MACHINE LEARNING FRAMEWORK TO PRICE-SETTING RISK-AVERSE DATA-DRIVEN NEWSVENDOR PROBLEM

A Thesis

by

Eren Atsız

Submitted to the
Graduate School of Sciences and Engineering
In Partial Fulfillment of the Requirements for
the Degree of

Master of Science

in the
Department of Data Science

Özyeğin University
December 2022

MACHINE LEARNING FRAMEWORK TO PRICE-SETTING RISK-AVERSE DATA-DRIVEN NEWSVENDOR PROBLEM

Approved by:

Assistant Professor Enis Kayış,
Advisor
Department of Industrial Engineering
Özyeğin University

Assistant Professor Dilek Günneç
Department of Industrial Engineering
Özyeğin University

Assistant Professor Erinç Albey,
Co-Advisor
Department of Industrial Engineering
Özyeğin University

Assistant Professor Mehmet Önal
Department of Industrial Engineering
Özyeğin University

Date Approved: 29 December 2022

Professor Mehmet Güray Güler
Department of Industrial Engineering
Yıldız Technical University

To my love...

ABSTRACT

Many business-owners, including manufacturers, retailers, distributors are in need to have decisions to make ordering for tons of products in their daily routines. This decision-making cannot be considered trivial since there is always a risk of ordering more or less then what is actually required, which may result in severe undesirable circumstances that are overstocking or shortages. The newsvendor problem in the literature focuses on this trade-off in its classical applications. In such approaches, there is an assumption of demand distribution is known. However, true demand is almost never known and constantly change due to being open to oodles of internal and external parameters; therefore, it is almost impossible to be known and hard to be predicted in real life. For this reason, traditional approaches to newsvendor model in the literature are very valuable yet it is possible to create different variations of this problem which are more appropriate for the real world instances. In addition to quantity of ordering, the determination of price is another serious concern of decision-makers. In order to satisfy the need of determining correct amount of order and regarding price positioning, there is a different approach to newsvendor problems, which is price-setting providing not only quantity but also price information. In this paper, we focus on this variant of newsvendor problem and create a data-driven price-setting newsvendor model, where demand is completely unknown. What is unique in this work is to integrate another element in addition to price decision, which is enabling a risk-averse decision-making perspective. In this way, the constructed model decides on quantity of order and price for products, as well as satisfies a certain risk constraint in order to a set a barrier of probability to not generate under a predetermined level of profit. The price-setting and risk-averse applications of newsvendor

model exist in the literature yet combining these two different perspectives is what makes our approach novel.

ÖZETÇE

Üreticiler, perakendeciler, distribütörler dahil olmak üzere birçok şirket sahibi, günlük rutinlerinde ürün siparişi oluştururken bir karar verme ihtiyacı içindedirler. Bu karar süreci, basit veya önemsiz değildir, çünkü her zaman gerekenden daha fazla veya daha az ürün sipariş verme veya üretme riski vardır. Bu durum da fazla envanter veya talebin altında gerçekleşen ürün temini gibi istenmeyen ciddi neticelere yol açabilir. Literatürdeki gazete satıcısı problemi, klasik uygulamalarında belirtilen durumdaki dengeye odaklanmaktadır. Bu tür yaklaşımlarda, talep dağılımının bilindiği varsayımı vardır. Ancak gerçek dünyada talep neredeyse hiç bilinmez ve iç ve dış birçok parametreye açık olması nedeniyle sürekli değişir; bu nedenle gerçek hayatta bilinmesi neredeyse imkansızdır ve tahmin edilmesi zordur. Bu yüzden literatürde yer alan gazete satıcısı modeline ait geleneksel uygulamalar, çok değerli olmakla birlikte gerçek dünyaya daha uygun varyasyonlarının çalışılması mümkündür. Sipariş miktarına ek olarak, fiyatın belirlenmesi karar vericilerin bir diğer ciddi endişesidir. Doğru sipariş veya üretim miktarını belirleme ve fiyat konumlandırma ihtiyacını karşılamak için, gazete satıcısı probleminde sadece miktar değil aynı zamanda fiyat bilgisi de sağlayan yaklaşımlar vardır. Bu yazıda, haber satıcısı probleminin fiyatın belirlenmesine de odaklandığı bu varyantına odaklanıyoruz ve talep dağılımının tamamen belirsiz olduğu, bütünüyle veriye dayalı bir haber satıcısı modeli oluşturuyoruz. Bu çalışmayı literatürdeki ilgililerinden ayıran nokta ise, riskten kaçınan bir karar verme perspektifi sağlaması ve böylece, fiyat ve miktar kararına ek bir başka unsurun da algoritmaya entegre edilmiş olmasıdır. Böylelikle oluşturulan model, sipariş miktarına ve ürünlerin fiyatına karar vermenin yanı sıra, bu değişkenlerin kullanılması neticesinde gerçekleşecek olan karın, önceden belirlenmiş / hedeflenmiş bir kar düzeyi altında

oluşmaması için bir olasılık kısıtlamasını karşılar; riskten kaçınmaya olanak tanır. Haber satıcısı modelinin fiyat belirleme ve riskten kaçınma uygulamaları literatürde ayrı yaklaşımlar halinde mevcuttur, ancak çalışmanın bu iki farklı bakış açısını aynı modelde birleştiriyor oluşu, yaklaşımımızı literatürü zenginleştiren, gerçek dünyadaki uygulamalara yaklaştıracak kompleksitede kılan özelliğidir.

ACKNOWLEDGEMENTS

I feel very lucky to have the opportunity to work with Assistant Professor Erinç Albey and Assistant Professor Enis Kayış, their support and guidance were everything to me. Their great help during the challenges and obstacles I faced throughout the thesis was invaluable. I am eternally grateful for their tolerance, understanding, and constructive approach, not only during this thesis period but also throughout the whole process we worked together. Above all, they always shared their precious time and ideas, which helped me to be who I am and what I have achieved now.

I also would like to express my gratitude to Professor Burcu Balçık. She never stopped encouraging me and was always by my side whenever I need through my all studies. She has great contributions to me; her support and advice will always be remembered, I will always be grateful.

I would like to thank to Professor Okan Örsan Özener, Assistant Professor Dilek Günneç, Ozan Gözbaşı and Tuba Gözbaşı for their support and valuable advices which helped me a lot during this extraordinary time I spent on the thesis. I cannot imagine being on this path without their guidance and indulgence.

I also want to thank my family and my dear friends Sena, Berk, Ceren, Birce, and Ahmet who accompanied and supported me during my master's degree; our valuable discussions have always pushed me forward. Special thanks to my dear friend Batuhan, I consider myself very lucky to work with him and to have his friendship.

Finally, Gözde, you know I love you beyond words. Nothing would be possible without your unconditional support and love. For me, you are a wonder of nature; every warm smile you have fills my lungs with the smell of earth after rain. Thank you for being my light.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Monitoring the inventory of a product and implementing appropriate managerial decisions are undoubtedly one of the major challenges in operation management. Misevaluation and poor decision making regarding production amount or order quantity are aim to be abstained since results such as overstocks bringing about undesired markdowns or shortages causing unsatisfied demand and lost customers would not be simplified and both result in diminishing profitability. In other words, trade-off between producing or ordering too much or too little is a critical decision to make, and searching for the optimum in between is the interest for all firms. Therefore, manufacturers, distributors, retailers and many other businesses put great efforts on inventory management with a large number of workforce in order to have more accurate decisions and to achieve predetermined targets in terms of stock management and profitability.

Nonetheless, reaching or approaching the optimal order quantity is an unconsummated success in today's highly competitive world; hence, the "price" of the product is the pivotal element accompanying the optimal order quantity. Despite of the fact that a wise inventory management plan and optimum order quantity levels bring about better profitability, "price" is the principal factor in today's profit-oriented, rapid-changing world full of fierce competition. For this reason, independently of type of the business, operation and inventory management can be conducted in an effective and profitable sense only if firms concentrate on the amount of the production or order quantity and the related selling price in the market as a joint decision.

Apart from that inventory management is one of the most crucial effort-areas for all firms today, history of this multi-faceted subject has also a long-history and elaborate study in the literature. One of the most common research study thoroughly representing inventory management problems is the framework of the newsvendor problem and its variations. A traditional newsvendor problem simply aims to decide the order quantities for products as having business sustainable with the greatest possible profitability. The trade-off analysis between having excessive and insufficient inventory is captured by the newsvendor model. This decision regarding the inventory level indeed grounds on demand information yet classical newsvendor problem is modeled based on a demand distribution assumption derived from historical sales data. Despite valuable history and extensive appliances of such theoretical newsvendor models, the hypothesis of a certain statistical distribution of demand does not hold water in real-life examples. In particular, the correct demand information and distribution is almost never known to business owner and open to changes constantly in real life. Therefore, having problem evolved into data-driven approach, in a setting where the related demand function is a random variable with an unknown probability distribution is more challenging but would lead to greater performance of inventory models in real-world situations.

Although deciding on order quantity and price of a product with a data-driven approach where demand distribution is uncertain leads to better accuracy, conducting inventory management in such setting can be acknowledged as complex and risky. To be more precise, a lack of threshold for profit may result in fluctuating profit through terms, which is not consistent and sustainable for the firms. Therefore, in such an approach, the fundamental responsibility of the decision-makers is to set a balance between profitability and financial risks. For this reason, in addition to the price-setting approach, including another dimension as being risk-averse, can create a more useful synergy for the firms aiming to keep risks in a certain level as well as to maximize their profits.

In this paper the price-setting newsvendor problem that is a well-known variant of a traditional newsvendor problem taking a significant place in literature is examined. Aim is to approach price-setting newsvendor problem in a fully data-driven approach where demand distribution is completely unknown and to integrate a different aspect that is creating a risk-averse decision-making strategy. In order to provide a risk-averse decision, the model contains a risk structure in the objective function and the aim is to monitor the risk and satisfy this chance constraint, which is "Value at Risk" (VaR) in this paper. VaR addresses and answers the question of how much a decision maker can lose with probability x% in a given time frame [1]. For example, if a business-owner is interested in 90% single-day VaR of 1 million USD, this means the business have the risk of gain less than 1 million USD within a day with only probability of 10%. In this way, the model purposes to provide not only price and quantity of orders to decision-maker but also satisfies a certain threshold for the expected profit by monitoring a VaR constraint. That approached newsvendor model is not only price-setting but also risk-averse brings about uniqueness that sets the work apart from the others in the literature.

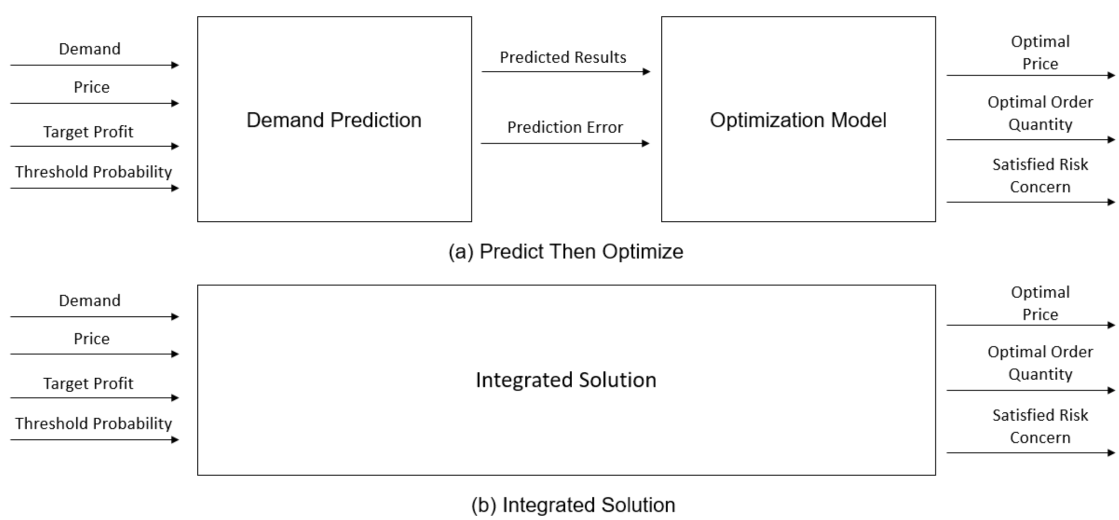

Figure 1: High-level representation of the proposed models

In the scope of this study, we propose an Integrated Solution by constructing a Mixed Integer Non-linear Programming (MINLP) model and also we develop a "Predict then Optimize" algorithm as shown in the Figure 1 above. Finally, we simulate the obtained results with Monte Carlo Simulation, which was constructed with the assumption that the real world is known. As a result, we find that the results are close to the best results in the real world in a great extent.

CHAPTER II

LITERATURE REVIEW

2.1 Previous work and Background

Inventory management can be treated as a multifaceted problem that includes various objectives, constraints and parameters; and one of the most classical problems related to this field can be acknowledged as the newsvendor problem. Analysis of this problem enables key insights which have wide range of inferences for many businesses including, hospitality, airline, fashion and sporting goods, at both manufacturing and sale levels [2].

The traditional approach to the newsvendor problem forms the backbone of the stochastic inventory theory due to its simple yet effective structure, which provides substantial insights for the firms to determine the the optimal ordering quantity for maximizing profit [3]. In this regard, majority of classical approaches to inventory management assume that the demand distribution and its parameters are known [4]. However, over the years, academic interest and studies in the area of this problem have progressed and evolved to have a different setting, where demand distribution is not known. The uncertainty of the relevant demand distribution has added a new dimension to this problem, increasing the quality of the implications and the richness of the approach to be more suitable for real-world applications. In this literature review section, we focus on inventory management problems where the demand distribution is unknown.

In addition to that demand uncertainty helps inventory management problem to become more realistic, it is also a major challenge while deciding on the appropriate quantity of order levels since actual demand distribution depends on a large data containing various parameters such as seasonality, branding, etc. and may change disruptively over time [5, 6]. Increasing access to large data sets and developing technologies and data-driven methods, which put historical information into a future perspective and create a road-map to inventory management, leastwise eliminate possible challenges and may provide an opportunity to have better-informed decisions. In this part, first, we focus on data-driven approaches to newsvendor problem and further, its price-setting and risk averse variants are discussed.

Firstly, data-driven models may be constructed with and without covariate information. In the case of having no covariate information, the data is available in the form of "n" independent historical information of only demand. The classical methodology to solve such cases is to "first predict then optimize", which depends on performance of the prediction model such that challenges and misguidance can emerge in case of having a prediction model with a poor performance. Under this setting, there are important works in the literature which involve parameter forecast and optimization using operational statistics methods. In these approaches, prediction of the demand distribution does not exist, and the quantity of order decision is only as a function of the historical demand data [7]. One example to common algorithms is sample average approximation (SAA) [8, 9], where the objective function is replaced by an average based on independent random instances which are derived from the real demand distribution [10].

Secondly, data-driven models to approach newsvendor problem can be created with covariate information. In such models, monitorable features such as seasonality contributed to estimation of future demand and foresee the appropriate amount for quantity of order. For instance, one study interprets the problem as a dynamic procurement model in order to satisfy unknown demand within a known time period [11]. Tree-based approach is used as applying regressions to predict demand for one product according to past track of the other products. Moreover, in an other great example, the problem is focused on considering linear programming in order to tackle the challenge resulting from combination of external variables and randomness which constitutes the linear demand function [12].

Besides data-driven perspective with uncertain demand function, standard newsvendor problem can be enriched with one step further focusing on variables to decide. In traditional ways, the order quantity is the only decision variable [13]. Adding "price" as a variable to decide would enrich the problem outputs and draw a better strategy to decision-makers expecting the maximum profit as deciding on quantities. This variant of data-driven approach is called "price-setting". The joint price and inventory problem when demand information not known is studied by several papers, one example is presented in the literature, where the price-setting newsvendor model is created based on optimal pricing and quantity of order decisions considering the minmax regret criterion [14]. Additionally, there are other works that used kernel-based and regularization methods for multifaceted problems [15]. The appliances of such approaches to the price-setting variant of newsvendor problem hold complexity due to requirement of conditional value at risk (CVAR) prediction.

In addition to data-driven approach and price-setting variants of newsvendor problems, incorporating financial risks as making managerial decisions regarding inventories is considered as another essential perspective in recent years. More precisely, decision-maker predetermines a certain level of profit for a single period of inventory problem and accept a certain probability of risk for actual profit to realize under that threshold [16]. A great work exist in literature presents an outstanding example for risk-averse newsvendor model and used value at risk (VaR) in order to measure the risk in the context of inventory management [17]. In a risk-averse newsvendor model, as deciding the quantity of the orders under an uncertain demand in order to maximize the expected profit, the model searches for the probability distribution of expected profit, both using analytical and simulation methods.

2.2 Contributions

Our study provides a data-driven framework for determining the price and quantity, uniquely in a risk-averse approach, which makes the work novel and different than the other variants of the newsvendor models in the literature. The main purpose of the study is to propose a new structure that monitors the risk of not being under a certain level of profit as providing an optimal ordering quantity and price determination for products. In this way, quantity and price deciding needs of decision-makers are handled jointly while also satisfying a risk constraint, which makes model appropriate to a real-world scenario, where business-owners are in need to decide not only correct amount of quantity but also price for products as also having concerns regarding a certain level of expected profit. In order to do so, we use value at risk (VaR) to define risk as a probabilistic constraint in objective function of the mathematical model. Briefly stated, in the scope of this work, the price-setting newsvendor problem is discussed together with risk structures with a completely data-driven approach.

Our contribution is to construct a data-driven model that combines two existing separate variants of the newsvendor problem, which are price-setting and risk-averse approaches, and to extend the newsvendor problem appliances in the literature. The proposed approach in this paper is more complex, which makes it closer to the real-life examples, and more holistic since since it proposes optimal quantity of orders and price under a risk structure that concentrates on the risk of profit-loss concern of the decision-maker. In this regard, we present our solutions under two strategies, which are integrated solution and predict then optimize;

- ***Integrated Solution:*** This part presents mixed integer non-linear mathematical model (MINLP). The model decides on optimal solutions for quantity and price simultaneously while also monitoring and satisfying VaR as a chance constraint. The significant contribution of this created MINLP model is to have a completely data-driven structure that does not include any prediction model in its framework.

- ***Predict then Optimize:*** This approach allows us to decide separately the near optimal price and order quantity pair, with closed-form newsvendor solutions over the t distribution. Despite deciding on price and quantity separately, we present quite competitive and robust results under favour of the created methodology.

Finally, we simulate obtained results with Monte Carlo Simulation which is built under the assumption that the real world is known. To be more specific, the optimal price and order quantity obtained with both approaches are applied over the demands where the real equation between demand-price is known (for example, the normal distribution with mean, standard deviation known); the expected profit value and the availability of chance constraint are examined. As a result, it is observed that the results are highly close to the best solutions of the real cases.

CHAPTER III

METHODOLOGY

In this section, we present our framework developed for the extended data-driven version of the newsvendor problem, which focuses on historical data in an environment where the demand distribution is unknown, works on deciding on the price and optimal order quantity simultaneously, and also presents a solution that can satisfy the decision maker's needs regarding risk perception. First of all, we present the single-period newsvendor model with a classical closed-form solution and its structure with Value at Risk (VaR) extension as preliminaries in Subsection 3.1. Next, we present the formulations of the solutions we propose, respectively: we present the algorithm of our *"Predict then Optimize"* approach in Section 3.2 and the mathematical model of the *"Integrated Solution"* structure we proposed in Section 3.3.

3.1 Preliminaries: Classical Newsvendor Model

3.1.1 Classical Single-Period Newsvendor Model

The Newsvendor Problem, which is one of the most studied problems due to its applicability to the most important and different subjects in the field of operations research and operations management, aims to maximize the profit to be earned by deciding on the optimal order quantity. Assuming that the state of the demand distribution, which is one of the most important parts of this model, is known in the classical newsvendor model, a closed-form optimal result can be reached. In the case of the Newsvendor problem, the sales price, unit cost, and salvage value notations are p, c_0, p_v respectively. It is assumed that the relationship between them will be $p > c_0 > p_v$, as foreseen.

While the decision maker is trying to decide on the optimal order quantity, if the determined quantity is more than the actual demand, it will dispose of the excess quantity with the salvage value; on the contrary, if the determined quantity is less than the realized demand, an opportunity of earning an extra profit cannot be captured and the profit diminishes. The solution of the problem should be designed to balance these two situations and aim to maximize profit. For this purpose, the profit function of the problem where the order quantity q and demand D is as follows:

$$\pi(q, D) = (p - c_0)q - (p - p_v)(q - D)^+ \tag{1}$$

with $(q - D)^+ = max(q - D, 0)$

The distribution function of D is F, and F^{-1} is assumed to exist. The closed-form solution of the problem is as follows:

$$\max\ E(\pi(q, D)) \tag{2}$$

derived by

$$q^* = F^{-1}\left(\frac{p - c_0}{p - p_v}\right) \tag{3}$$

where optimal order quantity (q^*), the decision maker can maximize the optimal profit when the demand distribution is known and the F distribution is invertible.

3.1.2 Newsvendor Model with VaR Constraint

One of the most common examples of situations where the decision maker finds the optimal order quantity by trying to maximize profits with the classical newsvendor model, but also has the desire to manage risk at the same time is the Newsvendor problem with VaR extension. In this structure, the target profit and confidence level predetermined by the decision-maker should be provided. With the VaR constraint, the decision maker wants to set a threshold probability in order not to have less than a certain profit level which is predetermined [18]. The target profit is denoted as v, and the probability of reaching this profit level v should be equal to α ($\alpha \in (0, 1]$).

The Newsvendor model with VaR constraint can be processed in various ways yet our major concentration regarding the formulation below that is parallel with the modeling approach of Gan et al. (2004) [19].

$$\max \; E(\pi(q, D)) \tag{4}$$

subject to

$$P(\pi(q, D) \leq v) \leq \alpha \tag{5}$$

The problem has a closed-form solution, but an admissible solution exists only if the equation $v \leq F^{-1}(\alpha)(p - p_v)$ is provided. The q^* value can be calculated by the equation given below.

$$q^*(\alpha, v) = \begin{cases} q^*, & \text{if} \quad q^* \leq \dfrac{F^{-1}(\alpha)(p - p_v) - v}{c_0 - p_v} \\ \dfrac{F^{-1}(\alpha)(p - p_v) - v}{c_0 - p_v}, & \text{if} \quad q^* \geq \dfrac{F^{-1}(\alpha)(p - p_v) - v}{c_0 - p_v} \end{cases} \tag{6}$$

3.2 Predict then Optimize

As one of the data-driven solutions, our first approach is to Predict then Optimize; prediction and optimization are applied separately. In this framework, although the performance of the optimization model depends on how well the previous prediction model performs, accurate and robust results are provided in a structure where the prediction model can explain the data.

We present a 5 step algorithm flow for the Predict then Optimize structure.

1. ***Demand Prediction.*** Creating a machine learning model and learning model parameters in order to predict demand ($\hat{y}_h$ - "predicted value") based on price using historical price-demand dataset.

    - For this stage, we focus on the linear regression model ($Y = \beta_0 + \beta_1 X + \epsilon$).

2. ***Standard Error Calculation.*** ML model standard error (see) calculation is performed at this stage as a prelude to the 3^{rd} stage. The equation used for linear regression:

$$\sqrt{MSE \times \left(1 + \frac{1}{n} + \frac{(x_h - \bar{x})^2}{\sum (x_i - \bar{x})^2}\right)} \tag{7}$$

where x_h used predictior's value and x_i, i^{th} predictior's value.

3. ***Critical Value Calculation.*** Finding the critical value of the newsvendor problem with the non-standard t distribution using the $\hat{y}_h$ obtained from the prediction model and the standard error:

$$F^{-1}(\alpha) = \Phi^{-1}(\alpha) \times se + \hat{y}_h \tag{8}$$

where α represents the distribution probabilities.

4. ***Newsvendor Problem with VaR Solution.*** For each price in the data, the optimal order quantity is calculated with the closed-form solution of the newsvendor model (classic or with VaR), and the expected profit is calculated over the data created with the t-distribution obtained over it.

5. ***Ultimate Solution with Response Curve Fitting.*** The response curve is fitted (e.g., polynomial regression) over the expected profits obtained according to the prices; near-optimal price calculation is performed on the curve and the price and demand are decided.

    - We present the polynomial regression structure for the response curve.

$$y = b_0 + b_1^{x_1} + b_2^{x_2} + b_3^{x_3} + ... + b_n^{x_n} \tag{9}$$

As a result of the 5-stage approach, we obtain the optimal solution for each historical price in the data with closed form solutions, and we aim to obtain near-optimal profit returns for different prices over the response curve.

3.3 Integrated Solution

In this section, we present a mixed integer nonlinear programming (MINLP) formulation for the price-setting risk-averse data-driven newsvendor problem. We describe the notation for the model first and then present the mathematical model below.

Sets
$\mathcal{T}$: Set of historical events (scenarios)
Parameters
c_o: Purchase cost
p_v: Salvage value
α: Threshold probability value
v: Target profit value
p_t: Historical selling price t
p_{min}: The price of the scenario with the lowest price in historical data
p_{max}: The price of the scenario with the highest price in historical data
p_{avg}: Average value of highest and lowest price
d_t: Historical demand for price t
c: Probability level determinant
δ'_{ti}: the demand adjustment coefficient of the scenario t
Decision Variables
q: Optimal order quantity
p: Optimal price
δ'_t: The realization probability of the scenario t under the optimal price
z_t: The distance of the historical price t relative to the optimal price
y_t: Auxiliary decision variables for price t
γ_t: Auxiliary binary decision variables for price t

Table 1: Set, parameter and variable definitions of the mathematical model

The formulation for the mathematical model is presented next.

$$\text{Maximize} \sum_{t \in T} \delta'_t[(p - c_o)q - (p - p_v)y_t] \tag{10}$$

subject to

$$y_t \geq q - \sum_{i \in T} \delta'_{ti} d_i \qquad \forall t \in T \tag{11}$$

$$(p - c_o)q - (p - p_v)y_t + M\gamma_t \geq v \qquad \forall t \in T \tag{12}$$

$$\sum_{t \in T} \delta'_t \gamma_t \leq \alpha \tag{13}$$

$$z_t = (p - p_t)^2 \qquad \forall t \in T \tag{14}$$

$$\delta'_t = \frac{e^{-cz_t}}{\sum_{t \in T} e^{-cz_t}} \qquad \forall t \in T \tag{15}$$

$$p \geq p_{min} \tag{16}$$

$$p \leq p_{max} \tag{17}$$

$$p \geq c_o \tag{18}$$

$$p, q \geq 0 \tag{19}$$

$$y_t, z_t, \delta_t \geq 0 \qquad \forall t \in T \tag{20}$$

$$\gamma_t \in 0, 1 \tag{21}$$

The objective function (10) of the model we present is one of the objective functions of the widely used standard newsvendor models and aims to maximize profit. In addition, the objective function aims to assign probabilities for the previously obtained scenarios considering the optimal price. Considering the similarity of the determined optimal price compared to prices of the previous scenarios, demand realized in the past is evaluated. Constraint (11) provide determining the amount of unsold product for salvage. Constraint (12 - 13) is the addition of the VaR structure to the model as a chance constraint reflecting the status and probability of the determined target profit to the model. Constraint (12 - 13) ensures that the situation and probability of the determined target profit are reflected in the model; paves the way for risk management by ensuring that the problem can be resolved in terms of target profit and threshold determined by the decision maker. Constraint (14) provides the calculation of the distance metric used for the model, and with Constraint (15), the similarity status of the scenarios in the data is created according to the optimum price over the defined distance metric. Constraint (16 - 17) ensures that the optimal price to be determined is within the historical price range, so that the model cannot provide a price beyond the previously lagged min-max price. Constraint (18), as a usual constraint of the newsvendor problem, ensures that the price to be determined is not lower than the cost. Lastly, constraints (19 - 21) define the decision variable domains.

The *Integrated Solution* we present with the mathematical model does not provide any demand forecasts or does not include such a structure in the mathematical model. It is aimed to create optimization by learning directly from the data. However, for this approach, the relationship between price and demand should be reflected in the model. For this purpose, the historical demand information in the scenarios is not presented as a direct input to the model, and adjustments are made in the demands based on the relationship of prices, similar to the structure in constraint (15). The coefficients that will provide these adjustments are calculated with the following equation.

$$\delta'_{ti} = \frac{e^{-c(p_t-p_i)^2}}{\sum_{i\in T} e^{-c(p_t-p_i)^2}} \qquad \forall i \in T, \forall t \in T \tag{22}$$

where

$$c = e^{p_t(p_t-p_{avg})^2} \qquad \forall t \in T \tag{23}$$

An important feature of the equations used for δ'_{ti} and δ'_t is that they distribute the similarities so that the total is 100%. The structure used appears in different forms in the pricing literature. In addition, the coefficient c in the equation stands out as a criterion that determines the aggressiveness of the price distribution and its proper determination takes a critical role in the problem. According to different values of c, all probabilities to be satisfied can be almost equal, or for constraint (15), only the value closest to p can take any probability. For this reason, we propose the following equation in order to establish a structure that will be price sensitive, distributing the probabilities more towards the extremes at average prices, while considering only the prices around them at extreme prices. In addition, the price multiplier in the equation supports the model to be more sensitive to high prices. Parameter c is set similarly for δ'_{ti} and δ'_t, and for δ'_t used as a decision variable in the mathematical model, c is calculated by the following equation.

$$c = e^{p(p-p_{avg})^2} \tag{24}$$

CHAPTER IV

NUMERICAL EXPERIMENTS

4.1 Dataset Generation

Within the scope of this study, we construct different stochastic-based price-demand relations in order to measure the performance of the models and approaches we propose. For the dataset generation methods we use, we base the dataset creation functions and values in the article "A Prescriptive Machine-Learning Framework to the Price-Setting Newsvendor Problem." [13]. Although this study does not include any VaR structure, it basically proposes three different functions in order to test the relationships between price and demand, and we focus on these three functions specified for the structures we propose. The functions used are basically such that the price-related error changes and evolves toward the use of more complex equations. The first function used is shown below as the G1 function (25).

$$\textbf{G1}. \quad Y = \beta_0 + \beta_1 p + (\gamma_0 + \gamma_1 p + \gamma_2 p^2)\epsilon \tag{25}$$

As a random variable with the error term (ϵ) normal distribution, its mean is 0 and standard deviation is 1. In addition, other parameters used in dataset creation: $\beta_0 = 200.0, \beta_1 = -35.0, \gamma_0 = 36.0, \gamma_1 = -12.0, \gamma_2 = 2.1$.

The main difference of the second dataset creation function G2 (26) used is that it contains 2 different error terms, positive and negative. Although the distribution of error terms can be explained with a certain distribution, the error distribution of demand created as a result of the G2 function can be explained with an empirical distribution.

$$\textbf{G2}.\quad Y = \beta_0 + \beta_1 p + \beta_1 p^2 + (\gamma_0 + \gamma_1 p)\epsilon^- + \gamma_2 p^2 \epsilon^+ \tag{26}$$

In the parameters used for G2, the error term (ϵ) has a normal distribution with mean 0 and standard deviation 1. The coefficients and application cases of the parameters used: $\epsilon^- = min(\epsilon, 0)$, $\epsilon^+ = max(0, \epsilon)$, $\beta_0 = 215.0$, $\beta_1 = -37.0$, $\beta_2 = -1.1968$, $\gamma_0 = 36.0$, $\gamma_1 = -4.0$, $\gamma_2 = 3.0$.

The equations G1 and G2 have demand functions with means which are diminishing linear functions of price factor and variances which are nonmonotonic quadratic functions of price as indicated by Raz and Porteus [20].

The third and final dataset creation function uses an extended version of G1. G3 uses different distributions in addition to the error term structure, which varies depending on price. The G3 equation is as shown below.

$$\textbf{G3}.\quad Y = \beta_0 + \beta_1 p + \beta_3 T + \sum_{k \in \mathcal{K}} X_k + (\gamma_0 + \gamma_1 p + \gamma_2 p^2 + \gamma_3 T + \sum_{k \in \mathcal{K}} X_k)\epsilon \tag{27}$$

The dataset allows variation of the equation by including possible covariate information. ϵ is a random variable with normal distribution N(0,1) as in G1 and G2. Other parameters: $\beta_0 = 200.0$, $\beta_1 = -35.0$, $\beta_3 = 10.0$, $\gamma_0 = 36.0$, $\gamma_1 = -12.0$, $\gamma_2 = 2.1$, and $\gamma_3 = 2.0$. T and X_k , which are considered as covariate information, has a different distribution from the error term; T has a uniform distribution between [-1, 1]. X_k takes values as different as $1, ..., K$. In this respect, the ones up to K/2 have normal distributions N(0,1) and the ones from $K/2$ to K are accepted as bernoulli distribution, which is the special case of the binomial distribution, and 0-1 is considered to create a result. The probability is assumed to be 0.5. In addition, K in the study is considered as 32 as in the applied study [13].

For the datasets that will be used in the study, the price is obtained with a uniform distribution, which varies between 1.5 and 4. Within the scope of the problem, the cost (c_o) is set to 1.0 and the salvage value (p_v) is set to 0.5. However, different data sets are created using the G1, G2 and G3 equations. There are different sizes determined so that the demand is 10, 20 and 30. In this problem, target profit is set as 30% more than the average cost in each instance, so the target profit varies according to the instances. At the same time, the threshold probability value, which is important for VaR, is studied in two different versions as 15% and 30%. In other words, a created instance is tried to be solved with both 15% threshold probability and 30% threshold probability without changing any information, and the results are examined. 30 instances for each function are created in 3 different sizes and for 2 different threshold probability values. Thus, 180 different instances are created for each function and 540 different instances examined in total as can be seen in Table 2.

	Total instance	Size	Threshold probability
	60	10	15% - 30%
G1	60	20	15% - 30%
	60	30	15% - 30%
	60	10	15% - 30%
G2	60	20	15% - 30%
	60	30	15% - 30%
	60	10	15% - 30%
G3	60	20	15% - 30%
	60	30	15% - 30%

Table 2: Types of main instances

To examine the heteroscedasticity case in addition to three different equations, we construct two different versions of the dataset similar to Harsha et al [13]. For this, based on the G1 function, we add a factor H to the parts of the price in the coefficient $(\gamma_0 + \gamma_1 p + \gamma_2 p^2)$, which is the multiplier of the error term, and we get the new multiplier $(\gamma_0 + H\gamma_1 p + H\gamma_2 p^2)$. We refer to the case where H is zero as the homoscedastic case and define the G1-v2 (28) function for this.

$$\mathbf{G1-v2}. \quad Y = \beta_0 + \beta_1 p + (\gamma_0)\epsilon \tag{28}$$

$H = 1$ is the G1 function itself that we have already defined, and we additionally create the $H = 2$ heteroscedastic case G1-v3 (29).

$$\mathbf{G1-v3}. \quad Y = \beta_0 + \beta_1 p + (\gamma_0 + 2\gamma_1 p + 2\gamma_2 p^2)\epsilon \tag{29}$$

In parallel with the number of creations of main instances, we create 30 instances for each size and threshold probability option in homoscedastic state analysis. We create 180 instances for G1-v2 and G1-v3 as can be seen in Table 3. Together with the main instances, a total of 900 instances were created for this study.

	Total instance	Size	Treshold probability
	60	10	15% - 30%
G1-v2	60	20	15% - 30%
	60	30	15% - 30%
	60	10	15% - 30%
G3-v3	60	20	15% - 30%
	60	30	15% - 30%

Table 3: Types of G1 variations instances

4.2 Benchmark: Groundtruth Solution

We consider the closed-form solutions of classical newsvendor and extensions with VaR, which we mentioned in Section 3.1, as Groundtruth solutions for evaluating and comparing the approaches we propose.

As applying the groundtruth solution, it is assumed that the parameters in the dataset creation and the distribution of the error term (standard deviation and mean values) are known. As is known, the closed-form classical newsvendor with VaR extension solution provides the optimal order quantity under the Var constraint, but it cannot provide any price decision within the scope of the problem. In our approach, we are trying profit maximization with enumeration and closed-form solutions within the price range of the problem using price, and we determine the price and order quantity that enables the maximum profit by satisfying the problem constraints for each price.

In the context of the problem, the optimal solution is calculated with Equation 6 for the relevant price value. As can be seen, the optimal solution can be obtained from time to time from a closed-form solution with VaR constraints, or it can also be obtained from the classic newsvendor solution depending on the constraints' status and inputs of the problem.

4.3 Monte Carlo Simulation

To measure the performance of our two proposed solutions, the *"Integrated Solution"* and the *"Predict then Optimize"* approach, we are constructing a problem-specific Monte Carlo simulation structure.

Each of the created instances is obtained for the groundtruth solution and the two suggested data-driven methods within the scope of the simulation setup, and it is aimed to compare their performances. In this context, simulation aims to show how the data, which are inputs for the results obtained, perform under the assumption that the creation equations and creation states are known. In other words, simulation is considered as a tool to reveal the real-world performances of the results obtained by different methods.

Within the scope of the simulation, the demands are created with the datasets used in the dataset creation section. At the same time, the parameter values obtained as a result of the distributions used while creating the instance for G3 are also used and these parameters are not recreated.

4.4 Computational Results

In this section, we present a comparison of the results obtained with the methods we have indicated. In Section 4.4.1, we show an example and details of a solution obtained with the Integrated Solution. In Section 4.4.2, we provide examples of setting the c parameter used in the MINLP formulation of the Integrated Solution's method. Finally, we show the solutions of the test instances created and the comparison of the methods with the two methods and groundtruth that is shown in Section 4.4.3. In this study, AMPL [21] was used for mathematical modeling and the BARON solver [22] ran on NEOS Server[23] in MINLP formulation. Although the server specifications vary, the average Dell PowerEdge R430 server is Intel Xeon E5 2.3GHz. Also, the Predict then Optimize structure and Monte Carlo Simulation are coded in Python 3.8.

4.4.1 Illustrative Example

We examine the Integrated Solution on the first instance containing 20 demand-price data of the G1 instance. When solving this instance, it is considered as $c_0 = 1.0$ and $p_v = 0.5$, and the target profit is set to 130 as 30% more than the average cost. Instance provides the optimal price and optimal order quantity using the MINLP and BARON solver offered by the Integrated Solution. For this problem, the optimal price is 3.39 and the optimal order quantity is 92.

Record	Demand	Price	y_t	$(p-c_o)q-(p-p_v)y_t$	z_t	γ_t	δ'_t
1	161.25	1.58	0.00	222.27	3.28	0	0.00
2	145.06	1.77	0.00	222.27	2.62	0	0.00
3	122.95	2.00	0.00	222.27	1.93	0	0.00
4	120.83	2.05	0.00	222.27	1.80	0	0.00
5	124.31	2.23	0.00	222.27	1.35	0	0.00
6	74.25	2.54	18.75	168.10	0.72	0	0.00
7	85.11	2.60	7.89	199.46	0.62	0	0.01
8	116.58	2.68	0.00	222.27	0.50	0	0.01
9	83.30	2.90	9.70	194.24	0.24	0	0.04
10	85.65	2.93	7.35	201.04	0.21	0	0.05
11	95.54	3.02	0.00	222.27	0.14	0	0.07
12	92.15	3.15	0.85	219.82	0.06	0	0.10
13	113.65	3.15	0.00	222.27	0.06	0	0.10
14	127.15	3.21	0.00	222.27	0.03	0	0.11
15	74.91	3.49	18.09	169.99	0.01	0	0.12
16	81.88	3.55	11.12	190.14	0.03	0	0.12
17	82.46	3.71	10.54	191.82	0.10	0	0.08
18	55.23	3.73	37.77	113.12	0.12	1	0.07
19	62.22	3.75	30.78	133.33	0.13	0	0.07
20	91.94	3.83	1.06	219.22	0.19	0	0.05

Table 4: First instance with 20 price-demand data of G1.

δ'_t which is a decision variable of the mathematical model, varies according to the price. The price for this solution is 3.39, while the highest rates are given to the records where the price is 3.49 and 3.55. In this case, at the same time, the records that the price fell below 2.54 had no effect on the solution of the model. The problem aims to provide a 30% threshold probability under the effect of the VaR constraint. The only record where the γ_t value is 1, that is, the targeted income cannot be achieved, is the record with a price of 3.75. However, since the probability given to this record is 7%, the problem is seen as feasible as can be seen from Table 4.

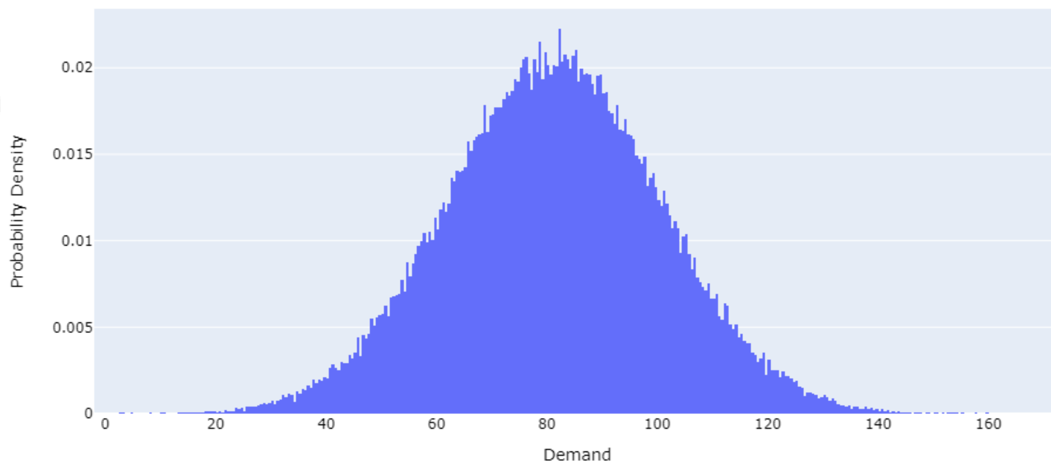

Figure 2: G1, Size=20 demand distribution when p=3.39 (100,000 demand)

The results obtained are tested in the validation phase with the Monte Carlo Simulation. For this, 100,000 new demands are created using the function in which the instance was created, and the distribution of the demands for the optimal price decided by the Integrated Solution is shown in Figure 2.

As a result of the simulation, according to the results obtained using the optimal price and quantity determined by the integrated solution, 14.9% of the tests were below the target profit, but since the threshold probability was 30%, the solution was feasible and a total profit of 179,000 was reached.

4.4.2 Examination of Similarity Status and Determination of Parameter c

In this part, we examine the importance of the c parameter, which has a critical importance for the performance of the mathematical model (Integrated Solution) and affects both δ'_{ti} and δ'_t, and its changes according to price.

We first present the sensitivity analysis for δ_{ti}. We examine the difference between the values of the c parameter used for the calculation of δ_{ti} compared to the expected demand values if it takes a fixed value from 1 to 9. Likewise, if the c parameter is set as suggested in equation (23) (indicated as implemented in the tables), we examine the sum of squared error and standard deviation obtained according to the expected demand average.

	Average			Standard Deviation		
	Size			Size		
c	10	20	30	10	20	30
1	1,352	2,716	4,386	845	1,737	1,909
2	952	1,630	2,556	722	1,599	1,573
3	938	1,404	2,106	700	1,506	1,367
4	981	1,348	1,963	716	1,454	1,260
5	1,033	1,345	1,925	741	1,425	1,206
6	1,083	1,365	1,933	764	1,408	1,181
7	1,129	1,394	1,962	785	1,399	1,172
8	1,173	1,427	2,002	805	1,395	1,171
9	1,214	1,462	2,049	824	1,395	1,176
Implemented	1,002	1,356	1,960	804	1,411	1,205

Table 5: Average SSE and Standard deviation values relative to expected demand for G1

Our aim here is to manage the price-demand relationship in the mathematical model prepared for the Integrated Solution and to bring the demand data provided as input to the model as close as possible to the expected demand values of the applied equations (based on dataset creation). For this purpose, adjusted values with $\sum_{i \in T} \delta'_{ti} d_i$

are used instead of direct demand information in constraint (12), and the c parameter directly affects δ_{ti}.

In Tables 5 and 6, we show the different c values for instances created with G1 and G2 and the performance of our proposed approach. In this case, the Implemented method seems to be the most appropriate choice for most sizes of G1 and G2, according to the average SSE values and standard deviations. While the lowest error values for some sizes can be obtained for $c = 3$ or $c = 4$, it is seen that the most robust application is the recommended (implemented) approach while monitoring all sizes and standard deviations of errors. Since the G3 dataset is created in a similar way over G1, and the expected values are similar, it is not shared separately.

	Average			Standard Deviation		
	Size			Size		
c	10	20	30	10	20	30
1	1,138,026	2,121,636	3,282,733	262,846	356,879	540,088
2	1,125,493	2,096,509	3,244,266	263,596	357,668	541,553
3	1,120,557	2,086,891	3,227,960	264,467	358,943	543,567
4	1,117,820	2,081,975	3,218,955	265,137	359,987	544,969
5	1,116,079	2,079,098	3,213,274	265,665	360,813	545,902
6	1,114,897	2,077,290	3,209,388	266,088	361,472	546,531
7	1,114,065	2,076,104	3,206,583	266,424	362,006	546,963
8	1,113,470	2,075,305	3,204,479	266,685	362,445	547,264
9	1,113,038	2,074,757	3,202,858	266,883	362,810	547,475
Implemented	1,112,505	2,080,536	3,188,675	263,632	363,790	540,980

Table 6: Average SSE and standard deviation values relative to expected demand for G2

In addition, as can be seen in the detailed boxplots representations shared in Figures 3, 4 and 5, the Implemented Solution approach is not only suitable for average values but also provides better results in terms of min error values than when c is set fixed.

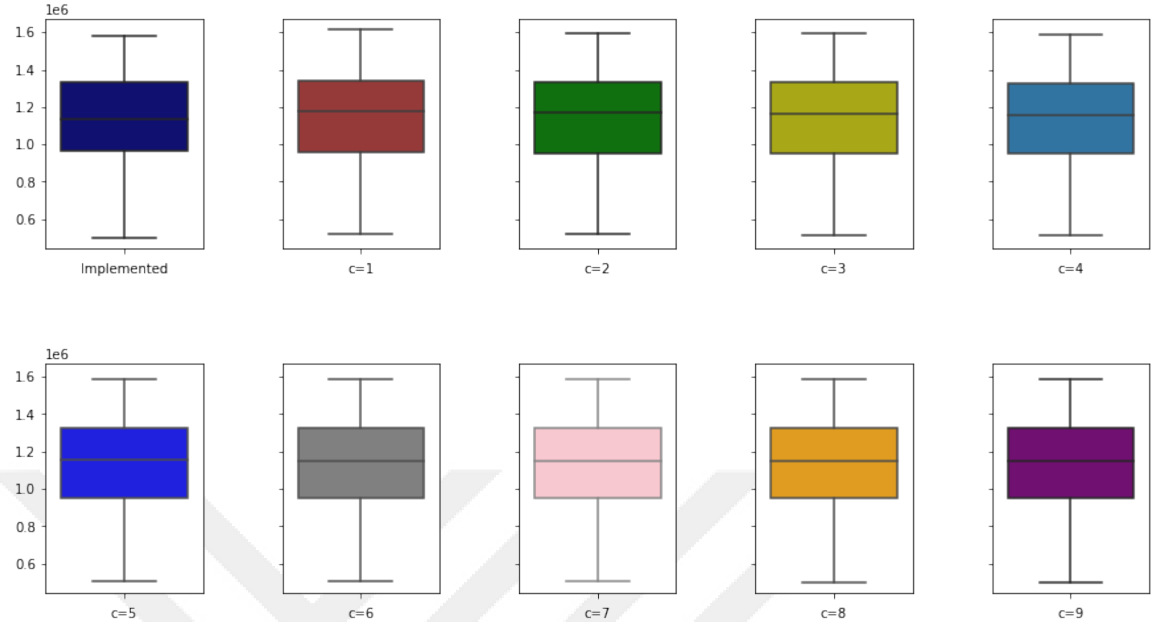

Figure 3: The SSE between the expected and the calculated demand values for G2 Size=10

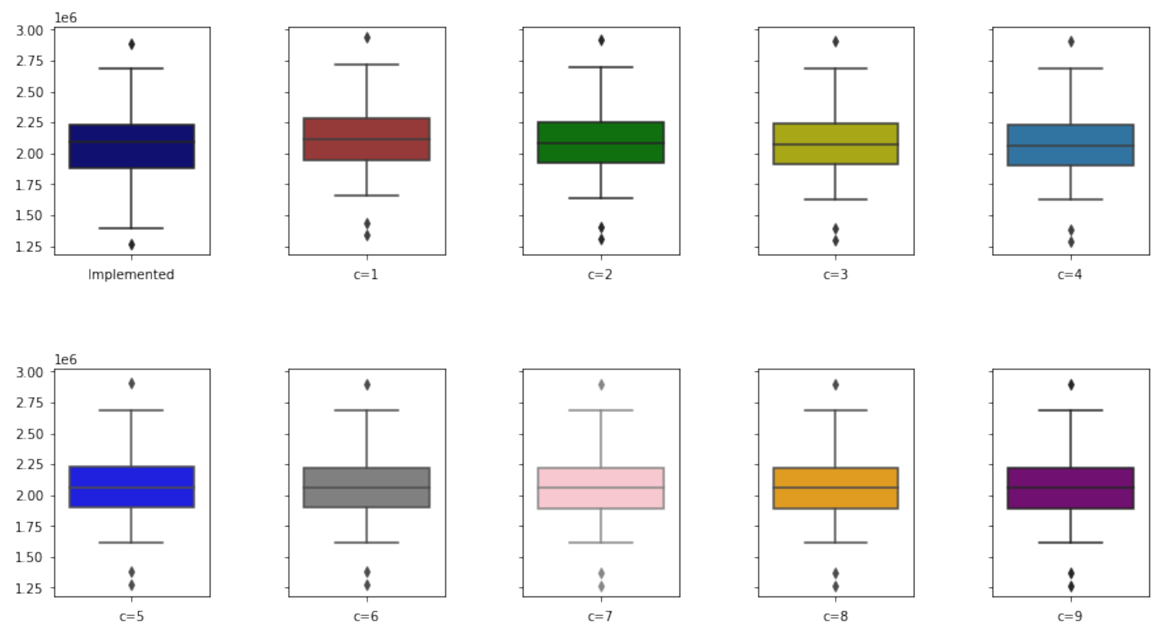

Figure 4: The SSE between the expected and the calculated demand values for G2 Size=20

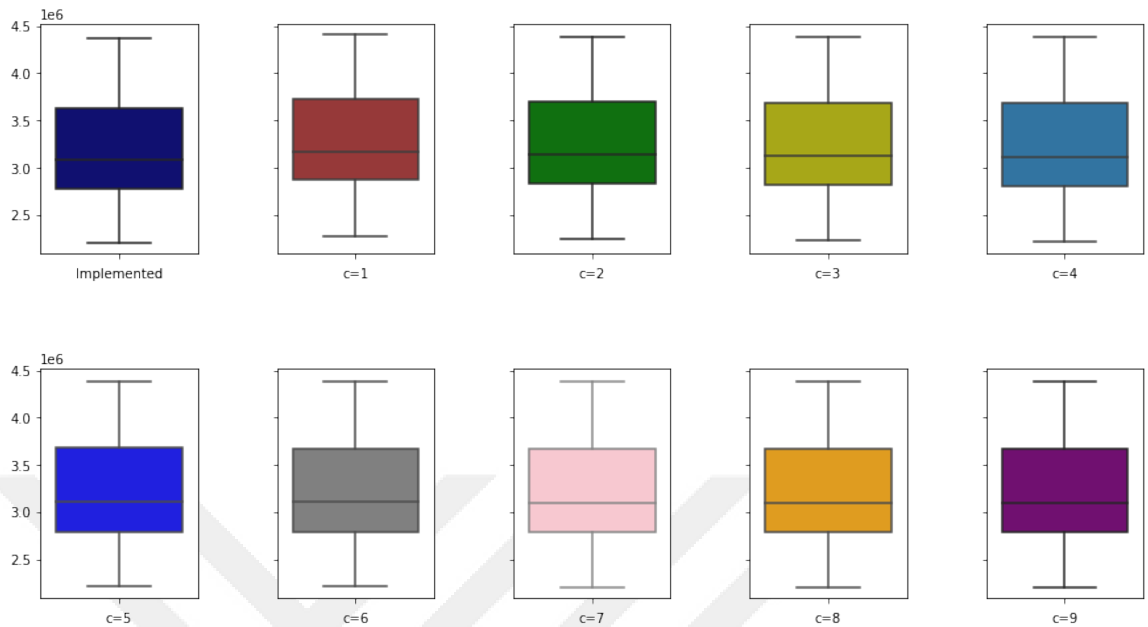

Figure 5: The SSE between the expected and the calculated demand values for G2 Size=30

In this section, we secondly present a representative example of the impact resulted from parameter c value on the δ_t' decision. As can be seen from Table 4 and Figure 6, the minimum price value for the relevant instance is 1.58 and the maximum price value is 3.83, and the optimal price found in the Integrated Solution result is 3.39. In order to see the variation of the probabilities according to the price, we present a review under four different prices, these are low, medium, high and optimal prices, respectively. We choose 1.80 for low price, for average price, the min-max average price for medium level 2.71, and 3.70 for high.

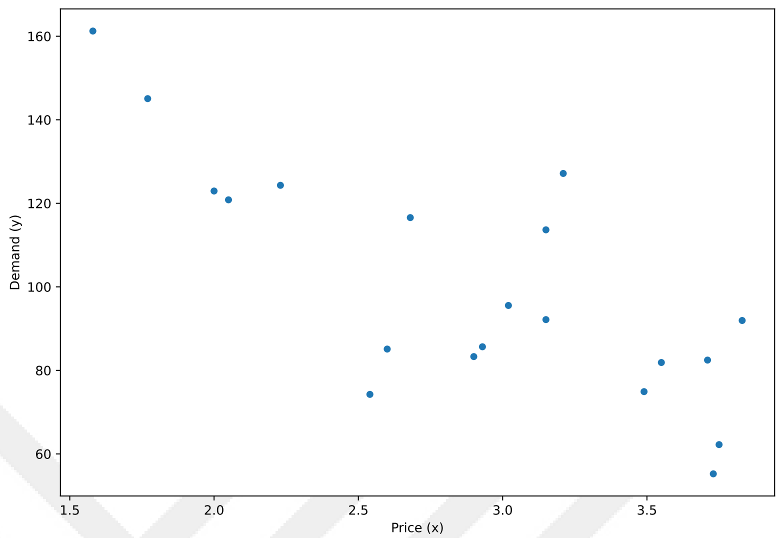

Figure 6: G1, Size=20 Demand-price relation

In Figure 7, we present the distribution of δ_t' values in the selection of low-mid-high and optimal prices. The x-axis represents the record data in Table 4, respectively, from left to right. In this case, for example, when the price is 1.80, the probabilities of the high-priced records are set to zero and these records have no effect on the model. Similarly, when the price is 3.70, the records below 2.70 get zero or near zero probabilities. On the other hand, in cases where the price is close to the average values within the price set (average prices), all the records take a probability value and the model considers each record under the condition of similarity to the determined one.

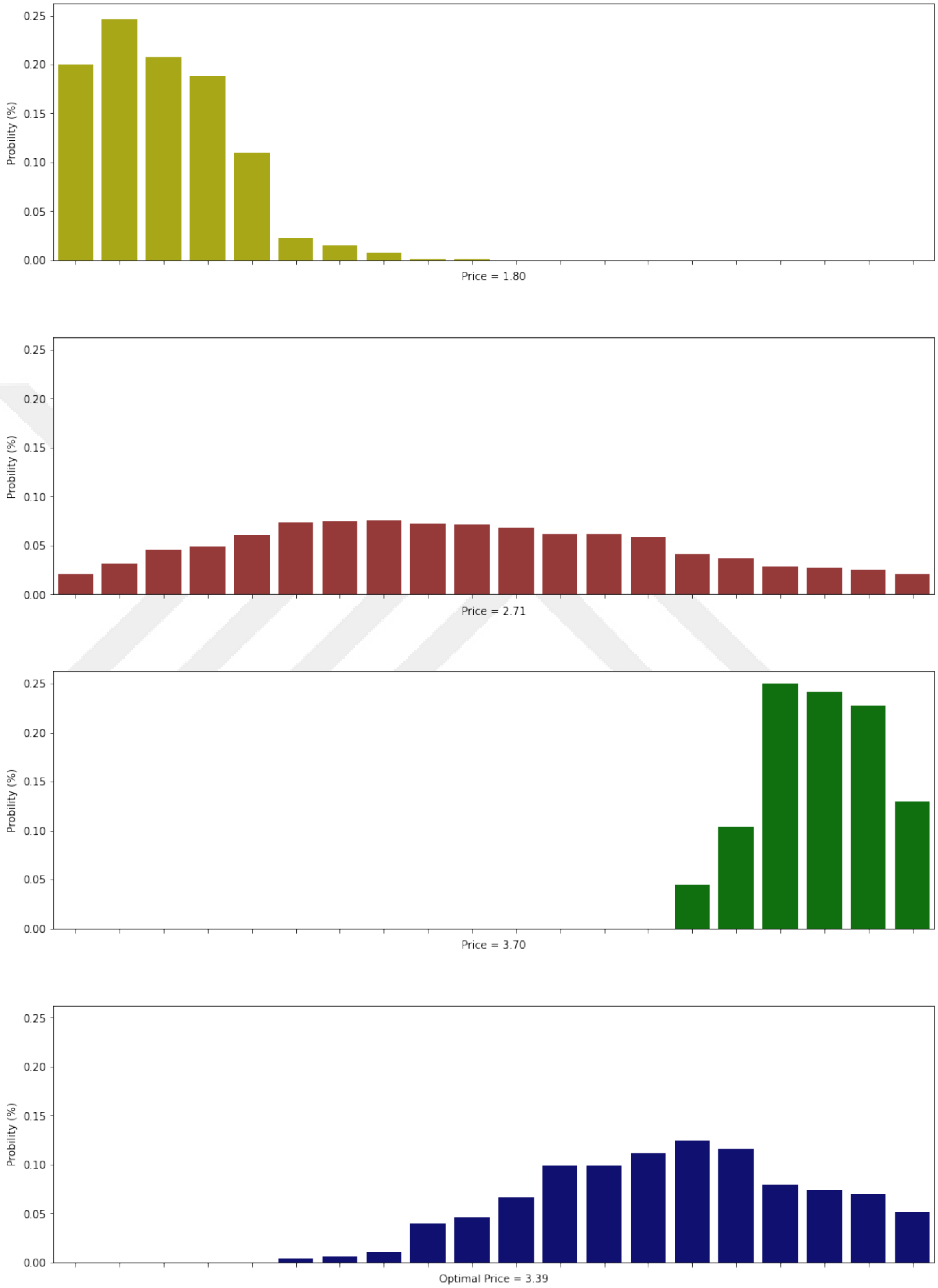

Figure 7: δ'_t probabilities of scenarios that vary by price

4.4.3 Numerical Results: Performance Comparison with Small and Large Datasets

In this section, we present the results of "Integrated Solution" and "Predict then Optimize", which we recommend for 5 different datasets created in the dataset generation section, together with Monte Carlo Simulation in order to understand the real performance and compare the results according to the groundtruth results. We test the generated datasets with two different threshold values (15% and 30%). For each size of the dataset, 30 runs are executed. In this study, the average cost is assumed as 1.0 and the salvage value is as 0.5. For the integrated solution, each run has a limit of 720 minutes via Baron solver.

First, we examine how many different instances the proposed models and groundtruth can enable feasible results in Table 7. Groundtruth shows that G2's 15% threshold value instances are feasible for a very low number of instances (6.33 on average). Apart from that, it is seen that there is a feasible result in an average of 19 out of 30 instances of G3 with 15% α value. It is understood that there are feasible results for more than 25 instances in most of the instances belonging to the datasets other than these.

Although the Integrated Solution provides results for many instances, the feasibility of the results after simulation varies. In this case, it is understood that for the 15% α value, the Integrated Solution has difficulty in finding feasible results in general and its solutions are largely ineffective. However, when the α value is 30%, there is a significant change in the feasibility performance of the Integrated Solution, and the post-simulation performance is close to the model performance in most of the datasets. At the same time, when the α value is 30%, the feasibility status of the Integrated Solution performs the same or in some cases better than the Predict then Optimize method.

			Groundtruth	Integrated Solution		Predict then Optimize	
Type	Threshold Probability	Size	Generated	Before Sim.	After Sim.	Before Sim.	After Sim.
G1	15%	10	25	30	4	25	11
		20	29	28	6	25	15
		30	29	30	9	28	17
	30%	10	30	29	26	28	23
		20	30	30	28	30	28
		30	30	28	27	30	26
G2	15%	10	11	30	0	9	4
		20	5	29	0	2	0
		30	4	27	0	2	1
	30%	10	26	29	14	22	16
		20	27	29	11	21	11
		30	29	29	14	24	17
G3	15%	10	19	29	2	18	4
		20	18	29	2	20	7
		30	20	29	6	19	5
	30%	10	29	30	21	30	20
		20	30	28	23	30	20
		30	30	26	21	30	23
G1-v2	30%	10	24	29	10	24	9
		20	28	29	10	26	13
		30	28	27	8	29	13
G1-v3	30%	10	30	30	30	30	30
		20	30	30	30	30	30
		30	30	30	30	30	30

Table 7: Feasibility status of created instances

In Table 8, we examine the differences of the two proposed methods according to the groundtruth solution and show how much the results obtained from the simulation deviate from the real world. The simulation setup operated with 100,000 different runs for each instance, but even so, as a result of randomness, negative values are seen in Table 8 with small differences. In these cases, it is seen that the results are the same as the groundtruth, but negative results can be obtained with small differences due to randomness.

In general, we see that both methods provide values close to groundtruth on average, in cases where they can remain feasible as a result of simulation. As the number of data increases for the Integrated Solution, it is seen that getting closer to the result of groundtruth in the average values improves. In case of having α as 15%, the average difference is 7% if the size is 10, and the difference diminishes to 2% if the size is increased to 30. A similar situation occurs when α is 30%, the average difference is 5% when the size is 10, and 4% when the size is 30.

The differences in instances where the results from the Integrated Solution are highest according to groundtruth are lower than those from Predict then Optimize. In this case, it is seen that there is a lower tendency to move away from the optimal result in cases where the Integrated Solution provides feasible results. It is also seen that the Integrated Solution performs better on average differences when the α is 15%, and the two proposed approaches have similar performance when the α is 30%.

In addition, the G1-v3 instance is very suitable to be explained with a linear line due to the low error term in its creation. Therefore, Predict then Optimize method has a great advantage for the G1-v3 dataset and the results show this advantage.

			Integrated Solution			Predict then Optimize		
Type	Treshold Probability	Size	Min (%)	Avg (%)	Max (%)	Min (%)	Avg (%)	Max (%)
G1	15%	10	0.02	0.06	0.13	0.00	0.09	0.29
		20	0.02	0.04	0.08	0.00	0.04	0.09
		30	0.00	0.03	0.13	0.00	0.07	0.24
	30%	10	0.00	0.04	0.13	0.01	0.05	0.24
		20	0.00	0.05	0.15	0.00	0.02	0.06
		30	0.01	0.03	0.08	0.00	0.03	0.09
G2	15%	10	-	-	-	0.08	0.11	0.18
		20	-	-	-	-	-	-
		30	-	-	-	0.08	0.08	0.08
	30%	10	0.01	0.04	0.24	0.00	0.05	0.18
		20	0.00	0.06	0.20	0.01	0.06	0.14
		30	0.01	0.05	0.21	0.01	0.03	0.14
G3	15%	10	0.03	0.08	0.12	-0.02	0.04	0.12
		20	0.07	0.13	0.19	0.01	0.08	0.15
		30	-0.09	-0.01	0.06	0.00	0.09	0.21
	30%	10	0.00	0.03	0.09	0.00	0.04	0.18
		20	0.01	0.06	0.32	0.01	0.05	0.25
		30	0.00	0.03	0.11	0.01	0.02	0.06
G1-v2	30%	10	-0.01	0.11	0.26	0.01	0.10	0.25
		20	0.02	0.09	0.18	0.01	0.10	0.20
		30	0.04	0.10	0.20	0.00	0.07	0.28
G1-v3	30%	10	0.00	0.04	0.11	0.00	0.01	0.04
		20	0.01	0.04	0.08	0.00	0.01	0.05
		30	0.01	0.04	0.15	0.00	0.00	0.01
Avg	15%	10	0.02	0.07	0.13	-0.02	0.08	0.29
		20	0.02	0.07	0.19	0.00	0.05	0.15
		30	-0.09	0.02	0.13	0.00	0.08	0.24
	30%	10	-0.01	0.05	0.26	0.00	0.04	0.25
		20	0.00	0.05	0.32	0.00	0.04	0.25
		30	0.00	0.04	0.21	0.00	0.03	0.28

Table 8: Comparison of objective values with respect to groundtruth for the proposed methods

Finally, we compare the performances of our two approaches, over the instances where they are feasible as a result of simulation. We present comparison results in two different ways. First of all, considering that the method providing a better result of objective values is also performing better, we present the number of instances in which better results occur. As can be seen from the "Exact %" column in Table 9, which result performs better in percentage in terms of dataset type and size is shown. In this case, it is indicated that Predict then Optimize provides better results for instances where α is 15%. For α 30%, although the results are similar, Predict then Optimize gives better results in 192 instances according to comparison in "Exact %" column, while Integrated Solution can only perform better in 96 instances.

In the "Interval %" column of Table 9, we aim to compare the two methods within a certain range instead of directly comparing them, and to reveal how the Integrated Solution differs from Predict then Optimize. In this context, Integrated Solution's performance is considered better if the result obtained by the Integrated Solution is greater than 95% of the result obtained from Predict then Optimize. When the results are analyzed from this point of view, it is seen that Integrated Solution provides better results in 234 of the instances while Predict then Optimize approach enables slightly better results in only 54 of them. Therefore, it is seen that the model results are quite close to each other and the majority of them are within 95% of each other.

			Integrated Solution		Predict then Optimize	
	Threshold Probability	Size	Exact % (100%)	Interval % (95%)	Exact % (100%)	Interval % (95%)
G1	15%	10	100.0	100.0	0.0	0.0
		20	25.0	100.0	75.0	0.0
		30	50.0	87.5	50.0	12.5
	30%	10	52.4	85.7	47.6	14.3
		20	30.8	65.4	69.2	34.6
		30	52.2	87.0	47.8	13.0
G2	15%	10	-	-	-	-
		20	-	-	-	-
		30	-	-	-	-
	30%	10	54.5	81.8	45.5	18.2
		20	55.6	77.8	44.4	22.2
		30	33.3	66.7	66.7	33.3
G3	15%	10	50.0	50.0	50.0	50.0
		20	50.0	50.0	50.0	50.0
		30	33.3	100.0	66.7	0.0
	30%	10	45.0	95.0	55.0	5.0
		20	33.3	77.8	66.7	22.2
		30	56.3	87.5	43.8	12.5
G1-v2	30%	10	16.7	50.0	83.3	50.0
		20	85.7	85.7	14.3	14.3
		30	28.6	57.1	71.4	42.9
G1-v2	30%	10	20.0	83.3	80.0	16.7
		20	0.0	80.0	100.0	20.0
		30	0.0	90.0	100.0	10.0
Total Record			96	234	192	54

Table 9: The ratio of providing better results where the proposed approaches are feasible

CHAPTER V

CONCLUSION

Many firms regardless of the area of work or their business models have two fundamental common problems. The first problem is to to decide production or order quantity of the products that have certain lifetime before the selling period. In this problem, if demand realizes lower than the order quantity of a product, then the firm-owner will experience underage costs. Reciprocally, if demand is higher than the quantity, the firm will incur overage cost. Therefore, trade-off and any wrong judgement between producing or ordering too much or too little is a crucial decision for the future of the business. The second common problem accompanying deciding the optimal order quantity is to determine the right price for a product since this joint decision of quantity and price can bring about the profit, which is the main objective to focus on for every business.

The classical newsvendor problem is one of the most famous example of such inventory management problems in the literature. Ultimate aim of the newsvendor problem is to decide optimal order quantities for products. In the traditional approach, a retailer for example, assumes that there is a specific distribution of demand derived from the historical sales data. Despite extensive appliances of such theoretical newsvendor models, hypothesis of a certain statistical distribution of demand does not hold water in real-life examples. Therefore, approaching the problem in a price-setting scenario where the related price-dependent demand function is a random variable with an unknown probability distribution is more challenging yet leading much more accurate results.

Although data-driven, price-setting approach to newsvendor problem is multifaceted and can lead to better accuracy, conducting inventory management in such setting is complex and risky. For example, lack of threshold for profit may result in fluctuating profit through terms, which is not consistent and sustainable for the firms. Therefore, in such approach, the fundamental responsibility of the decision-makers is to set a balance between profitability and financial risks. For this reason, in addition to price-setting approach, including another dimension as being risk-averse, can create a more useful synergy for the firms as maximizing profit and reducing risks.

The aim of this paper is to approach newsvendor problem in a scenario where demand is completely unknown; therefore the whole work is data-driven. The point that differentiates our work from the example studies in literature is that the model is not only a price-setting newsvendor problem but also provides a risk-averse approach, where the business-owner has opportunity to determine a target profit and a threshold probability of risk for not to gain less than the target level. In this regard, Integrated Solution using MINLP and Predict then Optimize algorithms are used.

According to analyses, Integrated Solution has shown valuable performance on different datasets without including any prediction methods and learning a certain price-demand relationship, and in many respects, it outperforms the Predict then Optimize method, which is hard to beat. The Integrated Solution provides a similar or better result than the Predict then Optimize method on most datasets when considering average differences with respect to groundtruth and yielding feasible results. It provides results only 2% away from the optimal value when α is 15% and size is 30.

REFERENCES

[1] P. Kouvelis and R. Li, “Integrated risk management for newsvendors with value-at-risk constraints,” *Manufacturing & Service Operations Management*, vol. 21, no. 4, pp. 816–832, 2019.

[2] C. C. Wei and L. T. Chen, “Supply chain replenishment decision for newsvendor products with multiple periods and a short life cycle,” *Sustainability*, vol. 13, no. 22, p. 12777, 2021.

[3] T. Wang, “The newsvendor problem with advertising,” in *IEEE International Conference on Service Operations and Logistics, and Informatics*, vol. 1, pp. 886–889, IEEE, 2008.

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

[5] J. Huber, S. Müller, M. Fleischmann, and H. Stuckenschmidt, “A data-driven newsvendor problem: From data to decision,” *European Journal of Operational Research*, vol. 278, no. 3, pp. 904–915, 2019.

[6] J. R. Trapero, M. Cardos, and N. Kourentzes, “Empirical safety stock estimation based on kernel and garch models,” *Omega*, vol. 84, pp. 199–211, 2019.

[7] S. Lin, Y. Chen, Y. Li, and Z.-J. M. Shen, “Data-driven newsvendor problems regularized by a profit risk constraint,” *Production and Operations Management*, vol. 31, no. 4, pp. 1630–1644, 2022.

[8] A. Shapiro, “Monte carlo sampling methods,” *Handbooks in operations research and management science*, vol. 10, pp. 353–425, 2003.

[9] A. J. 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.

[10] 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.

[11] G.-Y. Ban, J. Gallien, and A. J. Mersereau, “Dynamic procurement of new products with covariate information: The residual tree method,” *Manufacturing & Service Operations Management*, vol. 21, no. 4, pp. 798–815, 2019.

[12] A.-L. Beutel and S. Minner, “Safety stock planning under causal demand forecasting,” *International Journal of Production Economics*, vol. 140, no. 2, pp. 637–645, 2012.

[13] P. Harsha, R. Natarajan, and D. Subramanian, "A prescriptive machine-learning framework to the price-setting newsvendor problem," *Informs Journal on Optimization*, vol. 3, no. 3, pp. 227–253, 2021.

[14] C. Li and M. Lu, "Joint price and inventory optimization under minimax regret," *Available at SSRN 3060847*, 2017.

[15] G.-Y. Ban and C. Rudin, "The big data newsvendor: Practical insights from machine learning," *Operations Research*, vol. 67, no. 1, pp. 90–108, 2019.

[16] A. Özler, B. Tan, and F. Karaesmen, "Multi-product newsvendor problem with value-at-risk considerations," *International Journal of Production Economics*, vol. 117, no. 2, pp. 244–255, 2009.

[17] E. Luciano, L. Peccati, and D. M. Cifarelli, "Var as a risk measure for multiperiod static inventory models," *International Journal of Production Economics*, vol. 81, pp. 375–384, 2003.

[18] W. Jammernegg and P. Kischka, "Newsvendor problems with var and cvar consideration," in *Handbook of Newsvendor Problems*, pp. 197–216, Springer, 2012.

[19] X. Gan, S. P. Sethi, and H. Yan, "Coordination of supply chains with risk-averse agents," *Production and Operations Management*, vol. 13, no. 2, pp. 135–149, 2004.

[20] G. Raz and E. L. Porteus, "A fractiles perspective to the joint price/quantity newsvendor model," *Management Science*, vol. 52, no. 11, pp. 1764–1777, 2006.

[21] D. M. Gay, "The ampl modeling language: An aid to formulating and solving optimization problems," in *Numerical Analysis and Optimization*, pp. 95–116, Springer, 2015.

[22] N. V. Sahinidis, *BARON 17.8.9: Global Optimization of Mixed-Integer Nonlinear Programs,* User's Manual, 2017.

[23] J. Czyzyk, M. P. Mesnier, and J. J. Moré, "The neos server," *IEEE Computational Science and Engineering*, vol. 5, no. 3, pp. 68–75, 1998.

VITA

Eren Atsız is an M.Sc. candidate in Data Science at Özyeğin University under the supervision of Assistant Professor Enis Kayış and Assistant Professor Erinç Albey. He graduated from Beşiktaş Atatürk Anatolian High School and he holds a B.Sc. degree in Industrial Engineering with a minor degree in Computer Science from Özyeğin University. His primary research interests lie at the intersection of optimization and machine learning, as his ultimate vision is to create approaches toward multifaceted problems with the usage of accurate and interpretable models, including state-of-the-art analytics methods.