low marked-down price and select the price and order quantity combination that yields to the larger average total profit prior to the beginning of the first stage.

Table I.70: CovD results for Poisson demand

Data/Parameter Setting	Train Data Set Size	CoVd (Inverse of Coefficient of Variation for the difference)							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	50	0.04	0.03	0.72	0.72	0.12	0.55	0.09	0.04
	250	0.02	0.02	0.96	0.96	0.15	0.61	0.06	0.02
	450	0.01	0.01	0.93	0.93	0.13	0.57	0.05	0.01
2	50	0.29	0.28	0.72	0.73	0.45	0.49	0.47	0.29
	250	0.18	0.18	0.66	0.66	0.32	0.36	0.36	0.18
	450	0.17	0.17	0.65	0.65	0.33	0.41	0.34	0.17
3	50	0.10	0.10	1.02	1.02	0.29	0.67	0.22	0.10
	250	0.01	0.01	1.10	1.10	0.21	0.56	0.14	0.01
	450	0.00	0.00	1.07	1.07	0.18	0.56	0.11	0.00
4	50	0.35	0.34	0.80	0.80	0.53	0.59	0.63	0.35
	250	0.22	0.22	0.72	0.72	0.39	0.42	0.50	0.22
	450	0.21	0.21	0.72	0.72	0.44	0.48	0.50	0.21
5	50	0.08	0.08	0.43	0.43	0.13	0.19	0.10	0.08
	250	0.04	0.04	0.65	0.65	0.08	0.15	0.05	0.04
	450	0.03	0.03	0.65	0.65	0.08	0.14	0.04	0.03
6	50	0.25	0.25	0.53	0.53	0.32	0.33	0.34	0.25
	250	0.20	0.20	0.53	0.53	0.26	0.26	0.27	0.20
	450	0.18	0.18	0.52	0.52	0.26	0.26	0.27	0.18
7	50	0.13	0.13	0.69	0.70	0.24	0.26	0.18	0.13
	250	0.06	0.06	0.77	0.77	0.15	0.20	0.12	0.06
	450	0.04	0.04	0.76	0.76	0.13	0.19	0.11	0.04
8	50	0.29	0.29	0.59	0.59	0.39	0.40	0.49	0.29
	250	0.23	0.23	0.55	0.55	0.32	0.33	0.40	0.23
	450	0.21	0.21	0.55	0.55	0.32	0.33	0.41	0.21

Table I.71: CovD results for Uniform demand

Data/Parameter Setting	Train Data Set Size	CoVd (Coefficient of Variation for the difference)							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	50	0.12	0.12	0.15	0.15	0.19	0.24	0.15	0.12
	250	0.09	0.09	0.11	0.11	0.12	0.17	0.11	0.12
	450	0.07	0.06	0.10	0.10	0.10	0.14	0.10	0.13
2	50	0.13	0.12	0.15	0.15	0.19	0.24	0.15	0.12
	250	0.09	0.09	0.11	0.11	0.12	0.17	0.11	0.11
	450	0.07	0.07	0.10	0.10	0.10	0.14	0.10	0.11
3	50	0.16	0.16	0.19	0.19	0.24	0.24	0.19	0.16
	250	0.09	0.09	0.12	0.12	0.13	0.15	0.12	0.11
	450	0.09	0.10	0.11	0.11	0.10	0.14	0.11	0.18
4	50	0.15	0.16	0.19	0.19	0.25	0.25	0.19	0.16
	250	0.09	0.09	0.12	0.12	0.13	0.15	0.12	0.12
	450	0.09	0.08	0.11	0.11	0.10	0.13	0.11	0.14
5	50	0.11	0.11	0.13	0.13	0.13	0.14	0.12	0.11
	250	0.05	0.06	0.07	0.07	0.07	0.07	0.07	0.07
	450	0.06	0.07	0.07	0.07	0.06	0.06	0.07	0.07
6	50	0.11	0.11	0.13	0.13	0.13	0.14	0.13	0.11
	250	0.05	0.05	0.07	0.07	0.07	0.07	0.07	0.07
	450	0.05	0.06	0.07	0.07	0.06	0.06	0.07	0.07
7	50	0.13	0.13	0.15	0.15	0.13	0.15	0.15	0.14
	250	0.09	0.10	0.11	0.11	0.10	0.10	0.11	0.10
	450	0.09	0.11	0.10	0.10	0.10	0.10	0.10	0.10
8	50	0.13	0.13	0.15	0.15	0.14	0.15	0.15	0.14
	250	0.09	0.10	0.11	0.11	0.10	0.10	0.11	0.10
	450	0.09	0.10	0.10	0.10	0.10	0.10	0.10	0.10

Table I.72: CovD results for Triangular distributed demand

Data/Parameter Setting	Train Data Set Size	CoVd (Inverse of Coefficient of Variation for the difference)							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	50	0.10	0.10	0.18	0.18	0.18	0.29	0.18	0.10
	250	0.09	0.10	0.17	0.17	0.16	0.26	0.16	0.17
	450	0.08	0.11	0.16	0.16	0.15	0.24	0.16	0.59
2	50	0.10	0.10	0.18	0.18	0.19	0.29	0.18	0.10
	250	0.09	0.10	0.17	0.17	0.16	0.26	0.16	2.12
	450	0.08	0.12	0.16	0.16	0.15	0.24	0.15	0.49
3	50	0.17	0.16	0.24	0.24	0.26	0.34	0.24	0.16
	250	0.10	0.14	0.18	0.18	0.18	0.27	0.18	0.14
	450	0.09	0.25	0.17	0.17	0.17	0.26	0.17	0.45
4	50	0.16	0.16	0.24	0.24	0.25	0.33	0.24	0.16
	250	0.10	0.13	0.18	0.18	0.18	0.27	0.18	0.14
	450	0.09	0.25	0.17	0.17	0.17	0.26	0.17	0.48
5	50	0.10	0.10	0.15	0.15	0.15	0.17	0.15	0.09
	250	0.06	0.09	0.11	0.11	0.09	0.12	0.11	0.09
	450	0.06	0.12	0.11	0.11	0.09	0.11	0.11	0.09
6	50	0.10	0.10	0.15	0.15	0.15	0.17	0.15	0.09
	250	0.06	0.07	0.11	0.11	0.09	0.12	0.11	0.10
	450	0.06	0.11	0.11	0.11	0.09	0.12	0.11	0.10
7	50	0.16	0.15	0.20	0.20	0.19	0.22	0.20	0.16
	250	0.12	0.14	0.15	0.15	0.15	0.17	0.15	0.14
	450	0.11	0.27	0.15	0.15	0.15	0.16	0.14	0.15
8	50	0.16	0.16	0.20	0.20	0.19	0.22	0.20	0.16
	250	0.11	0.17	0.15	0.15	0.15	0.17	0.15	0.14
	450	0.11	0.43	0.15	0.15	0.15	0.16	0.14	0.16

Table I.73: CovD results for Right-Skewed Triangular distributed demand

Data/Parameter Setting	Train Data Set Size	CoVd (Inverse of Coefficient of Variation for the difference)							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	50	0.22	0.21	0.27	0.27	0.32	0.44	0.27	0.22
	250	0.18	0.21	0.24	0.24	0.26	0.41	0.24	0.24
	450	0.17	0.18	0.23	0.23	0.24	0.39	0.23	0.30
2	50	0.22	0.21	0.27	0.27	0.33	0.46	0.27	0.22
	250	0.19	0.19	0.24	0.24	0.26	0.41	0.24	0.23
	450	0.17	0.18	0.23	0.23	0.24	0.39	0.23	0.22
3	50	0.19	0.19	0.25	0.25	0.33	0.38	0.25	0.19
	250	0.12	0.13	0.17	0.17	0.20	0.30	0.17	0.16
	450	0.12	0.17	0.16	0.16	0.17	0.29	0.16	0.25
4	50	0.21	0.19	0.25	0.25	0.32	0.38	0.25	0.19
	250	0.12	0.12	0.17	0.17	0.20	0.30	0.17	0.16
	450	0.12	0.14	0.16	0.16	0.17	0.29	0.15	0.26
5	50	0.11	0.11	0.14	0.14	0.17	0.19	0.13	0.11
	250	0.06	0.06	0.09	0.09	0.09	0.11	0.09	0.09
	450	0.06	0.07	0.09	0.09	0.09	0.11	0.09	0.10
6	50	0.11	0.11	0.14	0.14	0.16	0.19	0.13	0.11
	250	0.06	0.08	0.09	0.09	0.09	0.11	0.09	0.09
	450	0.06	0.06	0.09	0.09	0.09	0.11	0.09	0.11
7	50	0.18	0.18	0.20	0.20	0.22	0.24	0.20	0.18
	250	0.13	0.16	0.15	0.15	0.16	0.17	0.15	0.15
	450	0.13	0.17	0.15	0.15	0.16	0.17	0.15	0.15
8	50	0.18	0.18	0.20	0.20	0.22	0.24	0.20	0.18
	250	0.13	0.13	0.15	0.15	0.16	0.17	0.15	0.15
	450	0.13	0.17	0.15	0.15	0.16	0.17	0.14	0.15

Table I.74: CovD results for Left-Skewed Triangular distributed demand

Data/Parameter Setting	Train Data Set Size	CoVd (Coefficient of Variation for the difference)							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	50	0.10	0.10	0.18	0.18	0.17	0.25	0.17	0.10
	250	0.09	0.08	0.16	0.16	0.14	0.21	0.16	0.16
	450	0.08	0.10	0.15	0.15	0.13	0.20	0.15	0.54
2	50	0.10	0.10	0.18	0.18	0.17	0.25	0.17	0.10
	250	0.09	0.09	0.16	0.16	0.14	0.21	0.16	0.14
	450	0.08	0.09	0.15	0.15	0.13	0.20	0.15	3.65
3	50	0.14	0.14	0.22	0.22	0.21	0.27	0.22	0.14
	250	0.09	0.16	0.17	0.17	0.15	0.20	0.17	0.12
	450	0.08	0.20	0.16	0.16	0.13	0.19	0.16	0.55
4	50	0.14	0.14	0.22	0.22	0.20	0.28	0.21	0.14
	250	0.08	0.10	0.16	0.16	0.15	0.20	0.16	0.14
	450	0.08	0.30	0.16	0.16	0.13	0.19	0.15	0.52
5	50	0.09	0.09	0.14	0.14	0.13	0.14	0.14	0.09
	250	0.06	0.10	0.11	0.11	0.08	0.09	0.11	0.09
	450	0.06	0.14	0.11	0.11	0.08	0.09	0.11	0.09
6	50	0.09	0.09	0.14	0.14	0.13	0.14	0.14	0.09
	250	0.05	0.08	0.11	0.11	0.08	0.09	0.11	0.09
	450	0.05	0.14	0.11	0.11	0.08	0.09	0.11	0.09
7	50	0.12	0.12	0.17	0.17	0.14	0.16	0.17	0.12
	250	0.08	0.17	0.14	0.14	0.10	0.12	0.14	0.11
	450	0.08	0.19	0.14	0.14	0.10	0.12	0.14	0.14
8	50	0.12	0.12	0.17	0.17	0.14	0.17	0.17	0.12
	250	0.08	0.10	0.14	0.14	0.10	0.12	0.14	0.12
	450	0.08	0.19	0.14	0.14	0.10	0.12	0.13	0.15

Table I.75: CovD results for Symmetric Triangular distributed and Uniform distributed demand

Data/Parameter Setting	Train Data Set Size	CoVd (Inverse of Coefficient of Variation for the difference)							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	50	0.10	0.11	0.17	0.17	0.17	0.25	0.18	0.11
	250	0.08	0.10	0.15	0.15	0.12	0.20	0.15	0.15
	450	0.08	0.10	0.14	0.14	0.13	0.19	0.14	0.31
2	50	0.11	0.11	0.17	0.17	0.17	0.26	0.17	0.11
	250	0.09	0.08	0.15	0.15	0.13	0.21	0.14	0.14
	450	0.08	0.09	0.14	0.14	0.13	0.19	0.14	0.19
3	50	0.16	0.16	0.22	0.22	0.24	0.29	0.23	0.16
	250	0.09	0.12	0.17	0.17	0.14	0.21	0.17	0.16
	450	0.09	0.10	0.16	0.16	0.13	0.21	0.16	0.27
4	50	0.16	0.16	0.22	0.22	0.23	0.29	0.22	0.16
	250	0.10	0.10	0.17	0.17	0.14	0.22	0.17	0.14
	450	0.09	0.15	0.16	0.16	0.14	0.21	0.16	0.25
5	50	0.10	0.09	0.15	0.15	0.14	0.16	0.15	0.09
	250	0.06	0.08	0.11	0.11	0.09	0.11	0.11	0.09
	450	0.05	0.18	0.11	0.11	0.08	0.09	0.11	0.09
6	50	0.10	0.09	0.15	0.15	0.13	0.16	0.15	0.09
	250	0.05	0.06	0.11	0.11	0.09	0.10	0.11	0.08
	450	0.05	0.13	0.11	0.11	0.08	0.10	0.11	0.09
7	50	0.15	0.15	0.20	0.20	0.19	0.21	0.19	0.15
	250	0.11	0.17	0.15	0.15	0.14	0.15	0.15	0.14
	450	0.11	0.23	0.15	0.15	0.14	0.14	0.15	0.15
8	50	0.15	0.15	0.20	0.19	0.19	0.21	0.19	0.15
	250	0.12	0.14	0.15	0.15	0.14	0.16	0.15	0.15
	450	0.11	0.25	0.15	0.15	0.14	0.15	0.14	0.15

I.8 Selection of Best Method

We record the count the method that gives the maximum for the j^th test period ($J = 1, \ldots, 50$) and for replication $r = 1, \ldots, 50$. Then we calculate the percentage success over 2500 testing data by formula #of times yielding the best result/2500) $*$ 100. The results are obtained for all parameter settings given in Table 4.4 and for all demand distributions selected as well as for each model in Table 3.1. The results for MIP-P and MIP-GP are also shown below. MIP-P is the model where mixed integer program is used to solve both stages to find two decisions simultaneously. Lastly, MIP-GP is a simple heuristic where we run a mixed integer program for the high marked-down price as well as the low marked-down price and select the price and order quantity combination that yields to the larger average total profit prior to the beginning of the first stage.

Table I.76: Selection of best method when demand is Poisson distributed

Data/Parameter Setting	Train Data Set Size	(# of times yielding the best result/2500)*100							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	50	43.32	42.56	18.72	18.36	29.32	44.56	31.24	43.32
	250	34.2	34.52	5.96	5.96	37.76	43.36	26.52	34.2
	450	34.96	34.96	6.68	6.68	36.28	43.12	27.6	34.96
2	50	42.6	43.88	34.12	32.96	41.68	43.44	34.68	42.6
	250	40.96	41.76	39.96	39.96	39.44	42.92	29.48	40.96
	450	45	45	46.48	46.48	43.64	42.64	36.56	45
3	50	28.24	28.56	16.96	16.96	23.08	54.64	20.24	28.24
	250	31.48	31.48	19.68	19.68	21.24	54.12	21.04	31.48
	450	33.72	33.72	25.2	25.2	22.64	53.84	23	33.72
4	50	38	38.56	36.64	36.64	33.84	54.52	36.48	37.88
	250	33.44	33.44	34.96	34.96	28	51.76	33.52	33.44
	450	32.64	32.68	36.48	36.48	32.32	50	33.6	32.64
5	50	35	34.08	21.56	21.56	31.24	61.88	30.68	35
	250	29	27.96	11.92	11.92	20.48	61.52	23.6	29
	450	30.36	31.2	17.24	17.24	22.48	61.88	29.2	30.36
6	50	54.64	55.8	41.64	41.64	58.04	61.56	55.88	54.64
	250	61.64	61.8	44.36	44.36	66.08	69.76	66.72	61.64
	450	63.36	63.32	43.04	43.04	64.68	68.8	65.52	63.4
7	50	36.2	35.96	24.72	25.48	26.8	71	38.88	36.2
	250	22.8	22.8	19.72	19.72	18.92	71.92	21.68	22.8
	450	24.36	23.96	21.32	21.32	19.08	72.4	22.4	24.36
8	50	63.6	63.6	39.24	39.24	61.68	68.16	58.08	63.6
	250	54.48	54.52	35.12	35.12	58.88	64.72	62	54.48
	450	54.52	55.88	34	34	55.6	64.48	60.84	54.52

Table I.77: Selection of best method when demand is Uniform distributed

Data/Parameter Setting	Train Data Set Size	(# of times yielding the best result/2500)*100							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	50	11.96	9	11	7.12	27.4	37.92	9.76	9.44
	250	16	3.6	18.08	15.6	19.4	39.44	16.6	4.12
	450	20.2	6.32	16.08	15.08	11.92	37.52	17.32	10.6
2	50	18.88	5.48	7.44	8.24	24.24	38.8	7.76	8.32
	250	17.36	9.4	14.76	13.4	16.52	40.24	14.72	3.2
	450	22.44	4.76	17.44	18.48	10.52	37.4	17.28	4.24
3	50	14.6	7.24	8	8.92	25.32	46.92	11.2	7.24
	250	10.28	6.32	10.8	9.84	21.6	48.44	8.6	3.8
	450	10.76	5.84	16.32	16.32	19.56	42.64	15.84	18.56
4	50	18.52	9.2	7.24	8.16	25.24	47.6	10.04	7.04
	250	8.72	8.48	8.76	10	22.36	50.76	8.96	2.8
	450	13.24	5.44	16.84	16.84	17.88	39.76	14.6	19.16
5	50	11.04	11.64	14.2	14.2	26.12	54.08	15.8	8.52
	250	8.08	19.56	7.4	6.84	25.12	41.48	6.24	2.04
	450	7.44	17.88	1.8	1.8	28.32	43.36	2.28	2.76
6	50	19.88	12.8	14.6	14.6	27.28	53.84	14.68	11.24
	250	3.68	13.04	6.12	5.64	22.96	52.2	7.64	4
	450	8.52	18.72	1.56	1.52	29.24	41.24	1.92	2.84
7	50	38.88	36.04	45.12	47.04	23.48	59.16	44.04	39.84
	250	5.92	53.48	3.68	3.68	20.52	13.88	3.84	8.6
	450	1.68	54.4	0.6	0.6	24.08	14.36	0.6	7.76
8	50	44.44	35.92	42.84	44.76	22.04	56.6	41.92	37.88
	250	6.32	35.6	2.36	2.36	21.08	35.4	2.64	10.96
	450	2.28	59.4	0.44	0.44	23.16	8.72	0.64	6.76

Table I.78: Selection of best method when demand is Symmetric Triangular distributed

Data/Parameter Setting	Train Data Set Size	(# of times yielding the best result/2500)*100							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	50	14.16	7.28	3.44	3.32	32.32	39.6	6.8	8.12
	250	21.12	13.08	10.24	9.32	15	36.44	8.92	4.44
	450	22.08	10.08	7.84	8.8	15.32	25.76	8.04	18.28
2	50	10.8	7.08	4.16	4.2	33.12	39.64	7.96	9.16
	250	16.64	11.2	14.04	13.16	15.2	8.92	13.88	30.44
	450	25.64	11.4	7.48	4.76	9.12	28.56	5.68	15.4
3	50	9.04	6.92	4.84	4.88	21.96	50.64	6.52	10.96
	250	22.04	12.48	6.88	6.04	10.64	41.08	8.88	4
	450	17.44	16.72	12.76	14.52	9.44	30.68	13.52	9.68
4	50	14.08	8.56	5.08	5.12	20.2	49.36	6.72	9.64
	250	18.96	13.56	8.12	8.08	10.96	44.16	10.24	2.32
	450	18.28	19.88	11.08	11.4	7.08	26.16	15.72	13.32
5	50	9.36	10.04	7.04	7.04	19.92	55.92	8.68	7.84
	250	10.12	24.12	6	6	15.4	39.4	6.2	3.96
	450	8.72	33.16	12.8	13.36	11.84	27.84	15.84	6.8
6	50	10.36	8.32	6.08	6.08	20	55.6	8.24	7.12
	250	11.52	21.56	9	9.76	15.6	40.24	10.8	2.32
	450	13.36	28.8	12.48	12.44	7.88	33.68	14.84	3.76
7	50	9.12	8.84	7.24	6.48	16.96	65.8	7.12	5.96
	250	6.84	34	5.76	5.76	9	40.44	8.96	2.8
	450	7.92	42.8	7.96	7.96	5.76	31.2	13.36	2.68
8	50	9.6	9	6.4	7.08	15	68.28	7.52	8.68
	250	9.76	39.08	4	4	9.48	35.92	6.24	1.96
	450	6.68	53.24	8.8	8.8	5.76	21.08	14.24	3.76

Table I.79: Selection of best method when demand is Right-Skewed Triangular distributed

Data/Parameter Setting	Train Data Set Size	(# of times yielding the best result/2500)*100							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	50	15.76	12	11.32	8.96	20.08	38.76	15.2	13.92
	250	23.2	9.8	5	4.96	8.56	37.16	22	5.6
	450	23.84	8.84	1.4	1.4	5.4	32.24	26.52	22.8
2	50	15.04	10.28	8.8	10	19.8	38.56	13.52	10.68
	250	19.88	13.04	5.84	4.72	6.44	38.32	24.32	5.2
	450	10.72	7.68	2.4	2.48	3.24	32.6	14.12	31.76
3	50	16.64	10.4	11	11.04	15.08	50	15.8	14.08
	250	21.48	4.44	14.96	14.96	3.76	49.28	22.48	2.36
	450	19.04	7.4	9.52	9.52	1.28	45.4	24.96	23
4	50	14.08	14.8	11.8	11.28	13.64	48.56	15.56	14.84
	250	16.32	12.52	12	14.12	3.16	50.36	17.24	2.52
	450	14.08	7.48	9.52	9.48	2	40.2	30.2	16.68
5	50	7.08	7.92	6.48	6.84	18.92	56.24	11.72	9.32
	250	14.44	8.84	17	17.68	7.16	51.36	13.48	3.68
	450	11.56	12.68	21.96	22.56	4.64	48.6	24.36	14.48
6	50	11.44	8.72	5.8	6.12	17.56	54.6	10.68	10
	250	15.2	11.72	16.48	15.8	6.48	53.4	12.6	2.72
	450	12.12	10.88	18	18	4.36	52.2	21.12	8.32
7	50	19.28	22.6	14.88	14.96	15.88	62.88	17.8	20
	250	9.12	22.52	9.68	9.08	10.08	51.8	7.16	1.16
	450	4.24	30.12	16.08	15.32	7.04	43.88	8.8	6.16
8	50	19.48	21.52	16.36	15.52	15.88	63.72	19.04	23.2
	250	9.24	17.16	6.4	6.2	10.36	56.32	4.2	1.92
	450	8	36.88	13.24	12.84	6.52	36.96	7.28	5.68

Table I.80: Selection of best method when demand is Left-Skewed Triangular distributed

Data/Parameter Setting	Train Data Set Size	(# of times yielding the best result/2500)*100							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	50	8.56	9.72	3.6	3.64	35.2	41.08	5.76	7.68
	250	17.68	14.24	6.08	5	19.84	35.44	7.84	3.08
	450	20.48	6.88	3.72	2.84	19.76	28.24	3.04	20.8
2	50	14.36	8	4.08	3.96	32.4	40.88	6.64	7.2
	250	21.8	3.88	6.32	5.28	19.6	40.36	8.28	4.32
	450	23.56	10.6	7.28	7.08	18.12	12.12	6	27.48
3	50	10.48	6.68	2.48	3.64	25.04	51.04	4.52	7.32
	250	19.08	14.36	3.68	3.68	20.4	40.92	3.48	2.56
	450	12.36	19.12	6.68	7.4	19.04	27.88	7.08	13.44
4	50	12.32	5.84	3.24	2.44	24.56	51.8	6.04	7.08
	250	17.32	9.56	2.76	3.84	18.6	47.24	4.8	3.48
	450	14.8	26.52	7.88	8.44	19.4	19.6	4.96	10.6
5	50	9.44	6.4	4.8	4.8	22.88	58.64	5.88	6.88
	250	8.64	25.56	1.12	0.2	23.64	37.08	2.76	1.72
	450	12.76	38.28	2.44	3	20.56	22.68	2.2	2.8
6	50	10.68	7.2	4.36	4.4	21.36	58.32	5.8	7.36
	250	11.04	21.04	1.04	0.08	22.76	40.16	2.2	2.24
	450	14.8	40.92	1.92	1.96	19.2	19.84	2.56	3.8
7	50	8	6.2	5.52	5.76	16.96	69.88	6.44	8.48
	250	7.04	45.52	3.96	3.68	12.44	29.28	2.72	2.88
	450	10.08	55.48	0.92	1.32	12.08	18.8	1.2	1.88
8	50	5.96	7.96	4.32	4.56	15.24	68.2	4.72	6.72
	250	8.8	36.84	3.68	4.04	10.64	38.2	2.84	2.52
	450	9.4	56.36	2.52	3.08	11.12	17.96	2.92	1.36

Table I.81: Selection of best method when demands are Symmetric Triangular distributed and Uniform distributed

Data/Parameter Setting	Train Data Set Size	# of times yielding the best result/2500)*100							
		SAA-P	MIP-SAA	SAA-QR	MIP-QR	SAA-LR	MIP-LR	MIP-GP	MIP-P
1	50	11.72	6.52	6.52	4.32	35.4	38.64	6.28	6.16
	250	22.04	14.56	1.96	2.88	21.48	36.64	3.28	3.36
	450	26.68	6.6	3.12	4.04	23.56	32	3.24	9.08
2	50	11.6	10.84	6.24	6.4	31.08	38.28	7.8	7.4
	250	19.88	15.68	7.04	6.2	18.68	37.68	6.52	2.72
	450	19.84	11.28	5.48	5.48	23.76	32.84	7.6	6
3	50	14.44	10.08	8.48	8.48	19.8	49.72	10.96	13.8
	250	15.04	10.68	10.12	10.88	16.56	44.48	12.24	1.84
	450	17	11.12	16.24	17.04	9.76	40.72	14.24	11.52
4	50	12.56	11.52	7.72	7.72	19.2	49.56	10.56	14.4
	250	20.68	15.56	3.08	3.08	16.4	42	4.32	3.4
	450	17	12.24	15.16	14.12	9.68	38.76	13.64	6.4
5	50	9.32	9.84	6.52	6.6	18.16	56.08	6.96	9.32
	250	9.68	20.92	9.28	10.08	17	40.96	9.68	2.56
	450	12.6	29.36	11.48	11.36	11.52	30.12	13.04	3.88
6	50	11.2	11.36	7.32	7.36	18.28	54.56	8.72	8.2
	250	12.32	11.68	9.44	9.32	14.96	49.6	8.64	4.64
	450	12.4	32.84	13.56	13	11	27.68	13.76	2.36
7	50	12.36	9.72	8.72	8.72	17.56	67.64	8.64	12.72
	250	6.88	33.68	9.4	9.04	8.84	40.52	10.4	1.84
	450	9.36	42.84	10	10.56	6.76	31.24	11.24	3.2
8	50	15.68	8.36	7.2	6.8	16.6	65.64	7.56	11.52
	250	7.32	35.24	8.68	9.52	10.24	40.24	9.2	2.84
	450	7.24	46.4	11.04	11.6	6.72	27.88	12.32	6.04