# ALTERNATIVE MATHEMATICAL MODELS FOR REVENUE MANAGEMENT PROBLEMS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ERMAN TERCİYANLI

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
INDUSTRIAL ENGINEERING

JULY 2009

Approval of the thesis:

**ALTERNATIVE MATHEMATICAL MODELS FOR REVENUE MANAGEMENT PROBLEMS**

submitted by **ERMAN TERCİYANLI** in partial fulfillment of the requirements for the degree of
**Master of Science in Industrial Engineering Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen ________________
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Nur Evin Özdemirel ________________
Head of Department, **Industrial Engineering**

Assist. Prof. Dr. Zeynep Müge Avşar ________________
Supervisor, **Industrial Engineering Department, METU**

**Examining Committee Members:**

Assoc. Prof. Dr. Esra Karasakal
Industrial Engineering Dept., METU ________________

Assist. Prof. Dr. Zeynep Müge Avşar
Industrial Engineering Dept., METU ________________

Assist. Prof. Dr. Serhan Duran
Industrial Engineering Dept., METU ________________

Assist. Prof. Dr. Ayşe Kocabıyıkoğlu
Faculty of Business Administration, Bilkent University ________________

Assist. Prof. Dr. Banu Yüksel Özkaya
Industrial Engineering Dept., Hacettepe University ________________

**Date:** ________________

**I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.**

Name, Last Name: ERMAN TERCİYANLI

Signature :

# ABSTRACT

ALTERNATIVE MATHEMATICAL MODELS FOR REVENUE MANAGEMENT PROBLEMS

Terciyanlı, Erman
M.S., Department of Industrial Engineering
Supervisor : Assist. Prof. Dr. Zeynep Müge Avşar

July 2009, 161 pages

In this study, the seat inventory control problem is considered for airline networks from the perspective of a risk-averse decision maker. In the revenue management literature, it is generally assumed that the decision makers are risk-neutral. Therefore, the expected revenue is maximized without taking the variability or any other risk factor into account. On the other hand, risk-sensitive approach provides us with more information about the behavior of the revenue. The risk measure we consider in this study is the probability that revenue is less than a predetermined threshold level. In the risk-neutral cases, while the expected revenue is maximized, the probability of revenue being less than such a predetermined level might be high. We propose three mathematical models to incorporate the risk measure under consideration. The optimal allocations obtained by these models are numerically evaluated in simulation studies for example problems. Expected revenue, coefficient of variation, load factor and probability of the poor performance are the performance measures in the simulation studies. According to the results of these simulations, it shown that the proposed models can decrease the variability of the revenue considerably. In other words, the probability of revenue being less than the threshold level is decreased. Moreover, expected revenue can be increased in some scenarios by using the proposed models. The approach considered in this thesis is especially proposed for small scale airlines because risk of obtaining revenue less than the threshold level is more for this type of airlines as compared to large scale airlines.

Keywords: Revenue management, Seat inventory control, Mathematical programming, Risk

# ÖZ

GELİR YÖNETİMİ PROBLEMLERİ İÇİN ALTERNATİF MATEMATİKSEL MODELLER

Terciyanlı, Erman
Yüksek Lisans, Endüstri Mühendisliği Bölümü
Tez Yöneticisi : Yrd. Doç. Dr. Zeynep Müge Avşar

Temmuz 2009, 161 sayfa

Bu çalışmada, havayolu ağları için koltuk envanter kontrolü problemi riskten kaçınan bir karar vericinin bakış açısıyla incelenmektedir. Gelir yönetimi literatüründe, genelde karar vericilerin risk-nötr olduğu varsayılmaktadır. Bundan dolayı, değişkenlik ya da başka bir risk faktörü göz önüne alınmadan beklenen gelir için en yüksek değer bulunmaya çalışılmaktadır. Öte yandan, riske duyarlı yaklaşım gelirin davranışı hakkında daha fazla bilgi sağlamaktadır. Bu çalışmada, gelirin belirlenmiş bir eşik değerinden düşük olması olasılığı risk ölçüsü olarak kullanılmaktadır. Risk-nötr durumlarda, beklenen gelir için en yüksek değer elde edilirken gelirin belirlenmiş bir değerden düşük olması olasılığı yüksek olabilmektedir. Bu çalışmada, belirtilen risk ölçüsünü dikkate alarak üç matematiksel model önerilmiştir. Örnek problemler için, modellerden elde edilen optimal dağıtımlar simülasyon çalışmalarında sayısal olarak değerlendirilmiştir. Beklenen gelir, değişkenlik katsayısı, yük faktörü ve kötü performans olasılığı simülasyon çalışmalarında kullanılan performans ölçütleridir. Bu simülasyon sonuçlarına göre, önerilen modellerin gelirdeki değişkenliği azaltabildiği gösterilmiştir. Başka bir deyişle, gelirin eşik değerinden düşük olma olasılığı azaltılmıştır. Bunun yanında, bazı senaryolarda beklenen gelir önerilen yöntemler kullanılarak arttırılabilmiştir. Eşik değerinin altında gelir elde etme riski küçük ölçekli havayolu şirketlerinde büyük ölçekli şirketlere oranla daha yüksek olduğundan, bu tezde çalışılan yaklaşım özellikle küçük ölçekli havayolu şirketleri için önerilmektedir.

Anahtar Kelimeler: Gelir yönetimi, Koltuk stok kontrolü, Matematiksel programlama, Risk

*To my family*

# ACKNOWLEDGMENTS

This thesis would not have been possible without the support of many people. First and foremost I would like to express my gratitude to my supervisor Assist. Prof. Dr. Zeynep Müge Avşar for her endless support, guidance and encouragement. During this study, I have learned a lot from her.

I would also like to thank the members of the examining committee; Assoc. Prof. Dr. Esra Karasakal, Assist. Prof. Dr. Serhan Duran, Assist. Prof. Dr. Ayşe Kocabıyıkoğlu and Assist. Prof. Dr. Banu Yüksel Özkaya for their advices and contributions to this study.

I would like to thank Demet Çetiner for valuable contributions and her guiding study, which draws a road map for this study. I am also grateful to Ahmet Melih Selçuk for his suggestions during the study.

Deepest gratitude are also due to my colleagues in TÜBİTAK SPACE Power Electronics Group. I would like to send my special thanks to Prof. Dr. Muammer Ermiş and Prof. Dr. Işık Çadırcı for their valuable support during my study.

Until the start of my undergraduate study, he was always near me with his endless support. Thanks to my brother Alper Terciyanlı.

Lastly, I would like to express my gratitude to my family for their continuous love and support throughout my life.

# TABLE OF CONTENTS

# LIST OF TABLES

TABLES

# LIST OF FIGURES

FIGURES

# CHAPTER 1

# INTRODUCTION

Revenue management (RM), also called as Yield Management (YM), is basically defined as a tool for maximizing revenue by using demand management decisions. The most common definition for RM in airline industry is due to American Airlines :*"Selling the right seats to the right people at the right time"*. More generally Pak and Piersma (2002) define revenue management as follows: *"The art of maximizing profit generated from a limited capacity of a product over a finite horizon by selling each product to the right customer at the right time for the right price"*.

Revenue management history starts with the deregulation of the airline industry in USA in 1970s. Therefore, airline industry is the main area where revenue management is successfully applied. Moreover, RM can be used in most of the industries where demand management decisions are critical as in the cases of hotels, car rental, retailing, media and broadcasting, natural gas storage and transmission, electricity generation and transmission, air cargo, theaters, sporting events and restaurants.

This thesis deals with the revenue management applications in airline industry. The applicability of revenue management in airline industry results from the following typical characteristics of the sector. First of all, the product is perishable; unsold seats at the departure time of the flight cannot be sold later. Secondly, the profit is maximized by finding the right combination of the passenger types on the flight. Since operating costs of a flight, such as airport costs, fuel costs, and personnel costs, are much higher than the total marginal costs of passengers, it is assumed that marginal cost of a passenger is zero. That is, objective of the airline revenue management problem is maximization of the revenue for the flights that are already scheduled. Scheduling of the flights is a structural decision for the airline companies

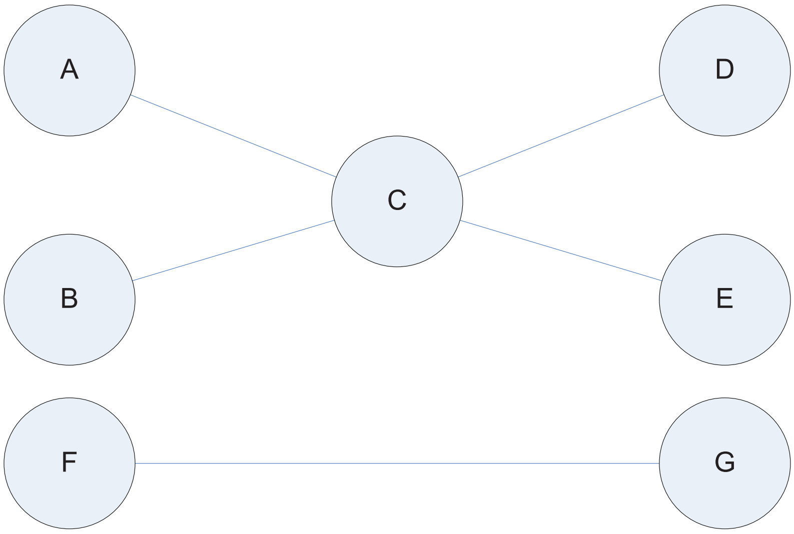


Figure 1.1: An example airline network.

and out of the scope of revenue management.

Briefly, revenue management problem in airline industry is defined as managing the capacities of the flights in a network, where flights can be in a connecting or local traffic. An example graph that contains a connecting network and a local traffic is given in Figure 1.1. The traffic between nodes F and G is called local traffic and the traffic among nodes A to E is called connecting network traffic. The nodes in this graph are the locations of the airports and they are connected by the flights shown with arcs. Node C in this graph is called as *hub*. A hub is an airport that an airline uses as a transfer point to get passengers to their intended destination. A journey between an origin node and a destination node is called an *itinerary* (*OD*). A product is an origin-destination-fare combination, that is abbreviated as *ODF*. In *single-leg* flights, because of the uniqueness of the origin-destination pair, the products are defined for the fare classes. On the other hand, because of the increasing number of origin-destination pairs, *network* traffic is more complex.

In airline revenue management, *fare classes* are determined according to the segments of customers. Customer segments differ according to the characteristics such as type of the customers and condition of the tickets. Customer types are generally considered in two groups; leisure traveler and business traveler. That is, customer types are not only classified according to locations of the seats on the aircrafts such as first, business and economy classes, but also according to the characteristics of the customers such as the arrival time of the demand. Leisure travelers, generally tend to arrive earlier than the business travelers. The conditions of the tickets are also important for defining customer segments. These are basically called as options for cancellation, overnight stay, refund and advance purchase option. Low price tickets *(discounted fares)* are offered for leisure travelers to attract them at the beginning of the booking horizon and increase the capacity used. The *booking horizon*, or *booking period*, is the period of time in which booking is available. High fare tickets *(full fare tickets)* are offered to the business travelers and they generally have some options such as cancellation and partial or full refund. The *load factor* which is defined as the ratio of seats filled on a flight to the total number of seats available would be increased in a network by offering discounted fares. However, revenue per passenger may be increased by offering full fare tickets. Hence, the performance of a network is not only measured by the earned revenue, but also by load factor and revenue per passenger. In order to decrease the risk of empty departure, overbooking is considered. *Overbooking* means selling more seats than the capacity. Overbooking is used for coping with the uncertainties regarding the sold tickets such as cancellation and no-show. *Cancellation* is an option for customer and the customer with that option can cancel the tickets and get a refund of partial or full fare. Moreover, some of the customers do not arrive at the time of departure without cancellation, which is called *no-show*. A revenue management glossary due to McGill and van Ryzin (1999) is given in the Appendix for clarifying the definitions of the terms in revenue management that may have different meanings in more general contexts.

There are two approaches for revenue management in airline industry : *capacity allocation*, also called as *seat inventory control*, and *dynamic pricing*. The corresponding classification due to Talluri and van Ryzin (2005) is as follows: *Quantity based RM* and *Price based RM*. In seat inventory control policy, the decision maker, who has the responsibility for determining the capacity allocations and/or fares makes the decisions of accepting or rejecting the ticket requests. There are multiple tickets that differ in options and fares. Availability of these tickets for the customer changes over booking horizon. At the beginning of the booking horizon, most of the tickets are open for sale and they are closed as the departure time of the flight gets closer. On the other hand, dynamic pricing offers only one product with a price that changes over time. Both of these approaches are used in real life according to the characteristics of the sector and the distinction between the approaches is not always sharp. Increasing the price of a good is not so different than closing a discounted class. In the sectors where changing prices is costly, it is logical to use capacity allocation instead of dynamic pricing. The cost of changing prices is due to the operational costs such as announcing them to the customers in print media or tariff books. In some sectors, firms have more price flexibility than quantity flexibility. In online retailing, the cost of changing prices is nearly zero and dynamic pricing is used in order to manage inventories and revenues. On the other hand, in restaurants, using dynamic pricing causes updating price lists frequently, which is impractical and costly. This situation is not so different for airline industry. Although some firms use dynamic pricing, most of the airline companies announce their prices over a given time interval and do not update them frequently because of the advertising reasons. Moreover, in seat inventory control, the only variable that must be stored and announced is the status of the product that is open or close. Because of the easy implementation of seat inventory control policy, it is widely used in the airline industry and this thesis focuses on seat inventory control policies.

Seat inventory control is considered under two headings depending on the number of flight legs. *Single-Leg Seat Inventory Control* is used to maximize revenue for a direct flight between an origin-destination pair, which is generally isolated from the other flights in the network. The main disadvantage of Single-Leg Seat Inventory Control is optimizing the booking limit locally, whereas in real life, companies want to maximize revenue for the whole network. *Network Seat Inventory Control* deals with all of the legs in a network simultaneously. The main disadvantage of the Network Seat Inventory Control is the complexity of the problem which increases as network gets larger.

The main approach in seat inventory control is setting booking limits and protection levels for each fare class in order to maximize revenue. The number of seats that are protected for a high-fare class and not available for low-fare classes is called the *protection level*. On the other hand, *booking limit* is defined as the number of seats that are allocated to a specific fare class. A control policy is called *partitioned booking limit control policy* when booking limits are used in such a way that each fare class has a separate booking limit. In other words, a seat that is allocated for a fare class cannot be booked for another class. The aircraft departs with empty seat when demand is lower than the booking limit in partitioned booking limit control policy. Another policy is *nested booking limit control policy* such that fare classes are ordered according to some criteria and seats that are available for a low ranked fare class are also available for a high ranked one. Nested control policy gives higher revenues and load factors than the partitioned policy.

*Bid price control policy* is another common policy for seat inventory control in airline revenue management problems. In bid price control policy, a request is accepted if fare of the class exceeds the opportunity cost of the corresponding itinerary. The opportunity cost of an itinerary is defined as the expected loss in the future revenue from using the capacity now rather than reserving it for future use. It is approximately calculated by summing the bid prices of the flight legs the itinerary uses. A *bid price* is the net value for an incremental seat on a particular flight leg in the airline network. The difference between bid price and opportunity cost is generally not clear. Although sum of bid prices of the flight legs the itinerary uses is an approximation for opportunity cost, there need not be one-to-one correspondence between the optimal bid prices and the opportunity costs. Talluri and van Ryzin (2005) explains this situation with an example. Consider a single-resource problem in which high-revenue products arrive before low-revenue products. In this case, the optimal bid price is zero. On the other hand, the opportunity cost at each point in time $t$ will in general not be zero. The main difference in bid price control policy compared to booking limit control policies is that the class is open without any limit when the fare of the $ODF$ exceeds the opportunity cost of the corresponding itinerary. The main advantage of the bid price approach is that it is very easy to implement. It requires only a comparison between bid prices of the legs and fare of the $ODF$.

The mathematical models developed for airline network revenue management problems are classified into two groups according to the assumptions for demand behavior: deterministic and stochastic models. Deterministic models assume that demand for a particular $ODF$ is equal to the expected value. On the other hand, in probabilistic models, probabilistic nature of the demand is incorporated into the models.

In the RM literature, it is generally assumed that the decision makers are risk-neutral. Therefore, the expected revenue is maximized without taking the variability or any other risk factor into account. Levin et al. (2008) state that the long term average revenues will be maximized as long as good risk-neutral strategies are employed because of the law of large numbers. In real life networks, there are hundreds or thousands of successive flight departures in a year and the impact of the revenue of a single case (a flight for a single-leg or network traffic) on the gross revenue is not severe. However, it is also important to manage the demand in the short term by incorporating the risk factors, especially for the small sized airlines or new flights. Small sized airlines change their routes according to the changes in demand especially because of the seasonal effects. In these cases, the number of flight departures in a year is small according to general network RM problems. Additionally, risk factors can be included in the revenue management problems for new flights with higher risks as compared to the existing ones.

In this thesis, a network seat inventory control problem is considered from the perspective of a risk-sensitive decision maker. This risk-sensitive approach provides us with more information about the behavior of the revenue. As it is given in the previous paragraph, the studies with risk-neural cases only aim to maximize the total expected revenue. This objective is not sufficient when the decision maker desires to keep the total revenue higher than a predetermined level. In the risk-neural cases, the expected revenue might be high, but the probability of revenue being less than that predetermined level might be high. This predetermined level is called the *threshold* level throughout the thesis. In contrary to the risk-neutral studies in the literature, our approach solves the dilemma between expected revenue and the probability of revenue being less than a predetermined level. In this study, overbooking, cancellation and no-show are not allowed. It is assumed that customers make their decisions for the classes that they request and a shift between classes does not occur. Moreover, customers arrive sequentially, which means batch booking is not allowed. The probabilistic nature of the demand is taken into account in the proposed models. In the real life problems, it is likely that the decision makers have a revenue threshold level and they want to minimize the probability that the revenue is less than this level. This situation is considered in this study. The risk factor is used in the proposed mathematical programming models by minimizing or limiting the probability that the revenue is less than a threshold level.

Three probabilistic mathematical models are proposed in this thesis. These are called *SLP-RM*, *PMP-RC* and *RRS*. *SLP-RM* and *RRS* are linear programming formulations, but *PMP-RC* is an integer programming formulation. There are two *SLP-RM* models that are used successively. *SLP-RM-1* minimizes probability that the revenue is less than a threshold level for a number of sample demands. *SLP-RM-2* model maximizes the expected revenue by using the output of the *SLP-RM-1* model. *PMP-RC* maximizes the expected revenue with an additional constraint on the probability that the revenue is less than a threshold level. *RRS* model maximizes the revenue by solving the model many times for different realizations of demand. The methodology we propose for the use of the *RRS* model is as follows: the bid price for an itinerary is approximated by taking the average of the bid prices for the instances with revenue less (or more) than the threshold level. It is assumed that the risk-sensitive (risk-taking) decision makers use the average of the bid prices for the instances with the revenue less (more) than the threshold level.

Although the concept of risk-sensitivity is extensively studied in the literature for different types of problems, there are only a few studies for risk-sensitive approaches in revenue management. These studies can be classified into two groups according to the types of the problems: pricing problems and seat inventory control problems. The dynamic pricing studies are similar to the ones proposed for general inventory models. In these studies, risk is formulated in the objective functions in order to find a price for the good throughout the selling horizon. In a recent study, Levin et al. (2008) introduce a risk measure by augmenting the expected revenue with a penalty term for the probability that total revenues fall below a desired level of revenue. This risk measure is equal to the probability that the total revenues fall below a threshold level and similar to one that is used in this thesis. Levin et al. (2008) propose this approach for optimal dynamic pricing of perishable services or products. There are only a few studies in the literature that incorporate risk-aversion into the classical seat inventory control problem for airline industry. Weatherford (2004) proposes a new concept called expected marginal seat utility *(EMSU)*. The *EMSU* is based on the expected marginal seat revenue *(EMSRa)* heuristic introduced by Belobaba (1989). In *EMSU*, the revenue gained from a ticket is substituted by the utility of its revenue. Barz and Waldmann (2007) extend the static and dynamic models for single-leg revenue management problem to introduce the risk using an exponential utility function. For seat inventory control, Çetiner (2007) proposes two mathematical programming models by including the variance of the revenue in addition to the expected revenue. To conclude, this thesis is the first study in the RM literature that uses seat inventory control for risk-sensitive cases without the need of estimating hardly defined parameters, which simplifies the implementation and the decision making procedure. In the previous studies, the estimation of utility function parameters and penalty parameters for variances are not straightforward. In our proposed approach, the only parameter that must be estimated by the decision maker is the threshold level for revenue. The main disadvantage of the *SLP-RM* and *PMP-RC* models are the computational complexities of the models.

The organization of the thesis is as follows: In Chapter 2, the related literature is reviewed for single-leg and network RM problems. Chapter 3 presents the network seat inventory models and control policies in detail. The alternative models and the control techniques we propose for the RM problems for risk aversion are in Chapter 4. Chapter 5 is devoted to the simulation models and estimation of the parameters. Chapter 6 is on the numerical analyses and comparisons of the proposed approach with the existing approaches in the literature. The thesis ends with concluding remarks and suggestions for future research in Chapter 7.

# CHAPTER 2

# LITERATURE REVIEW

RM problems can be classified into two groups: *Single-Leg Seat Inventory Control* and *Network Seat Inventory Control*. Single-leg seat inventory control is at a single flight-leg level. On the other hand, network seat inventory control optimizes the complete network that contains more than one leg simultaneously. Single-leg problems are covered in an important part of the literature up to 1990s. By the change in the structure of airline industry from single-leg flights to network traffics and the development in computational capacities to solve more complicated mathematical models, detailed studies on network RM have started both at the airline companies and at the research institutes. In this chapter, firstly the related studies on single-leg seat inventory control are summarized and, then, the studies on network seat inventory control are reviewed.

## 2.1 An Overview of Single-Leg Seat Inventory Control

Single-leg seat inventory control is the first problem studied in airline revenue management. The problem is allocating the seats on a single-leg flight to different types of customers. This problem is considered as static or dynamic depending on the assumption used for the arrival of customers. In static case, customer classes are assumed to arrive sequentially: a low-fare customer books earlier than all of the passengers from the classes with higher fares. In dynamic single-leg control, such an assumption is not made. All of the assumptions in static single-leg seat inventory control are listed by McGill and van Ryzin (1999) as follows: 1) booking classes are sequential; 2) low-before-high arrival pattern; 3) statistically independent demands of booking classes; 4) no cancellation, no-show and overbooking; 5) no batch booking; 6) single flight leg.

The first study for the use of mathematical models in airline industry to maximize expected revenue is due to Littlewood (1972). The main assumption he considers for the arrival of booking classes is the *low-before-high fare* booking. By using all of the six assumptions above for only two classes, the rule for accepting low-fare passengers is given by Littlewood (1972) as follows:

$$f_2 \geq f_1 P(D_1 \geq x),$$

where $f_i$ is the fare for class $i = 1, 2$ and $f_1 \geq f_2$. $P(\cdot)$ is the probability of the event under consideration. $D_1$ is the random variable denoting the total demand for class 1. $x$ is the number of seats allocated to class 1 and $P(D_1 \geq x)$ is the probability of selling all reserved seats for class 1. Therefore, basically $f_1 P(D_1 \geq x)$ is the expected marginal revenue of the $x^{th}$ seat reserved for class 1. It is also obvious that $P(D_1 \geq x)$ is a monotonically decreasing function of $x$ and so is $f_1 P(D_1 \geq x)$ . Now, optimal seat inventory control policy for a single-leg flight is determined by finding the smallest $x$ value such that $f_2 \geq f_1 P(D_1 \geq x)$. This $x$ equals to the number of seats protected for the high fare class and is known as the protection level. In other words, the demand request of a low fare customer is accepted as long as the remaining capacity of the flight is higher than this value.

Mayer (1976) extends Littlewood's work by using a simulation study and updating the rule more than once during the booking horizon before departure when low-before-high arrival assumption is relaxed. Belobaba (1987) develops a heuristic which is called Expected Marginal Seat Revenue ($EMSRa$) for maximizing flight leg revenues of multiple fare class inventories. In this heuristic, the protection levels for each higher class $i$ over lower class $j$, $S^i_j$, is the smallest value that satisfies the following equation.

$$f_i P(D_i \geq S^i_j) \leq f_j.$$

The total protection level for the $n-1$ highest fare classes, $\Pi_{n-1}$, is calculated by summing $n-1$ protection levels as follows:

$$\Pi_{n-1} = \sum_{i=1}^{n-1} S^i_n.$$

Hence, a request for a low fare class is accepted when the remaining capacity of the flight is higher than the total protection level for the fare classes that have higher fares than the requested class. $EMSRa$ heuristic is optimal only for two fare classes, but practical also for multiple fare class problems. The simulation studies of McGill (1989) and Wollmer (1992) show that the mean revenue from the seat inventory control policy of $EMSRa$ method is very close to that of the optimal policy for multiple fare class problems. However, a later study due to Robinson (1995) shows that the performance of the $EMSRa$ heuristic depends on the demand distribution and gives poor results for more general demand distributions.

Curry (1990), Wollmer (1992) and Brumelle and McGill (1993) propose alternative methods for obtaining optimal booking limits for the single-leg problems under the six assumptions above due to McGill and van Ryzin (1999). Curry (1990) also relaxes the assumption of single-leg flight and proposes an approximate model for network seat inventory control by assuming a continuous demand distribution. Wollmer (1992) presents a model with a discrete real life demand data to find booking limits for the fare classes. In the study of Brumelle and McGill (1993), both discrete and continuous demand distributions are considered. The approach of Brumelle and McGill (1993) maximizes expected revenue using a set of equations based on the partial derivatives of the expected revenue function. Moreover, they show that the optimal protection levels are expressed in terms of joint probability distributions as follows:

$$
\begin{aligned}
f_2 &= f_1 P(D_1 > \Pi_1) \\
f_3 &= f_1 P(D_1 > \Pi_1, D_1 + D_2 > \Pi_2) \\
&\dots \\
f_k &= f_1 P(D_1 > \Pi_1,\ D_1 + D_2 > \Pi_2,\ \dots,\ D_1 + D_2 + \dots + D_{k-1} > \Pi_{k-1}),
\end{aligned}
$$

where $\Pi_k$ is the protection level for fare class $k$ and $f_1 \geq f_2 \geq \dots \geq f_k$. If the remaining capacity of flight is higher than the protection level for a class, the demand request for that class is accepted. The method summarized above is called $EMSRb$.

The studies summarized above are for the assumption of low-before-high demand pattern. In dynamic models, this assumption is relaxed and a low-fare customer is allowed to arrive after a high-fare one. The first study on dynamic models is due to Lee and Hersh (1993). In this study, a discrete-time dynamic programming model is given. Moreover, batch arrival assumption is also relaxed by allowing multiple seat bookings. Unlike the previous models that use probability distributions, the demand is modeled as a stochastic process in this study. The demand intensity for a seat in booking class varies with time.

Lautenbacher and Stidham (1999) propose a discrete-time finite horizon Markov Decision

Process to solve the single-leg problem without cancellations, overbooking and no-shows. The dynamic models in this study do not differ from the one given by Lee and Hersh (1993). The main contribution of this study is using both static and dynamic approaches with Markov Decision Processes and showing the similarities between them.

Subramanian et al. (1999) extend the model of Lautenbacher and Stidham (1999) by including overbooking, cancellations and no-shows. In this model, the optimal booking policy is characterized by state and time dependent booking limits for the fare classes. The main results of the study can be summarized as follows: the booking limits need not be monotonic; it may be optimal to accept a low-fare class rather than high one because of the cancellation probabilities; an optimal policy depends on both the total capacity and the remaining available capacity.

Gosavi et al. (2002) suggest a stochastic optimization technique, called Reinforcement Learning, for the single-leg problem. They use the Semi-Markov Decision Process (SMDP) allowing overbooking, concurrent demand arrivals from different fare classes and class dependent, random cancellations.

## 2.2 An Overview of Network Seat Inventory Control

As it is mentioned in Chapter 1, it is hard to fly from an origin to a destination directly without using a transfer center in the hub and spoke systems. Hubs are the huge airports where most of the passengers are transferred from one flight to another and can be different for different airline companies. Because of these structural changes in airline industry, single-leg seat inventory control policy has lost its effectiveness significantly against network seat inventory control. In network seat inventory control, seats are allocated simultaneously for different customer segments and for different flight legs in a network.

Buhr (1982) is the first one who introduces a model for the seat inventory control problem with two legs and one fare class. In this model, it is allowed to board at the intermediate node. Buhr (1982) defines expected marginal revenue of the $S_{OD}^{th}$ seat for an origin-destination pair $OD$ as follows:

$$EMR_{OD}(S_{OD}) = f_{OD}\bar{P}_{OD}(S_{OD}),$$

where $f_{OD}$ is the fare of the origin-destination pair under consideration and $\bar{P}_{OD}(S_{OD})$ is the probability of selling the $S_{OD}^{th}$ seat to the passengers. Then, Buhr (1982) shows that total revenue is maximized by minimizing

$$\Delta EMR = EMR_{AC}(S_{AC}) - (EMR_{AB}(S_{AB}) + EMR_{BC}(S_{BC}))$$

subject to the capacity constraint. A-B-C is the network with legs AB and BC. Buhr (1982) also gives an approximation for the case of more than one fare class using a two-step approach, in which the allocation is determined for *OD* pairs first and then the allocation of seats among fare classes for a given *OD* itinerary is determined.

First study in the literature for a large network seat inventory control problem is due to Glover et al. (1982). The authors propose a maximum profit network flow model when demand is deterministic. The integer programming formulation of this network model is called Deterministic Mathematical Programming (*DMP*) model and the details of the model are given in Chapter 3. The linear relaxation of this model is called Deterministic Linear Programming (*DLP*). In the network flow formulation, two arc sets are used, one is for the flight legs in forward direction and the second set is for the *ODF*s in backward direction. The limitations on forward arcs are the aircraft capacities on a flight leg and the backward arcs are limited by the *ODF* demand estimates.

Wollmer (1986) introduces a mathematical programming formulation for a multi-leg multi-class network problem. A binary decision variable is defined by Wollmer (1986) for every possible *ODF* and seat $i$ as $x_{ODF,i}$. The objective function is the total expected marginal revenue and the only constraint in the formulation is the capacity constraint. Wollmer's formulation is computationally difficult to solve because of the large number of binary decision variables.

The studies for network RM problems reviewed up to this point are for booking limit controls. Bid price control policy in network seat control problems is firstly introduced by Simpson (1989) and developed by Williamson (1992). In bid price control, a seat is sold if fare of that *ODF* exceeds the sum of bid prices of the legs along the path. Simpson (1989) and Williamson (1992) use deterministic linear programming models to get dual prices of the capacity constraints which they propose to use as the bid prices. The study due to Williamson (1992) is a significant study on network seat control because it includes partitioned, nested and bid price policies, demand aggregation, simulation and comparisons of different methods.

The results of this study have been widely used in the following studies.

Also, Talluri and van Ryzin (1998) study on the bid price control for network RM problems. Because of the increasing popularity of the bid price control in RM, they analyze theoretical basis of the policy that leads to intuitive and practical use in the other studies. Moreover, they show that bid price control is not optimal in general especially when leg capacities and sales volumes are not large enough. Talluri and van Ryzin (1999) consider a randomized version of the *DLP* model which is called Randomized Linear Programming (*RLP*). In *RLP*, the realized demands of itineraries are considered in a deterministic linear programming model. *RLP* is obtained by replacing the expected demand figures in *DLP* by the realized demands. Then, the dual prices for different demand realizations are used to obtain a bid price approximation. The authors also give the conditions under which the *RLP* provides an unbiased estimator of the gradient of perfect information network. The *RLP* method is simple and has some slight improvements in revenue over *DLP*. The details of the *RLP* model are given in Chapter 3.

de Boer (1999) and de Boer et al. (2002) give an extensive study on the analysis of deterministic and stochastic network revenue management problems. There is a common phenomenon in the literature that the deterministic models outperform more advanced probabilistic methods. The authors argue that this is due to a booking process that includes nesting of the fare classes and the probabilistic models suffer more from ignoring the nesting in the model. The reason of this situation is explained in Chapter 3. Moreover, a stochastic model, Stochastic Linear Programming (*SLP*), is developed by de Boer (1999) using demand aggregation. *SLP* is given in Chapter 3.

Overbooking is also an important area for network revenue management. Overbooking is allowing total volume of the sales to be higher than the capacity of the flights. This way, capacity utilization of the flights is increased significantly. There is a wide research history for overbooking. The first study is due to Beckmann (1958) and on a non-dynamic optimization model. Some other related studies on overbooking for network airline revenue management are given by McGill and van Ryzin (1999). We only summarize important and recent studies in the literature. Shlifer and Vardi (1975) propose an overbooking model for three cases: single-leg flight carrying a single type of passenger; a single-leg flight carrying two types of passengers; two-leg flight. Biyalogorsky et al. (1999) propose using overbooking with opportunistic cancellations. In this study, the request of a high fare customer is accepted even if the remaining capacity is zero and then the ticket of a low fare customer is cancelled with a compensation. Ringbom and Shy (2002) suggest an *"adjustable-curtain"* strategy in order to determine the number of business and economy class bookings. In this strategy, the airline can adjust the size of the business and economy class sections before boarding. This way, overbooking for business class passengers is allowed. Karaesmen and van Ryzin (2004) suggest a two-period optimization model to determine the overbooking levels. In the first period, reservations are accepted with the probabilistic knowledge of cancellations. In the second period, surviving customers are assigned to the various inventory classes to minimize penalties of assignments.

## 2.3 An Overview of Other Studies in Revenue Management

The survey in Sections 2.1 and 2.2 summarizes major studies on mathematical programming models for airline seat inventory control policies. In this section, we summarize the studies on pricing in which prices rather than quantities are used as the primary demand management variables.

Pricing has been extensively studied in the literature because of the wide implementation area. The objective of pricing is determining the prices for various classes (customer segments) during the booking horizon. Weatherford (1991) presents a formulation for making pricing/allocation decisions simultaneously. Gallego and van Ryzin (1994) give an optimal dynamic pricing for the case of stochastic demand. Gallego and van Ryzin (1997) consider the pricing strategies of multiple firms in revenue management context. In this study, firms have finite capacities and the problem is revenue maximization over the finite horizon. Bitran and Candeltey (2003) and Elmaghraby and Keskinocak (2003) review the pricing studies in revenue management. Feng and Gallego (2000) and Aviv and Pazgal (2005) propose approaches for pricing problems to find optimal or approximate solutions.

Revenue management can be applied to many industries other than airline industry. The airlines, hotels and rental car industries are called as traditional industries and have similar characteristics such as perishability of the goods, varying demand over time and insignificant variable costs for customers. Talluri and van Ryzin (2005) and Chiang et al. (2007) give a list of areas where revenue management is successfully applied in the literature. In the study of

Chiang et al. (2007), the studies on revenue management for traditional areas, such as hotels and car rental, and non-traditional areas are listed. Some of the non-traditional areas are as follows: restaurants, hospitals and health care, sport events, cargo and freight, broadcasting and media, project management and retailing.

The studies summarized in this chapter are all to maximize the expected revenue without taking the risk factors into account. More clearly, all of the studies in the previous sections are proposed for a risk neutral decision maker. However, in real life problems, according to the changes in the market, one may prefer to work with risk-sensitive approaches.

Although the risk-aversion is widely used for a variety of inventory models, it is new for revenue management problems. Barz and Waldmann (2007) and Levin et al. (2008) summarize risk literature for inventory models. The studies due to Agrawal and Seshadri (2000), Feng and Xiao (1999), Chen et al. (2005), Chen and Federgruen (2000), Eeckhoudt et al. (1995), and Martínez-de Albéniz and Simchi-Levi (2006) are some of the studies on risk-sensitive inventory models. Risk is generally incorporated into revenue management problems in a dynamic pricing framework. Feng and Xiao (1999) use an objective function that incorporates a penalty function to reflect changes in the sales variance as a result of price changes. Lancaster (2003) uses a sensitivity analysis instead of directly incorporating risk aversion into revenue management models. Weatherford (2004) and Chen et al. (2006) use expected utility of the revenue instead of expected revenue. Barz and Waldmann (2007) study a single-leg revenue management problem from the perspective of risk-sensitive decision maker using exponential utility function. Levin et al. (2008) use dynamic pricing with a loss-probability as a risk measure which we also use in the proposed models in Chapter 4. The only study in the literature that uses mathematical models for risk-sensitive seat inventory control policies is due to Çetiner (2007). She proposes two models incorporating variance of the revenue in addition to the expected revenue. First model (*EMVLP*) is to use variance in the objective function and the second model (*CVLP*) is to use it in the constraints. Our work in this thesis is also for the use of mathematical models in risk-sensitive seat inventory control.

# CHAPTER 3

# NETWORK SEAT INVENTORY CONTROL

This chapter is devoted to the studies on network seat inventory control in the literature. A general dynamic programming model is introduced first. In Sections 3.1 and 3.2, deterministic and probabilistic mathematical programming models are summarized. Control policies for RM problems are given in Section 3.3. This chapter ends with the risk-sensitive models in the literature in Section 3.4. For network seat inventory control, the complete network is considered with all the dependence relations between different legs.

First of all, the notation used in the following sections for network RM problems is introduced. The network has $m$ resources, which are the legs between origins and destinations. There are $n$ products (origin, destination and fare combinations) offered on those legs. An itinerary is a trip from an origin to a destination. $\boldsymbol{A}$ is an $m \times n$ matrix and $j^{th}$ column of the matrix gives the resources needed for product $j$. $a_{lj}$ is the entry of matrix $\boldsymbol{A}$ for row $l$ and column $j$. It is used to relate resources and products: $a_{lj} = 1$ if resource $l$ is a part of the trip of product $j$ and 0 otherwise. $T$ is the length of the booking period horizon and $t$ is the remaining time until boarding. That is, the time indices run backward and departure of the flight is at time 0. In general, discrete-time models are considered in the literature. The time periods are generally assumed to be small enough and only one demand request arrives in a period. $\boldsymbol{c} = (c_1, ..., c_m)$ where $c_l$ is the capacity of the flight that flies through leg $l$. Letting $\boldsymbol{A}_j$ denote the $j^{th}$ column of matrix $\boldsymbol{A}$, the capacity vector is updated as $\boldsymbol{c}$-$\boldsymbol{A}_j$ when product $j$ is sold. $\boldsymbol{\mathfrak{F}(t)}$ is used as the demand vector for time period $t$ such that $\boldsymbol{\mathfrak{F}(t)} = (\mathfrak{F}_1(t), ..., \mathfrak{F}_n(t))$. $\boldsymbol{\mathfrak{F}(t)}$ is the random variable denoting the price of product $j$ requested in period $t$. $\mathfrak{F}_j(t) = f_j > 0$ when product $j$ is requested at a price of $f_j$ at time $t$, and $\mathfrak{F}_j(t) = 0$ when there is no request for product $j$. $\boldsymbol{\mathfrak{F}(t)} = \boldsymbol{0}$ when no request arrives for any of the products. The sequence $\{\boldsymbol{\mathfrak{F}(t)}; t \geq 1\}$ is assumed to be independent across time $t$ with known joint distribution in each period $t$.

$\boldsymbol{u}(t) = (u_1(t), ..., u_n(t))^\top$ denotes the decision made at time $t$: $u_j(t) = 1$ if request for product $j$ is accepted and $u_j(t) = 0$ if it is rejected. $\top$ denotes transpose of the vector. The main factors, which affect the decision at time $t$ upon arrival of a demand request, are the remaining capacity and price of the product requested. Therefore, the decision to be made is a function of these factors, which can be denoted by $\boldsymbol{u}(t) = \boldsymbol{u}(t, \boldsymbol{c}, \boldsymbol{f})$ where $\boldsymbol{f} = (f_1, ..., f_n)$.

Using the notation introduced above, a dynamic programming model is formulated by Talluri and van Ryzin (2005) to find the optimal accept/reject decisions for the requests. $V_t(\boldsymbol{c})$ is the maximum expected revenue for the last $t$ periods when the remaining capacity is $\boldsymbol{c}$ at time $t$. Then, the backward recursive function is

$$V_t(\boldsymbol{c}) = E[\max_{\boldsymbol{u}(t)}\{\boldsymbol{\mathfrak{F}(t)}^\top \boldsymbol{u}(t) + V_{t-1}(\boldsymbol{c} - \boldsymbol{A}\boldsymbol{u}(t))\}],$$

with the boundary condition

$$V_0(\boldsymbol{c}) = 0 \text{ for all } \boldsymbol{c}.$$

$E(\cdot)$ denotes expected value of the random variable under consideration. The optimal control for this formulation is given as follows:

$$u_j^*(t, \boldsymbol{c}, f_j) = \begin{cases} 1 & \text{if } f_j \geq V_{t-1}(\boldsymbol{c}) - V_{t-1}(\boldsymbol{c} - \boldsymbol{A}_j) \text{ and } \boldsymbol{A}_j \leq \boldsymbol{c}\ , \\ 0 & \text{otherwise.} \end{cases}$$

Hence, the optimal control policy for accepting a request is of the form: accept a booking request for product $j$ if the remaining capacity is sufficient and the price of product $j$ exceeds the opportunity cost of the reduction in resource capacities required to satisfy the request.

The displacement cost, $V_{t-1}(\boldsymbol{c}) - V_{t-1}(\boldsymbol{c} - \boldsymbol{A}_j)$, in the control policy given above leads to bid price control for network RM. Moreover, Talluri and van Ryzin (2005) show that the condition for accepting product $j$ in period $t$ can be approximated as follows when it is supposed that the optimal value function $V_{t-1}(\boldsymbol{c})$ has a gradient $\nabla \boldsymbol{V}_{t-1}(\boldsymbol{c})$ .

$$\begin{aligned} f_j &\geq V_{t-1}(\boldsymbol{c}) - V_{t-1}(\boldsymbol{c} - \boldsymbol{A}_j) \\ &\approx \nabla \boldsymbol{V}_{t-1}^T(\boldsymbol{c})\boldsymbol{A}_j \\ &= \sum_{l \in \boldsymbol{A}_j} \pi_l(t, c), \end{aligned}$$

where $\pi_l(t, c) = \frac{\partial}{\partial c_l} V_{t+1}(\boldsymbol{c})$.

Based on the approximation above, a request for a product can be accepted if the price exceeds the sum of the bid prices, $\pi_l(t, c)$, for all of the resources used by that particular product.

Because of the large dimensionality of the dynamic programming model, approximations should be used. There can be two basic approximations for the problem: using a simplified network model, e.g. solving the problem as a static mathematical programming problem such as Deterministic Linear Programming Model, or decomposing the network problem into single-leg problems to work with bid prices of subproblems determined independently. For both of these approximations, having good estimates of the optimal value function and bid prices is important.

## 3.1 Deterministic Mathematical Programming Models

The deterministic formulation of the network seat inventory control problem is such that probabilistic nature of the demand is ignored by working with expected value of the demand. The formulation is called $DMP$ that stands for Deterministic Mathematical Programming. The network formulation of this integer programming model is firstly introduced by Glover et al. (1982). $DMP$ model is given below.

$$DMP \quad : \quad Maximize \sum_{j=1}^{n} f_j x_j \tag{3.1}$$

$$subject\ to \tag{3.2}$$

$$\sum_{j \in S_l} x_j \leq C_l \quad \text{for } l = 1, ...., m, \tag{3.3}$$

$$x_j \leq E(D_j) \quad \text{for } j = 1, ..., n, \tag{3.4}$$

$$x_j \geq 0 \text{ and integer} \quad \text{for } j = 1, ...., n. \tag{3.5}$$

The decision variable $x_j$ in this model is the number of seats allocated for $ODF\ j$. Constraint (3.3) in the model is the capacity constraint for the legs where the summation of allocations for $ODF$s on a leg is smaller than the leg capacity. $S_l$ denotes the set of $ODF$ combinations using flight leg $l$, $S_l = \{j | a_{lj} = 1\}$. In (3.4), the upper bound on the allocation of an $ODF$ is the expected demand of $ODF\ j$, given by $E(D_j)$. The objective function is to maximize the total revenue. The main advantage of the $DMP$ model is the simplicity of the model. The output of the model is the set of allocations of the seats to the products. Use of these allocations without any nesting is called the partitioned booking policy. Partitioned policy is such that the seats are allocated to only one of the $ODF$s and the seats that are not sold to that $ODF$ will remain empty. This drawback can be handled by using a nesting heuristic for the implementation of the allocation obtained by the $DMP$ model. The major drawback of the $DMP$ model is ignoring stochastic nature of the demand requests. It is assumed that the demand is certain and equals to the mean demand. Although the allocations obtained by the $DMP$ model give higher revenues as compared to almost all of the other models in the literature, its performance is directly dependent on the demand forecasting quality.

Another problem to solve the $DMP$ model is the integrality constraint for $x_j$s as pointed out by Williamson (1992). In order to obtain integer seat allocations, decision variables must be integer. This would cause an increase in solution time. However, this problem was solved by Williamson (1992) by using linear relaxation of the model and working with integer demand estimates in the constraints. This relaxed model is called *DLP* that stands for Deterministic Linear Programming. The following statement is due to Williamson (1992): *"Under the integrality of the network problems (referring to the study of Bradley et al., 1977), if the upper and lower bounds on the decision variables are integers and the right hand side values of the flow balance constraints are integer, the solution will be integer. Thus, by requiring both the demand constraint values to be integer, an integer solution can be obtained."* Hence, Williamson (1992) claims that by rounding the expected demand values, integer solutions can be obtained by the *DLP* model. However, de Boer (1999) gives a counter example where the right hand sides and limits are integer, but the solution is not. de Boer (1999) argues that the theoretical background of Williamson (1992) is not clear, but it works in the real airline data surprisingly. Moreover, although de Boer (1999) and Williamson (1989) state that rounding the solution may give worse results than rounding the demand data, there is no theoretical and numerical ground of this statement to the best of our knowledge.

## 3.2 Probabilistic Mathematical Programming Models

The most general probabilistic model in the network RM is called as Probabilistic Mathematical Programming or Probabilistic Nonlinear Programming Model which is to maximize the expected revenue in terms of the protection levels. The abbreviation used for this model is

$PMP$ or $PNLP$, and the model is given below.

$$PMP \quad : \quad Maximize\ E(\sum_{j=1}^{n} f_j \min\{x_j, Dj\}) \tag{3.6}$$

$$subject\ to$$

$$\sum_{j \in S_l} x_j \leq C_l \quad \text{for } l = 1, ...., m,$$

$$x_j \geq 0\ and\ integer \quad \text{for } j = 1, ...., n,$$

where $min\{x_j, D_j\}$ is the random variable that gives minimum of demand for $ODF\ j$ and number of seats allocated to $ODF\ j$. The $PMP$ model is nonlinear because of the objective function and hard to solve. DeMiguel and Mishra (2006) state that $PMP$ can be reformulated as a stochastic program with simple recourse and can be solved numerically. Unfortunately, non-linear programming formulations are computationally hard to solve for large scale problems. de Boer et al. (1999) argue that, in general, the outcome of implementing the solution of the stochastic model exceeds the outcome of implementing the deterministic solution and it is called the value of the stochastic solution. However, Williamson (1992) shows that, in simulation studies, booking control policies using *DLP* performs better than the ones using $PMP$ for airline revenue management problems. de Boer (2002) argues that *DLP* outperforms probabilistic models because nesting is not considered in *DLP*. Although both deterministic and probabilistic models are non-nested, the adverse impact of not incorporating nesting is more for the probabilistic models. Probabilistic models assign more seats to high-fare classes in order to earn upward potential of high-fare demand, which in fact only aggravates the degree of overprotection. The deterministic model is unable to recognize this potential and this drawback turns out to be an advantage in nested environment. In the numerical studies in Chapter 6 of this thesis, it is shown that revenues for deterministic and probabilistic models are not so different when the difference among fares for different classes are low.

Second probabilistic model which is widely used in the literature is the Expected Marginal Revenue (*EMR*) Model. This model again incorporates the probability distribution of the demand. The objective function is the summation of expected marginal revenues of the seats. The only constraint in the model is the capacity constraint for the legs. $x_j(i)$ is the binary decision variable for a given $ODF\ j$ and seat $i$.

$$x_j(i) = \begin{cases} 1 & \text{if } i \text{ or more seats are allocated to } ODF\ j, \\ 0 & \text{otherwise.} \end{cases} \qquad \text{for } i = 1, ..., B_j,$$

$B_j$ is defined as the maximum number of seats that can be allocated to the *ODF j*. Three different ways to determine the value of $B_j$ are proposed as follows:

- $B_j = \min_l\{C_l : j \in S_l\}$,
- $B_j = B = \max_l\{C_l\}$,
- $B_j = \min_{l \in T_j}\{\vartheta_l, C_l\}$ *where* $\vartheta_l = \max\{\Upsilon : P(D_j \leq \Upsilon) \leq \psi,\ \Upsilon$ *is integer and* $j \in S_l\}$ *and* $T_j = \{l : j \in S_l\}$.

In the first way, $B_j$ is defined as the minimum capacity of the legs which are used by *ODF j*. Secondly, it is defined as the maximum of the flight capacities directly. In the third way, $B_j$ for *ODF j* is the minimum of $\vartheta_l$ and $C_l$ for leg $l$ where leg $l$ is a part of the trip of *ODF j*. Moreover, $\vartheta_l$ is equal to the maximum of $\Upsilon$ where probability of demand for *ODF j* being smaller than $\Upsilon$ is smaller than a predetermined level, $\psi$. The number of seats allocated to an *ODF j*, $x_j$, can be expressed in terms of $x_j(i)$s as follows:

$$x_j = \sum_{i=1}^{B_j} x_j(i). \tag{3.7}$$

Then, the expected marginal revenue of the $i^{th}$ seat on *ODF j* is

$$EMR_j(i) = f_j P(D_j \geq i) x_j(i).$$

Summation of $EMR_j(i)$ over all $i$ and $j$ is the total expected revenue. As a result, *EMR* model is given as follows:

$$\begin{aligned} EMR \quad : \quad & Maximize \sum_{j=1}^{n} \sum_{i=1}^{B_j} f_j P(D_j \geq i) x_j(i) \\ & subject\ to \\ & \sum_{j \in S_l} \sum_{i=1}^{B_j} x_j(i) \leq C_l \quad \text{for } l = 1, ...., m, \\ & x_j(i) \in \{0, 1\} \quad \text{for } i = 1, ..., B_j \text{ and } j = 1, ...., n. \end{aligned} \tag{3.8}$$

Since the decision variables in the model are binary, computational difficulties would be encountered for large networks. Williamson (1992) suggests to work with linear relaxation of this model because of the monotonically decreasing behavior of the complementary probability $P(D_j \geq i)$ in the objective function in $i$ for each $j$. The following remark is useful in showing that the linear relaxation of the *EMR* model gives also integer results.

**Remark 3.2.1** *(due to Williamson 1992) In the linear relaxation of the EMR model, a full seat is allocated to* $x_j(i)$ *before any portion of a seat is allocated to* $x_j(i+1)$. *That is,* $x_j(i+1)$

*can take a positive value only if $x_j(i) = 1$.*

**Proof.** *The probability of demand being larger than i, i.e., $P(D_j \geq i)$, in the objective function decreases as i increases. As a result of this, $EMR_j(i)$ decreases monotonically as i increases. This guarantees that a full seat is allocated to $x_j(i)$ before any portion of a seat is allocated to $x_j(i+1)$.* ■

Remark 3.2.1 only guarantees that $x_j(i+1)$ cannot take a positive value if $x_j(i) < 1$. Therefore, for any $j$, $x_j(i)$ may take a fractional value for at most one seat, $i$. Combining this remark with the integrality property of network problems, the following constraints can be used in the relaxed formulation: $0 \leq x_j(i) \leq 1$ for all $j$ and $i$. Then, the resulting allocations for the *ODF*s can be found by summing related decision variables as follows: $x_j = \sum_{i=1}^{B_j} x_j(i)$. Although large number of decision variables leads to computational burden for the applicability of linear programming (LP) relaxation of *EMR* model, technical improvements in computational capabilities might be expected to avoid this disadvantage.

Because of the computational difficulties to solve the *EMR* model, an aggregation method is proposed by de Boer et al. (2002) leading to an approximation of *EMR*: Stochastic Linear Programming (*SLP*). In *SLP*, suppose $D_j$ can only take some values $d_j(1) < d_j(2) < ... < d_j(K_j)$ such that $K_j$ is the number of aggregate demand groups.

$$
\begin{aligned}
SLP \quad : \quad & Maximize \sum_{j=1}^{n} f_j x_j - \sum_{j=1}^{n} f_j \sum_{k=1}^{K_j} P(D_j < d_j(k)) x_j(k) \qquad (3.9)\\
& subject\ to \\
& \sum_{j \in S_l} x_j \leq C_l \quad \text{for } l = 1, ...., m, \\
& x_j = \sum_{j=1}^{K_j} x_j(k) \quad \text{for } j = 1, ..., n, \\
& x_j(1) \leq d_j(1) \quad \text{for } j = 1, ..., n, \\
& x_j(k) \leq d_j(k) - d_j(k-1) \quad \text{for } j = 1, ..., n \text{ and } k = 2, ..., K_j, \\
& x_j(k) \geq 0 \quad \text{for } j = 1, ..., n \text{ and } k = 1, ..., K_j, \\
& x_j \geq 0 \quad \text{for } j = 1, ..., n,
\end{aligned}
$$

where each $x_j$ is split up in several smaller allocations $x_j(k)$, which represents the amount of seats allocated to the $k^t h$ partition of the demand, i.e., for the demand that falls in the interval $(d_j(k-1), d_j(k))$. In *EMR* model, $B_j$ is defined as the maximum capacity of the legs which are used by *ODF* $j$ and in *SLP* model, $K_j$ is defined as the maximum number of demand groups. Hence, the maximum value of $d_j(k)$ for all demand groups is equal to $B_j$, i.e., $d_j(K_j) = B_j$. First term in the objective function is total revenue gained when all of the seats which are allocated to the itineraries are sold. Second term is a correction for the uncertainty of demand.

The *SLP* model is linear and the solution of *SLP* gives integer results when integrality property due to Bradley et al. (1977) holds. Although the theoretical background of integrality is not clear as in Section 3.1, this formulation works in the real airline data again. The complexity of the *SLP* model can be decreased by aggregating demands more, but causing the results to get poor. Moreover, de Boer et al. (2002) show that the LP relaxation of *EMR* is only a special case of *SLP*. That is, objective function of *SLP* can be rewritten as

$$\sum_{j=1}^{n} \sum_{k=1}^{K_j} f_j P(D_j \geq d_j(k)) x_j(k)$$

by letting $d_j(k+1) - d_j(k) = 1$ and $d_j(1) = 1$ for all $j$ and $k$.

Talluri and van Ryzin (1999) propose a randomized version of the deterministic linear programming (*DLP*) model for computing network bid prices. This model is called the Randomized Linear Programming (*RLP*) model. In *RLP*, the bid prices are calculated by solving deterministic linear programs for each demand realization.

$$RLP \quad : \quad Maximize \sum_{j=1}^{n} f_j x_j \tag{3.10}$$

$$subject\ to$$

$$\sum_{j \in S_l} x_j \leq C_l \quad \text{for } l = 1, ...., m, \tag{3.11}$$

$$0 \leq x_j \leq d_j \quad \text{for } j = 1, ..., n, \tag{3.12}$$

where $x_j$ is the number of seats allocated for *ODF* $j$ and $d_j$ is the specific demand realization. Expected demands on the right-hand sides of the demand constraints in the *DLP* model are replaced with the specific demand values. Then, averages of the dual variables of (3.11) for the specified set of demand realizations are used as the approximate bid prices. Let $\mu_r(l)$ be the bid price for leg $l$ for demand realization $r$ and equals to the dual price of the capacity constraint. Then, the estimated bid price for leg $l$ is $\frac{1}{N} \sum_{r=1}^{N} \mu_r(l)$. The advantage of this method is the decrease in complexity of the problem as compared to *EMR*.

## 3.3 Control Methods for Network RM

In the previous section, the mathematical models that aim to obtain optimal allocations and bid prices are summarized. In this section, control methods or policies are given which are used to make the decision of accepting or rejecting a booking request based on the optimal allocations or bid prices. From this point on, the first step for RM to use mathematical programming models is called the *optimization* step and the second step for the implementation of the allocations or bid prices obtained from the models is called the *control* step. In this section, three types of control policies are summarized. These are partitioned booking limit control, nested booking limit control and bid price control.

### 3.3.1 Partitioned Booking Limit Control

This control method directly uses the booking limit allocations, which are obtained from the mathematical models in the optimization step. This is the basic and most straightforward control policy. Each of the booking limits is used only for the corresponding $ODF$ and unsold capacity of an $ODF$ cannot be used for other $ODF$s even if they have higher fares. Therefore, the revenue obtained using this control policy generally turns out to be lower than the revenue obtained by the other policies described in this section. Moreover, the load factor under this policy is generally less than the load factor under the other control policies. These drawbacks can be overcome by updating booking limits frequently. However, for each update, an additional optimization and re-forecasting of future demand are required. As a result, partitioned booking limit control policy is rarely used in airline industry.

### 3.3.2 Nested Booking Limit Control

Nested booking control policy is suggested to overcome the major drawback of partitioned policy, in which the unsold seat for an $ODF$ cannot be sold to a request with an higher fare. For nested booking limit control policy, fare classes are ranked. The booking limit for the $ODF$ with the lowest ranked fare class can be used for all of the classes, and the booking limit for the highest ranked fare class is equal to the capacity of the flight. The seats allocated to a fare class are allowed to be booked by a higher ranked class.

The main difficulty in nested booking limit control is the ranking of fare classes for a network RM problem. Williamson (1992) proposes three different ways of ranking. These are nesting by fare class, nesting by fares and nesting by shadow prices. In the first nesting strategy, fare classes are ranked according to the class types, where a full fare class is ranked higher than a low fare class of an origin-destination pair. In this ranking strategy, amounts of the fares do not have any impact on the ranking. Therefore, gain from a discounted fare passenger of a long path can be lost and a full fare class of a short path can be given priority although the gain from full fare of a short path can be less than the discounted one of a long path. Therefore, this strategy generally gives poor results in network RM problems.

Second strategy, nesting by fares, is proposed by Boeing Commercial Airplane Company and based on *ODF*s' fare values. The classes are ranked according to the fares of itineraries and itinerary with the highest fare is ranked in the first position. In this approach, the itineraries for long paths are generally ranked in the first positions although they may have less contributions to the revenue than the itineraries for short paths. In this strategy, it is possible for low-yield long-path itineraries which have small number of allocations in the optimization models to have access an important number of seats. Therefore, first and second nesting strategies are two extreme cases and both of them generally give poor results.

In the third nesting strategy, shadow prices are used, which is proposed by Williamson (1992). Note that shadow price is an increment in the revenue that would be gained when one more seat is allocated to a particular *ODF*, when other allocations are kept constant. The idea behind this approach is as follows: itineraries with higher shadow prices would most probably give more revenue than the ones that have lower shadow prices, therefore they should be ranked higher. In the *DLP* model, shadow price for an *ODF* is simply the dual price of the respective demand constraint. However, in the probabilistic models, there is no demand constraint associated with each *ODF*. Williamson (1992) proposes to calculate shadow prices as incremental change in revenue which is generated by forcing an additional seat to be allocated to a given *ODF* at the expense of another *ODF* or combination of *ODF*s. However, this would be time consuming especially for large-scale networks because the model must be used once for each possible *ODF*.

de Boer et al. (2002) suggest using dual prices of the capacity constraints in the probabilistic models in order to obtain an estimate of the dual prices of *ODF*s as follows: the opportunity cost of an $ODF$ is the sum of dual prices of the legs which are elements of that particular itinerary. This is only an approximation of the opportunity cost of this itinerary. Then, by subtracting this value from fare of the $ODF$, an estimate of an $ODF$'s net contribution to the network revenue is found. That is, by subtracting opportunity cost from the fare of the $ODF$, the profit that will be gained from that $ODF$ can be calculated.

It is hard to incorporate the nesting strategies considered above into mathematical models. Therefore, nesting heuristics are developed in the literature to be used with the allocations obtained from the mathematical models without considering nesting. In this thesis, we use the nesting heuristic that is proposed by de Boer et al. (2002). The notation used in this nesting heuristic is as follows:

$r_j$ : the number of booking requests for $ODF\ j$ that have been accepted,

$c_l$ : remaining capacity on leg $l$.

**Nesting Heuristic** (due to de Boer et al. 2002)

**Step 0.** Let $c_l = C_l\ for\ l = 1, ..., m$ and $r_j = 0\ for\ j = 1, ..., n$.

**Step 1.** A booking request for an $ODF\ j'$ arrives and should be considered for acceptance.

**Step 2.** Define $b_j = max\{x_j - r_j, 0\}$ for all $ODF\ j$.

**Step 3.** Define $b_l = \sum b_j$ for all legs $l \ni j' \in S_l$. The summation is over all $ODF\ j$ that outranks $j'$ and that also crosses leg $l$ .

**Step 4.** Define $b_{min} = min\{c_l - b_l \mid j' \in S_l\}$.

**Step 5.** If $b_{min} > 0$, accept the booking request and let $c_l = c_l - 1$ for all legs $l \ni j' \in S_l$ and let $r_{j'} = r_{j'} + 1$. Decline the request otherwise.

**Step 6.** If another booking request arrives within the booking horizon, go to step 1. Otherwise, stop.

Step 0 is the initialization step where capacities and number of booking requests that have been accepted so far are set to the initial values. A booking request arrives at step 1 and is taken into consideration for acceptance. Then, $b_l$ values are calculated for all legs. These limits are protected for fare classes that are ranked higher than the current $ODF\ j'$. In step 4, the number of available seats are found considering all legs on $ODF\ j'$. If the result of step 4 is higher than zero, then the booking request is accepted and capacity is decreased for the legs of $ODF\ j'$ by one. The number of booking requests that are accepted is increased by one at step 5.

### 3.3.3 Bid Price Control

As it is seen in Section 3.3.1 and Section 3.3.2, one somehow needs to avoid drawbacks of the booking limit control policies in order to improve the results. As an alternative to the booking limit control policies, a bid price control policy is proposed by Simpson (1989) and Williamson (1992) as a part of a research project in MIT. The main idea behind the policy is the following: if the fare of demand request is higher than a threshold level, then the request is accepted. Otherwise, it is rejected. The threshold level for an $ODF$ in LP formulations is calculated by summing dual prices of the capacity constraints of the legs that are on this particular $ODF$. If fare of the $ODF$ is lower than the bid price, the fare class for this origin-destination pair is called closed. Easy implementation of this control policy is the main advantage. Only the remaining capacity and the class status are necessary information that must be kept for use. The major drawback of the policy is the lack of the controls on the acceptance of requests. In this policy, there is no limit on the number that can be accepted for any of the itineraries when the class is open. Therefore, one class which has small contribution to the revenue can use most of the capacity. Williamson (1992) shows that with frequent updates, bid price and booking limit control policies give almost the same results. However, updating frequently is time consuming especially for the probabilistic models.

## 3.4 Risk-Sensitive Models in the RM Literature

Traditional RM models that are summarized up to this section are risk neutral. The objective of these models is to maximize expected revenues without considering variability of the revenue. In this section, three studies on risk aversion are summarized.

Barz and Waldmann (2007) study on static and dynamic single-leg revenue management problems from a perspective of risk sensitive decision maker using an exponential utility function. Moreover, they show that all well known structural results of the optimal policy that maximizes expected revenue hold for the risk-sensitive optimal policy as well. The authors give the dynamic model for a risk-neutral case and extend this model using a risk-sensitive approach. They consider a single-leg flight with a capacity of $C$ seats. Fare classes are again denoted by $j = 1, ..., n$ with associated fare of $f_j$. It is assumed that $0 < f_n < f_{n-1} < ... < f_1$. Cancellations, no-shows and batch bookings are not allowed in the model. The booking horizon is divided into $T$ intervals and the periods are indexed by $t$. For each period $t$, the probability of observing customer request for class $j$ is given by $p_{jt}$. Furthermore, $p_{0t} = 1 - \sum_{j=1}^{n} p_{jt}$ denotes the probability of no customer request in period $t$. Then, Barz and Waldman (2007) come up with a Markov Decision Process; namely, $MDP(T, \aleph, \Im, (q_t), (r_t), V_0)$. The state space, action space, transition law, reward function and boundary condition of the model are given by the authors as follows:

State Space: $\aleph = \{(c, j)|c \leq C, j \leq n\}$, $c$ is the remaining capacity and $j$ is the requested booking class, where $j = 0$ is the artificial class having $f_0 = 0$.
Action Space: $\Im = \{0, 1\} \equiv \{reject, accept\}$
Transition Law: $q_t((c, j), a, (c - a, j)) = p_{jn},\ where\ a \in A(c, i)\ and\ A(c, i) = \Im; for\ i > 0.$
Reward Function: $r_t((c, j), a) = ar_j$
Boundary Condition: $V_0(c, i) = 0$ for $c \geq 0$ and $V_0(c, i) = \bar{r}c$ for $c < 0$ with $\bar{r} > max_j\{f_j\}$.
Markov Policy: $\pi = (g_T, g_{T-1}, ..., g_1)$ is defined as a sequence $g_T, g_{T-1}, ..., g_1$ of decision rules $g_t$ specifying the action $a_t = g_t(c_t, j_t)$ to be taken at stage $t$ in state $(c_t, j_t)$. $G$ denote the set of all decision rules and $G^T$ is the set of all policies.
Then, the state process of the $MDP$ is denoted by $(x_T, x_{T-1}, ..., x_1)$. $V^*(c, i)$ is the maximum expected revenue starting with capacity $c$ and request $i$, i.e.,

$$V^*(c, i) = max_{\pi \in G^T} E_\pi[\sum_{t=1}^{T} f_t(x_t, g_t(x_t)) + V_0(x_0)|x_T = (c, j)].$$

Moreover, in dynamic programming $V^* \equiv V_T$ is the unique solution to the optimality equation

$$V_t(c, i) = max_{a \in A(c,j)} af_j + \sum_{j'=0}^{n} p_{j't} V_{t-1}(c - a, j).$$

The decision rule of a control limit is defined by,

$$g_t(c, j) = \begin{cases} 1 & \text{if } c > y_{j-1}(n), \\ 0 & \text{otherwise.} \end{cases} \tag{3.13}$$

To conclude, Barz and Waldmann (2007) give optimal protection level rule as

$$y^*_{j-1}(t) = max\{c : r_j < \sum_{j=0}^{n} p_{jt}(V_{n-1}(c, j) - V_{n-1}(c - 1, j))\}\ for\ j = 1, ..., n.$$

For the risk-sensitive approach, Barz and Waldmann (2007) define the expected utility of the revenue $R_\pi := \sum_{t=1}^{T} f_t(X_t, f_t(X_t)) + V_0(X_0)$. The expected utility function used in this study is exponential and given by $u_\gamma(x) = -exp(-\gamma x)$. The policy $\pi^{*\gamma} = (g_T^{*\gamma}, g_{T-1}^{*\gamma}, ..., g_1^{*\gamma})$ is called $\gamma - optimal$.

Let $V^{*\gamma}(c, j)$ be the maximal exponential utility,

$$V^{*\gamma}(c, j) = \max_{\pi \in G^T} E_\pi[-exp(-\gamma[\sum_{t=1}^{T} f_t(X_t, g_t(X_t)) + V_0(X_0)])|X_T = (c, j)].$$

Then, $V^{*\gamma} \equiv V_T^\gamma$ is the unique solution of

$$V_t^\gamma(c, j) = \max_{a \in A(c,j)} \{exp(-\gamma a f_j) \sum_{j=0}^{k} p_{jn} V_{n-1}(c - a, j)\},$$

where $V_0^\gamma(c, j) = -exp(-\gamma V_0(c, i))$. In order to simplify the notation, they define $L_t v(c) = \sum_{j=0}^{k} p_{jn} V(c, j)$ for an arbitrary real-valued function v. Moreover, it is found to be more convenient to work with $G_t^\gamma := -V_t^\gamma$ which is the unique solution of

$$\begin{aligned} G_t^\gamma(c, j) &= min_{a \in \{0,1\}}\{exp(-\gamma a f_j) L_n G_{n-1}^\gamma(c - a)\} \\ &= L_n G_{n-1}^\gamma(c - 1) min\{exp(-\gamma f_j), \frac{L_n G_{n-1}^\gamma(c)}{L_n G_{n-1}^\gamma(c - 1)}\}. \end{aligned}$$

Then, the optimal policy is

$$y_{i-1}(t) = max\{c \in \{0, ..., C - 1\} : exp(-\gamma f_j) > \frac{L_t C_{t-1}^\gamma(c)}{L_t C_{t-1}^\gamma(c - 1)}$$

such that $\pi^{*\gamma} = (g_T^{*\gamma}, g_{T-1}^{*\gamma}, ..., g_1^{*\gamma})$ defined by

$$g_t(c, j) = \begin{cases} 1 & \text{if } c > y_{j-1}^{*\gamma}(n), \\ 0 & \text{otherwise}, \end{cases} \tag{3.14}$$

is $\gamma$-optimal.

Barz and Waldmann (2007) experiment their approach with a single-leg example. In this example, there are four fare classes and the demand is normally distributed. Barz and Waldmann note that risk-sensitive decision maker prefers a lower certain revenue compared to future uncertain revenue. The main disadvantage of this study is that the proposed approach is usable only for single-leg flights. Therefore, extending this approach for network traffics is a pending matter.

Levin et al. (2008) present a model for optimal dynamic pricing of perishable services or products. This model is proposed for applications in which attainment of a revenue target is an important consideration for managers. Levin et al. (2008) study the problem of dynamically pricing items over a finite time horizon so that both the expected revenues and the risk of poor performance are taken into account. Risk is introduced into the model by augmenting the expected revenue objective with a penalty term for the probability that total revenues fall below a desired level of revenue, a loss-probability risk measure. The main contribution of this paper to the literature is explicit inclusion of the revenue process in the description of the system's state to formulate a class of simple Markovian models incorporating risk. Because of the similarities between risk measures of this approach and our approach, the details of this study are given below. The numerical studies due to Levin et al. (2008) are only for general inventory models. Therefore, there is not anything in their study about the applicability of the model to the airline RM problems.

In the study of Levin et al. (2008), the proposed model is presented as a stochastic optimal control problem over $[T, 0]$ in continuous time. Demand is distributed as a Non-Homogeneous Poisson Process (NHPP) which depends on time $t$ and current price $p$. The sale process $N(t)$ is limited by the initial inventory, $Y_T$, which is shown as $N(t) = min\{N'(t), Y_T\}$. The demand process $N'(t)$ is a nonhomogeneous Poisson process. The revenue is defined as a stochastic integral of the form $R(t) = \int_T^t p(T)dN(T)$. Therefore, the risk-neutral case is simply maximization of $E[R(0)]$. In the risk sensitive case, Levin et al. (2008) add the following constraint to the model: $P[R(0) \geq z] \geq \pi_0$, where $z$ is a minimum desired level of revenue and $\pi_0$ is the minimum acceptable probability with which the desired level to be reached. This is substantially similar to the constraint we use in the *PMP-RC* model in Chapter 4. Moreover, Levin et al. (2008) use risk factor in the objective function by multiplying risk measure with a penalty parameter as follows: $max\ E[R(0)] - \zeta P[R(0) < Z]$, where $\zeta$ is the penalty parameter and $\zeta \in [0, +\infty]$. Levin et al. (2008) interpret $\zeta$ as the maximum cost associated with not meeting the desired level of revenue $Z$.

The last risk-sensitive approach reviewed in this section for network RM problems is due to Çetiner (2007). Çetiner (2007) proposes two models: *EMVLP* and *CVLP*. In these models, variance of the revenue is chosen as a control variable for risk sensitivity of the decision maker. The first model, *EMVLP*, is to penalize variance of the revenue by a given factor while maximizing the expected revenue. In the second model, *CVLP*, the total expected revenue is maximized under a constraint on the ratio of the expectation and variance of the total revenue.

The models are given below:

$$EMVLP \quad : \quad Maximize \sum_{j=1}^{n}\sum_{k=1}^{K_j} f_j P(D_j \geq k) x_j(k)$$

$$-\theta \sum_{j=1}^{n}\sum_{k=1}^{K_j} x_j(k) f_j^2 P(D_j \geq k) P(D_j < k)$$

$subject\ to$

$$\sum_{j \in S_l} x_j \leq C_l \quad \text{for } l = 1, ..., m,$$

$$x_j = \sum_{k=1}^{K_j} x_j(k) \quad \text{for } j = 1, ..., n,$$

$$x_j(1) \leq d_j(1) \quad \text{for } j = 1, ..., n,,$$

$$x_j(k) \leq d_j(k) - d_j(k-1) \quad \text{for } j = 1, ..., n \text{ and } k = 2, ..., K_j \ ,$$

$$x_j(k) \geq 0 \quad \text{for } j = 1, ..., n \text{ and } k = 2, ..., K_j.$$

$$CVLP \quad : \quad Maximize \sum_{j=1}^{n}\sum_{k=1}^{K_j} f_j P(D_j \geq k) x_j(k)$$

$subject\ to$

$$\sum_{j \in S_l} x_j \leq C_l \quad \text{for } l = 1, ..., m,$$

$$x_j = \sum_{k=1}^{K_j} x_j(k) \quad \text{for } j = 1, ..., n,$$

$$x_j(1) \leq d_j(1) \quad \text{for } j = 1, ..., n,,$$

$$x_j(k) \leq d_j(k) - d_j(k-1) \quad \text{for } j = 1, ..., n \text{ and } k = 2, ..., K_j,$$

$$\sum_{j=1}^{n}\sum_{k=1}^{K_j} x_j(k) f_j^2 P(D_j \geq k) P(D_j < k) \leq$$

$$\rho \sum_{j}\sum_{k} x_j(k) f_j P(D_j \geq k),$$

$$x_j(k) \geq 0 \quad \text{for } j = 1, ..., n \text{ and } k = 2, ..., K_j.$$

In the first model, *EMVLP*, the first term of the objective function is the expected marginal revenue and the second term represents the variance of the marginal revenue penalized by $\theta$. The constraints in this model are the same as the ones used in the $SLP$ model proposed by de Boer et al. (2002). The numerical studies on *EMVLP* model show that the expected revenue and variance of the revenue can be controlled by changing the $\theta$ values. The main disadvantage of this model is that setting $\theta$ is not a straightforward task for decision makers. An additional effort is required to convert the risk-sensitivity measure of the decision maker to the $\theta$ values.

In the second model, *CVLP*, a measure which is the ratio of the expectation and variance of the revenue is used in the constraints. This measure is bounded as follows: $\frac{Var(MR)}{E(MR)} \leq \rho$. Moreover, Çetiner (2007) uses the $EMVLP$ and $CVLP$ models together. The *EMVLP* model is solved first and then the *CVLP* model is solved by taking the ratio of the expectation and variance of the total revenue corresponding to the optimal *EMVLP* solution as the right hand side of the constraint in the *CVLP* model.

To summarize, Barz and Waldmann (2007) propose a Markov Decision Process in order to solve single-leg airline revenue management problems. In the study of Levin et al. (2008), a general model for optimal dynamic pricing of perishable services or products is given. The only model that is proposed for risk-sensitive applications in airline network revenue management problems is due to Çetiner (2007). The numerical results of the risk-sensitive models due to Çetiner (2007) and the models proposed in this thesis are analyzed and compared in Chapter 6.

# CHAPTER 4

# THE PROPOSED RISK-SENSITIVE APPROACH

Recall from the previous chapters that there are only a few studies on the risk-sensitive approaches in RM literature. The approach that is proposed by Çetiner (2007) is the only one that uses seat inventory control to solve the network revenue management problems for risk-sensitive cases. Other risk-sensitive approaches cited in Chapter 3 are proposed for dynamic pricing problems and/or single-leg problems. In this chapter, we propose models by restricting or minimizing the probability of revenue being less than a predetermined threshold level for airline network revenue management problems. This probability is calculated by working with sample demand realizations and used in the objective function or in one of the constraints in the proposed mathematical models. The derivations to formulate this probability are given in Section 4.1. In Section 4.2, general multi-objective optimization methods are summarized. A lexicographic optimization model is developed in Section 4.3 to minimize the probability of revenue being less than a threshold level. Section 4.4 is devoted to the approximations of this proposed lexicographic optimization model. In Section 4.5, a mathematical model is given to use the probability measure under consideration in a constraint while maximizing the expected revenue. Finally, a method is in Section 4.6 given to solve the *RLP* model due to Talluri and van Ryzin (1999) for risk-sensitive cases. The resulting seat allocations or bid prices obtained from the models are used with the control policies in Chapter 5 and Chapter 6. In order to improve the solution quality of the models, a Bayesian update approach is used in Chapter 5. The seat allocations or bid prices are updated at the predetermined times throughout the booking horizon.

## 4.1 Risk Measure

In real life, attainment of a revenue target is an important consideration for managers. Decision makers, who deal with risk in their decision making process, want to minimize the probability of poor performance. In this thesis, the risk of a certain decision is measured by the probability of earning revenue that is less than a predetermined threshold level. The interpretation of this measure is easier for decision makers than the interpretation of the variance or any other measure considered in Chapter 3 for the existing studies in the literature. In this section, the derivations for this measure are given. The risk measure is formulated as $P(R < L)$, where $L$ is the predetermined threshold level for revenue denoted by random variable $R$.

Let $Z_j = min\{D_j, x_j\}$, where $D_j$ is the demand of $ODF\ j$ and $x_j$ is the number of seats allocated for $ODF\ j$. Recall from Chapter 3 that the total revenue for $PMP$ is expressed in terms of $Z_j$s. Let random variable $R_{PMP}(\boldsymbol{x})$ denote the total revenue earned in $PMP$ for given allocation $(x_1, ..., x_n)$, that is,

$$R_{PMP}(x_1, ..., x_n) = \sum_{j=1}^{n} f_j Z_j. \tag{4.1}$$

Now, consider the total marginal seat revenue in the objective function of the $EMR$ model. As shown by Çetiner (2007), the marginal seat revenue is expressed in terms of random variable $F_j(i)$ which is the marginal seat revenue obtained when the $i^{th}$ seat is allocated as an additional seat for $ODF\ j$. That is,

$$F_j(i) = \begin{cases} f_j & \text{if } D_j \geq i, \\ 0 & \text{otherwise.} \end{cases} \tag{4.2}$$

If the $i^{th}$ seat is allocated to $ODF\ j$ and the demand for that $ODF$ is greater than $i$, the gain for that seat is equal to the fare of the $ODF\ j$. Otherwise, the gain is zero. Then, the total revenue that can be obtained from the network can be calculated by summing up the marginal seat revenues of the seats of all itineraries. Let total revenue be denoted by random variable $R_{EMR}(\boldsymbol{x})$ in this case.

$$R_{EMR}(\boldsymbol{x}) = \sum_{j=1}^{n} \sum_{i=1}^{B_j} F_j(i) x_j(i), \tag{4.3}$$

where

$$x_j(i) = \begin{cases} 1 & \text{if } i \text{ or more seats are allocated to } ODF\ j, \\ 0 & \text{otherwise,} \end{cases} \qquad \text{for } i = 1, ..., B_j. \quad (4.4)$$

Here, $\boldsymbol{x}_j = (x_j(1), ..., x_j(B_j))$ for $j = 1, ..., n$ and $\boldsymbol{x} = (\boldsymbol{x}_1, ..., \boldsymbol{x}_n)$. Note that $x_j = |\boldsymbol{x}_j|$. The expected revenue for a network in the objective function of the *EMR* model is

$$E(R_{EMR}(\boldsymbol{x})) = \sum_{j=1}^{n} \sum_{i=1}^{B_j} f_j Pr(D_j \geq i) x_j(i). \quad (4.5)$$

By using the equations (4.1), (4.2), (4.3), the equality of the revenue functions used in *PMP* and *EMR* models for a given allocation is shown in Lemma 4.1.1.

**Lemma 4.1.1** $R_{PMP}(\boldsymbol{x}) = R_{EMR}(\boldsymbol{x})$ *for a given allocation* $\boldsymbol{x} = (x_1, ..., x_n)$.

**Proof.** *$R_{EMR}(\boldsymbol{x})$ in (4.3) is rewritten as shown below using the definition of $F_j(i)$ in (4.2).*

$$R_{EMR}(\boldsymbol{x}) = \sum_{j=1}^{n} (\sum_{i=1}^{D_j} f_j x_j(i) + \sum_{i=D_j+1}^{B_j} 0 \cdot x_j(i)) \quad (4.6)$$

$$= \sum_{j=1}^{n} f_j \sum_{i=1}^{D_j} x_j(i). \quad (4.7)$$

*From (3.7), $x_j = \sum_{i=1}^{B_j} x_j(i)$. Also, recall Remark 3.2.1 for allocating a full seat to $x_j(i)$ before any portion of a seat is allocated to $x_j(i+1)$. Then, in (4.7) above,*

$$\sum_{i=1}^{D_j} x_j(i) = \begin{cases} D_j & \text{if } D_j \leq x_j, \\ x_j & \text{otherwise.} \end{cases}$$

*That is, $\sum_{i=1}^{D_j} x_j(i) = min\{D_j, x_j\}$ which is $Z_j$ defined for $R_{PMP}(x_1, ..., x_n)$ and the proof is complete.* ■

In Lemma 4.1.1, it is shown that the random revenue functions for given allocations in *PMP* and *EMR* are equal. Moreover, Çetiner (2007) shows that $E(R_{EMR}(\boldsymbol{x})) = E(R_{PMP}(\boldsymbol{x}))$ and also gives the formulation for $Var(R_{EMR}(\boldsymbol{x}))$. Note that objective functions of *PMP* and *EMR* are the same. Then, equivalence of the *EMR* and *PMP* models can be claimed if the feasible regions of these two models are also the same. In Lemma 4.1.2, it is shown that the feasible regions for *EMR* and *PMP* are the same.

Let $\Phi^e$ be the feasible region for *EMR* model and $\Phi^p$ be the feasible region for *PMP* model.

**Lemma 4.1.2** $\Phi^e = \Phi^p$.

**Proof.** *The only constraint other than the integrality constraints in EMR and PMP models is the capacity constraint. Left hand side of the capacity constraint for EMR model is*

$$\sum_{j \in S_l} \sum_{i=1}^{B_j} x_j(i) = \sum_{j \in S_l} x_j \qquad for\ l = 1, ..., m,$$

*since* $x_j = \sum_{i=1}^{B_j} x_j(i)$. *Hence, the capacity constraints for EMR and PMP models are the same.* ■

Based on the observations in Lemma 4.1.1 and 4.1.2, total revenue is denoted by random variable $R$ throughout the thesis without specifying the model in the subscript. The approach we propose in this thesis is based on the use of the following risk measure: probability of total revenue being less than $L$, $P(R(\boldsymbol{x}) < L)$. This measure is limited or minimized by using the models given in the following sections. Formulation for the probability of the revenue being smaller than $L$ is given in Lemma 4.1.3. $p(\boldsymbol{d})$ and $v(\boldsymbol{x}, \boldsymbol{d})$ in Lemma 4.1.3 are introduced next. For a given sample demand $\boldsymbol{d} = (d_1, ..., d_j, ..., d_n)$ and seat allocation $\boldsymbol{x}$, $v(\boldsymbol{x}, \boldsymbol{d})$ is defined as follows:

$$\begin{aligned} v(\boldsymbol{x}, \boldsymbol{d}) &= P(\sum_{(i,j)\, \ni\, D_j \geq i} f_j x_j(i) < L \mid \boldsymbol{D} = \boldsymbol{d}) \\ &= I_{\{\sum_{(i,j)\, \ni\, d_j \geq i} f_j x_j(i) < L\}}, \end{aligned} \tag{4.8}$$

where $\boldsymbol{D} = (D_1, ..., D_n)$ is the vector of random variables denoting demands and

$$I_{\{\sum_{(i,j)\, \ni\, d_j \geq i} f_j x_j(i) < L\}} = \begin{cases} 1 & \text{if } \sum_{(i,j)\, \ni\, d_j \geq i} f_j x_j(i) < L, \\ 0 & \text{otherwise.} \end{cases}$$

Also,

$$\begin{aligned} p(\boldsymbol{d}) &= P(\boldsymbol{D} = \boldsymbol{d}) \\ &= \prod_{j=1}^{n} P(D_j = d_j). \end{aligned} \tag{4.9}$$

For a given seat allocation $\boldsymbol{x}$, $v(\boldsymbol{x}, \boldsymbol{d})$ is the (conditional) probability that total revenue is less than $L$ given that the demand is equal to $\boldsymbol{d} = (d_1, ..., d_n)$. $v(\boldsymbol{x}, \boldsymbol{d})$ turns out to be an indicator function because total revenue for the specified $\boldsymbol{x}$ would be constant for given $\boldsymbol{d}$, it is either smaller or larger than or equal to $L$. Therefore, in the proposed models, a binary decision variable is defined for $v(\boldsymbol{x}, \boldsymbol{d})$. $p(\boldsymbol{d})$ is the probability that the demand vector is equal to $\boldsymbol{d}$

and it is assumed that demands of different booking classes are independent according to the studies due to de Boer (1999). $\sum_{\boldsymbol{d}}$ denotes the summation over all $\boldsymbol{d}$, note that the random variable $D_j$ can only take integer values for every $j$.

**Lemma 4.1.3** *For a given* $\boldsymbol{x}$*,*

$$P(R(\boldsymbol{x}) < L) = \sum_{\boldsymbol{d}} v(\boldsymbol{x}, \boldsymbol{d})p(\boldsymbol{d}).$$

**Proof.**

$$P(R(\boldsymbol{x}) < L) \quad = \quad \sum_{\boldsymbol{d}} P(R(\boldsymbol{x}) < L \mid \boldsymbol{D} = \boldsymbol{d})P(\boldsymbol{D} = \boldsymbol{d}).$$

*From (4.3),*

$$P(R(\boldsymbol{x}) < L) \quad = \quad \sum_{\boldsymbol{d}} P(\sum_{j=1}^{n} \sum_{i=1}^{B_j} F_j(i)x_j(i) < L \mid \boldsymbol{D} = \boldsymbol{d})P(\boldsymbol{D} = \boldsymbol{d}).$$

*Then, by definition of* $F_j(i)$ *in (4.2),*

$$\begin{aligned} P(R(\boldsymbol{x}) < L) &= \sum_{\boldsymbol{d}} P(\sum_{j=1}^{n} f_j \sum_{i=1}^{D_j} x_j(i) < L \mid \boldsymbol{D} = \boldsymbol{d})P(\boldsymbol{D} = \boldsymbol{d}) \\ &= \sum_{\boldsymbol{d}} P(\sum_{(i,j)\, \ni\, D_j \geq i} f_j x_j(i) < L \mid \boldsymbol{D} = \boldsymbol{d})P(\boldsymbol{D} = \boldsymbol{d}) \\ &= \sum_{\boldsymbol{d}} v(\boldsymbol{x}, \boldsymbol{d})p(\boldsymbol{d}). \end{aligned}$$

■

## 4.2 Multi-Objective Optimization

The risk measure given in Section 4.1 can be used for minimizing risks. However, in RM problems, the main aim is maximization of the expected revenue. In fact, our approach is to work with two objectives: expected revenue maximization and risk minimization. This requires a multi-objective optimization formulation. Multi-objective optimization is defined as the simultaneously optimizing two or more conflicting objectives subject to certain constraints. A general maximization problem is formulated by Kosmidou and Zopounidis (2004) as follows:

$$\begin{aligned} &\max \{g_1(\boldsymbol{x}), g_2(\boldsymbol{x}), ..., g_n(\boldsymbol{x})\} \\ &subject\ to \\ &\boldsymbol{x} \in \mathcal{F}, \end{aligned}$$

where $\mathcal{F}$ is the feasible set and $g_i(\boldsymbol{x})$, $i = 1, ..., n$, are the objective functions. There exist several solution methods to solve this kind of multi-objective problems and the ones described by Kosmidou and Zopounidis (2004) are given in the following paragraphs.

**Lexicographic Optimization**:

1. Classify the objectives from the most important to the least important: $g_1, g_2, ...g_n$.
2. Maximize $g_i$ over the set of feasible solutions and let $\mathcal{F}^i$ be the set of optimal solutions in $\mathcal{F}^{i-1}$ corresponding to the maximization of $g_i$. Note that $\mathcal{F}^i \subseteq \mathcal{F}^{i-1} ... \subseteq \mathcal{F}^1 \subseteq \mathcal{F}$.
3. If $|\mathcal{F}^i| = 1$ or $i = n$, there is only one solution, otherwise set $i = i + 1$ and go to step 2.

The main advantage of this method is its simplicity.

In our models, there are only two objectives: minimizing $P(R(\boldsymbol{x}) < L)$ and expected revenue maximization. The optimal solutions of the expected revenue maximization models are generally singleton. That is, there is not in general any alternative optimal solution. Therefore, lexicographic optimization method can be used in our case by minimizing $P(R(\boldsymbol{x}) < L)$ first and then maximizing expected revenue over the set of $\boldsymbol{x}$ minimizing $P(R(\boldsymbol{x}) < L)$. The main disadvantage of the method stated by Kosmidou and Zopounidis (2004) is that independent examination of each criterion from the others degrades partially the multi-objective character of the problem.

**Global Criterion Method**: In global criterion method, the multi-objective problem is converted to a simple optimization problem by using a general function. This function is denoted by $u(\boldsymbol{x})$ and defined as $u(\boldsymbol{x}) = g(g_1(\boldsymbol{x}), g_2(\boldsymbol{x}), ..., g_n(\boldsymbol{x}))$. Then, the problem turns into a classical maximization of $u(\boldsymbol{x})$ subject to $\boldsymbol{x} \in \mathcal{F}$. The main disadvantage of the model is the difficulties in determining the $u(\boldsymbol{x})$ function. In simplest cases, this function is considered to be linear expressed as a weighted average of the goals. In our case, this method can be used by working with the following objective: $Maximize\ E(R(\boldsymbol{x})) - \delta P(R(\boldsymbol{x}) \leq L)$. However, determination of $\delta$ is not a straightforward work.

**Interactive Procedures**: This method is defined by Kosmidou and Zopounidis (2004) as follows:*"In the first stage of such procedures, an initial efficient solution is obtained and it is presented to the decision maker. If this solution is considered acceptable by the decision maker, then the solution procedure stops. If this is not the case, then the decision maker is asked to provide information regarding his preferences on the pre-specified objectives. This* *information involves the objectives that need to be improved as well as the trade-offs that he is willing to undertake to achieve these improvements. The objective of defining such information is to specify a new search direction for the development of a new improved solution. This process is repeated until a solution is obtained that is in accordance with the decision maker's preferences or until no further improvement of the current solution is possible."* This solution method requires important amount of effort. In airline revenue management problems, it is necessary to update solutions frequently and this increases the amount of required effort more. Therefore, it does not seem appropriate for airline revenue management problems with multi-objectives.

**Goal Programming**: Goal programming (GP) is given as an extension of linear programming to include multiple objectives. In goal programming, goal target values are set for objectives and then different GP variants are used to minimize the deviations from these target levels. The deviation is characterized by a difference underlying distance metric or utility function. Three of these variants are weighted goal programming, lexicographic goal programming and minmax goal programming. Goal programming is widely used in the literature because of the easy implementation. However, in weighted GP and minmax GP, there are some difficulties in setting weights. Lexicographic goal programming is not so different from lexicographic optimization and have the same disadvantages of independent examination.

All of the methods described above are analyzed for our case and lexicographic optimization method is chosen to be used in this study.

## 4.3 The Proposed Lexicographic Optimization Approach

The proposed lexicographic optimization method is used by working with two optimization models; namely, *PMP-RM-1* and *PMP-RM-2*. The abbreviation *PMP-RM* stands for *Probabilistic Mathematical Programming with Risk Measure*. The first model, *PMP-RM-1*, minimizes the probability of gaining revenue less than threshold level over the feasible region. Then, the feasible region in *PMP-RM-2* is set to the set of optimal solutions found by *PMP-RM-1*. The second model, *PMP-RM-2*, maximizes the expected revenue over this restricted feasible region. These two models are given next. At this point, $v(\boldsymbol{x}, \boldsymbol{d})$ is replaced with $v(\boldsymbol{d})$ in the models as seen below. Note that the dependence relation between $\boldsymbol{v}$ and $\boldsymbol{x}$ results from the constraints (4.12) and (4.13).

As in the *EMR* model, (4.11) is the capacity constraint used for limiting the number of seats allocated to *ODF*s using leg $l$, and (4.14) is the constraint for setting $x_j(i)$s to 0 or 1. Since $v(\boldsymbol{d})$ is equal to an indicator function for each $\boldsymbol{x}$, it is defined as a binary decision variable in (4.15). Three constraints in addition to the constraints of *EMR* model are (4.12), (4.13) and (4.15). Here, $M$ is a big number. By using (4.12) and (4.13), we guarantee that $v(\boldsymbol{d}) > 0$ when total revenue is less than $L$. In other words, if $\sum_{(i,j) \ni d_j \geq i} f_j x_j(i) < L$, then (4.13) is redundant and (4.12) implies that $v(\boldsymbol{d}) > 0$, and thus $v(\boldsymbol{d}) = 1$. If $\sum_{(i,j) \ni d_j \geq i} f_j x_j(i) \geq L$, then (4.12) is redundant and (4.13) implies that $v(\boldsymbol{d}) = 0$.

$$\textit{PMP-RM-1} \quad : \quad \textit{Minimize} \sum_{\boldsymbol{d}} p(\boldsymbol{d})v(\boldsymbol{d}) \tag{4.10}$$

*subject to*

$$\sum_{j \in S_l} \sum_{i=1}^{B_j} x_j(i) \leq C_l \quad \text{for } l = 1, ...., m, \tag{4.11}$$

$$Mv(\boldsymbol{d}) \geq - \sum_{(i,j) \ni d_j \geq i} f_j x_j(i) + L \quad \text{for all } \boldsymbol{d}, \tag{4.12}$$

$$M(1 - v(\boldsymbol{d})) \geq \sum_{(i,j) \ni d_j \geq i} f_j x_j(i) - L \quad \text{for all } \boldsymbol{d}, \tag{4.13}$$

$$x_j(i) \in \{0, 1\} \quad \text{for } j = 1, ...., n \text{ and } i = 1, ..., B_j, \tag{4.14}$$

$$v(\boldsymbol{d}) \in \{0, 1\} \quad \text{for all } \boldsymbol{d}, \tag{4.15}$$

The *PMP-RM-2* model maximizes the expected revenue when the feasible region is the set of optimal solutions corresponding to the minimization of risk measure in *PMP-RM-1*. The $v(\boldsymbol{d})$ values found in *PMP-RM-1* model are used as parameters in *PMP-RM-2*. That is, $x_j(i)$s are the only decision variables in *PMP-RM-2* model. As a result, the model is given for specified $\boldsymbol{v} = (..., v(\boldsymbol{d}), ...)$. That is why the notation used for the model is *PMP-RM-2*$|_{\boldsymbol{v}}$. The model is given below.

$$PMP\text{-}RM\text{-}2|_{\boldsymbol{v}} \quad : \quad Maximize \sum_{j=1}^{n}\sum_{i=1}^{B_j} f_j P(D_j \geq i) x_j(i) \tag{4.16}$$

$$subject\ to$$

$$\sum_{j \in S_l}\sum_{i=1}^{B_j} x_j(i) \leq C_l \quad \text{for } l = 1, ...., m, \tag{4.17}$$

$$Mv(\boldsymbol{d}) \geq -\sum_{(i,j)\ \ni\ d_j \geq i} f_j x_j(i) + L \quad \text{for all } \boldsymbol{d}, \tag{4.18}$$

$$M(1 - v(\boldsymbol{d})) \geq \sum_{(i,j)\ \ni\ d_j \geq i} f_j x_j(i) - L \quad \text{for all } \boldsymbol{d}, \tag{4.19}$$

$$x_j(i) \in \{0, 1\} \quad \text{for } j = 1, ...., n \text{ and } i = 1, ..., B_j. \tag{4.20}$$

With the use of *PMP-RM-1* and *PMP-RM-2* models, the lexicographic optimization method simply gives the optimal seat allocations that minimizes $P(R(\boldsymbol{x}) < L)$ first and then maximizes expected revenue over the set of seat allocations with minimum risk. The procedure is analyzed in detail considering the following compact representations of the *PMP-RM-1* and *PMP-RM-2* models.

$$PMP\text{-}RM\text{-}1 : Min\ \{g(\boldsymbol{v}) \mid (\boldsymbol{x}, \boldsymbol{v}) \in \Phi\},$$

$$PMP\text{-}RM\text{-}2|_{\boldsymbol{v}^*} : Max\ \{h(\boldsymbol{x}) \mid \boldsymbol{x} \in \Phi_{|\boldsymbol{v}^*}\},$$

where $g(\boldsymbol{v})$ is the objective function given in (4.10) and $h(\boldsymbol{x})$ is the objective function given in (4.16). $\Phi$ is the set of constraints in (4.11)-(4.15) and $\Phi_{|\boldsymbol{v}^*}$ is the set of constraints in (4.17)-(4.20) with $v(\boldsymbol{d}) = v^*(\boldsymbol{d})$ for all $\boldsymbol{d}$. Note that $\Phi_{|\boldsymbol{v}^*} \subseteq \Phi$. The *PMP-RM* Procedure is given below using the compact representation.

***PMP-RM*** **Procedure:**

**Step 1.** Solve *PMP-RM-1*: $Min\ \{g(\boldsymbol{v}) \mid (\boldsymbol{x}, \boldsymbol{v}) \in \Phi\}$ for $(\boldsymbol{x}^*, \boldsymbol{v}^*)$. Set $P(R(\boldsymbol{x}^*) < L) = \sum_{\boldsymbol{d}} p(\boldsymbol{d}) v^*(\boldsymbol{d})$.

**Step 2.** Solve *PMP-RM-2*$|_{\boldsymbol{v}^*}$: $Max\ \{h(\boldsymbol{x}) \mid \boldsymbol{x} \in \Phi_{|\boldsymbol{v}^*}\}$ for the optimal allocation $\boldsymbol{x}^{**}$. Set $P(R(\boldsymbol{x}^{**}) < L) = \sum_{\boldsymbol{d}} p(\boldsymbol{d}) v^*(\boldsymbol{d})$ and $E(R(x^{**})) = \sum_{j=1}^{n}\sum_{i=1}^{B_j} f_j P(D_j \geq i) x_j^{**}(i)$.

It must be noted here that the resulting seat allocations for *PMP-RM-1* may have a conflict with the definition of $x_j(i)$ in (4.4). In other words, there may exist an optimal allocation $\boldsymbol{x}^*$ for *PMP-RM-1* such that some $x_j^*(i') = 0$ and $x_j^*(i'') = 1$ for some $j$ although $i'' > i'$. However, it is shown in the following lemmas, propositions and remarks that there exists at least one optimal allocation for *PMP-RM-1* model that is also a proper allocation in terms of the definition of $x_j(i)$ in (4.4).

In Lemma 4.3.1, it is shown that the optimal solution for *PMP-RM-1* is feasible for *PMP-RM-2*$|_{\boldsymbol{v}^*}$ and all of the feasible solutions of *PMP-RM-2*$|_{\boldsymbol{v}^*}$ are optimal for *PMP-RM-1*. That is, *PMP-RM-2*$|_{\boldsymbol{v}^*}$ is constructed in such a way that the feasible region of *PMP-RM-2*$|_{\boldsymbol{v}^*}$, namely $\Phi_{|v^*}$, consists of optimal solutions of *PMP-RM-1*. This is true in general for lexicographic optimization, but it is shown here for our models.

**Lemma 4.3.1** $(\boldsymbol{x}, \boldsymbol{v}^*)$ *for any* $\boldsymbol{x} \in \Phi_{|v^*}$ *is optimal for PMP-RM-1.*

**Proof.** *Proof follows from* $\{\boldsymbol{x} \mid (\boldsymbol{x}, \boldsymbol{v}^*) \in \Phi\} = \Phi_{|v^*}$. ■

The lemmas, remarks and the proposition in this section aim to show that the optimal seat allocation for the proposed lexicographic optimization satisfies Condition 4.3.2 below. Condition 4.3.2 is given for the rule of allocating a full seat to $x_j(i)$ before any portion of a seat is allocated to $x_j(i+1)$.

**Condition 4.3.2** *For* $j = 1, ..., n$*, if* $x_j(i+1) = 1$*, then* $x_j(i) = 1$.

For the case $\boldsymbol{x}$ does not satisfy Condition 4.3.2, $\boldsymbol{x}'$ is defined below in Definition 4.3.3. This definition is used for arranging $\boldsymbol{x}_j(i)$s in a row. The use of Condition 4.3.2 and Definition 4.3.3 is shown in Example 4.3.4.

**Definition 4.3.3** *For an allocation* $\boldsymbol{x}$*,* $\boldsymbol{x}'$ *satisfies Condition 4.3.2 and the equation below:*

$$\sum_{i=1}^{B_j} x'_j(i) = \sum_{i=1}^{B_j} x_j(i) \qquad \text{for } j = 1, ..., n. \tag{4.21}$$

Equation (4.21) is used for equating the total number of seat allocations in $\boldsymbol{x}$ and $\boldsymbol{x}'$ for each $j = 1, ..., n$.

**Example 4.3.4** *Suppose that the capacity of a sample single leg flight is 10 and an example allocation is* $\boldsymbol{x}$ *=[1 1 0 0 1 0 1 0 0 1]. Then, the corresponding allocation* $\boldsymbol{x}'$ *given by Definition 4.3.3 is* $\boldsymbol{x}'$ *=[1 1 1 1 1 0 0 0 0 0].*

In Lemma 4.3.5 below, it is shown that the objective function value of *PMP-RM-2*$|_{\boldsymbol{v}^*}$ for $\boldsymbol{x}'$ is greater than the one for $\boldsymbol{x}$.

**Lemma 4.3.5** *Let* $\boldsymbol{x}$ *be an allocation that does not satisfy Condition 4.3.2 and* $\boldsymbol{x}'$ *be given as in Definition 4.3.3. Then,* $h(\boldsymbol{x}) \leq h(\boldsymbol{x}')$.

**Proof.** *From (4.21),*

$$f_j \sum_{i=1}^{B_j} x'_j(i) \quad = \quad f_j \sum_{i=1}^{B_j} x_j(i) \quad \text{for } j = 1, ..., n.$$

*Since* $\boldsymbol{x}'$ *satisfies Condition 4.3.2 and (4.21), and* $P(D_j \geq i)$ *is nonincreasing in i,*

$$\sum_{i=1}^{B_j} f_j P(D_j \geq i) x'_j(i) \quad \geq \quad \sum_{i=1}^{B_j} f_j P(D_j \geq i) x_j(i) \quad \text{for } j = 1, ..., n.$$

*As a result,*

$$\sum_{j=1}^{n} \sum_{i=1}^{B_j} f_j P(D_j \geq i) x'_j(i) \quad \geq \quad \sum_{j=1}^{n} \sum_{i=1}^{B_j} f_j P(D_j \geq i) x_j(i).$$

■

Lemma 4.3.5 is given for the expected total revenue. In Lemma 4.3.6, total revenue figures for $\boldsymbol{x}'$ and $\boldsymbol{x}$ are compared for a particular given demand realization $\boldsymbol{d} = (d_1, ..., d_n)$. This lemma is given in order to show that not only the expected revenue but also the revenue for a given demand vector for allocation $\boldsymbol{x}'$ is greater than or equal to the revenue for allocation $\boldsymbol{x}$.

**Lemma 4.3.6** *Let* $\boldsymbol{x}$ *be an allocation that does not satisfy Condition 4.3.2 and* $\boldsymbol{x}'$ *be given as in Definition 4.3.3. Then,*

$$\sum_{(i,j) \ni d_j \geq i} f_j x_j(i) \quad \leq \quad \sum_{(i,j) \ni d_j \geq i} f_j x'_j(i) \quad \text{for any } \boldsymbol{d}.$$

**Proof.** *Let* $\boldsymbol{D} = \boldsymbol{d}$. *For* $\boldsymbol{x}$, *total revenue is equal to*

$$\sum_{(i,j) \ni d_j \geq i} f_j x_j(i) \quad = \quad \sum_{j=1}^{n} f_j (\sum_{i=1}^{d_j} x_j(i)). \tag{4.22}$$

*From Definition 4.3.3,*

$$\sum_{i=1}^{d_j} x_j(i) \quad \leq \quad \sum_{i=1}^{d_j} x'_j(i) \quad \text{for } j = 1, .., n.$$

*Then, using the relation in (4.22),*

$$\begin{aligned} \sum_{(i,j) \ni d_j \geq i} f_j x_j(i) \quad &\leq \quad \sum_{j=1}^{n} f_j (\sum_{i=1}^{d_j} x'_j(i)) \\ &= \quad \sum_{(i,j) \ni d_j \geq i} f_j x'_j(i). \end{aligned}$$

■

In Lemma 4.3.7, it is shown that $\boldsymbol{x}^{*\prime}$ that satisfies Condition 4.3.2 and (4.21) is a feasible solution of *PMP-RM-2*$|_{\boldsymbol{v}^*}$ for optimal solution, $(\boldsymbol{x}^*, \boldsymbol{v}^*)$, of *PMP-RM-1*. Then, from Lemma 4.3.1, $(\boldsymbol{x}^{*\prime}, \boldsymbol{v}^*)$ is an optimal solution of *PMP-RM-1*.

**Lemma 4.3.7** *Let* $(\boldsymbol{x}^*, \boldsymbol{v}^*)$ *be an optimal solution for the PMP-RM-1 model and* $\boldsymbol{x}^{*\prime}$ *be given as in Definition 4.3.3. Then,* $\boldsymbol{x}^{*\prime} \in \Phi|_{\boldsymbol{v}^*}$.

**Proof.** *The proof is given by showing that all of the constraints in PMP-RM-2*$|_{\boldsymbol{v}^*}$ *are satisfied by* $\boldsymbol{x}^{*\prime}$. $\boldsymbol{x}^*$ *satisfies (4.11) in PMP-RM-1. Then, due to (4.21) in Definition 4.3.3, constraint (4.17) in PMP-RM-2*$|_{\boldsymbol{v}}$ *for* $\boldsymbol{v} = \boldsymbol{v}^*$ *is satisfied by* $\boldsymbol{x}^{*\prime}$. *Constraint (4.18) in PMP-RM-2*$|_{\boldsymbol{v}}$ *for* $\boldsymbol{v} = \boldsymbol{v}^*$ *is also satisfied by* $\boldsymbol{x}^{*\prime}$ *as shown below.*

$$\begin{aligned} Mv^*(\boldsymbol{d}) &\geq -\sum_{(i,j)\,\ni\, d_j \geq i} f_j x_j^*(i) + L \quad \text{for all } \boldsymbol{d}, \\ &\geq -\sum_{(i,j)\,\ni\, d_j \geq i} f_j x_j^{*\prime}(i) + L \quad \text{for all } \boldsymbol{d}, \end{aligned}$$

*where the first inequality is (4.12) in PMP-RM-1 under the optimal policy and the second inequality results from Lemma 4.3.6.*

*Now, consider constraint (4.19) in PMP-RM-2*$|_{\boldsymbol{v}}$ *for* $\boldsymbol{v} = \boldsymbol{v}^*$.

- *If* $\sum_{(i,j)\,\ni\,\bar{\boldsymbol{d}}_j \geq i} f_j x_j^{*\prime}(i) - L \leq 0$ *for some* $\bar{\boldsymbol{d}}$, *then*

$$M(1 - v^*(\bar{\boldsymbol{d}})) \geq \sum_{(i,j)\,\ni\,\bar{\boldsymbol{d}}_j \geq i} f_j x_j^{*\prime}(i) - L$$

*is redundant. That is, in this case,* $\boldsymbol{x}^{*\prime}$ *satisfies (4.19) when* $v = v^*$.

- *If* $\sum_{(i,j)\,\ni\,\bar{\boldsymbol{d}}_j \geq i} f_j x_j^{*\prime}(i) - L > 0$ *for some* $\bar{\boldsymbol{d}}$, *then consider the two cases:* $v^*(\bar{\boldsymbol{d}}) = 0$ *and* $v^*(\bar{\boldsymbol{d}}) = 1$.

*If* $v^*(\bar{\boldsymbol{d}}) = 0$, *then*

$$M(1 - v^*(\bar{\boldsymbol{d}})) \geq \sum_{(i,j)\,\ni\,\bar{\boldsymbol{d}}_j \geq i} f_j x_j^{*\prime}(i) - L$$

*is redundant because the left hand side above is equal to M.*

*If* $v^*(\bar{\boldsymbol{d}}) = 1$, *then*

$$M(1 - v^*(\bar{\boldsymbol{d}})) \ngeq \sum_{(i,j)\,\ni\,\bar{\boldsymbol{d}}_j \geq i} f_j x_j^{*\prime}(i) - L,$$

*because the left hand side is equal to zero while the right hand side is positive. (Note that, in this case,* $\sum_{(i,j)\,\ni\,\bar{d}_j\geq i} f_j x_j^*(i) - L \leq 0$ *from (4.13).) That is, constraint (4.19) in PMP-RM-2*$|_{\boldsymbol{v}}$ *for* $\boldsymbol{v} = \boldsymbol{v}^*$ *is not satisfied by* $\boldsymbol{x^{*\prime}}$*. In this case, letting*

$$\bar{v}(\boldsymbol{d}) = \begin{cases} 0 & \textit{for } \boldsymbol{d} \ni \sum_{(i,j)\,\ni\,d_j\geq i} f_j x_j^{*\prime}(i) - L > 0 \textit{ and } v^*(\boldsymbol{d}) = 1, \\ v^*(\boldsymbol{d}) & \textit{otherwise}, \end{cases}$$

*it is observed that (*$\boldsymbol{x^{*\prime}}$*,* $\bar{\boldsymbol{v}}$*) is feasible for PMP-RM-1. (*$\bar{\boldsymbol{v}}$ *is defined in order for* $\boldsymbol{x^{*\prime}}$ *to satisfy (4.19) also for* $\boldsymbol{d}$ *such that* $\sum_{(i,j)\,\ni\,d_j\geq i} f_j x_j^{*\prime}(i) - L > 0$ *and* $\boldsymbol{v}^*(\boldsymbol{d}) = 1$*.) Note that (4.12) is not violated by (*$\boldsymbol{x^{*\prime}}$*,* $\bar{\boldsymbol{v}}$*) in PMP-RM-1. This is because the right (left) hand side in (4.12) below*

$$M\bar{v}(\boldsymbol{d}) \geq -\sum_{(i,j)\,\ni\,d_j\geq i} f_j x_j^{*\prime}(i) + L$$

*is negative (0) for each* $\boldsymbol{d}$ *such that* $\sum_{(i,j)\,\ni\,d_j\geq i} f_j x_j^{*\prime}(i) - L > 0$ *and* $v^*(\boldsymbol{d}) = 1$*. Then,* $(\boldsymbol{x^{*\prime}}, \bar{\boldsymbol{v}}) \in \Phi$*. Now, consider the objective function* $g(\boldsymbol{v}^*)$ *of PMP-RM-1 for* $(\boldsymbol{x^*}, \boldsymbol{v}^*)$ *and* $g(\bar{\boldsymbol{v}})$ *for* $(\boldsymbol{x^{*\prime}}, \bar{\boldsymbol{v}})$*. Here, letting* $\mathfrak{D} = \left\{\boldsymbol{d} \mid \sum_{(i,j)\,\ni\,d_j\geq i} f_j x_j^{*\prime}(i) - L > 0 \textit{ and } v^*(\boldsymbol{d}) = 1\right\}$*,*

$$\begin{aligned} g(\boldsymbol{v}^*) &= \sum_{\boldsymbol{d}\in\mathfrak{D}} p(\boldsymbol{d})v^*(\boldsymbol{d}) + \sum_{\boldsymbol{d}\notin\mathfrak{D}} p(\boldsymbol{d})v^*(\boldsymbol{d}) \\ &> \sum_{\boldsymbol{d}\in\mathfrak{D}} p(\boldsymbol{d})\bar{v}(\boldsymbol{d}) + \sum_{\boldsymbol{d}\notin\mathfrak{D}} p(\boldsymbol{d})v^*(\boldsymbol{d}) \quad \textit{by definition of } \bar{\boldsymbol{v}}, \\ &= g(\bar{\boldsymbol{v}}). \end{aligned}$$

*This contradicts to the optimality of* $(\boldsymbol{x}^*, \boldsymbol{v}^*)$ *for PMP-RM-1. As a result, there does not exist a demand vector* $\bar{\boldsymbol{d}}$ *causing* $\boldsymbol{x^{*\prime}}$ *to be infeasible for PMP-RM-2*$|_{\boldsymbol{v}^*}$*.* ■

**Remark 4.3.8** *Let* $(\boldsymbol{x}^*, \boldsymbol{v}^*)$ *be an optimal solution for the PMP-RM-1 model and* $\boldsymbol{x^{*\prime}}$ *be given as in Definition 4.3.3. From Lemma 4.3.7,* $(\boldsymbol{x^{*\prime}}, \boldsymbol{v}^*)$ *is also optimal for PMP-RM-1. As a result, there exists at least one optimal allocation for PMP-RM-1 (feasible allocation for PMP-RM-2*$|_{\boldsymbol{v}^*}$*) that satisfies Condition 4.3.2.*

*If* $\boldsymbol{x}^*$ *satisfies Condition 4.3.2, then the observation above immediately follows. If* $\boldsymbol{x}^*$ *does not satisfy Condition 4.3.2, then* $\boldsymbol{x^{*\prime}}$ *given in Definition 4.3.3 is in* $\Phi|_{\boldsymbol{v}^*}$ *(feasible for PMP-RM-2*$|_{\boldsymbol{v}^*}$*) from Lemma 4.3.7, and, thus,* $\boldsymbol{x^{*\prime}}$ *is optimal for PMP-RM-1 from Lemma 4.3.1. As a result,* $\Phi|_{\boldsymbol{v}^*}$ *is not restricted to the allocations (optimal allocations for PMP-RM-1) that do not satisfy Condition 4.3.2. Furthermore,* $\boldsymbol{x^{*\prime}}$ *is not dominated by* $\boldsymbol{x}^*$ *in PMP-RM-2*$|_{\boldsymbol{v}^*}$ *as shown in Lemma 4.3.5.* ■

In order to conclude that $\boldsymbol{x}^{**}$ as the optimal solution of *PMP-RM-2*$|_{\boldsymbol{v}^*}$ satisfies Condition 4.3.2, the final observation used in Proposition 4.3.9 below is that $P(D_j \geq i)$ in $h(\boldsymbol{x})$ decreases monotonically in $i$ for each $j$.

**Proposition 4.3.9** *Let $\boldsymbol{v}^*$ and $\boldsymbol{x}^{**}$ be optimal solutions for PMP-RM-1 and PMP-RM-2$|_{\boldsymbol{v}^*}$, respectively. Then, $\boldsymbol{x}^{**}$ satisfies Condition 4.3.2.*

**Proof.** *Let $\boldsymbol{x}^*$ denote optimal solution of PMP-RM-1. From Lemma 4.3.7 and Remark 4.3.8, $\boldsymbol{x}^{*\prime}$ that satisfies Condition 4.3.2 and (4.21) is in $\Phi|_{\boldsymbol{v}^*}$ for every $\boldsymbol{x}^*$. Also, $\Phi|_{\boldsymbol{v}^*}$ is restricted to the optimal solutions of PMP-RM-1. Then, $\boldsymbol{x}^{**}$ satisfies Condition 4.3.2 due to Lemma 4.3.5. Recall that $P(D_j \geq i)$ in $h(\boldsymbol{x})$ decreases monotonically as i increases for each $j$.* ■

**Remark 4.3.10** *Let $\boldsymbol{x}^*$ and $\boldsymbol{x}^{**}$ denote optimal solutions of PMP-RM-1 and PMP-RM-2$|_{\boldsymbol{v}^*}$, respectively.*

***a)*** *$h(\boldsymbol{x}^*) \leq h(\boldsymbol{x}^{**})$.*

*Let $\boldsymbol{x}^{*\prime}$ be given as in Definition 4.3.3. $\boldsymbol{x}^*$, $\boldsymbol{x}^{*\prime}$ and $\boldsymbol{x}^{**}$ are in $\Phi|_{\boldsymbol{v}^*}$. $h(\boldsymbol{x}^*) \leq h(\boldsymbol{x}^{*\prime}) \leq h(\boldsymbol{x}^{**})$ where the first inequality is due to Lemma 4.3.5 and the second inequality follows from the optimality of $\boldsymbol{x}^{**}$ for PMP-RM-2$|_{\boldsymbol{v}^*}$. Note that if $\boldsymbol{x}^* \neq \boldsymbol{x}^{**}$, then $\boldsymbol{x}^*$ and $\boldsymbol{x}^{**}$ are alternative optima for PMP-RM-1.*

***b)*** *$P(R(\boldsymbol{x}^*) < L) = P(R(\boldsymbol{x}^{**}) < L) = \sum_{\boldsymbol{d}} p(\boldsymbol{d})v^*(\boldsymbol{d})$.* ■

## 4.4 The Proposed Approximation

The main difficulty for the procedure given in Section 4.3 is that linear relaxation cannot be avoided for the Integer Programming (IP) formulations with binary decision variables $x_j(i)$ and $v(\boldsymbol{x}, \boldsymbol{d})$. Because of the computational burden of *PMP-RM* procedure, an approximate procedure for *PMP-RM* is suggested in this section. Consider the *PMP-RM* models for an example network. Let the number of legs be $m = 3$ and number of *ODF*s be $n = 18$ and the bound for $i$ be $B_j = 100$ for all $j$. Then, there are 1800 real decision variables, $x_j(i)$s, and $(100 + 1)^{18} \approx 10^{36}$ integer decision variables, $\boldsymbol{v}$s, and $2 \times (100 + 1)^{18} + 1 \approx 2 \times 10^{36}$ constraints in *PMP-RM-1* model. Although this sample network is smaller than real life networks, use of the approximations is unavoidable because of the computational time and memory limits.

The approximation we propose in this thesis is underlined below.

(1) The *PMP-RM* models are solved only for a small number of demand realizations. The idea behind this approximation is not so different from the *RLP* model. Talluri and van Ryzin (1999) solve *RLP* model many times, which is suggested to be more than 20, each for a demand realization and, then, the average of the bid prices obtained for each run is used in the control policy. In this thesis, a given number of demand realizations but not all of them are considered simultaneously while solving the *PMP-RM* models just once and the resulting allocations or bid prices are used in control policies.

(2) Based on the integrality property of the network problems in Section 3.2, Williamson (1992) claims that the linear relaxation of the *EMR* model gives integer solutions. In other words, the claim is the following: if the upper and lower bounds on the decision variables are integer and the right hand side values of the flow balance constraints are integer, then the solution will be integer. However, de Boer (1999) gives a counter example to show that, under these conditions, the models can give non-integer solutions. de Boer (1999) also states that, in real life problems, the solutions of the models are surprisingly integer. In our numerical experiments, only one of all seats of an *ODF* can take a fractional value (as in Condition 4.4.1) when the integrality constraints for $x_j(i)s$ are relaxed. Based on these observations, integrality constraints for $\boldsymbol{x}_j(i)s$ are relaxed for both *PMP-RM-1* and *PMP-RM-2*. According to Condition 4.4.1, only one $x_j(i)$ value can be fractional for each *ODF* $j$. If $x_j(\varpi)$ is fractional, then all of the $x_j(i)s$ are zero for $i > \varpi$ and $x_j(i)s$ are equal to 1 for $i < \varpi$.

**Condition 4.4.1** *For* $j = 1, ..., n$, *if* $x_j(i+1) = 1$, *then* $x_j(i) = 1$; *if* $x_j(i) < 1$, *then* $x_j(i+1) = 0$.

(3) The *PMP-RM* models are solved by relaxing the integrality constraints for $x_j(i)$ and $v(\boldsymbol{d})$ in the *PMP-RM-1* model and, then, rounding the positive $v(\boldsymbol{d})$ values less than 1 to 1. Because of the definition of $v(\boldsymbol{d})$ in (4.8), if $v(\boldsymbol{d}) = 0$, then the probability that revenue is less than the threshold level $L$ is equal to zero for given $\boldsymbol{d}$ and the associated seat allocation. On the other hand, this probability is positive if $v(\boldsymbol{d}) > 0$. The approximation we consider is rounding $v(\boldsymbol{d})$ values to 1 when they are greater than 0 but smaller than 1. $v(\boldsymbol{d})$ values are rounded considering the definition of $v(\boldsymbol{d})$s. For the risk measure given in Section 4.1, $P(R(\boldsymbol{x}) < L)$ is calculated by using $v(\boldsymbol{x}, \boldsymbol{d})$, which is an indicator function. If fractional $v(\boldsymbol{d})$ values are used to calculate $P(R(\boldsymbol{x}) < L)$, the value found for $P(R(\boldsymbol{x}) < L)$ would be questioned.

These approximations make our models linear and easy to solve. Although the solution quality gets worse as a result of the approximations, the numerical results are acceptable according to the experiments reported in Chapter 5 and Chapter 6.

First, consider approximations (2) and (3) above. The relaxed models are called *PLP-RM*, which stands for *Probabilistic Linear Programming with Risk Measure*. $(x^*, v^*)$ denotes optimal solution of *PMP-RM-1*. $\tilde{\boldsymbol{v}}$ is obtained by rounding $\boldsymbol{v}^*$ as explained for the approximation (3) above. The compact representations of *PLP-RM* models are as follows:

$$PLP\text{-}RM\text{-}1 : Min\ \{g(\boldsymbol{v}) \mid (\boldsymbol{x}, \boldsymbol{v}) \in \tilde{\Phi}\},$$

$$PLP\text{-}RM\text{-}2|_{\tilde{\boldsymbol{v}}} : Max\ \{h(\boldsymbol{x}) \mid \boldsymbol{x} \in \tilde{\Phi}_{|\tilde{\boldsymbol{v}}}\},$$

where $\tilde{\Phi}$ is the set of constraints in (4.11)-(4.13) with $0 \le x_j(i) \le 1$ and $0 \le v(\boldsymbol{d}) \le 1$ instead of (4.14) and (4.15), respectively. $\tilde{\Phi}_{|\tilde{\boldsymbol{v}}}$ is the set of constraints in (4.17)-(4.19) with $v(\boldsymbol{d}) = \tilde{v}(\boldsymbol{d})$ and $0 \le x_j(i) \le 1$ instead of (4.20). $\boldsymbol{x}^{**}$ is the optimal solution of *PMP-RM-2*$|_{\tilde{\boldsymbol{v}}}$. Then, *PLP-RM-1* and *PLP-RM-2*$|_{\boldsymbol{v}}$ models are given as follows:

$$PLP\text{-}RM\text{-}1 \quad : \quad Minimize \sum_{\boldsymbol{d}} p(\boldsymbol{d})v(\boldsymbol{d}) \tag{4.23}$$

$$subject\ to$$

$$\sum_{j \in S_l} \sum_{i=1}^{B_j} x_j(i) \le C_l \quad \text{for } l = 1, ...., m, \tag{4.24}$$

$$Mv(\boldsymbol{d}) \ge - \sum_{(i,j)\, \ni\, d_j \ge i} f_j x_j(i) + L \quad \text{for all } \boldsymbol{d}, \tag{4.25}$$

$$M(1 - v(\boldsymbol{d})) \ge \sum_{(i,j)\, \ni\, d_j \ge i} f_j x_j(i) - L \quad \text{for all } \boldsymbol{d}, \tag{4.26}$$

$$0 \le x_j(i) \le 1 \quad \text{for } j = 1, ...., n \text{ and } i = 1, ..., B_j,$$

$$0 \le v(\boldsymbol{d}) \le 1 \text{ for all } \boldsymbol{d}.$$

$$PLP\text{-}RM\text{-}2|_{\boldsymbol{v}} \quad : \quad Maximize \sum_{j=1}^{n} \sum_{i=1}^{B_j} f_j P(D_j \ge i) x_j(i) \tag{4.27}$$

$$subject\ to$$

$$\sum_{j \in S_l} \sum_{i=1}^{B_j} x_j(i) \le C_l \quad \text{for } l = 1, ...., m, \tag{4.28}$$

$$Mv(\boldsymbol{d}) \ge - \sum_{(i,j)\, \ni\, d_j \ge i} f_j x_j(i) + L \quad \text{for all } \boldsymbol{d}, \tag{4.29}$$

$$M(1 - v(\boldsymbol{d})) \ge \sum_{(i,j)\, \ni\, d_j \ge i} f_j x_j(i) - L \quad \text{for all } \boldsymbol{d}, \tag{4.30}$$

$$0 \le x_j(i) \le 1 \quad \text{for } j = 1, ...., n \text{ and } i = 1, ..., B_j.$$

The *PMP-RM* Procedure is revised by using these relaxed models above and the resulting *PLP-RM* Procedure is given next with the use of approximations (2) and (3).

***PLP-RM* Procedure:**

**Step 1.** Solve *PLP-RM-1*: $Min\ \{g(\boldsymbol{v}) \mid (\boldsymbol{x}, \boldsymbol{v}) \in \tilde{\Phi}\}$ for $(\boldsymbol{x}^*, \boldsymbol{v}^*)$. For all $\boldsymbol{d}$, set

$$\tilde{\boldsymbol{v}}(\boldsymbol{d}) = \begin{cases} 1 & \text{if } \boldsymbol{v}^*(\boldsymbol{d}) > 0, \\ 0 & \text{if } \boldsymbol{v}^*(\boldsymbol{d}) = 0. \end{cases} \tag{4.31}$$

Set $P(R(\boldsymbol{x}^*) < L) = \sum_{\boldsymbol{d}} p(\boldsymbol{d})\tilde{v}(\boldsymbol{d})$.

**Step 2.** Solve *PLP-RM-2*$|_{\tilde{\boldsymbol{v}}}$: $Max\ \{h(\boldsymbol{x}) \mid \boldsymbol{x} \in \tilde{\Phi}_{|\tilde{\boldsymbol{v}}}\}$ for the optimal allocation $\boldsymbol{x}^{**}$. Set $P(R(\boldsymbol{x}^{**}) < L) = \sum_{\boldsymbol{d}} p(\boldsymbol{d})\tilde{v}(\boldsymbol{d})$ and $E(R(x^{**})) = \sum_{j=1}^{n} \sum_{i=1}^{B_j} f_j P(D_j \geq i) x_j(i)$.

For any fractional $x_j^{**}(i)$ value obtained by solving *PMP-RM-2*$|_{\tilde{\boldsymbol{v}}}$, the fractional value is rounded up (down) when $x_j^{**}(i) \geq 0.5$ ($x_j^{**}(i) < 0.5$). This is for the use of allocation is practice with real life partitioned or nested booking policies.

Next, Definition 4.3.3 and Lemmas 4.3.5 and 4.3.6 are revised for the relaxed cases as in Definition 4.4.2 and in Lemmas 4.4.4 and 4.4.5.

**Definition 4.4.2** *For an allocation* $\boldsymbol{x}$, $\boldsymbol{x}'$ *satisfies Condition 4.4.1 and the equation below.*

$$\sum_{i=1}^{B_j} x'_j(i) \quad = \quad \sum_{i=1}^{B_j} x_j(i) \qquad \text{for } j = 1, ..., n. \tag{4.32}$$

**Example 4.4.3** *Suppose that the capacity of a sample single leg flight is 10 and an example allocation is* $\boldsymbol{x}$ *=[0.5 0.3 0 0 0.8 0 0.2 0 0 0.9]. Then, the corresponding allocation* $\boldsymbol{x}'$ *given by Definition 4.3.3 is* $\boldsymbol{x}'$ *=[1 1 0.7 0 0 0 0 0 0 0].*

**Lemma 4.4.4** *Let* $\boldsymbol{x}$ *be an allocation that does not satisfy Condition 4.4.1 and* $\boldsymbol{x}'$ *be given as in Definition 4.4.2. Then,* $h(\boldsymbol{x}) \leq h(\boldsymbol{x}')$. ■

**Lemma 4.4.5** *Let* $\boldsymbol{x}$ *be an allocation that does not satisfy Condition 4.4.1 and* $\boldsymbol{x}'$ *be given as in Definition 4.4.2. Then,*

$$\sum_{(i,j)\,\ni\, d_j \geq i} f_j x_j(i) \quad \leq \quad \sum_{(i,j)\,\ni\, d_j \geq i} f_j x'_j(i) \qquad \text{for any } \boldsymbol{d}.$$

■

In Lemmas 4.4.4 and 4.4.5, the relation between an optimal allocation $\boldsymbol{x}$ and a rearranged allocation $\boldsymbol{x}'$ are studied. The observation in Lemma 4.3.7 is revised in Lemma 4.4.6 for the proposed approach with approximations (2) and (3).

**Lemma 4.4.6** *Let $(\boldsymbol{x}^*, \boldsymbol{v}^*)$ be an optimal solution for PLP-RM-1 model and $\boldsymbol{x}^{*\prime}$ be given as in Definition 4.4.2. Let*

$$\tilde{\boldsymbol{v}}(\boldsymbol{d}) = \begin{cases} 0 & \text{if } \boldsymbol{v}^*(\boldsymbol{d}) = 0, \\ 1 & \text{if } \boldsymbol{v}^*(\boldsymbol{d}) > 0. \end{cases} \tag{4.33}$$

*Then, $\boldsymbol{x}^{*\prime} \in \tilde{\Phi}_{|\tilde{\boldsymbol{v}}}$.*

**Proof.** *The proof is given by showing that all of the constraints in PLP-RM-2$|_{\tilde{\boldsymbol{v}}}$ are satisfied by $\boldsymbol{x}^{*\prime}$. $\boldsymbol{x}^*$ satisfies (4.24) in PLP-RM-1. Then, due to (4.32) in Definition 4.4.2, constraint (4.28) in PLP-RM-2$|_{\boldsymbol{v}}$ for $\boldsymbol{v} = \tilde{\boldsymbol{v}}$ is satisfied by $\boldsymbol{x}^{*\prime}$. Constraint (4.29) in PLP-RM-2$|_{\boldsymbol{v}}$ for $\boldsymbol{v} = \tilde{\boldsymbol{v}}$ is also satisfied by $\boldsymbol{x}^{*\prime}$ as shown below.*

$$\begin{aligned} M\tilde{v}(\boldsymbol{d}) \geq Mv^*(\boldsymbol{d}) \quad &\geq \quad -\sum_{(i,j)\,\ni\, d_j \geq i} f_j x_j^*(i) + L \qquad for\ all\ \boldsymbol{d}, \\ &\geq \quad -\sum_{(i,j)\,\ni\, d_j \geq i} f_j x_j^{*\prime}(i) + L \qquad for\ all\ \boldsymbol{d}, \end{aligned}$$

*where the first inequality results from the definition of $\tilde{\boldsymbol{v}}(\boldsymbol{d})$ and the second inequality is (4.25) in PLP-RM-1 under the optimal policy and the third inequality results from Lemma 4.4.5. Now, consider constraint (4.30) in PLP-RM-2$|_{\boldsymbol{v}}$ for $\boldsymbol{v} = \tilde{\boldsymbol{v}}$.*

- *If $\sum_{(i,j)\,\ni\,\bar{d}_j \geq i} f_j x_j^{*\prime}(i) - L \leq 0$ for some $\bar{\boldsymbol{d}}$, then, $\boldsymbol{x}^{*\prime}$ satisfies (4.30) when $\boldsymbol{v} = \tilde{\boldsymbol{v}}$ as shown below.*

$$M(1 - v^*(\bar{\boldsymbol{d}})) \geq M(1 - \tilde{v}(\bar{\boldsymbol{d}})) \geq 0 \geq \sum_{(i,j)\,\ni\,\bar{d}_j \geq i} f_j x_j^{*\prime}(i) - L.$$

- *If $\sum_{(i,j)\,\ni\,\bar{d}_j \geq i} f_j x_j^{*\prime}(i) - L > 0$ for some $\bar{\boldsymbol{d}}$, then consider the two cases: $v^*(\bar{\boldsymbol{d}}) = 0$ and $v^*(\bar{\boldsymbol{d}}) > 0$.*

*If $v^*(\bar{\boldsymbol{d}}) = 0$, then $\tilde{v}(\bar{\boldsymbol{d}}) = 0$. In this case, $\boldsymbol{x}^{*\prime}$ satisfies (4.30) when $\boldsymbol{v} = \tilde{\boldsymbol{v}}$ as shown below.*

$$M = M(1 - v^*(\bar{\boldsymbol{d}})) = M(1 - \tilde{v}(\bar{\boldsymbol{d}})) \quad \geq \quad \sum_{(i,j)\,\ni\,\bar{d}_j \geq i} f_j x_j^{*\prime}(i) - L.$$

*If* $v^*(\bar{\boldsymbol{d}}) > 0$*, then* $\tilde{v}(\bar{\boldsymbol{d}}) = 1$*. In this case,*

$$M(1 - \tilde{v}(\bar{\boldsymbol{d}})) \quad \not\geq \quad \sum_{(i,j)\,\ni\,\bar{d}_j \geq i} f_j x_j^{*\prime}(i) - L$$

*because the left (right) hand side above is zero (positive). Here,* $M(1 - v^*(\bar{\boldsymbol{d}})) \geq M(1 - \tilde{v}(\bar{\boldsymbol{d}}))$*. (Note that, in this case,* $\sum_{(i,j)\,\ni\,\bar{d}_j \geq i} f_j x_j^*(i) - L \leq M(1 - \boldsymbol{v}^*(\bar{\boldsymbol{d}}))$ *from (4.26).) That is, constraint (4.30) in PLP-RM-2*$|_{\boldsymbol{v}}$ *for* $\boldsymbol{v} = \tilde{\boldsymbol{v}}$ *is not satisfied by* $\boldsymbol{x}^{*\prime}$*.*

*In this case, letting*

$$\bar{v}(\boldsymbol{d}) = \begin{cases} 0 & \text{for } \boldsymbol{d} \ \ni \ \sum_{(i,j)\,\ni\,d_j \geq i} f_j x_j^{*\prime}(i) - L > 0 \text{ and } \boldsymbol{v}^*(\boldsymbol{d}) > 0, \\ v^*(\boldsymbol{d}) & \text{otherwise,} \end{cases}$$

*it is observed that* $(\boldsymbol{x}^{*\prime}, \bar{\boldsymbol{v}})$ *is feasible for PLP-RM-1. (*$\bar{\boldsymbol{v}}$ *is defined in order for* $\boldsymbol{x}^{*\prime}$ *to satisfy (4.30) also for* $\boldsymbol{d}$ *such that* $\sum_{(i,j)\,\ni\,d_j \geq i} f_j x_j^{*\prime}(i) - L > 0$ *and* $\tilde{\boldsymbol{v}}(\boldsymbol{d}) = 1$*.) Note that (4.25) is not violated by* $(\boldsymbol{x}^{*\prime}, \bar{\boldsymbol{v}})$ *in PLP-RM-1. This is because the right (left) hand side in (4.25) below*

$$M\bar{\boldsymbol{v}}(\boldsymbol{d}) \geq -\sum_{(i,j)\,\ni\,d_j \geq i} f_j x_j^{*\prime}(i) + L$$

*is negative (0) for each* $\boldsymbol{d}$ *such that* $\sum_{(i,j)\,\ni\,d_j \geq i} f_j x_j^{*\prime}(i) - L > 0$ *and* $v^*(\boldsymbol{d}) > 0$*. Then,* $(\boldsymbol{x}^{*\prime}, \bar{\boldsymbol{v}}) \in \tilde{\Phi}$*. Now, consider the objective function* $g(\boldsymbol{v}^*)$ *of PLP-RM-1 for* $(\boldsymbol{x}^*, \boldsymbol{v}^*)$ *and* $g(\bar{\boldsymbol{v}})$ *for* $(\boldsymbol{x}^{*\prime}, \bar{\boldsymbol{v}})$*. Here, letting* $\mathfrak{D} = \{\boldsymbol{d} | \sum_{(i,j) \ni d_j \geq i} f_j x_j^{*\prime}(i) - L > 0 \text{ and } \boldsymbol{v}^*(\boldsymbol{d}) > 0\}$*, we obtain*

$$\begin{aligned} g(\boldsymbol{v}^*) &= \sum_{\boldsymbol{d} \in \mathfrak{D}} p(\boldsymbol{d}) v^*(\boldsymbol{d}) + \sum_{\boldsymbol{d} \notin \mathfrak{D}} p(\boldsymbol{d}) v^*(\boldsymbol{d}) \\ &> \sum_{\boldsymbol{d} \in \mathfrak{D}} p(\boldsymbol{d}) \bar{v}(\boldsymbol{d}) + \sum_{\boldsymbol{d} \notin \mathfrak{D}} p(\boldsymbol{d}) v^*(\boldsymbol{d}) \quad \text{by definition of } \bar{\boldsymbol{v}}, \\ &= g(\bar{\boldsymbol{v}}). \end{aligned}$$

*This contradicts to the optimality of* $(\boldsymbol{x}^*, \boldsymbol{v}^*)$ *for PLP-RM-1. As a result, there does not exist a demand vector* $\bar{\boldsymbol{d}}$ *such that* $\sum_{(i,j)\,\ni\,d_j \geq i} f_j x_j^{*\prime}(i) - L > 0$ *and* $v^*(\bar{\boldsymbol{d}}) > 0$*. Then, there does not exist any* $\bar{\boldsymbol{d}}$ *such that* $\sum_{(i,j)\,\ni\,\bar{d}_j \geq i} f_j x_j^{*\prime}(i) - L > 0$ *and* $\tilde{v}(\bar{\boldsymbol{d}}) = 1$*, causing* $\boldsymbol{x}^{*\prime}$ *to be infeasible for PLP-RM-2*$|_{\tilde{\boldsymbol{v}}}$*. Thus,* $\boldsymbol{x}^{*\prime} \in \tilde{\Phi}|_{\tilde{\boldsymbol{v}}}$*.* ■

**Lemma 4.4.7** *Let* $(\boldsymbol{x}^*, \boldsymbol{v}^*)$ *and* $\boldsymbol{x}^{**}$ *denote optimal solutions of PLP-RM-1 and PLP-RM-2*$|_{\tilde{\boldsymbol{v}}}$*, respectively, where* $\tilde{\boldsymbol{v}}$ *is defined as in (4.33).*

***a)*** $\boldsymbol{x}^* \in \tilde{\Phi}|_{\tilde{\boldsymbol{v}}}$*.*

***b)*** $\boldsymbol{x}^{*\prime} \in \tilde{\Phi}|_{\boldsymbol{v}^*}$*.*

***c)*** $\tilde{\Phi}|_{\boldsymbol{v}^*} \subseteq \tilde{\Phi}|_{\tilde{\boldsymbol{v}}}$*.*

***d***) $h(\boldsymbol{x}^*) \leq h(\boldsymbol{x}^{**})$.

***e***) $P(R(\boldsymbol{x}^*) < L) = P(R(\boldsymbol{x}^{**}) < L) = \sum_{\boldsymbol{d}} p(\boldsymbol{d})\tilde{v}(\boldsymbol{d})$.

**Proof.**

*Proofs for (a) and (b) are very similar to the proof of Lemma 4.4.6. Here, the proof is given for (a) but skipped for (b).*

***a***) *(4.24) and (4.28) are the same, and* $\boldsymbol{x^*}$ *satisfies (4.24). That is, (4.28) is satisfied by* $\boldsymbol{x^*}$.

*The second inequality below is due to feasibility of* $\boldsymbol{x^*}$ *for PLP-RM-1, constraint (4.25) and the first inequality below is due to definition of* $\tilde{\boldsymbol{v}}$.

$$M\tilde{v}(\boldsymbol{d}) \geq Mv^*(\boldsymbol{d}) \quad \geq \quad - \sum_{(i,j)\, \ni\, d_j \geq i} f_j x_j^*(i) + L \qquad \text{for all } \boldsymbol{d}.$$

*Then,* $\boldsymbol{x^*}$ *satisfies (4.29) in PLP-RM-2*$|_{\boldsymbol{v}}$ *for* $\boldsymbol{v} = \tilde{\boldsymbol{v}}$.

*Now, consider (4.30) for* $\boldsymbol{v} = \tilde{\boldsymbol{v}}$.

• *If* $\sum_{(i,j)\, \ni\, \bar{d}_j \geq i} f_j x_j^*(i) - L \leq 0$ *for some* $\bar{\boldsymbol{d}}$, *then*

$$M(1 - v^*(\bar{\boldsymbol{d}})) \geq M(1 - \tilde{v}(\bar{\boldsymbol{d}})) \geq 0 \geq \sum_{(i,j)\, \ni\, \bar{d}_j \geq i} f_j x_j^*(i) - L.$$

*That is,* $\boldsymbol{x^*}$ *satisfies (4.30) when* $\boldsymbol{v} = \tilde{\boldsymbol{v}}$.

• *If* $\sum_{(i,j)\, \ni\, \bar{d}_j \geq i} f_j x_j^*(i) - L > 0$ *for some* $\bar{\boldsymbol{d}}$, *then* $v^*(\bar{\boldsymbol{d}}) = 0$ *by definition of* $\boldsymbol{v}$. *However, for the sake of rigour and completeness of the proof both* $v^*(\bar{\boldsymbol{d}}) = 0$ *and* $v^*(\bar{\boldsymbol{d}}) > 0$ *are considered.*

*If* $v^*(\bar{\boldsymbol{d}}) = 0$, *then* $\tilde{v}(\bar{\boldsymbol{d}}) = 0$. *In this case,*

$$M = M(1 - v^*(\bar{\boldsymbol{d}})) = M(1 - \tilde{v}(\bar{\boldsymbol{d}})) \quad \geq \quad \sum_{(i,j)\, \ni\, \bar{d}_j \geq i} f_j x_j^*(i) - L.$$

*That is,* $\boldsymbol{x^*}$ *satisfies (4.29) when* $\boldsymbol{v} = \tilde{\boldsymbol{v}}$.

*If* $v^*(\bar{\boldsymbol{d}}) > 0$, *then* $\tilde{v}(\bar{\boldsymbol{d}}) = 1$. *In this case,*

$$M(1 - \tilde{v}(\bar{\boldsymbol{d}})) \quad \ngeq \quad \sum_{(i,j)\, \ni\, \bar{d}_j \geq i} f_j x_j^*(i) - L.$$

*That is,* $\boldsymbol{x^*}$ *does not satisfy (4.29) when* $\boldsymbol{v} = \tilde{\boldsymbol{v}}$.

*However, letting*

$$\bar{v}(\boldsymbol{d}) = \begin{cases} 0 & \text{for } \boldsymbol{d} \ \ni \ \sum_{(i,j)\, \ni\, d_j \geq i} f_j x_j^*(i) - L > 0 \text{ and } \boldsymbol{v}^*(\boldsymbol{d}) > 0, \\ v^*(\boldsymbol{d}) & \text{otherwise}, \end{cases}$$

*it is observed that* $(\boldsymbol{x^*}, \bar{\boldsymbol{v}}) \in \Phi$ *but* $g(\boldsymbol{v^*}) > g(\bar{\boldsymbol{v}})$*, which contradicts to the optimality of* $(\boldsymbol{x^*}, \boldsymbol{v^*})$ *for PLP-RM-1. Then, there does not exist any* $\boldsymbol{d}$ *such that* $\sum_{(i,j)\,\ni\, d_j \geq i} f_j x^*_j(i) - L > 0$ *and* $v^*(\bar{\boldsymbol{d}}) > 0$ *(and thus* $\tilde{v}(\bar{\boldsymbol{d}}) = 1$*).*

***c****) From part (b) and by definition of* $\boldsymbol{x^*}$*, we have* $\boldsymbol{x^{*\prime}} \in \tilde{\Phi}|_{\boldsymbol{v^*}}$ *and* $\boldsymbol{x^*} \in \tilde{\Phi}|_{\boldsymbol{v^*}}$*, respectively. On the other hand, form part (a) and Lemma 4.4.6, we have* $\boldsymbol{x^*} \in \tilde{\Phi}|_{\tilde{\boldsymbol{v}}}$ *and* $\boldsymbol{x^{*\prime}} \in \tilde{\Phi}|_{\tilde{\boldsymbol{v}}}$*. Then,* $\tilde{\Phi}|_{\boldsymbol{v^*}} \subseteq \tilde{\Phi}|_{\tilde{\boldsymbol{v}}}$.

***d****)* $h(\boldsymbol{x^*}) \leq h(\boldsymbol{x^{*\prime}}) \leq h(\boldsymbol{x^{**}})$*, where the first inequality is due to Lemma 4.4.4 and the second inequality follows form the optimality of* $\boldsymbol{x}^{**}$ *for PLP-RM-2*$|_{\tilde{\boldsymbol{v}}}$*. Note that* $\boldsymbol{x^*}$*,* $\boldsymbol{x^{*\prime}}$ *and* $\boldsymbol{x^{**}}$ *are in* $\tilde{\Phi}|_{\tilde{\boldsymbol{v}}}$.

***e****) For* $x^*$ *and* $x^{*\prime}$*,* $P(R(\boldsymbol{x}^*) < L) = \sum_{\boldsymbol{d}} p(\boldsymbol{d})\tilde{v}(\boldsymbol{d})$ *from the first step of PLP-RM Procedure. Note that* $P(R(\boldsymbol{x}^*) < L)$ *is equated to the expression in terms of* $\tilde{\boldsymbol{v}}$ *but not* $\boldsymbol{v^*}$*. This is because the risk measure is defined in terms of binary* $v(\boldsymbol{d})$*s. (Recall Lemma 4.1.3) For any* $\boldsymbol{x}$ *in PLP-RM-2,* $\sum_{(i,j)\,\ni\, d_j \geq i} f_j x_j(i) - L \leq 0$ *or* $\sum_{(i,j)\,\ni\, d_j \geq i} f_j x_j(i) - L \geq 0$ *for given* $\tilde{\boldsymbol{v}}$ *and all* $\boldsymbol{d}$*. Then, for optimal solution* $x^{**}$*,* $P(R(\boldsymbol{x}^{**}) < L) = \sum_{\boldsymbol{d}} p(\boldsymbol{d})\tilde{v}(\boldsymbol{d})$*.* ■

Then, Proposition 4.3.9 can be revised as in Proposition 4.4.8 for the proposed approach with approximations (2) and (3).

**Proposition 4.4.8** *Let* $\boldsymbol{v}^*$ *be an optimal solution of PLP-RM-1 and* $\tilde{\boldsymbol{v}}$ *be defined as in (4.33). Let* $\boldsymbol{x}^{**}$ *be an optimal allocation for PLP-RM-2*$|_{\tilde{\boldsymbol{v}}}$*. Then,* $\boldsymbol{x}^{**}$ *satisfies Condition 4.4.1.*

**Proof.** *From Lemma 4.4.7 (a), (c) and Lemma 4.4.6,* $\tilde{\Phi}|_{\tilde{\boldsymbol{v}}}$ *includes not only* $\boldsymbol{x^*}$ *but also* $\boldsymbol{x^{*\prime}}$ *that satisfies Condition 4.4.1. If* $\boldsymbol{x}^{**}$ *is not in* $(\tilde{\Phi}|_{\tilde{\boldsymbol{v}}}$-$\tilde{\Phi}|_{\boldsymbol{v^*}})$*, then one of* $\boldsymbol{x^{*\prime}}$ *would be* $\boldsymbol{x}^{**}$*; otherwise,* $\boldsymbol{x}^{**}$ *would also satisfy Condition 4.4.1. Because* $P(D_j \geq i)$ *in* $h(\boldsymbol{x})$ *decreases in i.* ■

Next, the approximation (1) for sample demand realizations is used. The *PLP-RM-1* and *PLP-RM-2*$|_{\tilde{\boldsymbol{v}}}$ models are solved for a number of sample demand realizations. Let $\Psi$ be the set of sample demand realizations considered in the *PLP-RM-1* and *PLP-RM-2*$|_{\tilde{\boldsymbol{v}}}$ models and $|\Psi|$ be the total number of demand realizations. The objective function of the approximate *PLP-RM-1* model would, then, be the normalized probability that the revenue is less than the threshold level, $L$. That is, it would be $\frac{\sum_{\boldsymbol{d}\in\Psi} p(\boldsymbol{d})v(\boldsymbol{d})}{\sum_{\boldsymbol{d}\in\Psi} p(\boldsymbol{d})}$ corresponding to $P(R(\boldsymbol{x}) < L|\boldsymbol{D} \in \Psi)$. Instead, the average of $v(\boldsymbol{d})$ values, $\sum_{\boldsymbol{d}\in\Psi} \frac{v(\boldsymbol{d})}{|\Psi|}$, is used in the objective function of the approximate *PLP-RM-1* model. If the sample size, $|\Psi|$, is sufficiently large, then the average of $v(\boldsymbol{d})$ would work well to approximate $P(R(\boldsymbol{x}) < L)$. This is because $|\Psi_{\boldsymbol{d}}|/|\Psi|$ would be expected to converge to $p(\boldsymbol{d})$ when $|\Psi_{\boldsymbol{d}}|$ is the number of demand realization $\boldsymbol{d}$ in the sample $|\Psi|$. The other alternative for the objective function, i.e, $\frac{\sum_{\boldsymbol{d}\in\Psi} p(\boldsymbol{d})v(\boldsymbol{d})}{\sum_{\boldsymbol{d}\in\Psi} p(\boldsymbol{d})}$, does not work well numerically because dividing very small $p(\boldsymbol{d})$ values by very small $\sum_{\boldsymbol{d}\in\Psi} p(\boldsymbol{d})$ values causes computational problems in solving the optimization models.

Then, the resulting approximate *PLP-RM-1* and *PLP-RM-2* models are given as follows:

$$
\begin{aligned}
\textit{Approximate PLP-RM-1} \quad : \quad & \textit{Minimize} \sum_{\boldsymbol{d}\in\Psi} \frac{v(\boldsymbol{d})}{|\Psi|} \\
& \textit{subject to} \\
& \sum_{j\in S_l} \sum_{i=1}^{B_j} x_j(i) \leq C_l \quad \text{for } l = 1, ...., m, \\
& Mv(\boldsymbol{d}) \geq - \sum_{(i,j) \ni d_j \geq i} f_j x_j(i) + L \quad \text{for } \boldsymbol{d} \in \Psi, \\
& M(1 - v(\boldsymbol{d})) \geq \sum_{(i,j) \ni d_j \geq i} f_j x_j(i) - L \quad \text{for } \boldsymbol{d} \in \Psi, \\
& 0 \leq x_j(i) \leq 1 \quad \text{for } j = 1, ...., n \text{ and } i = 1, ..., B_j, \\
& 0 \leq v(\boldsymbol{d}) \leq 1 \text{ for } \boldsymbol{d} \in \Psi.
\end{aligned}
\tag{4.34}
$$

$$
\begin{aligned}
\textit{Approximate PLP-RM-2}|_{\boldsymbol{v}} \quad : \quad & \textit{Maximize} \sum_{j=1}^{n} \sum_{i=1}^{B_j} f_j P(D_j \geq i) x_j(i) \\
& \textit{subject to} \\
& \sum_{j\in S_l} \sum_{i=1}^{B_j} x_j(i) \leq C_l \quad \text{for } l = 1, ...., m, \\
& Mv(\boldsymbol{d}) \geq - \sum_{(i,j) \ni d_j \geq i} f_j x_j(i) + L \quad \text{for } \boldsymbol{d} \in \Psi, \\
& M(1 - v(\boldsymbol{d})) \geq \sum_{(i,j) \ni d_j \geq i} f_j x_j(i) - L \quad \text{for } \boldsymbol{d} \in \Psi, \\
& 0 \leq x_j(i) \leq 1 \quad \text{for } j = 1, ...., n \text{ and } i = 1, ..., B_j.
\end{aligned}
\tag{4.35}
$$

Note that the analytical results given in this section for the *PLP-RM* Procedure are also valid when a sample of demand realizations is used.

The computational time for the use of the approximate *PLP-RM-1* and *PLP-RM-2*$|_{\tilde{\boldsymbol{v}}}$ models with the *PLP-RM* procedure decreases even more by reformulating these models with aggregate demands as in the $SLP$ model due to de Boer (1999). Then, the resulting models that are called *SLP-RM* are as given below. The abbreviation *SLP-RM* stands for *Stochastic Linear*

*Programming with Risk Measure.*

$$SLP\text{-}RM\text{-}1 \quad : \quad Minimize \sum_{\boldsymbol{d}\in\Psi} \frac{v(\boldsymbol{d})}{|\Psi|}$$

$$subject\ to$$

$$\sum_{j\in S_l}\sum_{k=1}^{\kappa_j} x_j(k) \le C_l \quad \text{for } l = 1, ...., m,$$

$$x_j = \sum_{k=1}^{\kappa_j} x_j(k) \quad \text{for } j = 1, ...., n,$$

$$x_j(1) \le d_j(1),$$

$$x_j(k) \le d_j(k) - d_j(k-1) \quad \text{for } k = 2, ...., \kappa_j,$$

$$Mv(\boldsymbol{d}) \ge -\sum_{(k,j)\,\ni\, d_j \ge d_j(k)} f_j x_j(k) + L \quad \text{for } \boldsymbol{d} \in \Psi,$$

$$M(1 - v(\boldsymbol{d})) \ge \sum_{(k,j)\,\ni\, d_j \ge d_j(k)} f_j x_j(k) - L \quad \text{for } \boldsymbol{d} \in \Psi,$$

$$x_j(k) \ge 0 \quad \text{for } j = 1, ...., n \text{ and } k = 1, ..., \kappa_j,$$

$$x_j \ge 0 \quad \text{for } j = 1, ...., n,$$

$$0 \le v(\boldsymbol{d}) \le 1 \quad \text{for } \boldsymbol{d} \in \Psi.$$

Recall that, in *SLP*, $D_j$ is assumed to take a value in $\{d_j(i), ..., d_j(\kappa_j)\}$ where $d_j(1) < d_j(2) < ... < d_j(\kappa_j)$ for $j = 1, ..., n$. $x_j(k)$ is defined as the number of seats allocated for demand group $k$ between $d_j(k-1)$ and $d_j(k)$ and $x_j$ is the total seats allocated to *ODF* $j$.

$$SLP\text{-}RM\text{-}2|_v \quad : \quad Maximize\ \mathfrak{h}(\boldsymbol{x}) = \sum_{j=1}^{n} f_j x_j - \sum_{j=1}^{n} f_j \sum_{k=1}^{\kappa_j} P(D_j < d_j(k)) x_j(k)$$

$$subject\ to$$

$$\sum_{j\in S_l}\sum_{k=1}^{\kappa_j} x_j(k) \le C_l \quad \text{for } l = 1, ...., m,$$

$$x_j = \sum_{k=1}^{\kappa_j} x_j(k) \quad \text{for } j = 1, ...., n,$$

$$x_j(1) \le d_j(1),$$

$$x_j(k) \le d_j(k) - d_j(k-1) \quad \text{for } k = 2, ...., \kappa_j,$$

$$Mv(\boldsymbol{d}) \ge -\sum_{(k,j)\,\ni\, d_j \ge d_j(k)} f_j x_j(k) + L \quad \text{for } \boldsymbol{d} \in \Psi,$$

$$M(1 - v(\boldsymbol{d})) \ge \sum_{(k,j)\,\ni\, d_j \ge d_j(k)} f_j x_j(k) - L \quad \text{for } \boldsymbol{d} \in \Psi,$$

$$x_j(k) \ge 0 \quad \text{for } j = 1, ...., n \text{ and } k = 1, ..., \kappa_j,$$

$$x_j \ge 0 \quad \text{for } j = 1, ...., n.$$

*SLP-RM-2*$|_v$ model is used to maximize the expected revenue by setting $v(\boldsymbol{d})$ values as parameters which are found by solving *SLP-RM-1* model. As in the *SLP* model due to de Boer (1999), the first term in the objective function is the total revenue gained when all of the allocated seats are sold and the second term is a correction for uncertainty of demand. Although the computational time decreases significantly when *SLP-RM* models are used, the solution time is still a disadvantage of *SLP-RM* model as compared to *EMR* and *DLP*. The numerical studies on *SLP-RM* models are given in Chapter 6.

The derivations of the expected revenue used in *SLP-RM* models are given next. Define random variable $K_j$: $K_j = k$ if $D_j \in [d_j(k), d_j(k+1))$. Let $\boldsymbol{K} = (K_1, ..., K_n)$ and

$$F_j(k) = \begin{cases} f_j & \text{if } D_j \geq d_j(k), \\ 0 & \text{otherwise.} \end{cases} \tag{4.36}$$

Note that

$$F_j(k) = \begin{cases} f_j & \text{if } K_j \geq k, \\ 0 & \text{otherwise.} \end{cases} \tag{4.37}$$

Let $\boldsymbol{x}_j = (x_j(1), ..., x_j(\kappa_j))$ for $j = 1, ..., n$ and $\boldsymbol{x} = (\boldsymbol{x}_1, ..., \boldsymbol{x}_n)$. Then, the revenue for a given allocation $\boldsymbol{x}$ is

$$\begin{aligned} R_{SLP}(\boldsymbol{x}) &= \sum_{j=1}^{n} \sum_{k=1}^{\kappa_j} F_j(k) x_j(k) && (4.38) \\ &= \sum_{j=1}^{n} (\sum_{k=1}^{K_j} f_j x_j(k) + \sum_{k=K_j+1}^{\kappa_j} 0 \cdot x_j(0)) && (4.39) \\ &= \sum_{j=1}^{n} f_j \sum_{k=1}^{K_j} x_j(k) && (4.40) \\ &= \sum_{(k,j) \ni K_j \geq k} f_j x_j(k) && (4.41) \\ &= \sum_{(k,j) \ni D_j \geq d_j(k)} f_j x_j(k). && (4.42) \\ & && (4.43) \end{aligned}$$

The expected revenue for a given allocation $\boldsymbol{x}$ is also rewritten as follows:

$$E(R_{SLP}(\boldsymbol{x})) = \sum_{j=1}^{n}\sum_{k=1}^{\kappa_j} E(F_j(k))x_j(k) \tag{4.44}$$

$$= \sum_{j=1}^{n}\sum_{k=1}^{\kappa_j}(f_j \cdot P(K_j \geq k) + 0 \cdot P(K_j < k))x_j(k) \tag{4.45}$$

$$= \sum_{j=1}^{n}\sum_{k=1}^{\kappa_j}(f_j \cdot P(D_j \geq d_j(k)) + 0 \cdot P(D_j < d_j(k)))x_j(k) \tag{4.46}$$

$$= \sum_{j=1}^{n}\sum_{k=1}^{\kappa_j} f_j \cdot P(D_j \geq d_j(k))x_j(k). \tag{4.47}$$

$$\tag{4.48}$$

For the formulation, the demand $D_j$ is assumed to take a value in $\{d_j(1), ..., d_j(\kappa_j)\}$ for $j = 1, ..., n$. However, for evaluating the expected revenue as a function of a given $\boldsymbol{x}$, the original demand distribution is used. Note that the evaluation would be the same even if the original demand distribution is aggregated according to the formulation.

Next, it is shown that the analytical results given in this section for *PLP-RM-1* and *PLP-RM-2* models are also valid for *SLP-RM-1* and *SLP-RM-2* models. First, Condition 4.4.1 is revised as given in Condition 4.4.9. Next, Definition 4.4.2 and Lemmas 4.4.4 and 4.4.5 are revised.

**Condition 4.4.9** *For $j = 1, ..., n$, if $x_j(k+1) > 0$, then $x_j(k) = d_j(k) - d_j(k-1)$; if $x_j(k) < d_j(k) - d_j(k-1)$, then $x_j(k+1) = 0$.*

Then, the definition of $\boldsymbol{x}'$ is

**Definition 4.4.10** *For an allocation $\mathbf{x}$, $\mathbf{x}'$ satisfies Condition 4.4.9 and the equation below.*

$$\sum_{k=1}^{\kappa_j} x'_j(k) = \sum_{k=1}^{\kappa_j} x_j(k) \quad \textit{for } j = 1, ..., n. \tag{4.49}$$

**Example 4.4.11** *Suppose that $\kappa_j = 10$ for some $j$ and an example allocation is $\mathbf{x}$ =[2.5 0.5 0 0 3.8 0 4 0 0 3]. Then, the corresponding allocation $\mathbf{x}'$ given by Definition 4.3.3 is $\mathbf{x}'$ =[4 4 4 1.8 0 0 0 0 0 0].*

The allocation $x_j(k)$ obtained by solving the *SLP-RM-2*$|_\nu$ model is

$$x_j(k) = \begin{cases} 0 & \text{if } x_j \le d_j(k-1), \\ x_j - d_j(k-1) & \text{if } d_j(k-1) < x_j < d_j(k), \\ d_j(k) - d_j(k-1) & \text{if } x_j \ge d_j(k). \end{cases} \qquad \text{for } k = 1, ..., \kappa_j \textit{ and } j = 1, ..., n.$$

That is, the solution of *SLP-RM-2* would satisfy Condition 4.4.9 because $P(D_j \ge d_j(k))$ in $\mathfrak{h}(\boldsymbol{x}) = \sum_{j=1}^{n} \sum_{k=1}^{\kappa_j} f_j P(D_j \ge d_j(k)) x_j(k)$ decreases in $k$.

**Lemma 4.4.12** *Let* $\boldsymbol{x}$ *be an allocation that does not satisfy Condition 4.4.9 and* $\boldsymbol{x}'$ *be given as in Definition 4.4.10. Then,* $\mathfrak{h}(\boldsymbol{x}) \le \mathfrak{h}(\boldsymbol{x}')$.

**Proof.** *From (4.49),*

$$f_j \sum_{k=1}^{\kappa_j} x'_j(k) \;=\; f_j \sum_{k=1}^{\kappa_j} x_j(k) \quad \textit{for } j = 1, ..., n.$$

*Since* $\boldsymbol{x}'$ *satisfies Condition 4.4.9 and (4.49), and* $P(D_j \ge d_j(k))$ *is nonincreasing in* $k$,

$$\sum_{i=1}^{\kappa_j} f_j P(D_j \ge d_j(k)) x'_j(k) \;\ge\; \sum_{k=1}^{\kappa_j} f_j P(D_j \ge d_j(k)) x_j(k) \quad \textit{for } j = 1, ..., n.$$

*As a result,*

$$\sum_{j=1}^{n} \sum_{k=1}^{\kappa_j} f_j P(D_j \ge d_j(k)) x'_j(k) \;\ge\; \sum_{j=1}^{n} \sum_{k=1}^{\kappa_j} f_j P(D_j \ge d_j(k)) x_j(k).$$

■

**Lemma 4.4.13** *Let* $\boldsymbol{x}$ *be an allocation that does not satisfy Condition 4.4.9 and* $\boldsymbol{x}'$ *be given as in Definition 4.4.10. Then,*

$$\sum_{(k,j)\,\ni\, d_j \ge d_j(k)} f_j x_j(k) \;\le\; \sum_{(k,j)\,\ni\, d_j \ge d_j(k)} f_j x'_j(k) \quad \textit{for all } \boldsymbol{d}.$$

**Proof.** *Let* $\boldsymbol{D} = \boldsymbol{d}$ *and* $\boldsymbol{K} = (k_1, ..., k_n)$. *For* $\boldsymbol{x}$, *total revenue is equal to*

$$\sum_{(k,j)\,\ni\, d_j \ge d_j(k)} f_j x_j(k) \;=\; \sum_{j=1}^{n} f_j (\sum_{k=1}^{k_j} x_j(k)) \tag{4.50}$$

*from (4.36) and (). From Definition 4.4.10,*

$$\sum_{k=1}^{k_j} x_j(k) \;\le\; \sum_{k=1}^{k_j} x'_j(k) \quad \textit{for } j = 1, .., n.$$

*Then, using the relation in (4.50),*

$$\begin{aligned}\sum_{(k,j)\,\ni\, d_j \geq d_j(k)} f_j x_j(k) &\leq \sum_{j=1}^{n} f_j(\sum_{k=1}^{k_j} x'_j(k)) \\ &= \sum_{(k,j)\,\ni\, d_j \geq d_j(k)} f_j x'_j(k).\end{aligned}$$

■

Then, Lemma 4.4.6 and 4.4.8 can be revised for the *SLP-RM* models. As compared to the *PLP-RM* models, the only additional constraints in the *SLP-RM* models are

$$\begin{aligned}x_j(1) &\leq d_j(1), \\ x_j(k) &\leq d_j(k) - d_j(k-1) \quad \text{for } j = 2, ..., \kappa_j.\end{aligned}$$

Note that the constraint for $x_j$ can be removed in *SLP-RM-1* (*SLP-RM-2* by replacing the objective function with $\sum_{j=1}^{n} f_j P(D_j \geq d_j(k)) x_j(k)$). All of the other constraints and the objective function in the *PLP-RM* and *SLP-RM* models are directly comparable. The revision of Lemma 4.4.6 for the additional constraints in the *SLP-RM* models would be straightfroward because the additional constraints aforementioned above would already be satisfied by an allocation satisfying Condition 4.4.9.

## 4.5 Risk-Constrained Mathematical Programming Model

The approach considered in the previous section is to minimize risk first and then to maximize expected revenue for the risk level determined in the first step. In this section, we propose a model to maximize the expected revenue under a constraint on the same risk measure as the one considered in the previous section, i.e., the probability of total revenue being less than a threshold level. The model is formulated by using the same assumptions and approximations used in the previous section. The proposed *PMP-RC* model that stands for *Probabilistic*

*Mathematical Programming with Risk Constraint* is as follows:

$$PMP\text{-}RC \quad : \quad Maximize \sum_{j=1}^{n} f_j x_j - \sum_{j=1}^{n} f_j \sum_{k=1}^{\kappa_j} P(D_j < d_j(k)) x_j(k) \tag{4.51}$$

$$subject\ to$$

$$\sum_{j \in S_l} \sum_{k=1}^{\kappa_j} x_j(k) \leq C_l \text{ for } l = 1, ...., m,$$

$$x_j = \sum_{k=1}^{\kappa_j} x_j(k) \quad \text{for } j = 1, ..., n,$$

$$x_j(1) \leq d_j(1),$$

$$x_j(k) \leq d_j(k) - d_j(k-1) \quad \text{for } j = 2, ..., \kappa_j,$$

$$\sum_{\boldsymbol{d} \in \Psi} \frac{v(\boldsymbol{d})}{|\Psi|} < \rho,$$

$$Mv(\boldsymbol{d}) \geq - \sum_{(k,j) \ni d_j \geq d_j(k)} f_j x_j(k) + L \quad \text{for } \boldsymbol{d} \in \Psi,$$

$$M(1 - v(\boldsymbol{d})) \geq \sum_{(k,j) \ni d_j \geq d_j(k)} f_j x_j(k) - L \quad \text{for } \boldsymbol{d} \in \Psi,$$

$$x_j(k) \geq 0 \quad \text{for } j = 1, ...., n \text{ and } k = 1, ..., \kappa_j,$$

$$x_j \geq 0 \quad \text{for } j = 1, ...., n,$$

$$v(\boldsymbol{d}) \in \{0, 1\} \quad \text{for } \boldsymbol{d} \in \Psi, \tag{4.52}$$

where $\rho$ is a predetermined constant number between 0 and 1. Both the objective function and the constraints are similar to those in the *SLP-RM-2* model except the constraint with $\rho$ and the integrality constraint for the decision variables $v(\boldsymbol{d})$. Recall that we propose to solve *SLP-RM-2* for given specified $v(\boldsymbol{d})$ values. $\rho$ in *PMP-RC* is a threshold level to be determined by the decision maker.

There are some disadvantages of *PMP-RC* model as compared to the *SLP-RM* models. These disadvantages are itemized as follows:

• *PMP-RC* is an integer programming formulation due to (4.52) and solution time is more than that of the *SLP-RM* models for the same networks according to the results of the numerical experiments. It is numerically observed that it is not computationally efficient to solve the *PMP-RC* model for demand realizations more than 100.

• Dual prices can not be obtained by the *PMP-RC* model to be used as the bid prices because the model is an integer programming formulation.

• In *SLP-RM* models only $L$ values are set, but in *PMP-RC* model $\rho$ value must also be set. Setting $\rho$ values is not a straightforward task for different $L$ values.

## 4.6 Randomized Risk Sensitive Method

As stated before, main disadvantage of the models proposed in Sections 4.3-4.5 to solve network problems is the computational complexity. The computational time to solve these models increases as the number of demand realizations increases. In this section, a model that is based on Randomized Linear Programming Model (*RLP*) is proposed. *RLP* is proposed by Talluri and van Ryzin (1999) and a randomized version of Deterministic Linear Programming (*DLP*) model. The *RLP* model is given below for a given demand realization $\boldsymbol{d} = (d_1, ..., d_n)$.

$$RLP \quad : \quad Maximize \sum_{j=1}^{n} f_j x_j \tag{4.53}$$

$$subject\ to$$

$$\sum_{j \in S_l} x_j \leq C_l \quad \text{for } l = 1, ...., m, \tag{4.54}$$

$$0 \leq x_j \leq d_j \quad \text{for } j = 1, ..., n. \tag{4.55}$$

Recall the related discussion in Section 3.2.

The following procedure is proposed for both risk sensitive and risk taking decision makers. *RRS* stands for *Randomized Risk Sensitive*.

***RRS* Procedure:**

**Step 0.** Set the threshold level for revenue, $L$, and the number of demand realizations, $|\Psi|$.

**Step 1.** Solve *RLP* model $|\Psi|$ times. Let $\mu_r(l)$ be the bid price for leg $l$ and run $r = 1, ..., |\Psi|$. $\mu_r(l)$ is equal to the dual price of the capacity constraint (4.54). Let $\boldsymbol{x}^{(r)}$ be the optimal allocation for run $r$.

**Step 2.** Classify runs into two sets: $\mathfrak{G}^-$ and $\mathfrak{G}^+$.

$\mathfrak{G}^- = \{r | \sum_{j=1}^{n} f_j x_j^{(r)} < L\}$,

$\mathfrak{G}^+ = \{r | \sum_{j=1}^{n} f_j x_j^{(r)} \geq L\}$.

**Step 3.** Then, take the average of the bid prices for both of the sets.

$\mu^-(l) = \frac{1}{|\mathfrak{G}^-|} \sum_{r \in \mathfrak{G}^-} \mu_r(l)$ for $l = 1, ..., m$,

$\mu^+(l) = \frac{1}{|\mathfrak{G}^+|} \sum_{r \in \mathfrak{G}^+} \mu_r(l)$ for $l = 1, ..., m$.

**Step 4.** Use bid prices $\mu^-(l)$ and $\mu^+(l)$ in bid price control for risk-sensitive and risk-taking decision makers, respectively.

In this procedure, two bid prices are found: $\mu^-(l)$ for risk-sensitive decision makers and $\mu^+(l)$ for risk-taking decision makers. The risk sensitivity of a decision maker is directly related with the optimism of the decision maker. For the pessimistic decision makers, who think that the expected revenue would be less than the threshold level, using $\mu^-(l)$ is proposed. On the other hand, for the risk-taking decision makers, who think that the expected revenue would be more than the threshold level, using $\mu^+(l)$ is proposed. The level of risk-sensitivity can be changed by changing the threshold level. In our numerical studies, it is seen that $\mu^-(l) \leq \mu^+(l)$. Hence, risk sensitive decision makers tend to use lower bid prices as compared to risk-taking decision makers.

# CHAPTER 5

# SIMULATION MODELS

In Chapter 4, the proposed models are given. This chapter is devoted to the simulation models and parameter estimation for a small-scale sample network. For this purpose, the airline data given by de Boer (1999) for a three-leg flight network is used. In this sample network, nodes are connected with three legs and the total number of itineraries among these nodes is six. It is assumed that flights are identical, each with a capacity of 200. The number of fare classes for all of the *OD* pairs is three. Therefore, total number of *ODF*s in this sample network is 18. Booking requests are accepted during 150 days before departure. The figure of the three-leg network can be seen in Figure 5.1.

In this thesis, MATLAB programming language is used for optimization and simulation studies. Moreover, in optimization part, YALMIP (Löfberg, 2004) and GLPK are used to solve linear and mixed-integer mathematical models. YALMIP is a modeling language for defining and solving optimization problems, and implemented as a free toolbox for MATLAB. GLPK (GNU Linear Programming Kit) is freely distributed solver for linear and mixed integer programming. The MATLAB programming codes are given in Appendix C. The advantages and disadvantages of using Matlab programming language are listed below.

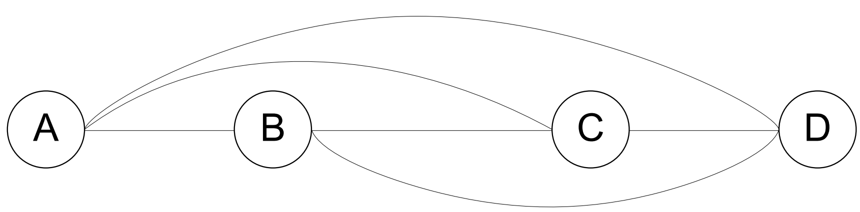


Figure 5.1: Sample network with three legs

**Disadvantages**

1. CPU time for mathematical models is higher than most of the optimization programs.
2. Inefficient for large scale problems.
3. The memory requirement is high.

**Advantages**

1. Simple and flexible codes.
2. Both optimization and simulation codes are in one programming language.
3. Easy output analysis (statistical toolbox, graphical interface, etc.).

This chapter is organized as follows: In Section 5.1, the simulation model is described and, then, Bayesian update strategy is given in Section 5.2. Section 5.3 is devoted to the numerical analysis of the mathematical models to determine the number of demand realizations, threshold level and aggregation size of the demand groups to be used in the models and procedures proposed in Chapter 4.

## 5.1 Simulation Models

In general, performances of the policies obtained from optimization models in RM are evaluated by using simulation models. In Chapter 6, performances of the approaches proposed in Chapter 4 are evaluated by simulation models and this section is devoted to the construction of the simulation models. The demand and network data for the simulation models are taken from the study of de Boer (1999) and given in Appendix B.

It is common in the literature that arrival process of the booking requests for RM problems is modeled using Poisson Processes. The arrival rate of this arrival process is a random variable that is fitted with the Gamma distribution. Moreover, in RM problems, arrival rate is not constant throughout the booking horizon and generally low-fare customers tend to arrive earlier than the high-fare customers. In order to introduce this situation into simulation models, a Non-Homogeneous Poisson Process (NHPP) is used. The random arrival rate for NHPP is defined as the product of Beta and Gamma distributed random variables. For our example problem, Beta density functions for different fare classes can be seen in Figure 5.2. Then, the resulting arrival rate at time $t$ for an $ODF$ is given by de Boer (1999) as follows;

$$\Lambda_{ODF}(t) = \varphi_{ODF} \cdot \beta_{ODF}(t) \cdot A_{OD}, \tag{5.1}$$

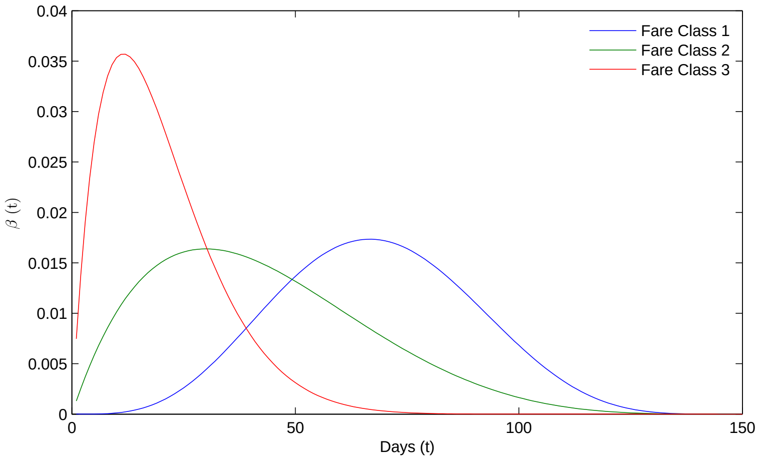


Figure 5.2: Beta density functions for three fare classes.

where

| | |
|---|---|
| $A_{OD}$ | : Stochastic expected number of booking requests for $OD$ pair in the NHPP, |
| $\beta_{ODF}(t)$ | : Standardized Beta distribution, |
| $\varphi_{ODF}$ | : Fraction of $OD$ bookings for fare class $F$. |

$A_{OD}$ is the expected total number of demand for itinerary $OD$ and has a Gamma distribution. $\varphi_{ODF}$ is the fixed fraction of $OD$ bookings for fare class $F$. $\beta_{ODF}(t)$ represents the Beta distribution used for spreading the total demand over the booking horizon. Then, the arrivals of the demands are modeled as a NHPP with arrival rate $\Lambda_{ODF}(t)$. Moreover, it is assumed that the expected number of arrivals $A_{OD}$ for each fare class is independent. Then, the factor $\varphi_{ODF}(t)$ can be left out and the $A_{OD}$ can be replaced with $A_{ODF}$; more clearly, $A_{ODF} = \varphi_{ODF}(t) \cdot A_{OD}$. The independence assumption is also suitable for real life problems because high fare customers rarely want to buy low fare tickets and low fare customers rarely want to buy high fare tickets. Therefore, we can easily claim that demands are independent for each $ODF$ in the RM problems, which is also tested with real booking data by de Boer et al. (1999). According to the non-parametric Spearman test, the null hypothesis of no correlation between demand for different fare classes is not rejected. Moreover, by using the same notation used in Chapter 4, the following equation is obtained.

$$\Lambda_j(t) = \beta_j(t) \cdot A_j \; for \; j = 1, ..., n. \tag{5.2}$$

The formulation given in (5.2) is not a pure Poisson process (Non-Homogeneous Poisson Process) and the demand distribution over the whole reservation period is given in Remark 5.1.1.

**Remark 5.1.1** *(de Boer 1999, Popovic 1987) The unconditional distribution of $D_j$, demand for ODF j, over the whole booking horizon is Negative Binomial with parameters $p_j$ and $\frac{\gamma_j}{\gamma_j+1}$. This is shown as $D_j \sim NB(p_j, \frac{\gamma_j}{\gamma_j+1})$.*

**Proof.** *$A_j$ in equation (5.2) is the stochastic expected number of ODF booking requests. This variable is distributed by Gamma($p_j$,$\gamma_j$), where $p_j$ is the number of phases and $\gamma_j$ is the rate. Suppose that $A_j = \lambda_j$ for $j = 1, ..., n$. Then, the arrival is Poisson distributed and the probability of observing a demand for k units throughout the booking horizon is given as follows for ODF j:*

$$P(D_j = k|A_j = \lambda_j) = \frac{\lambda_j^k}{k!}e^{-\lambda_j} \text{ for } k = 0, 1, .... \tag{5.3}$$

*We know that $\lambda_j$ is a random variable distributed according to Gamma distribution with parameters $p_j$ and $\gamma_j$ and the probability density function of $A_j = \lambda_j$ is*

$$f_{A_j}(\lambda_j) = \frac{\gamma^{p_j}\lambda^{p_j-1}e^{-\gamma\lambda_j}}{\Gamma(p_j)} \quad \text{for } \lambda_j > 0, \tag{5.4}$$

*where $\Gamma$ is the gamma function given as follows:*

$$\Gamma(x) = \int_0^\infty y^{x-1}e^{-y}dy. \tag{5.5}$$

*Using Equations (5.3) and (5.4), the distribution function of the total booking requests, $D_j$, is calculated as follows:*

$$\begin{aligned}
P(D_j = k) &= \int_0^\infty P(D_j = k|A_j = \lambda_j)f_{A_j}(\lambda_j)d\lambda_j \\
&= \int_0^\infty (\frac{\lambda^k}{k!}e^{-\lambda_j})(\frac{\lambda_j^{p_j-1}e^{-\gamma_j\lambda_j}\gamma_j^{p_j}}{\Gamma(p_j)})d\lambda_j \\
&= (\frac{\gamma_j}{\gamma_j+1})^{p_j}(\frac{1}{\gamma_j+1})^k\frac{\Gamma(k+p_j)}{\Gamma(p_j)k!} \\
&= \binom{p_j+k-1}{k}(\frac{\gamma_j}{\gamma_j+1})^{p_j}(\frac{1}{\gamma_j+1})^k \quad for\ k = 0, 1, ....
\end{aligned}$$

*Then, the unconditional distribution of demand for ODF j is Negative Binomial with parameters $p_j$ and $\frac{\gamma_j}{\gamma_j+1}$. For Negative Binomial Distribution, the expectation and the variance* *are*

$$E[D_j] = \frac{p_j}{\gamma_j},$$
$$Var[D_j] = \frac{p_j(\gamma_j+1)}{\gamma_j^2}.$$

■

Also, Beta distribution is used in (5.2) in order to model the arrival process. $\beta_j(t)$ is the standardized Beta distribution with parameters $\alpha_j > 0$ and $\beta_j > 0$ in booking horizon $[0, T]$. $T$ is the length of the booking period. Time $t$ for $0 \leq t \leq T$ denotes the remaining time until the departure.

$$\beta_j(t) = \frac{1}{T}(\frac{1}{T})^{\alpha_j-1}(1-\frac{t}{T})^{\beta_j-1}\frac{\Gamma(\alpha_j+\beta_j)}{\Gamma(\alpha_j)\Gamma(\beta_j)}, \tag{5.6}$$

Then, the following is the formulation for the number of arrivals for $ODF\ j$ from time $t$ to the departure when $A_j = \lambda_j$. Recall 5.2.

$$\int_0^t \lambda_j\beta_j(x)dx = \int_0^t \frac{\lambda_j}{T}(\frac{1}{T})^{\alpha_j-1}(1-\frac{x}{T})^{\beta_j-1}\frac{\Gamma(\alpha_j+\beta_j)}{\Gamma(\alpha_j)\Gamma(\beta_j)}dx. \tag{5.7}$$

Note that, here, the expected number of booking requests, $\lambda_j$, is multiplied by the cumulative density function of Beta distribution.

All of the required distribution functions for simulation models have been derived up to now. However, the simulation of NHPP is not straightforward unlike the case of Poisson Process with constant rate. The simulation techniques for generating arrivals from NHPP are described by Law and Kelton (2000). The authors give the following two algorithms. Algorithm 1 uses stationary Poisson Process, e.g., with a given rate $\Lambda_j$. In other words, the inter-arrival time between two arrivals is distributed exponentially with the cumulative distribution function $1 - e^{-\Lambda_j t}$. The algorithm is summarized as follows:

***Algorithm 1***

*1. Generate random variate t from Uniform(0,1).*

*2. Return the $i^{th}$ customer's arrival at $t_i = t_{i-1} - \frac{1}{\Lambda_j}\log t$.*

By substituting $\Lambda_j(t)$ for $\Lambda_j$, arrival times for NHPP can be obtained. However, Law and Kelton (2000) argue that this algorithm gives incorrect results especially for small rates.

In Algorithm 2, the expectation function $\mu_j$ is used, where $\mu_j(t)$ is the expected number of arrivals for $ODF\ j$ in the interval $[0, t]$;

$$\mu_j(t) = \int_0^t \Lambda_j(x)dx \quad for\ j = 1, ..., n. \tag{5.8}$$

Subscript $j$ of $\mu$ and $\Lambda$ is dropped in the algorithms below to keep the notation simpler. The algorithms should be used for each $ODF\ j$.

***Algorithm 2***

*1. Generate random variate $\iota$ from Uniform(0,1).*

*2. Generate arrival times with rate $y_i = y_{i-1} - log(\iota)$.*

*3. Return the $i^{th}$ customer's arrival at $t_i = \mu^{-1}(y_i)$.*

Law and Kelton (2000) show that the simulated arrival rates when applying Algorithm 2 agree with real (given) arrival rate. However, taking inverse of the expectation function given in (5.8) would involve some difficulties. In order to solve this problem, Algorithm 3 is proposed and used in this study as an approximate numerical method.

***Algorithm 3***

**Step 0.** Divide the booking horizon into small disjoint time intervals $(t_{\eta-1}, t_\eta)$ in such a way that union of the intervals is $(0, T)$. Generate a table of $\mu(t_\eta) = \int_0^{t_\eta} \Lambda(x)dx$ for all $\eta$.

**Step 1.** Generate random variate $\iota$ from Uniform(0,1).

**Step 2.** Generate arrival times with rate $y_i = y_{i-1} - log(\iota)$.

**Step 3.** Find the interval $[t_{\eta-1}, t_\eta]$ such that $\mu(t_{\eta-1}) \le y_i \le \mu(t_\eta)$.

**Step 4.** Set $\alpha = \frac{y_i - \mu(t_{\eta-1})}{\mu(t_\eta) - \mu(t_{\eta-1})}$. Then, the $i^{th}$ customer's arrival is at $t_i = t_{\eta-1} \cdot \alpha + t_\eta \cdot (1 - \alpha)$.

In Step 0, booking horizon is divided into small time intervals and a table for expectation function, $\mu$, is generated. The aim here is that the intervals should be very short to have the probability of having two customer arrivals in an interval is as small as possible. As the interval lengths get smaller, the approximate inverse of $\mu$ used in Algorithm 3 gets closer to the exact inverse. Note that the numerical approximation in this algorithm is $\mu(s) = \mu(t_\eta)$ for all $s$ in interval $\eta$. Step 1 and Step 2 are the same as the first two steps in Algorithm 2. In Step 3, the interval for the $y_i$ value is found from the table for $\mu$ values. Interpolation is used in Step 4 as an approximation to find the inverse of $\mu$.

## 5.2 Bayesian Forecasting Method

The second step for our simulation studies is updating the demand distribution at certain points in time during the booking horizon. In order to re-optimize the booking limits and bid prices upon every update of the demand distribution, Bayes' rule is used. The derivations of de Boer (1999) and Popovic (1987) are used to adapt Bayes' formula for RM problems. Talluri and van Ryzin (2005) note that *"A large class of forecasting methods use the Bayes' formula to merge a prior belief about forecast values with information obtained from observed data"*. Bayesian update is generally used when historical data are not sufficient to estimate distributions exactly, e.g., in the case of a new route in airlines. In other words, a new route is added to the network and there is no sufficient prior belief on the demand distribution of this new route.

In this paragraph, basic Bayesian forecasting is reviewed and its use is considered for RM. Let $Z_1, Z_2, ...$ be i.i.d. random variables representing a data-generation process and $Z_t$ has a density function $f(z|\theta)$, which is a function of single, unknown parameter $\theta$. Since $\theta$ is unknown, it is assumed that the prior density of $\theta$ is $g(\theta)$ and this density represents our current belief about $\theta$. As the data is observed, the new distribution, which is called posterior distribution, $g_1(\theta)$, is calculated by using Bayes' rule as follows:

$$g_1(\theta) \quad = \quad \frac{g_0(\theta) f(z_1|\theta)}{\int_\theta g_0(\theta) f(z_1|\theta) d\theta},$$

where $z_1$ is the first observation, $g_0(\theta)$ is the initial prior distribution. Note that $g_1$ is the conditional density function of $\theta$ given that $Z_1 = z_1$. The Bayes estimator of $\theta$ is then the expected value of $\theta$.

$$\theta^* \quad = \quad E[\theta] = \int_\theta \theta g_1(\theta) d\theta.$$

Talluri and van Ryzin (2005) note that, for a pair of distributions, the posterior distributions have the same distributional form as the prior, which is called conjugate family of prior distributions. Poisson-Gamma is a well known pair of conjugate families of prior distributions. This is pointed out by Talluri and van Ryzin (2005) as follows: "$Z_1, Z_2, ..., Z_N$ *have a Poisson distribution with mean $\lambda$ and $\lambda$ has a Gamma distribution with parameters $p, \gamma$. After observing $z_1, z_2, ..., z_N$, $\lambda$ has a Gamma distribution with parameter $p + \sum_{k=1}^{N} z_k$ and $\gamma + N$.*" The use of this observation above by de Boer (1999) and Popovic (1987) is explained in the remaining part of this section. Recall that $t$ is the remaining time to the departure and $T$ is the length of the booking horizon. Let $D_j(t_1,t_2)$ denote the demand for $ODF\ j$ in time interval $(t_2,t_1)$ where $t_1 > t_2$. From Remark 5.1.1 $D_j(T,0) \sim NB(p_j, \frac{\gamma_j}{\gamma_j+1})$. That is, $D_j(T,0) = D_j$. Let $d_j(T,t)$ be the realized demand for $ODF\ j$ in time interval $(t,T)$. Then, the conditional demand distribution of demand $D_j(t,0)$ for the remaining reservation period $(0,t)$ is again Negative Binomial with parameters $p_j + d_j(T,t)$ and $\frac{\gamma_j+\beta_j(T,t)}{\gamma_j+1}$.

**Remark 5.2.1** *(de Boer 1999, Popovic 1987)*
$D_j(t,0)|D_j(T,t) = d_j(T,t) \sim NB(p_j + d_j(T,t), \frac{\gamma_j+\beta_j(T,t)}{\gamma_j+1})$*, where* $\beta_j(T,t) = \int_t^T \beta_j(x)dx$*.*
**Proof.** *Assume that the booking horizon is divided into disjoint intervals* $(t_{\omega-1}, t_\omega)$ *in such a way that the union of intervals is* $(0,T)$*, and* $A_j = \lambda_j$*. Then, the expected number of arrivals in the* $\omega^{th}$ *interval is as follows:*

$$\mu_{j\omega} = \int_{t_{\omega-1}}^{t_\omega} \Lambda_j(t)dt = \lambda_j \int_{t_{\omega-1}}^{t_\omega} \beta_j(t)dt = \lambda_j c_{j\omega},$$

*where* $c_{j\omega} = \int_{t_{\omega-1}}^{t_\omega} \beta_j(t)dt$ *and second equality results from (5.2). Talluri and van Ryzin (2005) note that Gamma distribution has the following feature:*

$$\text{If } A_j \sim Gamma(p_j,\gamma_j) \text{ and } \mu_{j\omega} = A_j c_{j\omega}, \text{ then } \mu_{j\omega} \sim Gamma(p_j, \frac{\gamma_j}{c_{j\omega}}). \tag{5.9}$$

*For our case, since* $c_{j\omega}$ *is constant and prior distribution of* $A_j$ *is Gamma,* $\mu_{j\omega}$ *is Gamma distributed. Then, the probability of demand* $D_{j\omega}$ *for* $ODF\ j$ *in interval* $\omega$ *for given* $A_j = \lambda_j$ *is Poisson distributed with rate* $\mu_{j\omega} = \lambda_j c_{j\omega}$ *(Recall (5.3) in Remark 5.1.1.).*

$$P(D_{j\omega} = k|A_j = \lambda_j) = \frac{(\lambda_j c_{j\omega})^k}{k!} e^{-\lambda_j c_{j\omega}} \text{ for } k = 0,1,....$$

*An alternative notation for* $D_{j\omega}$ *is* $D_j(t_{\omega-1}, t_\omega)$*.*

*For the first time interval,*

$$P(D_{j1} = k|A_j = \lambda_j) = \frac{(\lambda_j c_{j1})^k}{k!} e^{-\lambda_j c_{j1}} \text{ for } k = 0,1,..., \tag{5.10}$$

*where* $A_j$ *is* $Gamma(p_j,\gamma_j)$ *distributed. Recall that the density function of* $A_j$ *is*

$$f_{A_j}(\lambda_j) = \frac{\gamma_j^{p_j} \lambda_j^{p_j-1} e^{-\lambda_j \gamma_j}}{\Gamma(p_j)} \qquad for\ \lambda_j > 0. \tag{5.11}$$

*From Equations (5.10) and (5.11), the probability of demand being equal to* $d_{j1}$ *in the first* *period is*

$$\begin{aligned} P(D_{j1} = d_{j1}) &= \int_0^{\infty} P(D_{j1} = d_{j1}|A_j = \lambda_j) f_{A_j}(\lambda_j) d\lambda_j \\ &= \int_0^{\infty} \frac{(\lambda_j c_{j1})^{d_{j1}}}{d_{j1}!} e^{-\lambda_j c_{j1}} f(\lambda_j) d\lambda_j \\ &= \binom{p_j + d_{j1} - 1}{d_{j1}} (\frac{\gamma_j}{\gamma_j + c_{j1}})^{p_j} (\frac{c_{j1}}{\gamma_j + c_{j1}})^{d_{j1}}, \end{aligned}$$

*which is again a Negative Binomial distribution. That is,* $D_{j1} = D_j(T, t_1) \sim NB(p_j, \frac{\gamma_j}{\gamma_j + c_{j1}})$.

*Now suppose* $D_{j1} = d_{j1}$. *From Bayes' rule, the posterior distribution of* $A_j$ *given that* $D_{j1} = d_{j1}$ *is*

$$\begin{aligned} f_{A_j|D_{j1}}(\lambda_j|d_{j1}) &= \frac{P(D_{j1} = d_{j1}|A_j = \lambda_j) f_{A_j}(\lambda_j)}{\int_0^{\infty} P(D_{j1} = d_{j1}|A_j = \lambda_j) f_{A_j}(\lambda_j) d\lambda_j} \\ &= \frac{\frac{(\lambda_j c_{j1})^{d_{j1}}}{d_{j1}!} e^{-\lambda_j c_{j1}} \frac{\gamma_j^{p_j} \lambda_j^{p_j - 1} e^{-\gamma_j \lambda_j}}{\Gamma(p_j)}}{\binom{p_j + d_{j1} - 1}{d_{j1}} (\frac{\gamma_j}{\gamma_j + c_{j1}})^{p_j} (\frac{c_{j,1}}{\gamma_j + c_{j1}})^{d_{j1}}} \\ &= \frac{\frac{\lambda_j^{p_j + dj1 - 1} e^{-\lambda_j(\gamma_j + c_{j1})} \gamma_j^{p_j} c_{j1}^{d_{j1}}}{d_{j1}! \Gamma(p_j)}}{\frac{(p_j + d_{j1} - 1)! \gamma_j^{p_j} c_{j1}^{d_{j1}}}{(p_j - 1)! d_{j1}! (\gamma_j + c_{j1})^{p_j + d_{j1}}}} \\ &= \frac{\lambda_j^{p_j + dj1 - 1} e^{-\lambda_j(\gamma_j + c_{j1})} (\gamma_j + c_{j1})^{p_j + d_{j1}}}{\Gamma(p_j + d_{j1})}. \end{aligned}$$

*That is,* $A_j|D_{j1} = d_{j1}$ *is again Gamma distributed with distribution function* $Gamma(p_j + d_{j1}, \gamma_j + c_{j1})$. *Then, from (5.9),* $\mu_{j1}$ *is also Gamma distributed.*

$$\mu_{j1} = A_j c_{j1} \Rightarrow \mu_{j1}|D_{j1} = d_{j1} \sim Gamma(p_j + d_{j1}, \frac{\gamma_j + c_{j1}}{c_{j1}}).$$

*The posterior distribution of* $\mu_{j2}|D_{j1} = d_{j1}$ *is also calculated in the same way as shown below:*

$$\mu_{j2} = A_j c_{j2} \Rightarrow \mu_{j2}|D_{j1} = d_{j1} \sim Gamma(p_j + d_{j1}, \frac{\gamma_j + c_{j1}}{c_{j2}}).$$

*Proceeding as in the case of deriving the distributions for the first interval, the posterior distributions for the second interval can be calculated as follows:*

$$\begin{aligned} D_{j2}|D_{j1} = d_{j1} &\sim NB(p_j + d_{j1}, \frac{\gamma_j + c_{j1}}{\gamma_j + c_{j1} + c_{j2}}), \\ A_j|D_{j1} = d_{j1}, D_{j2} = d_{j2} &\sim Gamma(p_j + d_{j1} + d_{j2}, \gamma_j + c_{j1} + c_{j2}), \\ \mu_{j2}|D_{j1} = d_{j1}, D_{j2} = d_{j2} &\sim Gamma(p_j + d_{j1} + d_{j2}, \frac{\gamma_j + c_{j1} + c_{j2}}{c_{j2}}). \end{aligned}$$

*By continuing with the following intervals, the results are generalized as follows for any interval n:*

$$D_{jn}|D_{j1} = d_{j1}, ..., D_{j,n-1} = d_{j,n-1} \sim NB(p + \sum_{\omega=1}^{n-1} d_{j\omega}, \frac{\gamma + \sum_{\omega=1}^{n-1} c_{j\omega}}{\gamma + \sum_{\omega=1}^{n} c_{j\omega}}),$$

$$A_j|D_{j1} = d_{j1}, ..., D_{jn} = d_{jn} \sim \Gamma(p + \sum_{\omega=1}^{n} d_{j\omega}, \gamma + \sum_{\omega=1}^{n} c_{j\omega}),$$

$$\mu_{jn}|D_{j1} = d_{j1}, ..., D_{jn} = d_{jn} \sim \Gamma(p + \sum_{\omega=1}^{n} d_{j\omega}, \frac{\gamma + \sum_{\omega=1}^{n} c_{j\omega}}{c_{jn}}).$$

*By using these equations, the demand distribution for the remaining reservation period, (0,t), is again Negative Binomial when the number of arrivals up to time t is* $d_j(T,t)$.

$$D_j(t,0)|D_j(T,t) = d_j(T,t) \sim NB(p_j + d_j(T,t), \frac{\gamma_j + \beta_j(T,t)}{\gamma_j + 1}). \tag{5.12}$$

*Then, the expectation of demand* $D_j(t,0)$ *in the remaining booking horizon is*

$$E(D_j(t,0)) = \frac{d_j(T;t) + p_j}{\gamma_j + \beta_j(T;t)} \beta_j(t,0). \tag{5.13}$$

■

## 5.3 Analyses of Mathematical Models

This section is devoted to the numerical analyses on proposed models in order to determine the following in the simulation studies: Number of demand realizations, threshold level for revenue and aggregation size of the demand groups for the models and procedures proposed in Chapter 4. In the following sections, these three values are determined by analyzing the sample network. The data for this network is given in Appendix B. Another issue that must be clarified at this point is determining the upper bounds for the number of seats available, $B_j$. The alternative ways to determine $B_j$ are given in Section 3.2. In the proposed models, third alternative is used with $\psi = .98$ and 80 is found as the $max_l\{\Upsilon_l : l = 1, 2, 3\}$. In order to simplify the mathematical models and simulation studies, 80 is used as the $B_j$ value for each $j$.

### 5.3.1 Studies on the Number of Demand Realizations

In the models proposed in Chapter 4, a sample number of demand realizations is used. It is obvious that generating all possible demand realizations is better than using a subset of demands. However, because of the computational limits, the size of the subset must be limited at an acceptable level. In this section, the number of demand realizations is determined by analyzing expected revenue, seat allocation and bid price as a function of the number of demand realizations.

First of all, the relation between the number of demand realizations, $\Theta$, and expected revenues is analyzed for *PLP-RM* models. The model is solved 30 times for $\Theta = 10, 20, ..., 100, 250$. Working with 30 runs can be explained by referring to Central Limit Theorem (Montgomery et al. 1997).

> *Central Limit Theorem (CLT): If $X_1, X_2, ..., X_n$ is a random sample of size n taken from a population with mean $\mu$ and $\sigma^2$, and if $\overline{X}$ is the sample mean, then the limiting form of the distribution $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ as $n \to \infty$, is the standard normal distribution.*

In general, if $n \geq 30$, the Normal approximation is assumed to be satisfactory regardless of the shape of the distribution. The formulas for sample mean, sample standard deviation and confidence intervals are given below. These formulas are given for a sample of size $n$, which is sufficiently large and taken from a population with unknown mean and variance. Let $X_1, X_2, ..., X_n$ be the random variables in this sample.

$$\text{Sample Mean} \quad : \quad \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}.$$

$$\text{Sample Standard Deviation} \quad : \quad S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}.$$

$$\text{Confidence Interval for } X \quad : \quad \bar{X} - z_{\alpha/2}\frac{S}{\sqrt{n}} \leq E(X) \leq \bar{X} + z_{\alpha/2}\frac{S}{\sqrt{n}}$$

where $z$ is the Standard Normal score to construct $(1-\alpha)\%$ confidence interval. In Figure 5.3, the sample mean varies between 71400 and 71600 without an increasing or decreasing behavior.

The main aim of this analysis is to find the minimum number of demand realizations such that seat allocations and expected revenue are almost stable beyond this number. Therefore, when sample mean for revenue is almost stable, the only thing that is important for this analysis is sample standard deviation. The sample mean, sample standard deviation and 95% confidence interval for revenue are given in Figure 5.3. The sample standard deviation decreases as $\Theta$ increases when $\Theta \leq 30$. For numbers greater than 30, sample standard deviations are almost stable. The conclusions according to these analyses are as follows:

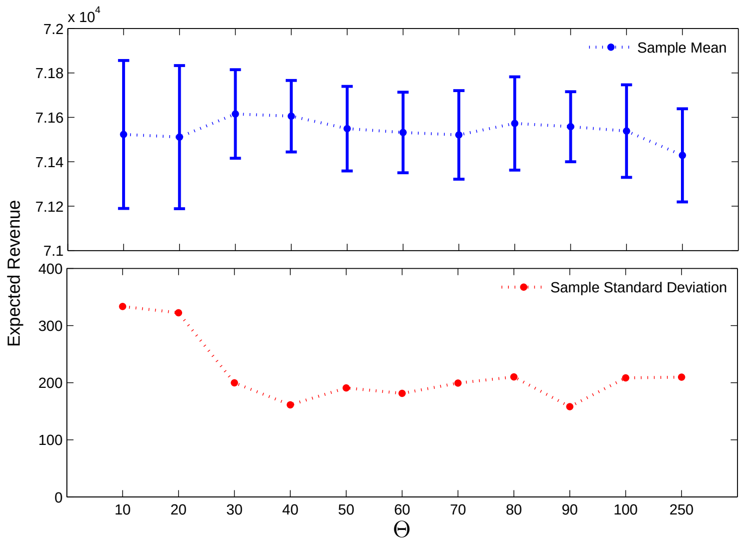


Figure 5.3: Expected revenue as a function of the number of demand realizations for *PLP-RM* procedure.

- The sample mean for expected revenue is nearly the same for different $\Theta$ values.
- The sample standard deviation decreases in $\Theta$ for $\Theta \leq 30$ and almost stable for $\Theta > 30$.
- It is expected that the sample mean and sample standard deviation would not change significantly when $\Theta > 30$.

The second analysis for *PLP-RM* procedure is for the relation between the number of demand realizations and seat allocations. As in the case of the analysis for revenue, sample mean for seat allocations are almost stable for different number of demand realizations. Therefore, only the sample standard deviations for nine sample *ODF*s are given in Figure 5.4. The following are the observations according to these analyses:

- The sample standard deviations are generally high for fare classes 1 and 3 which are the lowest and highest fare classes, respectively. Moreover, risk defined in terms of $P(R < L)$ is minimized by shifting seat allocations between high and low fare classes. Impact of the number of demand realizations on the allocations for the medium fare classes are generally less than the other classes.
- Increasing $\Theta$ generally decreases the sample standard deviations of the seat allocations.
- For $\Theta \geq 100$, the sample standard deviations are less than 2 seats.

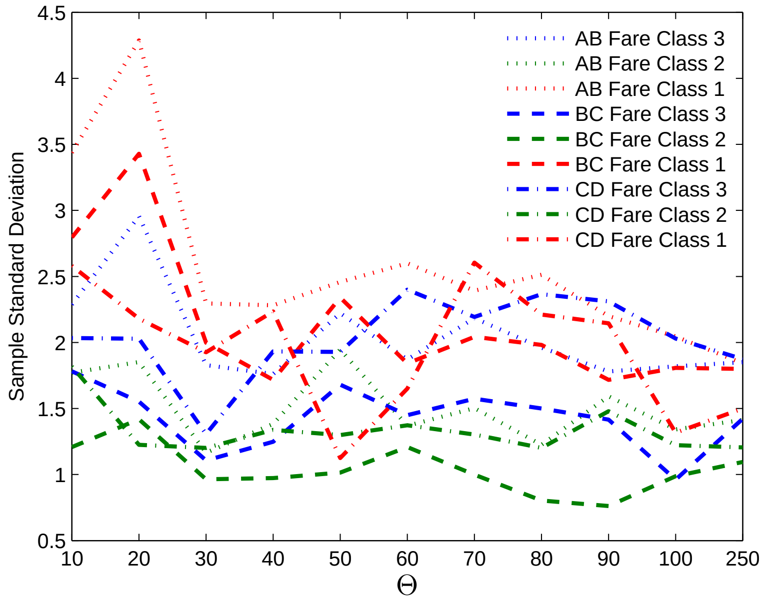


Figure 5.4: Seat allocations as a function of the number of demand realizations for *PLP-RM* procedure.

- The seat allocations are almost stable for higher $\Theta$ values, especially for $\Theta > 100$.

Another analysis for *PLP-RM* model is for bid prices using the same data given in Appendix B. Bid prices for all of the three legs are used in order to calculate sample mean and sample standard deviation values. In Figure 5.5, the graphs of these results and a sample 95% confidence interval for population mean is given. As it can be seen from the figure, even for 250 demand realizations, bid prices for leg *AC* may be significantly different for different runs. The confidence interval of population mean for leg *AC* is between 40 and 100 even for 250 demand realizations. More clearly, we are 95% confident that population mean of bid price is between 40 and 100, which is an unacceptable interval in bid price control. As a result, we conclude that using bid prices as a control mechanism is not reasonable when *PLP-RM* model is used.

Last analysis for *PLP-RM* model is for the relation between the number of demand realizations and $P(R < L)$. In Figure 5.6, the sample means and sample standard deviations for $L = 70000$ and different $\Theta$s are given. Moreover, 95% confidence interval for population mean of $P(R < L)$ is given in the figure. The observations are the same for different $L$ values; here, only the results for $L = 70000$ are given. The following are the observations.

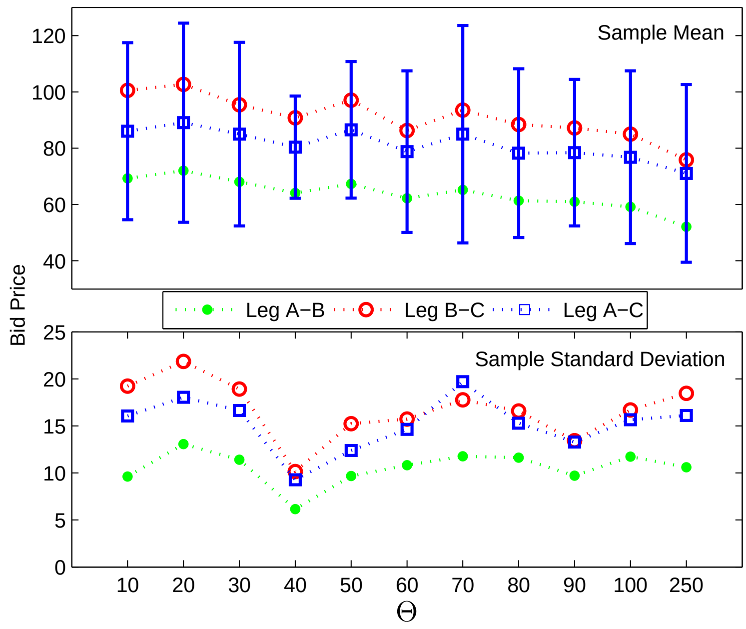


Figure 5.5: Bid price as a function of the number of demand realizations for *PLP-RM* procedure.

- The sample standard deviation decreases as number of demand realizations increases.
- As a result of the decrease in the standard deviations, the confidence intervals get narrow as $\Theta$ increases.
- Although confidence intervals are not small enough for the probability measure under consideration, they are considered acceptable for $\Theta \geq 50$.

Studies on the number of demand realizations for *PLR-RM* are completed. To conclude, $\Theta$ is set to the maximum of the values found in these analyses, $\Theta = max\{30, 100, 50\} = 100$.

Next the studies on the number of demand realizations are considered for *PMP-RC* model. For this model, $\Theta = 10, 20, ..., 100$ are used in the mathematical models and all of them are solved 30 times. In order to analyze $\Theta$ values, $L$ and $\rho$ are fixed at 61000 and 0, respectively. These values are chosen arbitrarily. Recall that $\rho$ is a predetermined constant number and used in the constraint $P(R < L) \leq \rho$. The sample mean and standard deviation of expected revenue are given as in Table 5.1. The main difficulty in these analyses is the infeasible solutions, which are excluded in the calculations of mean and standard deviation. As it can be seen in Table 5.1, the number of infeasible solutions out of 30 runs slightly increases as

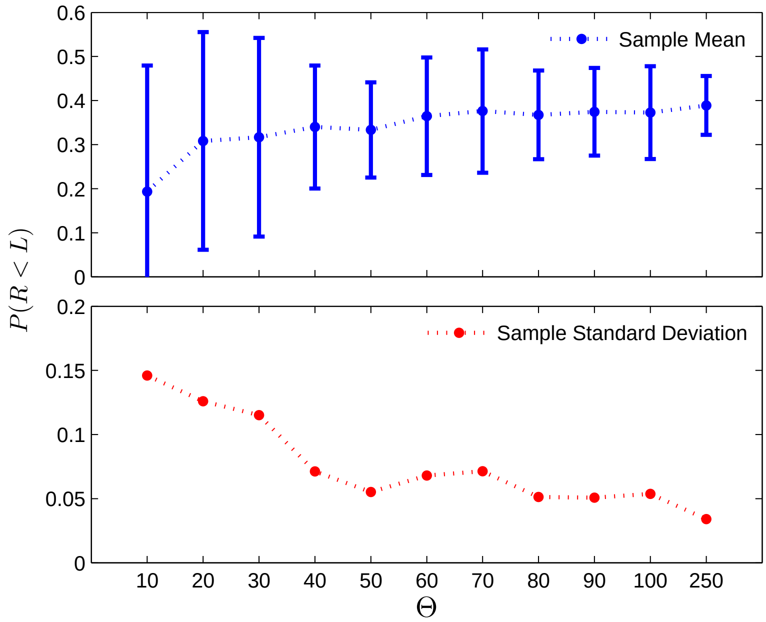


Figure 5.6: $P(R < L)$ as a function of the number of demand realizations for *PLP-RM* model, $L$=70000.

the number of demand realizations increases. Increasing the number of demand realizations generally increases the number of constraints and the feasible region gets smaller. Therefore, it is more probable to see more infeasible solutions when the number of demand realizations gets larger and the results become more realistic than the ones with small number of demand realizations. Although the results in Table 5.1 do not show a steady state, results for $\Theta$ =80, 90 and 100 are not so different. Because the feasible region gets smaller when $\Theta$ gets larger, the variability between different runs increases, which results an increase in sample standard deviation. According to these results, we conclude that

- By increasing the number of demand realizations, more realistic results are obtained in terms of sample mean and sample standard deviation.
- $\rho$ values for the risk constraint have a significant impact on the quality of the results. Low $\rho$ values result in more infeasible solutions and high values result in the same allocations as the classical *EMR* model.

The last analysis on the number of demand realizations is for the *RRS* procedure. In these analyses, *RLP* model is solved 1000 times and the solutions are divided into two groups with

Table 5.1: Revenues in *PMP-RC* Model for Different Demand Realizations, $L$=61000, $\rho = 0$.

| | **Number of Demand Realizations Θ** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **10** | **20** | **30** | **40** | **50** | **60** | **70** | **80** | **90** | **100** |
| **Sample Mean** | 71759,42 | 71622,23 | 71483,41 | 71090,73 | 71118,05 | 70867,28 | 70535,34 | 70221,06 | 70071,16 | 70386,55 |
| **Sample St. Dev.** | 21,50 | 276,17 | 353,83 | 594,82 | 845,83 | 1306,49 | 1146,35 | 1659,47 | 1416,30 | 1167,93 |
| **Infeasible Samples out of 30 runs** | 0 | 1 | 1 | 2 | 2 | 3 | 4 | 6 | 5 | 6 |

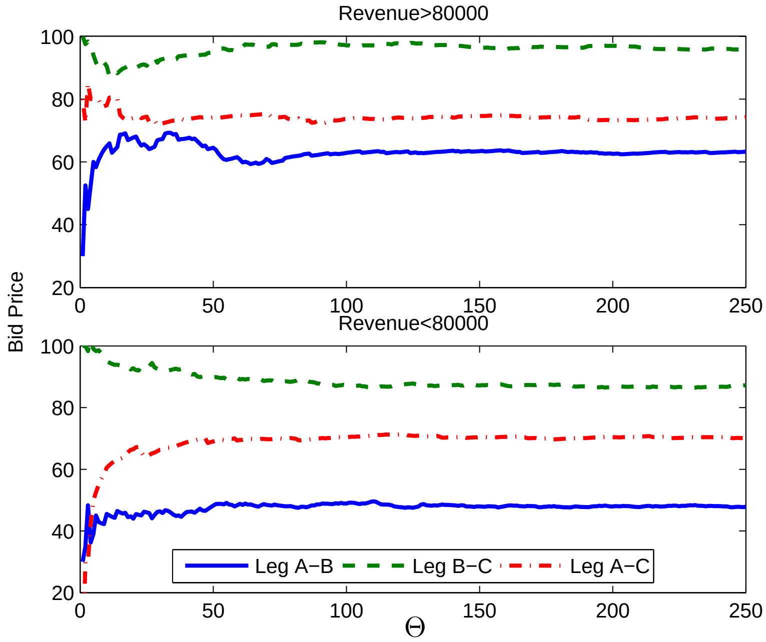


Figure 5.7: Bid prices as a function of the number of demand realizations for *RRS* procedure.

the proposed procedure in Chapter 4. Afterwards, these solutions are used to calculate mean bid prices for the number of demand realizations between 1 and 250. Mean bid prices are given in Figure 5.7 for three legs when the objective function values of the *RLP* models are greater than 80000 and smaller than 80000. 80000 is an arbitrarily chosen threshold level. It can be easily seen in Figure 5.7 that bid prices converge to a value when number of demand realizations are greater than 100. Therefore, the minimum number of Θ can be considered as 100 in our simulation studies.

As a result of the analyses in this section, it is determined that Θ value must be at least 100 in the simulation studies in Chapter 6 for *PLP-RM*, *PMP-RC* and *RRS* procedures.

### 5.3.2 Studies on Threshold Levels

This section is devoted to the analyses on threshold levels for revenue. Expected revenue, seat allocation and probability of revenue being less than the threshold level are taken into consideration in order to determine the interval for threshold level. These analyses are done in order to use the appropriate threshold levels in the simulation studies in Chapter 6. In real life, decision makers generally set their threshold level according to their experiences or other decision making mechanisms.

First analysis is again for *PLP-RM* models, which is solved 30 times for different threshold levels with $\Theta = 100$, which is set in Section 5.3.1. 30 different runs are shown with different colors in the figure. According to Figure 5.8, revenue is less than the one gained by using the *EMR* model between $L$=50000 and $L$=85000. The decrease in expected revenue is because of incorporating the risk measure. The probability of expected revenue being less than 50000 is nearly 0, and expected revenue is lower than 85000 with probability 1. $L$ values which are not in this range give nearly the same results with *EMR* model. The dashed lines in the graphs show a basic fitting, average values, for expected revenue and probability of revenue being less than the threshold level.

As a second analysis for *PLP-RM* models, the effect of threshold levels on booking limits is investigated. Figure 5.9 shows average booking limits for itineraries AB and AD, respectively. For both of the itineraries, booking limits again vary when the predetermined revenue level $L$ is between 50000 and 85000. Beyond these values, i.e., for values smaller than 50000 or greater than 85000, the seat allocations for *PLP-RM* model are the same as the allocation of the *EMR* model. According to Figures 5.8 and 5.9, the observations are as follows:

- When risk minimization is taken into account, a decrease in the expected revenue is seen.
- As it is expected, the probability graph has an $S$ shape, which is similar to the cdf of the Normal distribution.
- For our specific sample problem, risk minimization can only be considered in the range between 50000 and 85000. Out of this range, $P(R < L)$ is 0 or 1 and the optimal seat allocation for *EMR* model is also optimal for *PLP-RM* procedure.

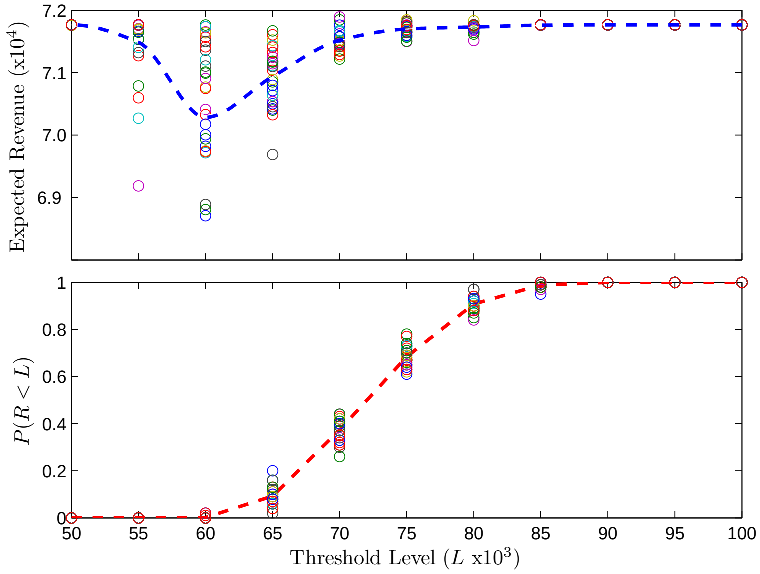


Figure 5.8: Revenue as a function of threshold levels for *PLP-RM* procedure.

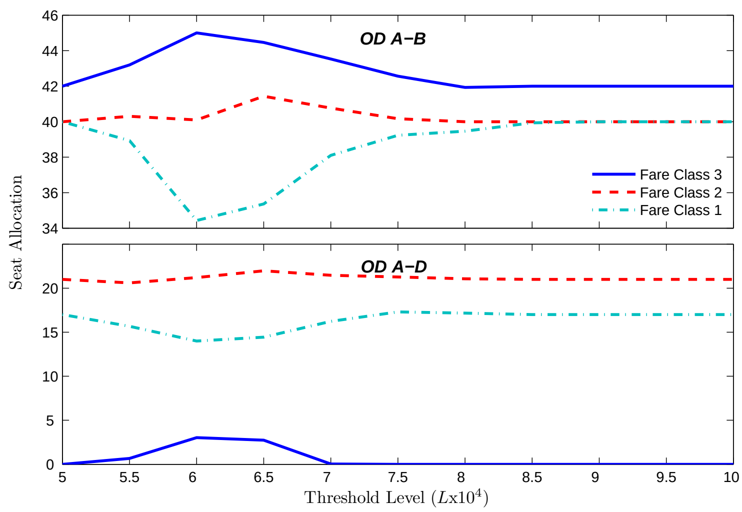


Figure 5.9: Seat allocations as a function of threshold levels for *PLP-RM* procedure.

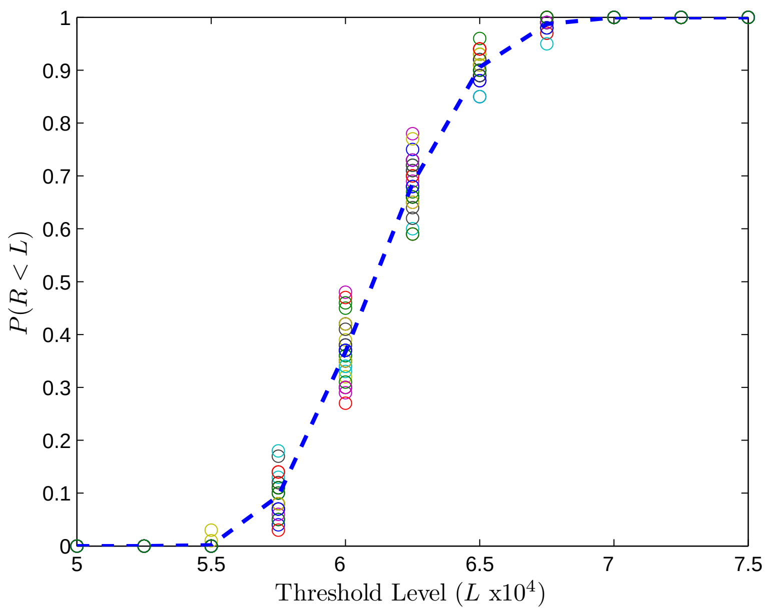


Figure 5.10: $P(R < L)$ as a function of $L$ when $\rho = 1$.

For the *PMP-RC* model, the constraint that is used to limit the probability of revenue being less than the threshold level is dependent on both the threshold level, $L$, and right hand side of the constraint, $\rho$. Hence, a three dimensional analysis is required for this model which is computationally impractical. However, the intervals for $\rho$ values for different $L$ values can be found without risk minimization as in Figure 5.10. When $\rho = 1$, the optimal solution of the *PMP-RC* model is equal to the optimal solution of *EMR* model because the constraint on $P(R < L)$ in the *PMP-RC* model becomes redundant. By solving *PMP-RC* model more than once, different $P(R < L)$ values can be obtained and a decision on the interval of $\rho$ values can be made according to these $P(R < L)$ values. Figure 5.10 gives us these intervals for $\rho$ values for different threshold levels. For threshold levels lower than 55000 or greater than 85000, the probabilities are almost the same for different runs. In other words, for a threshold level lower than 55000, we have almost no risk of obtaining revenue less than 55000. On the other hand, for a threshold level greater than 85000, it is almost certain that revenue will be less than 85000. As a result of this observation, the interval for threshold level in *PMP-RC* model is set as [55000, 85000].

The last analysis on the threshold levels is for the *RRS* procedure. The analyses are for limits between 70000 and 95000 with increments of 100. 1000 demand realizations are used as in

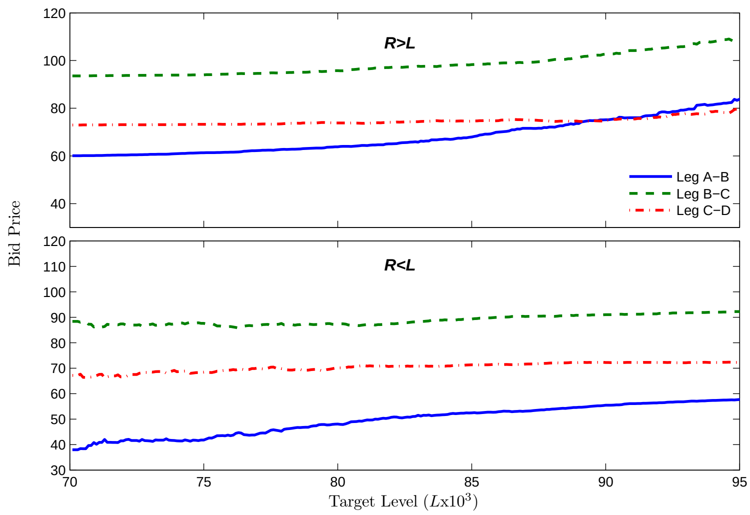


Figure 5.11: Bid prices as a function of threshold levels for *RRS* procedure.

Section 5.3.1. The expected revenue values obtained by *RLP* model used in the *RRS* procedure turn out to be similar to the expected revenue obtained by *DLP* model which is 84915. The revenue is generally greater than 70000 and smaller than 95000 in our experiments. Because of inadequate number of data for $L < 70000$ and $L < 95000$, bid prices are not calculated for those cases. The results for two groups with $R > L$ and $R < L$ are given in Figure 5.11. The observations are summarized as follows:

- Bid prices of the legs considered increase as the threshold level, $L$, increases.

- *RRS* procedure shows the difference between risk taking and risk-averse decision makers for adjusting bid prices. In the first group with $R > L$, customers are accepted with higher bid prices as compared to the bid prices of the second group. Briefly, decision makers, who use the bid prices of the first group, expect to get more revenue from high fare customers and reject low fare ones.

According to the analyses in this section, it is concluded that threshold levels for *PLP-RM* and *PMP-RC* procedures in simulation studies must be in the interval of [55000, 85000] and threshold levels for *RRS* procedure must be in the interval of [70000, 95000].

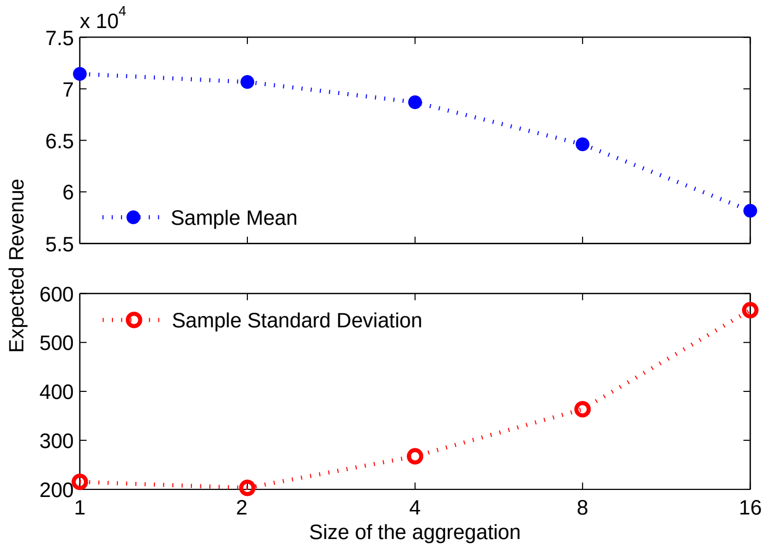


Figure 5.12: Revenue as a function of demand aggregation for *PLP-RM* procedure.

### 5.3.3 Studies on Demand Aggregation

The last study on mathematical models is for demand aggregation. As it is stated before, increasing the number of demand realizations, $\Theta$, is time consuming. However, larger $\Theta$ values give stronger results than the smaller ones give as it is examined in Section 5.3.1. Therefore, the dilemma between increasing the number of demand realizations and aggregating demands for the *SLP-RM* and *PMP-RC* models must be analyzed.

In this section, the size of the aggregated groups is changed from 2 to 16 and all of them are experimented 30 times. Recall from the first paragraph of Section 5.3 that the upper bounds for the number of seats available for all *ODF*s are 80. Size 1 means that all seats are taken into consideration separately as in the *EMR* model and size 16 means that demands are aggregated into 5 groups, which is $80/16$. The sample mean and standard deviation of the revenue are calculated for all of the demand groups as in Figure 5.12. The following are observed according to this analysis:

- Sample mean (sample standard deviation) decreases (increases) in the size of the groups.
- The decrease in mean and increase in standard deviation is almost negligible when size of the groups is smaller than or equal to 4.

As a second step, the impact of demand aggregation on seat allocation is analyzed. The

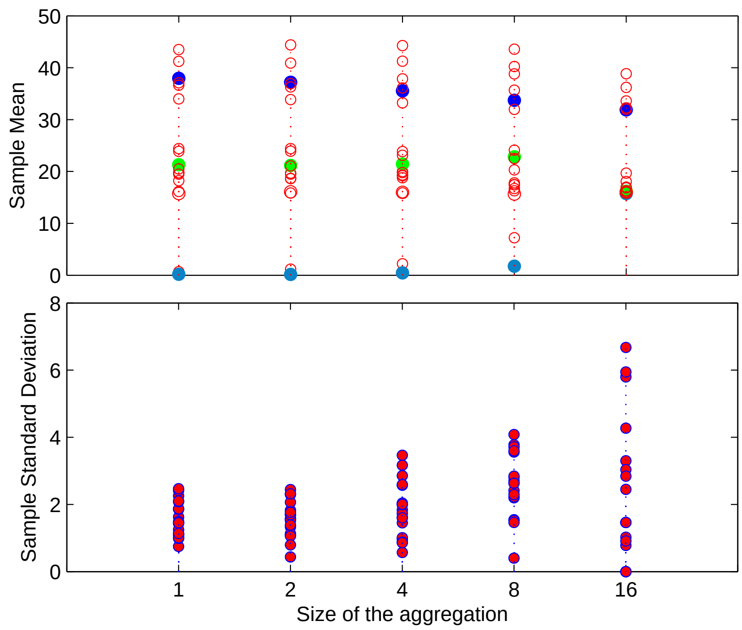


Figure 5.13: Seat allocation as a function of demand aggregation for *PLP-RM* procedure.

sample mean and standard deviation are used to analyze results as in Figure 5.13. Bullets are used to show sample means and standard deviations for a specific number of groups and $ODF$. In graph at the top, bullets for three different $ODF$s are shaded in order to show the behavior of booking limits as a function of demand aggregation. Dark blue bullets are for a high-fare class, green bullets are for a standard-fare class and light blue bullets are for a low-fare class. These graphs lead to the following observations:

- Sample means are almost the same for groups of 1 to 8.
- Sample standard deviation increases in the size of the demand groups. The sample standard deviations for a size up to 4 are smaller than 4 which is acceptable especially for nested booking policy.

For the *PMP-RC* model, the impact of demand aggregation on expected revenue is analyzed when $L = 60,000$ and $\rho = 0$. Again the values are chosen arbitrarily. According to the results given in Figure 5.14, both sample mean and standard deviation decrease as the size of demand groups increase. It is obvious that expectation should decrease as size of aggregated groups increases; note that there is no approximation in the model when the aggregation size is 1. However, in contrary to *SLP-RM* model, aggregating demands decreases variance. The

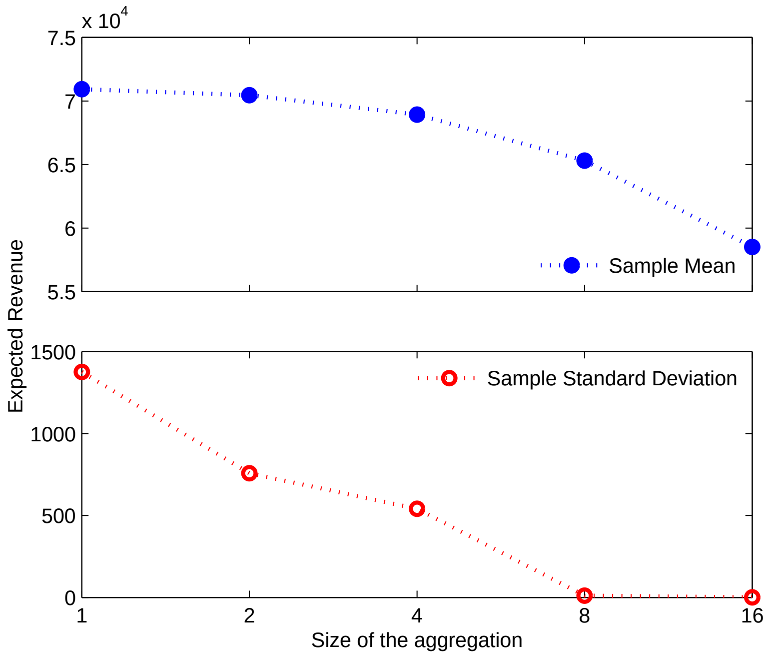


Figure 5.14: Expected revenue as a function of demand aggregation for *PMP-RC* model.

reason of this situation is that aggregating demands makes it easier to satisfy the risk constraint in the model. Hence, for high demand aggregation sizes, *PMP-RC* gives the same results as the *SLP* model in terms of seat allocations. Hence, it is not suggested to aggregate demands in the *PMP-RC* model.

As a result of all of the analyses in this section, the following are the conclusions for all of the models proposed in Chapter 4:

- Demand aggregation worsens the quality of the solutions in terms of expected revenue. The deterioration is acceptable when the number of the groups is smaller than or equal to 4 for our sample network.
- It is not suggested to use demand aggregation in the *PMP-RC* model.

# CHAPTER 6

# SIMULATION STUDIES

This chapter is devoted to the numerical studies on simulation models for different booking policies given in Chapter 4. The sample three-leg network due to de Boer (1999) that is given in Chapter 5 is used in this thesis for the numerical studies. Five different scenarios are taken into consideration. First one is the base problem given in Chapter 5. The second one is the case with low-before-high arrival pattern and the third one is the case with increased variance of low-fare demand. In the fourth case, the differences between fares are lowered. In the last scenario, the realistic coefficient of variations of demand and relatively close fares are used and the order of the arrivals is not specified. The data for the scenarios are given in Appendix B.1. These scenarios are taken from de Boer (1999) and they are studied in order to see the performance of the proposed models in different cases. In simulation studies, a Bayesian update is used for demand distributions in order to see the effect of update on the performances of the models. The booking limits (allocations) and bid prices obtained by solving the models are used in the simulation models. In simulation studies, the *EMR*, *DLP*, *SLP*, *EMVLP* models in the literature are compared with the proposed models *SLP-RM*, *PMP-RC* and *RRS* procedure.

According to the analyses in Chapter 5 for the proposed models, the number of demand realizations should be higher than 100 and aggregations should be for groups of four. However, it is not clear yet which one of the following is better for the proposed approach: using a small number of realizations without aggregation or using more realizations with aggregation. The analysis on this dilemma is in Section 6.1. Different than the analyses in Chapter 5, the seat allocations obtained by the models proposed in Chapter 4 are used in the simulation models in this chapter. The analyses and comparisons are based on the simulation results.

The sample mean (*SM*) and sample standard deviation (*SSD*) for revenue, sample coefficient of variation (*SCV*) and load factor (*LF*) are calculated using the simulation results. The "sample" in these analyses is the collection of data and its magnitude is equal to the replication number in simulation studies which is 10000 in our case. The sample coefficient of variation is defined as

$$SCV = \frac{S}{\bar{X}} \tag{6.1}$$

where $S$ is the sample standard deviation and $\bar{X}$ is the sample mean given in Section 5.3.1.

In addition to expected revenue, probability of poor performance, which is defined as the probability that the revenue is less than a threshold level, is the other performance measure considered in this thesis. Towards the end of evaluating simulation results to calculate this probability, let $r$ be the replication number and $k$ of these replications satisfy the following condition: $R < \tau$, where $\tau$ is a specified revenue level. Then,

$$\begin{aligned} \overline{p} &= \frac{k}{r}, \\ E(\overline{p}) &= p, \\ V(\overline{p}) &= \frac{p(1-p)}{r}, \end{aligned}$$

where $p$ is the proportion of success and $\overline{p}$ is the sample proportion of success. For $rp \geq 5$ and $r(1-p) \geq 5$,

$$\overline{p} \sim N(p, \frac{p(1-p)}{r})$$

due to Nelson et al. (2003). For our example problem, the simulation results are used in calculations of $\overline{p}$ values and 95% confidence interval for $\overline{p}$.

Organization of this chapter is as follows: Section 6.1 is devoted to the analysis on the dilemma between aggregation and the number of demand realizations used. In Section 6.2, simulation results for the base problem are given with nested booking control policy, partitioned booking control policy, bid price control policy and Bayesian update. The scenario for low-before-high arrival pattern is analyzed in Section 6.3. Section 6.4 is devoted to the scenario with increased low-fare demand variance and the scenario with smaller differences between fares is considered in Section 6.5. In Section 6.6, the last scenario with realistic variations and close fares is studied. This chapter ends with concluding remarks in Section

6.7. The optimal allocations used in the following sections and simulation codes are given in Appendices B.2 and C, respectively.

## 6.1 Dilemma Between Aggregation and Number of Demand Realization

The dilemma on using a small number of demand realizations without aggregation or using more realizations with aggregation is analyzed by solving base problem four times when $L = 70000$ and $\Theta = 100$ without aggregation and $\Theta = 1000$ with an aggregation of 4. The results are given in Table 6.1. In the second and third columns, the results for standard *DLP* and *EMR* models are given, respectively, in order to compare these models with the proposed ones. The following four columns are devoted to the results of simulation studies for *SLP-RM* procedure when $\Theta = 1000$, $L = 70000$ and demands are aggregated into groups of four. Each of the four columns is a separate run with the same $\Theta$ and $L$ parameters and the replication number for each run is 10000. Simulation studies are completed four times with the same data in order to see the difference between different simulation runs. In the last four columns, the results of simulation studies are given for *PLP-RM* model when $\Theta = 100$ and $L = 70000$. As it is seen in Table 6.1, both sample standard deviations and sample coefficient of variations for *SLP-RM* procedure are lower than the ones for the *DLP*, *EMR* and *PLP-RM* models. Moreover, the *SLP-RM* procedure gives better expected revenues than the *EMR* model. Our numerical observations justify the use of higher $\Theta$ values with demand aggregations instead of using lower $\Theta$ values without aggregation.

Table 6.1: *SLP-RM* procedure and *PLP-RM* model, simulation results with nested booking policy, L=70000.

| | *DLP* | *EMR* | *SLP-RM* ($\Theta$ = 1000) | | | | *PLP-RM* ($\Theta$ = 100) | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | run 1 | run 2 | run 3 | run 4 | run 1 | run 2 | run 3 | run 4 |
| $SM$ | **75,619** | 74,227 | 74,343 | 74,257 | 74,508 | 74,732 | **75,091** | 74,537 | 74,666 | 74,615 |
| $SSD$ | 6,785 | 6,950 | **5,582** | **5,594** | 5,642 | 5,792 | 6,291 | 6,411 | 6,021 | 6,640 |
| $SCV$ | 0.0897 | 0.0936 | **0.0751** | **0.0753** | 0.0757 | 0.0775 | 0.0838 | 0.0860 | 0.0806 | 0.089 |
| $LF$ | 0.8949 | 0.8785 | **0.9164** | **0.9186** | 0.9162 | 0.9131 | 0.9008 | 0.8938 | 0.9044 | 0.8875 |

In Figure 6.1, results for *PLP-RM* and *SLP-RM* models are given with confidence intervals. Because of the huge amount of replications (10000 replications in each run), the confidence intervals are so narrow, which shows that using $\overline{p}$ values without confidence intervals also

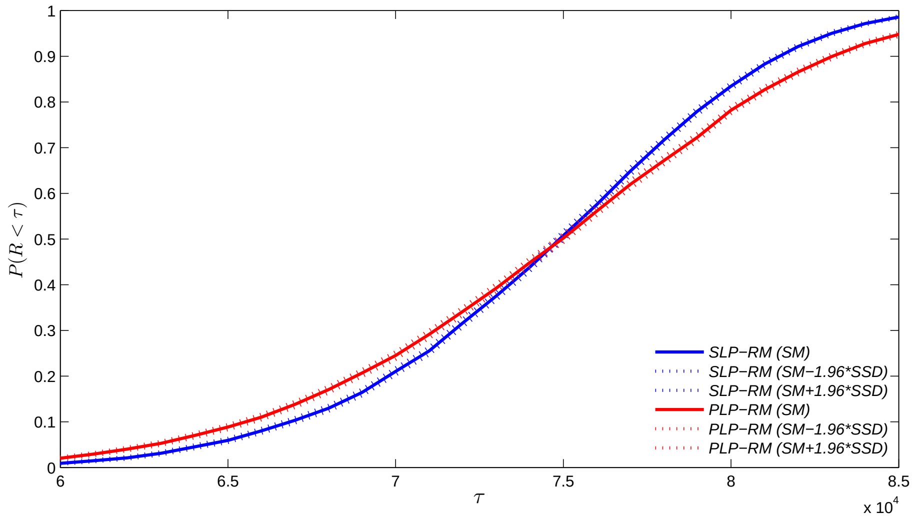


Figure 6.1: P($R < \tau$) for *SLP-RM* and *PLP-RM* models.

gives reasonable and sufficient information about the simulation results. The probability that the revenue is less than the threshold values is better for *SLP-RM* model when the threshold is less than 75000. In other words, a risk-sensitive decision maker can decrease the probability of revenue being less than the threshold level more by working with *SLP-RM* model when $\tau < 75000$. For $\tau > 75000$, *PLP-RM* model gives better results. Hence, *SLP-RM* model works better for risk-sensitive cases. In the remaining part of this chapter, *SLP-RM* and *PMP-RC* models and *RRS* procedure are analyzed with partitioned, nested and bid price control policies. The *SLP-RM* model is used for $\Theta = 1000$ with an aggregation size of four.

## 6.2 Base Problem

Base problem is the case that is close to the real life problems because of the following characteristics. In this case, long-haul flights are cheaper than the single-leg flights, there are important differences between fare classes and low-fare customers generally tend to arrive before high-fare customers, but not all of the low-fare customers arrive before high-fare customers. The data used in the base problem are given in Appendix B.1.1. The Beta distributions in Figure 5.2 are used in this section.

### 6.2.1 *SLP-RM* Model

The results are given in Tables 6.2 and 6.3 for partitioned and nested policies respectively. The mean revenues for nested booking policy are higher than the revenues for partitioned booking policy for all of the models. Moreover, some seats that are allocated to low-fare classes remain empty in partitioned booking policy although there are unsatisfied demands for high-fare classes, which causes a reduction in load factors. However, using partitioned policy decreases standard deviation and coefficient of variation slightly. Another important remark is that *EMR*-based models, which are *SLP*, *EMVLP* and *SLP-RM*, outperform *DLP* model in terms of expected revenue when partitioned booking policy is used. In nested booking policy, *DLP* gives better expected revenue than *EMR* and *EMR*-based models. Both in nested and partitioned seat inventory control policy, the models we propose give lower variances and coefficient of variations than *DLP* and *EMR* models. Also, load factors are increased when the proposed approach is used. In order to compare our approach with other risk averse approaches in the literature, we consider $EMVLP$ model due to Çetiner (2007). According to Table 6.2 and Table 6.3 , the standard deviation and coefficient of variation change in the $\theta$ value that is used in *EMVLP* model. Although *EMVLP* model gives lower coefficient of variation values when $\theta = 0.005$, it is not easy to make a generalization that *EMVLP* outperforms other models in terms of risk minimization.

Table 6.2: Simulation results with partitioned booking policy for the base problem.

| | | | | *EMVLP* ($\theta$) | | *SLP-RM* (*L*x1000) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *DLP* | *EMR* | *SLP* | 0.001 | 0.005 | 60 | 65 | 70 | 75 | 80 |
| *SM* | 70,567 | **71,698** | **71,535** | 71,515 | 67,337 | 70,189 | 70,908 | 71,522 | 71,550 | 71,547 |
| *SSD* | 5,598 | 6,268 | 5,951 | 5,641 | **2,951** | **4,301** | 5,077 | 5,691 | 5,860 | 6,305 |
| *SCV* | 0.07933 | 0.08742 | 0.08319 | 0.07887 | **0.04383** | **0.06128** | 0.07161 | 0.07957 | 0.08190 | 0.08813 |
| *LF* | 0.86247 | 0.86533 | 0.87808 | 0.88326 | **0.95188** | **0.91728** | 0.89019 | 0.87917 | 0.87444 | 0.86652 |

Table 6.3: Simulation results with nested booking policy for the base problem.

| | | | | *EMVLP* ($\theta$) | | *SLP-RM* (*L*x1000) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *DLP* | *EMR* | *SLP* | 0.001 | 0.005 | 60 | 65 | 70 | 75 | 80 |
| *SM* | **75,854** | 74,468 | 74,314 | 74,692 | 70,534 | 74,119 | 74,999 | **75,052** | 74,929 | 73,944 |
| *SSD* | 6,812 | 6,973 | 6,525 | 6,249 | **3,270** | **4,996** | 5,853 | 6,458 | 6,616 | 6,950 |
| *SCV* | 0.0898 | 0.0936 | 0.0878 | 0.0837 | **0.0464** | **0.0674** | 0.0780 | 0.0860 | 0.0883 | 0.0940 |
| *LF* | 0.8971 | 0.8807 | 0.8930 | 0.9019 | **0.9652** | **0.9338** | 0.9144 | 0.8995 | 0.8941 | 0.8779 |

Next, an analysis on the probability of poor performance is considered for the casess given in Table 6.2 and Table 6.3. In this analysis, $\overline{p}$ values for $\tau = 50000, 51000, ..., 100000$ are calculated and they are used in estimation of $P(R < \tau)$, where $\tau$ is a revenue level but not the threshold level $L$. Because of the huge number of replications used in the simulations, the standard deviation for $\overline{p}$ is very small. Hence, confidence intervals are not given in Figure 6.2. Table 6.4 is given for $60,000 < \tau < 70,000$. The best two results in each row are given with bold style in this table. Note that at least one of the best two values in each row in Table 6.4 belongs to the *SLP-RM* model. In Table 6.4 and Figure 6.2, it can be seen that $P(R < \tau)$ is almost the same for different models when $\tau$ is less than 65000. Recall from Table 6.3 that the expected revenue for *EMR* and *DLP* models are about 75000 and in Table 6.4, $P(R < 65000)$ for these models is less than 0.1. Especially for $\tau$ values between 65000 and 70000, the proposed model outperforms all other models in terms of probability values. *EMVLP* model gives reasonable results when $\theta = 0.005$, but the performance of this model is worse when $\theta = 0.001$. Moreover, *DLP* model works well when nested booking policy is used and $P(R < \tau)$ for *DLP* model is less than the values for other models when $\tau > 70,000$. However, risk-sensitive decision makers are generally concerned about the risk with respect to smaller threshold levels than the expected revenue. Hence, it is not meaninful to use higher $\tau$ values as performance measures in our studies. The concluding remarks in terms of sample mean, standard deviation, coefficient of variation and probability of poor performance for these analyses are as follows:

- Nested booking policy gives better results in terms of expected revenue than partitioned booking policy.
- *DLP* works better than *EMR* and *EMR*-based models in terms of expected revenue when nested booking policy is used.
- *SLP-RM* model generally gives better results than the other models in terms of probability of poor performance for $\tau$ values between 60000 and 70000.

### 6.2.2 *PMP-RC* Model

Second model that is proposed in Chapter 4 is *PMP-RC* model. A detailed analysis of this model is for the partitioned booking policy. As it is explained in Chapter 5, the use of *PMP-*

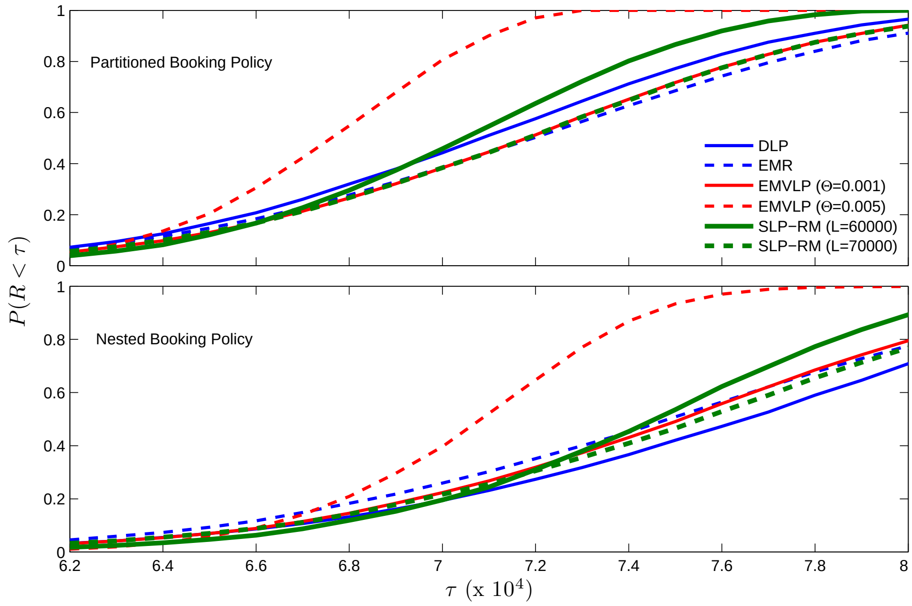


Figure 6.2: $P(R < \tau)$ for the base problem.

*RC* model is more complicated than the other models we propose because of the parameters (the threshold level, $L$, probability limit, $\rho$) to be specified. In our analysis, threshold level is set at 70000 with target probabilities $\rho = 0.028, 0.029, 0.030$ and $0.031$. The reason of selecting threshold level as 70000 is that the effect of risk sensitive models on expected revenue can be seen more clearly when threshold level is around mean expected revenue. The $\rho$ values are chosen by using the analyses in Figure 5.10. The seat allocations of these models are used in simulation studies with partitioned booking policy. The sample mean and standard deviation of the total revenue for *PMP-RC* model is not so different than the values for the *EMR* model as seen in Table 6.5. This shows us that *PMP-RC* model does not work better than *SLP-RM* model with $L$=70000 and *EMVLP* model with $\Theta = 0.005$. Because of the computational difficulties, the number of demand realizations for *PMP-RC* model is set at 100, which is small as compared to the one, 1000, used in *SLP-RM* model. Moreover, setting both threshold level, $L$, and probability limit, $\rho$, is not practical to use in real life problems. However, the results of the *PMP-RC* model can be improved by increasing the number of demand realizations. Solving this model with higher number of demand realizations in an acceptable computational time remains as a future work.

Table 6.4: $P(R < \tau)$ for the base problem.

| | | | | *EMVLP* ($\theta$) | | *SLP-RM* (*L*x1000) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $\tau$ | *DLP* | *EMR* | *SLP* | 0.001 | 0.005 | 60 | 65 | 70 | 75 | 80 |
| 61000 | 0.0253 | 0.0357 | 0.0290 | 0.0231 | **0.0064** | **0.0107** | 0.0171 | 0.0244 | 0.0269 | 0.0382 |
| 62000 | 0.0326 | 0.0453 | 0.0382 | 0.0319 | **0.0107** | **0.0175** | 0.0242 | 0.0329 | 0.0355 | 0.0493 |
| 63000 | 0.0414 | 0.0594 | 0.0494 | 0.0416 | **0.0195** | **0.0248** | 0.0328 | 0.0422 | 0.0451 | 0.0629 |
| 64000 | 0.0550 | 0.0736 | 0.0656 | 0.0538 | **0.0348** | **0.0346** | 0.0422 | 0.0556 | 0.0602 | 0.0816 |
| 65000 | 0.0695 | 0.0931 | 0.0846 | 0.0696 | 0.0587 | **0.0476** | **0.0569** | 0.0701 | 0.0749 | 0.1047 |
| 66000 | 0.0868 | 0.1173 | 0.1080 | 0.0887 | 0.0902 | **0.0633** | **0.0727** | 0.0886 | 0.0961 | 0.1308 |
| 67000 | 0.1070 | 0.1488 | 0.1350 | 0.1150 | 0.1404 | **0.0872** | **0.0933** | 0.1116 | 0.1209 | 0.1654 |
| 68000 | 0.1325 | 0.1832 | 0.1725 | 0.1459 | 0.2094 | **0.1192** | **0.1204** | 0.1421 | 0.1520 | 0.2003 |
| 69000 | 0.1614 | 0.2179 | 0.2092 | 0.1836 | 0.2962 | **0.1532** | **0.1537** | 0.1782 | 0.1888 | 0.2414 |
| 70000 | **0.1940** | 0.2593 | 0.2511 | 0.2235 | 0.3966 | 0.1961 | **0.1949** | 0.2170 | 0.2252 | 0.2837 |

Table 6.5: *PMP-RC* model, simulation results with partitioned booking policy for the base problem, *L*=70000.

| | | | | *EMVLP* ($\theta$) | | *PMP-RC* ($\rho$) | | | |
|---|---|---|---|---|---|---|---|---|---|
| | *DLP* | *EMR* | *SLP* | 0.001 | 0.005 | 0.28 | 0.29 | 0.30 | 0.31 |
| *SM* | 70,595 | **71,789** | 71,626 | 71,596 | 67,385 | **71,777** | 71,307 | 71,663 | 71,707 |
| *SSD* | 5,604 | 6,273 | 5,942 | 5,631 | **2,939** | 6,238 | **5,479** | 5,992 | 6,352 |
| *SCV* | 0.0794 | 0.0874 | 0.0830 | 0.0787 | **0.0436** | 0.0869 | **0.0768** | 0.0836 | 0.0886 |
| *LF* | 0.8622 | 0.8656 | 0.8784 | 0.8835 | **0.9521** | 0.8670 | **0.8848** | 0.8737 | 0.8632 |

### 6.2.3 *RRS* Procedure

Since we have no seat allocation output of the *RRS* model, the analysis is for bid price control policy. Although we know that updating bid prices frequently improves the solution quality, bid price control policy without update is used firstly. The results are given in Table 6.6. It can be seen from the table that the results are the same for some of the models as in the case of *DLP* and *EMR* models and *RRS* procedure with both $L < 75000$ and $L < 80000$. Also, the results are the same for *RLP* model and *RRS* procedure with both $L > 75000$ and $L > 80000$. Hence, it is not easy to make a conclusion about risk sensitivity.

### 6.2.4 Bayesian Update

Bayesian update is considered in this section for *DLP*, *EMR*, *EMVLP* and *SLP-RM* models and *RRS* procedure. Firstly, the impact of updating demand data, $p$ and $\gamma$ values in Gamma

Table 6.6: *RRS* model, simulation results with bid price control policy for the base problem.

| | | | | | *RRS* | | | |
|---|---|---|---|---|---|---|---|---|
| | *DLP* | *EMR* | *EMVLP* | *RLP* | $L < 75000$ | $L > 75000$ | $L < 80000$ | $L > 80000$ |
| *SM* | 73,460 | 73,460 | 66,750 | 76,629 | 73,460 | 76,629 | 73,460 | 76,629 |
| *SSD* | 4,684 | 4,684 | 3,550 | 5,818 | 4,684 | 5,818 | 4,684 | 5,818 |
| *SCV* | 0.064 | 0.064 | 0.053 | 0.076 | 0.064 | 0.076 | 0.064 | 0.076 |
| *LF* | 0.960 | 0.960 | 0.980 | 0.935 | 0.960 | 0.935 | 0.960 | 0.935 |

Distribution, on revenue and expected revenue is analyzed for *DLP* model by changing the number of updates when bid price control policy is used. The booking horizon consisting of $T$ time units is divided into equal update periods. For example, for two updates, the booking horizon is divided into three periods, each with a length of $\frac{T}{3}$. Updating bid prices improves the solution quality significantly for sample mean as seen in Table 6.7 and Figure 6.3. However, sample standard deviation and coefficient of variation also increase, which is not a desired situation for a risk-sensitive decision maker. On the other hand, it is seen in Figure 6.3 that updating bid prices frequently decreases the probability that revenue is less than the (specified) levels. Moreover, $P(R < \tau)$ decreases also for higher $\tau$ levels, which is a desired situation for the decision makers becasue this means that probability of higher revenue values increases. Although increasing the number of updates improves the revenue while $SSD$ and $SCV$ are staying at acceptable levels, its effect decreases beyond a certain value. For example, the difference between results of 10 and 20 updates is smaller than the difference between results of 2 and 3 updates. Updating frequently has not a significant effect on load factor, which is almost the same for 0, 2, 3 and 5 updates, but increases slightly for 10 and 20 updates.

Table 6.7: *DLP* model, simulation results with bid price control policy and Bayesian update for the base problem.

| | *Number of Updates* | | | | | |
|---|---|---|---|---|---|---|
| | 0 | 2 | 3 | 5 | 10 | 20 |
| $SM$ | 73,460 | 75,266 | 77,153 | 78,235 | 79,322 | 79,929 |
| $SSD$ | 4,684 | 4,899 | 5,499 | 5,719 | 5,859 | 5,951 |
| $SCV$ | 0.064 | 0.0651 | 0.0713 | 0.0731 | 0.0739 | 0.0744 |
| $LF$ | 0.960 | 0.9544 | 0.9553 | 0.9550 | 0.9613 | 0.9640 |

According to the analysis on the number of updates for *DLP* model, the number of updates for

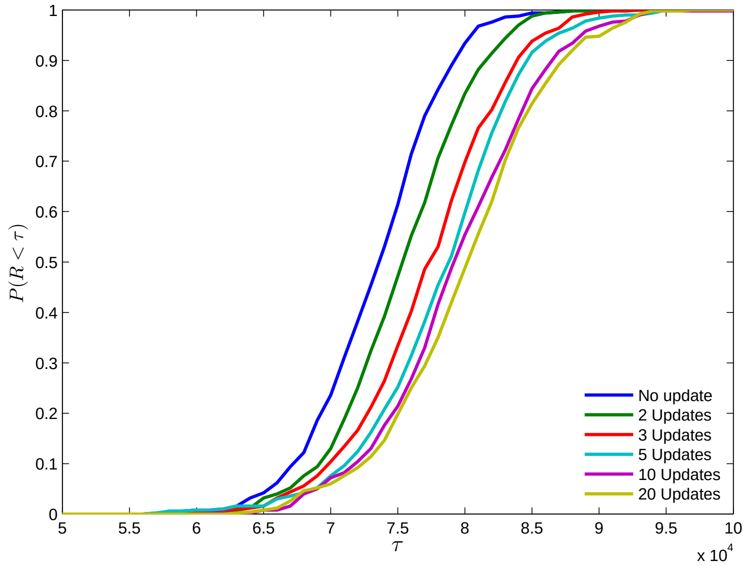


Figure 6.3: $P(R < \tau)$ for *DLP* model with Bayesian update.

the remaining part of the analysis is set at 3. Although it is known that increasing the number of updates improves the solution quality, the computational time limits this number. Three is the number which can be used in an acceptable computational time and gives significantly better results than the situation with no update. The seat allocations and bid prices for *DLP*, *EMR*, *EMVLP* and *SLP-RM* models and *RRS* procedure are used in the simulation model with Bayesian update policy. Because of the computational difficulties, the replication number is decreased to 500. The results for the models are given in Table 6.8 and Figure 6.4. The simulation studies for *RRS* procedure with Bayesian update show us that updating only demand data in *RRS* procedure does not improve the solution quality; moreover, it can even worsen the solutions. This situation occurs because of the requirement to update the policy parameter $L$ in *RRS* procedure. In this procedure, a threshold level, $L$, is set at the beginning of the booking horizon and bid prices are divided into two groups according to this threshold level. In Bayesian update, only demand data is updated, but this update may significantly affect the bid prices that belong to the groups $R < L$ or $R > L$. Hence, in *RRS* procedure, not only demand data but also policy parameter $L$ must be updated to obtain better results. However, a detailed study is required to update policy parameter based on the demand observations. Bayesian update or other update policies can be considered for this purpose, which remains as a future work. The following remarks result from the analysis in this section.

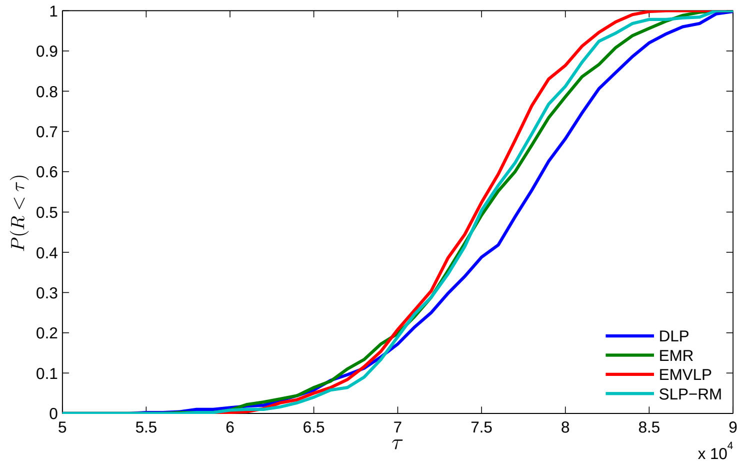


Figure 6.4: $P(R < \tau)$ when Bayesian update is used, 3 updates.

- *DLP* works better than *EMR* model both with nested and bid price control policies when Bayesian update is used.

- Load factor in simulation results with Bayesian update policy is significantly higher than the results without update under bid price control policy. As it is described in Chapter 3, there is no limit on the number that can be accepted for any of the itineraries when the class is open in this policy. Hence, customers with small contibutions to the revenue may be accepted more in this policy, which increases the load factors of the flights.

- Updating frequently decreases the probability that the revenue is less than the (specified) levels, $P(R < \tau)$.

Table 6.8: Simulation results for the base problem with Bayesian update, 3 updates.

| | *Nested Control* | | | | *Bid Price Control* | |
|---|---|---|---|---|---|---|
| | *DLP* | *EMR* | *EMVLP* | *SLP-RM* | *DLP* | *EMR* |
| *SM* | 76,360 | 75,000 | 74,272 | 74,994 | 77,153 | 74,146 |
| *SSD* | 6,674 | 6,089 | 5,049 | 5,423 | 5,499 | 4,421 |
| *SCV* | 0.0874 | 0.0812 | 0.0680 | 0.0723 | 0.0713 | 0.0596 |
| *LF* | 0.9104 | 0.9043 | 0.9311 | 0.9175 | 0.9553 | 0.97 |

## 6.3 Low-Before-High Arrival Pattern

In this section, the effect of change in arrival process on total revenue and risk is analyzed. In the base case, although low-fare class bookings tend to arrive earlier than high-fare class bookings, an important amount of customers of low-fare class arrive later than the arrival of the first high-fare class bookings. In this section, the arrival parameters for Beta distribution, graphed in Figure 5.2, are changed in order to see the effect of using low-before-high arrival pattern more strictly. The new arrival rates are given in Appendix B.1.2 and graphed in Figure 6.5. Because input data is the same as that of the base problem except for the arrival rates, mathematical programming solutions do not change. The simulation results are given in Table 6.10.

Because of the use of nested booking control policy in this scenario, the simulation results are almost the same as the ones for the base problem in Table 6.3. In nested booking policy, the seats that are allocated for low-fare classes can be used by a passenger with high-fare class request. In Figure 6.6 and Table 6.10, $P(R < \tau)$ values for this scenario are given. They are again not significantly different from the results given for the base problem. One of the best two results in Table 6.10 is for the *SLP-RM* model for all of the considered $\tau$ values. The other good result is for $EMVLP$ or for $DLP$ model. $DLP$ gives better results especially for greater $\tau$ values. It can be concluded that arrival patterns of the fare classes do not affect the revenues significantly when nested booking policy is used.

Table 6.9: Simulation results for low-before high arrival pattern with nested booking policy.

| | | | | $EMVLP$ ($\theta$) | | $SLP-RM$ ($L$x1000) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $DLP$ | $EMR$ | $SLP$ | 0.001 | 0.005 | 60 | 65 | 70 | 75 | 80 |
| *SM* | **75,378** | 74,149 | 74,009 | 74,309 | 69,780 | 73,548 | 74,506 | **74,662** | 74,561 | 73,639 |
| *SSD* | 6,499 | 6,928 | 6,455 | 6,135 | **3,029** | **4,818** | 5,662 | 6,329 | 6,503 | 6,919 |
| *SCV* | 0.0862 | 0.0934 | 0.0872 | 0.0826 | **0.0434** | **0.0655** | 0.0760 | 0.0848 | 0.0872 | 0.0940 |
| *LF* | 0.8985 | 0.8777 | 0.8911 | 0.9003 | **0.9643** | **0.9319** | 0.9134 | 0.8979 | 0.8923 | 0.8749 |

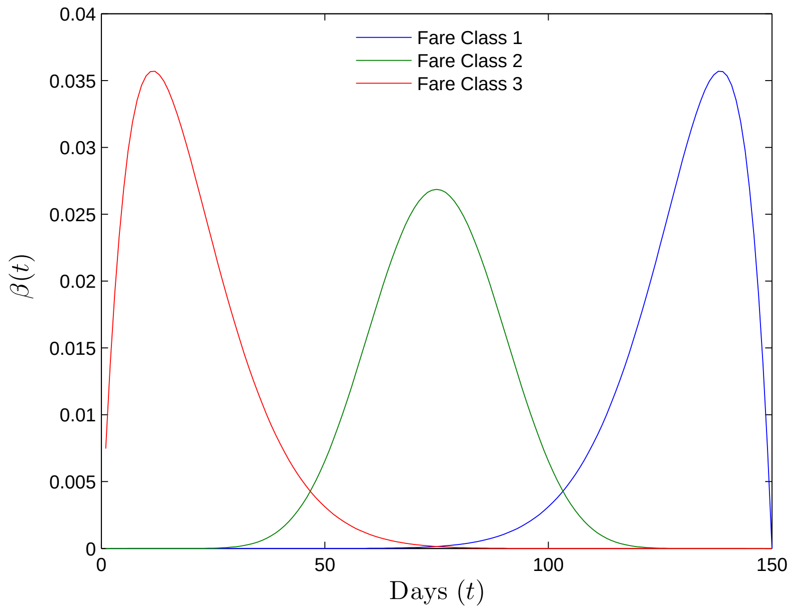


Figure 6.5: Low-before-high arrival rates

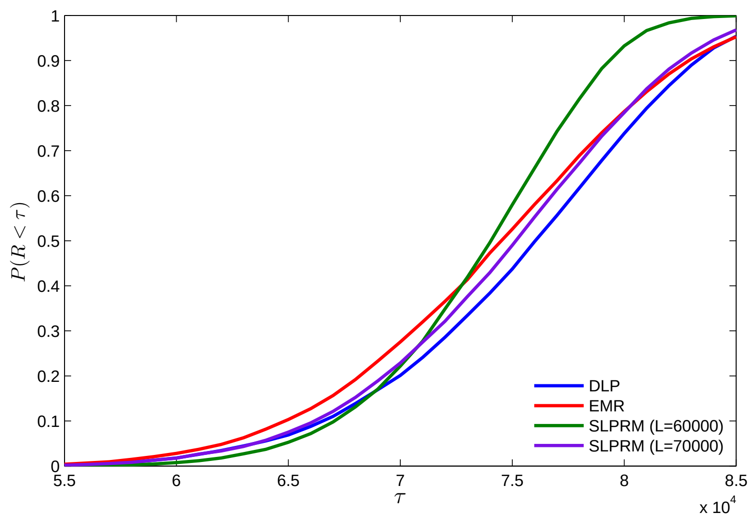


Figure 6.6: $P(R < \tau)$ for the case with low-before-high arrival rates.

Table 6.10: $P(R < \tau)$ for the case with low-before-high arrival rates.

| | | | | *EMVLP* ($\theta$) | | *SLP* − *RM* (*L*x1000) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $\tau$ | *DLP* | *EMR* | *SLP* | 0.001 | 0.005 | 60 | 65 | 70 | 75 | 80 |
| 60000 | 0.0180 | 0.0282 | 0.0217 | 0.0167 | **0.0030** | **0.0076** | 0.0124 | 0.0177 | 0.0203 | 0.0301 |
| 61000 | 0.0259 | 0.0372 | 0.0303 | 0.0236 | **0.0067** | **0.0119** | 0.0171 | 0.0265 | 0.0278 | 0.0392 |
| 62000 | 0.0344 | 0.0481 | 0.0389 | 0.0318 | **0.0125** | **0.0180** | 0.0261 | 0.0335 | 0.0369 | 0.0534 |
| 63000 | 0.0446 | 0.0628 | 0.0536 | 0.0428 | **0.0244** | **0.0271** | 0.0336 | 0.0436 | 0.0473 | 0.0707 |
| 64000 | 0.0559 | 0.0819 | 0.0701 | 0.0576 | **0.0435** | **0.0373** | 0.0446 | 0.0572 | 0.0630 | 0.0896 |
| 65000 | 0.0689 | 0.1034 | 0.0908 | 0.0757 | 0.0726 | **0.0529** | **0.0585** | 0.0754 | 0.0828 | 0.1137 |
| 66000 | 0.0881 | 0.1275 | 0.1153 | 0.0999 | 0.1157 | **0.0722** | **0.0803** | 0.0957 | 0.1031 | 0.1399 |
| 67000 | 0.1104 | 0.1566 | 0.1439 | 0.1234 | 0.1805 | **0.0978** | **0.1014** | 0.1215 | 0.1311 | 0.1732 |
| 68000 | 0.1389 | 0.1920 | 0.1805 | 0.1563 | 0.2616 | **0.1312** | **0.1308** | 0.1523 | 0.1605 | 0.2129 |
| 69000 | **0.1695** | 0.2331 | 0.2207 | 0.1937 | 0.3616 | 0.1708 | **0.1658** | 0.1890 | 0.1995 | 0.2554 |
| 70000 | **0.2012** | 0.2754 | 0.2685 | 0.2391 | 0.4776 | 0.2220 | **0.2078** | 0.2289 | 0.2405 | 0.3005 |

## 6.4 Increased Low-Fare Demand Variance

In the base problem, demand variance for high-fare class (class 1) is relatively higher than the demand variances for low-fare classes (class 2 and 3). de Boer (1999) considers the case of increased low-fare demand variance by changing these variances according to the study of Belobaba (1987), in which he states that 0.33 is a common coefficient of variation for demand in airline industry. This value is found by working with the real data of Western Airlines. In the base problem given in Section 6.2, the coefficient of variations are between 0.18 and 0.75. In this section, only the variances of the low-fare demands are increased which have coefficient of variations smaller than 0.22 in base problem. The data with increased low-fare demand variances are given in Appendix B.1.3. Increasing low-fare demand variance results in an increase in deviations from the mean for low-fare demands. Therefore, we expect that there must be a decrease in the number of seats allocated for low-fare classes. The optimal seat allocations of the models used in the simulation studies are given in Appendix B.11. As it is expected, there is a slight decrease in seat allocations for low-fare classes. These allocations are used in the simulation models for nested booking policy, the results are given in Table 6.11. The decrease in low-fare class seat allocations results in slight decreases in revenue and load factors. Moreover, an increase in sample standard deviation is seen because of the increase in variances. According to Figure 6.7 and Table 6.12, *SLP-RM* shows similar behavior as in the base problem. One of the best two results always is for the *SLP-RM* model. However, the results are not significantly different for *DLP*, *EMVLP* and *SLP-RM* models.

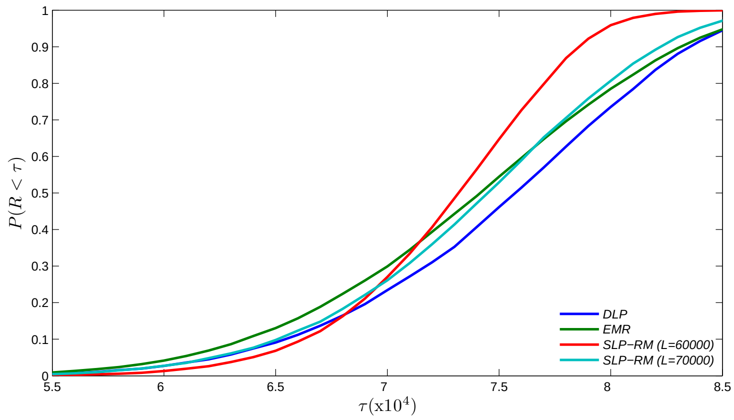


Figure 6.7: $P(R < \tau)$ for the case with increased low-fare demand variance.

Thus, a strong conclusion about comparison on the performances of these three models cannot be considered in this low-fare demand variance scenario.

Table 6.11: Simulation results for increased low-fare demand variance case with nested booking policy.

| | | | | *EMVLP* ($\theta$) | | *SLP-RM* ($L$x1000) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *DLP* | *EMR* | *SLP* | 0.001 | 0.005 | 60 | 65 | 70 | 75 | 80 |
| *SM* | **75,046** | 73,722 | 73,701 | 74,326 | 70,030 | 72,726 | **74,351** | 73,997 | 73,707 | 73,589 |
| *SSD* | 7,028 | 7,468 | 6,857 | 6,773 | **3,452** | **4,884** | 6,618 | 6,614 | 7,433 | 7,666 |
| *SCV* | 0.0936 | 0.1013 | 0.0930 | 0.0911 | **0.0493** | **0.0672** | 0.0890 | 0.0894 | 0.1008 | 0.1042 |
| *LF* | 0.8851 | 0.8614 | 0.8789 | 0.8889 | **0.9575** | **0.9292** | 0.8907 | 0.8879 | 0.8620 | 0.8541 |

Table 6.12: $P(R < \tau)$ for the case with low-before-high arrival rates.

| | | | | *EMVLP* ($\theta$) | | *SLP-RM* (*L*x1000) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $\tau$ | *DLP* | *EMR* | *SLP* | 0.001 | 0.005 | 60 | 65 | 70 | 75 | 80 |
| 60000 | 0.0269 | 0.0415 | 0.0320 | 0.0266 | **0.0072** | **0.0130** | 0.0247 | 0.0272 | 0.0424 | 0.0471 |
| 61000 | 0.0360 | 0.0540 | 0.0424 | 0.0358 | **0.0113** | **0.0193** | 0.0328 | 0.0354 | 0.0545 | 0.0603 |
| 62000 | 0.0450 | 0.0690 | 0.0539 | 0.0470 | **0.0205** | **0.0262** | 0.0439 | 0.0480 | 0.0685 | 0.0753 |
| 63000 | 0.0582 | 0.0864 | 0.0693 | 0.0591 | **0.0327** | **0.0377** | 0.0568 | 0.0612 | 0.0869 | 0.0963 |
| 64000 | 0.0746 | 0.1088 | 0.0895 | 0.0764 | **0.0528** | **0.0509** | 0.0722 | 0.0766 | 0.1072 | 0.1173 |
| 65000 | 0.0909 | 0.1305 | 0.1118 | 0.0968 | **0.0828** | **0.0684** | 0.0907 | 0.0979 | 0.1311 | 0.1401 |
| 66000 | **0.1120** | 0.1572 | 0.1370 | 0.1183 | 0.1258 | **0.0935** | 0.1137 | 0.1236 | 0.1573 | 0.1694 |
| 67000 | **0.1371** | 0.1887 | 0.1670 | 0.1473 | 0.1840 | **0.1220** | 0.1397 | 0.1482 | 0.1888 | 0.2013 |
| 68000 | **0.1646** | 0.2241 | 0.2040 | 0.1784 | 0.2579 | **0.1626** | 0.1706 | 0.1830 | 0.2228 | 0.2347 |
| 69000 | **0.1962** | 0.2607 | 0.2442 | 0.2127 | 0.3475 | 0.2113 | **0.2076** | 0.2211 | 0.2570 | 0.2690 |
| 70000 | **0.2343** | 0.2992 | 0.2854 | 0.2535 | 0.4598 | 0.2704 | **0.2457** | 0.2613 | 0.2968 | 0.3093 |

## 6.5 Smaller Differences Between Fares

de Boer (1999) states that decreasing differences between fare classes improves the results of the stochastic solutions (*EMR*-based models) as compared to the deterministic ones (*DLP* model). In this section, this situation is taken into consideration with smaller differences between fares. The input data for these cases are given in Appendix B.1.4. Decreasing differences between fares by decreasing the high-fares affects the seat allocations for high-fare classes in a negative way, the mathematical model results are given in Appendix B.12. In other words, it is now less attractive to allocate a seat to the high-fare class with decreased fares. This situation causes a reduction not only in revenue but also in standard deviation. More seats are allocated to low-fare classes which have small variances of demand compared to the high-fare classes. The results are summarized in Table 6.13 and Table 6.14. In base problem, because of the higher differences between fares, risk-sensitive models results in a considerable amount of decrease in revenue as compared to risk-neutral models. This is because seat allocations obtained by the risk-sensitive models are to decrease variability unlike risk-neutral models. Therefore, the risk can only be decreased by decreasing the revenue significantly. However, in the problem scenario with small differences between fares, changing seat allocations affects revenue less, but improves the solution quality significantly in terms of variability. Thus, the results for the proposed models in this section in Figure 6.8 and Table 6.14 are better than the results of other models as compared to the results in Sections 6.2, 6.3 and 6.4.

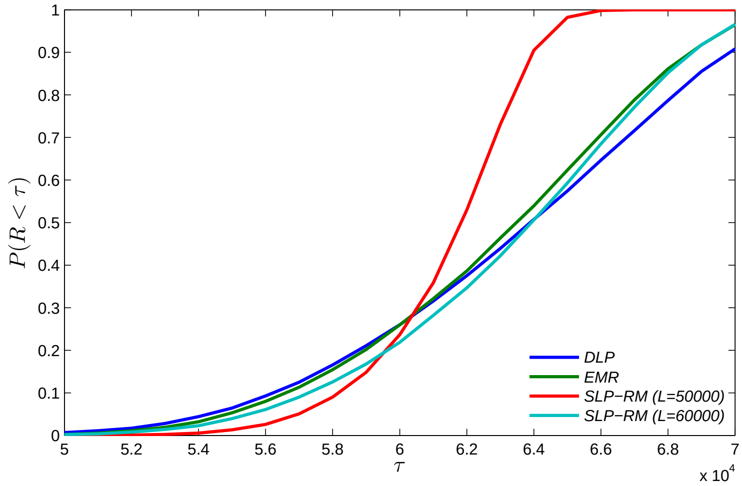


Figure 6.8: $P(R < \tau)$ for the case with smaller differences between fares.

Table 6.13: Simulation results for the case with smaller differences between fares, nested booking policy.

| | | | | *EMVLP* ($\theta$) | | *SLP-RM* (*L*x1000) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *DLP* | *EMR* | *SLP* | 0.001 | 0.005 | 50 | 55 | 60 | 65 | 70 |
| *SM* | **63,418** | 63,007 | 62,697 | 62,791 | 61,010 | 61,417 | 62,896 | **63,424** | 62,681 | 62,697 |
| *SSD* | 5,162 | 4,550 | 3,844 | 3,798 | **2,090** | **2,302** | 3,306 | 4,336 | 4,978 | 3,844 |
| *SCV* | 0.0814 | 0.0722 | 0.0613 | 0.0605 | **0.0343** | **0.0375** | 0.0526 | 0.0684 | 0.0794 | 0.0613 |
| *LF* | 0.9011 | 0.9162 | 0.9359 | 0.9365 | **0.9710** | **0.9664** | 0.9478 | 0.9227 | 0.9021 | 0.9359 |

Table 6.14: $P(R < \tau)$ for the case with smaller differences between fares, nested booking policy.

| | | | | *EMVLP* ($\theta$) | | *SLP-RM* ($L$x1000) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $\tau$ | *DLP* | *EMR* | *SLP* | 0.001 | 0.005 | 50 | 55 | 60 | 65 | 70 |
| 55000 | 0.0642 | 0.0531 | 0.0368 | 0.0336 | **0.0100** | **0.0133** | 0.0212 | 0.0395 | 0.0758 | 0.0368 |
| 56000 | 0.0927 | 0.0802 | 0.0607 | 0.0568 | **0.0235** | **0.0259** | 0.0354 | 0.0608 | 0.1075 | 0.0607 |
| 57000 | 0.1251 | 0.1131 | 0.0914 | 0.0865 | **0.0496** | **0.0505** | 0.0589 | 0.0899 | 0.1445 | 0.0914 |
| 58000 | 0.1656 | 0.1545 | 0.1317 | 0.1242 | 0.0936 | **0.0900** | **0.0920** | 0.1257 | 0.1907 | 0.1317 |
| 59000 | 0.2106 | 0.2019 | 0.1791 | 0.1721 | 0.1673 | **0.1483** | **0.1327** | 0.1677 | 0.2431 | 0.1791 |
| 60000 | 0.2600 | 0.2598 | 0.2417 | 0.2311 | 0.2759 | 0.2369 | **0.1911** | **0.2189** | 0.3004 | 0.2417 |
| 61000 | 0.3156 | 0.3212 | 0.3083 | 0.2983 | 0.4267 | 0.3585 | **0.2595** | **0.2819** | 0.3603 | 0.3083 |
| 62000 | 0.3752 | 0.3863 | 0.3883 | 0.3769 | 0.6250 | 0.5296 | **0.3477** | **0.3467** | 0.4278 | 0.3883 |
| 63000 | **0.4392** | 0.4638 | 0.4817 | 0.4713 | 0.8455 | 0.7306 | 0.4518 | **0.4221** | 0.4952 | 0.4817 |
| 64000 | **0.5073** | 0.5392 | 0.5783 | 0.5699 | 0.9708 | 0.9047 | 0.5723 | **0.5063** | 0.5673 | 0.5783 |
| 65000 | **0.5744** | 0.6232 | 0.6794 | 0.6752 | 0.9965 | 0.9822 | 0.7044 | **0.5930** | 0.6399 | 0.6794 |

## 6.6 Realistic Variations and Close Fares

In this section, the case with realistic coefficients of variation of demand, relatively close fares and no specific order of arrivals is analyzed. This case is a combination of the cases in Sections 6.4 and 6.5. As it is seen in Section 6.4, decreasing differences between fares improves the solution quality of risk averse models in terms of sample coefficient of variation, load factor and probability of poor performance. In this section, in addition to the lowered difference between fare classes, the coefficient of variation for demand is set almost at 0.33 which is considered as the realistic case by Belobaba (1987) and the arrivals are assumed to be the same for all fare classes. de Boer (1999) states that no specific order of arrivals favors the occurence of nesting. This scenario decreases the differences between fares, which is a realistic case in small sized airlines. In small sized airlines, the number of options is less than the regular cases and arrival times of the requests are not so different. As it is stated before, minimizing risk is more important for this type of airline companies because of the scale of the network. According to the results given in Table 6.15, Figure 6.9 and Table 6.16, the risk-sensitive models improve the results significantly. Both *EMVLP* and *SLP-RM* models give almost the same results and it is not easy to make a comparison between them. However, note that the solutions of these two models depend on the parameters selected. This scenario shows us that risk sensitive models work well for small scale airline networks that have the properties of relatively close fares and no specific order of arrivals.

Table 6.15: Simulation results for the case with smaller differences between fares, nested booking policy

| | | | | $EMVLP$ ($\theta$) | | $SLP-RM$ ($L$x1000) | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $DLP$ | $EMR$ | $SLP$ | 0.001 | 0.005 | 60 | 65 | 70 | 75 |
| $SM$ | 70,994 | 70,894 | 71,210 | 71,272 | **71,429** | **71,449** | 71,225 | 70,452 | 71,210 |
| $SSD$ | 3,995 | 3,552 | 2,947 | 3,099 | **2,416** | **2,729** | 3,084 | 3,821 | 2,947 |
| $SCV$ | 0.0563 | 0.0501 | 0.0414 | 0.0435 | **0.0338** | **0.0382** | 0.0433 | 0.0542 | 0.0414 |
| $LF$ | 0.9110 | 0.9235 | 0.9458 | 0.9407 | **0.9613** | **0.9524** | 0.9412 | 0.9102 | 0.9458 |

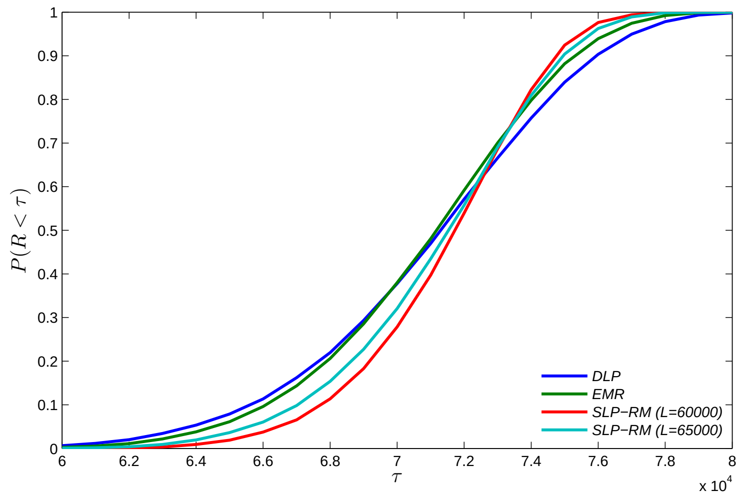


Figure 6.9: $P(R < \tau)$ for the case with smaller differences between fares.

Table 6.16: $P(R < \tau)$ for the case with smaller differences between fares.

| | | | | $EMVLP$ ($\theta$) | | $SLP-RM$ ($L$x1000) | | |
|---|---|---|---|---|---|---|---|---|
| | $DLP$ | $EMR$ | $SLP$ | 0.001 | 0.005 | 60 | 65 | 70 |
| 60000 | 0.0064 | 0.0023 | 0.0005 | 0.0007 | **0.0001** | **0.0003** | 0.0008 | 0.0056 |
| 61000 | 0.0117 | 0.0056 | 0.0011 | 0.0016 | **0.0002** | **0.0005** | 0.0015 | 0.0115 |
| 62000 | 0.0200 | 0.0110 | 0.0020 | 0.0039 | **0.0005** | **0.0014** | 0.0039 | 0.0218 |
| 63000 | 0.0344 | 0.0217 | 0.0064 | 0.0088 | **0.0018** | **0.0031** | 0.0090 | 0.0352 |
| 64000 | 0.0533 | 0.0380 | 0.0154 | 0.0189 | **0.0039** | **0.0091** | 0.0196 | 0.0579 |
| 65000 | 0.0790 | 0.0612 | 0.0295 | 0.0357 | **0.0106** | **0.0190** | 0.0364 | 0.0894 |
| 66000 | 0.1133 | 0.0960 | 0.0521 | 0.0591 | **0.0241** | **0.0374** | 0.0605 | 0.1300 |
| 67000 | 0.1621 | 0.1433 | 0.0901 | 0.0973 | **0.0472** | **0.0654** | 0.0984 | 0.1860 |
| 68000 | 0.2197 | 0.2062 | 0.1451 | 0.1516 | **0.0886** | **0.1138** | 0.1538 | 0.2534 |
| 69000 | 0.2933 | 0.2854 | 0.2221 | 0.2249 | **0.1550** | **0.1828** | 0.2273 | 0.3350 |
| 70000 | 0.3779 | 0.3799 | 0.3201 | 0.3168 | **0.2612** | **0.2784** | 0.3205 | 0.4281 |

## 6.7 Concluding Remarks

The numerical results reported in this chapter show that the performances of the mathematical models are directly related with the scenario considered. Although deterministic model, *DLP*, works surprisingly well for most of the cases, in the last two cases in Sections 6.5 and 6.6, the stochastic models *EMR*, *SLP*, *EMVLP* and *SLP-RM* give better results. The risk-averse models generally improve the solution qualities in terms of probability of poor performances in all of the cases and the *SLP-RM* model proposed in this thesis gives almost the same results as the *EMVLP* model due to Çetiner (2007). However, *SLP-RM* model simplifies the parameter selection in decision making process as compared to *EMVLP* model in real life problems. The conclusions related with risk sensitivity according to the analyses in Sections 6.1 to 6.5 are given below:

- Decreasing risks causes an increase in load factors.
- Although risk-sensitive models generally give less revenue and variability, simulation results of these models can give higher revenues than risk-neutral models.
- The performance of risk-sensitive models are better for small sized airline companies that have small differences among fares.

# CHAPTER 7

# CONCLUSION

In this study, network seat inventory problems without overbooking, cancellations and no-shows are taken into consideration from the perspective of the risk-sensitive decision maker. The main aim of the study is incorporating risk factors into the proposed models in order to guide risk-sensitive decision makers in developing their own policies for accepting or rejecting the booking requests. This risk-sensitive approach is the difference of this study from most of the studies in the literature.

Three probabilistic models and procedures are introduced in this study, which are called *SLP-RM*, *PMP-RC* and *RRS*. In *SLP-RM* procedure, which is composed of two linear programming models. The probability that the revenue is less than a specified threshold level is minimized in the first model and then the expected revenue is maximized in the second model by using the output of the first model. In *PMP-RC* model, the expected revenue is maximized with an additional constraint on the probability that the revenue is less than a threshold level. This model is an integer programming formulation. In *RRS* procedure, different than *SLP-RM* procedure and *PMP-RC* model, the output of the procedure is used only in the bid price control policy. In this procedure, revenue is maximized by solving the *RLP* model, that is proposed by Talluri and van Ryzin (1999), many times for different realizations of demand. Then, the bid price for an itinerary is approximated with a pessimistic (optimistic) or risk-sensitive (risk-taking) perspective by taking the average of the bid prices for the instances with revenue less (or more) than the threshold level. *RRS* procedure is simple to use and decreases the computational time as compared to *SLP-RM* procedure and *PMP-RC* model.

The booking limits and bid prices, taken from the proposed models, are used in the simulation studies in order to compare the proposed models with the existing models in the literature.

Different scenarios are considered and the performances of the proposed models are evaluated for these scenarios. Our risk-sensitive approach decreases the variability of the revenue. The risk-sensitivity of the decision maker is evaluated by changing the threshold level, which is an easily determined and controlled parameter. The proposed models and procedures do not only improve our solutions in terms of risk-sensitivity but also give better expected revenues in some problem scenarios than the well-known models in the literature, *DLP* and *EMR*. Especially for the scenario with close fares and no specific order of arrivals, which is more realistic for small-scaled airlines, the results of the proposed approach outperform the results of *DLP* and *EMR* models significantly.

The following remains as future work. Since *PMP-RC* model is an integer programming formulation, the computational complexity restricts the number of demand realizations in this model. Moreover, the approximations in *SLP-RM* procedure worsens the quality of the results in terms of expected revenue and $P(R < \tau)$. Some other linear programming techniques or heuristic search methods may be used in order to improve the solution quality of the models.

The *RRS* procedure cannot be effectively used because of the requirement to update the policy parameter (threshold level) based on the demand observations. A Bayesian update mechanism can be developed for *RRS* procedure. Furthermore, this update mechanism may be adapted in order to use it also in the other models, *SLP-RM* and *PMP-RC*.

The simulation studies show us that there is strong correlation between the load factor and variability of the revenue. The load factor increases significantly in risk-sensitive models and procedures. Therefore, using load factor in the models in order to manage the risk can be considered as a future research direction.

In this study, there are many assumptions: overbooking and cancellations are excluded, demands for different fare classes are assumed to be independent, batch booking is not allowed. The relaxation of these assumptions for risk-sensitive models is another important research direction.

# REFERENCES


[1] Aviv, Y. and Pazgal, A., 2005. A partially observed Markov decision process for dynamic pricing. *Management Science* 51, 1400-1416.

[2] Barz, C. and Waldmann, K.H., 2007. Risk-sensitive capacity control in revenue management. *Mathematical Methods in Operations Research* 5, 43-51.

[3] Beckman, M.J. and Bobkoski, F., 1958. Airline demand: An analysis of some frequency distributions. *Naval Research Logistics Quarterly* 5, 43-51.

[4] Belobaba, P.P., 1987. Air travel demand and airline seat inventory management. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA.

[5] Bitran, G. and Candeltey, R., 1993. An overview of pricing models for revenue management. *Manufacturing & Service Operations Management* 5, 203-230.

[6] Bradley, S.P., Hax, A.C. and Magnanti, T.L., 1977. *Applied Mathematical Programming*. Addison-Wesley Publishing Company, Inc.

[7] Brumelle, S.L. and McGill, J.I., 1993. Airline seat allocation with multiple nested fare classes. *Operations Research* 41, 127-137.

[8] Buhr, J., 1982. Optimal Sales Limits for 2-Sector Flights. *AGIFORS Symposium Proceedings* 22, 291-303.

[9] Chen, X., Sim, M., Simchi-Levi, D. and Sun., P., 2007. Risk aversion in inventory management. *Operations Research* 55, 828-842.

[10] Chiang, w.C., Chen, J., and Xu, X. 2006. An overview of research on revenue management: current issues and future research. *International Journal of Revenue Management* 1, 97-128.

[11] Çetiner, D., 2007. A risk-sensitive approach for airline revenue management problems. Master's Thesis, The Graduate School of Natural and Applied Sciences of Middle East Technical University, Turkey.

[12] Curry, R.E., 1990. Optimal airline seat allocation with fare classes nested by origins and destinations. *Transportation Sci.* 24, 193-204.

[13] de Boer, S.V., 1999. Mathematical Programming in Airline Seat Inventory Control. Master's Thesis, Vrije Universiteit Amsterdam, The Netherlands.

[14] de Boer, S.V., Freling, R. and Piersma, N., 2002. Mathematical programming for network revenue management revisited. *European Journal of Operations Research* 137, 72-92.

[15] DeMiguel, V. and Mishra, N., 2006. A multistage stochastic programming approach to network revenue management. Working paper, London Business School, 2006.

[16] Elmaghraby, W. and Keskinocak, P., 2003. Dynamic pricing in the presence of inventory considerations: research overview, current practices, and future directions. *Management Science* 49, 1287-1309.

[17] Feng, Y., B. Xiao., 1999. Maximizing revenues of perishable assets with a risk factor. *Operations Research* 47, 337-341.

[18] Feng, Y. and Gallego, G., 2000. Perishable asset revenue management with Markovian time dependent demand intensities. *Management Science* 46, 941-956.

[19] Gallego, G., van Ryzin, F., 1994. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science* 40, 999-1020.

[20] Gallego, G., van Ryzin, F., 1997. A multi-product dynamic pricing problem and its applications to network yield management. *Operations Research* 45, 24-41.

[21] Glover, F., Glover, R., Lorenzo, J. and McMillan, C., 1982. The passenger mix problem in the scheduled airlines. *Interfaces* 12, 73-79.

[22] Gosavi, A., Bandla, N. and Das, T.K., 2002. A reinforcement learning approach to a single leg airline revenue management problem with multiple fare classes and overbooking. *IIE Transactions* 34, 729-742.

[23] Karaesmen, I. and van Ryzin, G., 2004. Overbooking with substitutable inventory classes. *Operations Research* 52, 83-104.

[24] Lancaster, J., 2003. The financial risk of airline revenue management. *Journal of Revenue and Pricing Management* 2, 158-165.

[25] Lautenbacher, C.J. and Stidham, S.J., 1999. The underlying Markov decision process in the single-leg airline yield management problem. *Transportation Science* 33, 136-146.

[26] Law, A.M. and Kelton, W.D., 2000. *Simulation modelling and analysis*. Boston: McGraw-Hill.

[27] Lee, T.C. and Hersh, M., 1993. A model for dynamic airline seat allocation control with multiple seat bookings. *Transportation Science* 27, 252-265.

[28] Levin, Y., McGill, J., and Nediak, M., 2008. Risk in revenue management and dynamic pricing. *Operations Research* 56, 326-343.

[29] Littlewood, K., 1972. Forecasting and control of passenger bookings. *AGIFORS Sympos. Proc.* 12, 95-117.

[30] Mayer, M., 1976. Seat allocation, or a simple model of seat allocation via sophisticated ones. *AGIFORS Symposium Proc.* 16.

[31] McGill, J.I. and van Ryzin, G.J., 1999. Revenue management: Research overview and prospects. *Transportation Science* 33, 233-256.

[32] Nelson, P.R. and Coffin, M., Copeland, K.A.F., 2003. *Introductory statistics for engineering experimentation.* Academic Press. 233-256.

[33] Pak, K. and Piersma, N., 2002. Overview of OR techniques for airline revenue management. *Statistica Neerlandica.* 56, 480-496.

[34] Popovic, J.B., 1987. Decision making on stock levels in cases of uncertain demand rate. *European Journal of Operational Research.* 32, 276-290.

[35] Ringbom, S. and Shy, O., 2002. The 'adjustable-curtain' strategy: overbooking of multiclass service. *Journal of Economics* 77, 73-90.

[36] Robinson, L.W., 1995. Optimal and Approximate Control Policies for Airline Booking with Sequential Nonmonotonic Fare Classes. *Operations Research* 43, 252-263.

[37] Shlifer, R. and Vardi, Y., 1975. An Airline Overbooking Policy. *Transportation Science* 9, 101-114.

[38] Simpson, R.W., 1989. Using network flow techniques to find shadow prices for market and seat inventory control, MIT Flight Transportation Laboratory Memorandum M 89-1, Cambridge, Massachusetts.

[39] Subramanian, J., Stidham, S.J. and Lautenbacher C.J., 1999 Airline yield management with overbooking, cancellations and no-shows. *Transportation Science* 33, 147-167.

[40] Talluri, K.T. and van Ryzin, G.J., 1998. An analysis of bid-price controls for network revenue management. *Management Science* 44, 1577-1593.

[41] Talluri, K.T. and van Ryzin, G.J., 1999. A randomized linear programming method for computing network bid prices. *Transportation Science* 33, 207-216

[42] Talluri, K.T. and van Ryzin, G.J., 2005. *The theory and practice of revenue management.* New York: Springer.

[43] Weatherford, L.R., 1991. Perishable asset revenue management in general business situations, Ph.D. thesis, Darden Graduate School of Business Administration, University of Virginia, Charlottesville, VA.

[44] Weatherford L.R., 2004. EMSR versus EMSU: revenue or utility?. *Journal of Revenue and Pricing Management* 3, 277-284.

[45] Williamson, E.L., 1992. Airline network seat control: Methodologies and revenue impacts. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

[46] Wollmer, R.D., 1986. A hub-spoke seat management model. Unpublished Internal Report, Mc Donnell Douglas Corporation, Long Beach, CA.

[47] Wollmer, R.D., 1992. An airline seat management model for a single leg route when lower fare classes book first. *Operations Research* 40, 26-37

# APPENDIX A

# REVENUE MANAGEMENT GLOSSARY

# (due to McGill and van Ryzin, 1999)

**Aggregation of demand:** The level of summarization of passenger demand data. The trend has been toward increasing levels of disaggregation in seat inventory optimization; however, pricing, forecasting, and booking control processes often operate at different levels of aggregation. Possible dimensions for disaggregation include: market, season, month, week, section of week (e.g., midweek versus weekend), day of week, time of day, flight number, booking class, fare, flight leg, segment, and itinerary.

**Arrival pattern:** The pattern of arrivals of booking requests. In the airline context, some possible arrival patterns are: sequential booking classes, low before-high fares, or interspersed arrivals. Batch booking: (also multiple booking, or bulk arrival) A booking request that arrives through normal reservation channels for two or more seats to be booked for the same itinerary. Contrast with group bookings.

**Bid price:** A net value (bid-price) for an incremental seat on a particular flight leg in the airline network. Also referred to as minimum acceptable fare, hurdle price, probabilistic shadow price, displacement cost, or probabilistic dual cost.

**Bid price control:** A method of network seat inventory control that assesses the value of an ODF itinerary as the sum of the bid-prices assigned to individual legs in the itinerary. Typically, an ODF request is accepted if its fare exceeds the total bid prices. Also referred to as continuous nesting.

**Booking class:** A category of bookings that share common features (e.g., similar revenue values or restrictions) and are controlled as one class. This term is often used interchangeably with fare class or bucket.

**Booking limit:** The maximum number of seats that can be sold to a particular booking class. In nested booking systems, booking limits apply to the total number of seats sold to a particular booking class and any lower fare booking classes.

**Booking policy:** A booking policy is a set of rules that specify at any point during the booking process whether a booking class should be open. In general, such policies may depend on the pattern of prior demands or be randomized in some manner and must be generated dynamically as the booking process unfolds for each flight. In some circumstances, optimal or approximately optimal booking policies can be defined by a set of fixed protection levels or threshold curves.

**Buckets:** This term is used in two related ways. First, in older reservations systems, seats for different fare classes or groups of classes are pre-assigned to distinct buckets. These seats are available exclusively to bookings in that fare class. This method simplifies reservations control but is clearly undesirable from a revenue standpoint because seats could fly empty in a discount bucket even if there is higher fare demand available to fill them. Second, buckets also refer to clusters of different fare classes or ODFs that are grouped together for control purposes in a virtual nesting system. A single booking limit is set for all classes in the bucket or lower value buckets.

**Bulk arrival:** See batch booking.

**Bumping:** See denied boarding.

**Cabin:** The physical compartment of an aircraft containing a particular type of seating. For example, an aircraft may be equipped with a first class cabin and a coach cabin, each with different seating and separated by a partition. Multiple fare classes are usually available in each cabin of the aircraft.

**Cancellations:** Returns or changes in bookings that occur early enough in the booking period to permit subsequent rebooking through the reservations system.

**Coefficient of variation:** The standard deviation expressed as a proportion of the mean of a probability or relative frequency distribution. Thus a demand distribution with mean demand 100 and standard deviation 40 would exhibit a coefficient of variation of 0.40. Airline demand data typically display coefficients of variation in the range 0.25 to over 1.0, depending on the level of aggregation of the data.

**Connectivity** (in reservations systems): The degree to which the elements of the reservations system are electronically interconnected. See seamless availability.

**Continuous nesting:** (see bid-price control)

**Controllable booking classes:** All early reservations systems and many existing systems offer only a small number of distinct booking categories (five to ten) that can actually be controlled at booking outlets. Thus, regardless of the number of booking classes or distinct passenger itineraries that can be handled by the revenue management optimization process, the controls in such systems can only be applied to a small number of aggregate booking classes or buckets.

**Control limit policy:** A structural solution that specifies an upper bound (limit) on the number of seats sold in each fare class (or collection of fare classes) for each time before flight departure.

**CRS**: Computer reservations system.

**Defections:** It can occur that a confirmed passenger who shows up for a flight switches to a flight with another airline (usually because of a delay in the original flight departure). Defections constitute a relatively small component of lost passengers and are normally counted as part of no-shows. However, they are distinct from no-shows, and any attempt to predict their occurrence requires an estimation of the probability distribution for departure delays.

**Demand distribution:** An assignment of probabilities (probability distribution) to each possible level of demand for a flight or booking class. A preliminary estimate of such a demand distribution can be obtained by calculating the proportion of each demand level seen on comparable past flights; i.e., a relative frequency distribution.

**Demand factor:** The ratio of demand over capacity for a flight or booking class. (Contrast with load factor.)

**Denied boarding:** Turning away ticketed passengers when more passengers show-up at flight time than there are seats available on the flight, usually as a result of overbooking practices.

Denied boarding can be either voluntary, when passengers accept compensation for waiting for a later flight, or involuntary, when an insufficient number of passengers agree to accept compensation. In the latter case, the airline will be required to provide compensation in a form mandated by civil aviation law.

**Disaggregate:** See aggregation of demand. Displacement cost: In revenue management, the displacement (or opportunity) cost of a booking includes all future revenues that may be lost if the booking is accepted. Taken to the extreme, these include the revenue value of potential displaced future bookings anywhere in the airline network and goodwill costs from those displacements. Assessment of the costs and probabilities of such displacements should allow for the dynamics of cancellations and overbooking and the expected costs of oversold conditions.

**Diversion:** The booking of a customer at a fare level lower than one they would have been prepared to pay. This occurs, for example, when a business traveler has sufficient advance notice of a trip to book in a discount class intended primarily for leisure travelers. Restrictions are designed to inhibit such diversion.

**Dual prices (also shadow prices):** The marginal value of one additional unit of a constrained resource, as determined by a mathematical programming solution to an optimization model. Dual prices are one source of the marginal seat values used in bid-price control.

**Dynamic models:** Models that take into account future possible booking decisions in assessing current decisions. Most revenue management problems are properly modeled as dynamic programming problems.

**Expected marginal seat revenue (EMSR):** The expected revenue of an incremental seat if held open. This is a similar concept to that of bid-price but generally used in a simpler context. Expected revenue: The statistical expected revenue; that is, the sum of possible revenue values weighted by their probabilities of occurrence. Fare basis code: An alphanumeric encryption of the conditions and restrictions associated with a given fare. Usually several fare basis codes are contained in a single fare class.

**Fare class:** A category of booking with a (relatively) common fare. Typical labels for such classes [seeVinod (1995)] are: F for first class (separate compartment); J for business class, U for business class frequent flyer redemption (often separate compartment); Y for full fare coach; B, M, Q, V for progressively more discounted coach bookings; and T for frequent flyer coach cabin redemptions. Often other fare products (such as travel agent or company travelers) are categorized under one of these designations for control purposes.

**Fare product:** The full set of attributes associated with a specific transportation service. The set includes the fare as well as any restrictions or benefits that apply to that service at that fare.

**Fleet assignment:** Most airlines have a variety of aircraft types and sizes in their fleets. The fleet assignment process attempts to allocate aircraft to routes in the airline network to maximize contribution to profit. There are strong potential linkages between fleet assignment and revenue management processes because aircraft assignments determine leg capacities in the network.

**Flight leg:** A section of a flight involving a single takeoff and landing (or no boarding or deplaning of passengers at any intermediate stops). Also leg.

**Flight number:** A numeric or alphanumeric label for a flight service that involves (generally) a single aircraft departing from an origin airport, possibly making additional scheduled stops at one or more intermediate airports, and terminating at a destination airport.

**Full Nesting:** See nested booking.

**Global distribution system (GDS):** Computer and communications systems for linking booking locations with the computer reservation systems of different airlines. Examples are SABRE, Galileo, and Amadeus.

**Goodwill costs:** An airline's rejection of a booking request can affect a customer's propensity to seek future bookings from that airline. This cost is difficult to assess but is considered particularly acute in competitive markets and with customers who are frequent air travelers. An approximate assessment of the cost of a permanently lost customer is the expected net present value of all future bookings from the customer minus the opportunity costs of those bookings.

**Go-show:** Passengers who appear at the time of flight departure with a valid ticket for the flight but for whom there is no record in the reservation system. This no-record situation can occur when there are significant time lags in transferring booking information from reservations sources (e.g., travel agent's offices) to the CRS or when there are transmission breakdowns.

**Group bookings:** Bookings for groups of passengers that are negotiated with sales representatives of airlines; for example, for a large group from one company travelling to a trade show. These should be distinguished from batch bookings.

**Hub-and-spoke network:** A configuration of an airline's network around one or more major hubs that serve as switching points in passengers' itineraries to spokes connected to smaller centers. The proliferation of these networks has greatly increased the number of passenger itineraries that include connections to different flights.

**Hub bank:** A collection of inbound and outbound flights that are scheduled to arrive or depart within a time span that enables convenient passenger connections among flights. An airline hub will typically operate with several hub banks throughout the day.

**Incremental seat:** One additional seat, given the number of seats already booked. Independence of demands: The assumption that demands in one customer category (e.g., booking class or ODF) are statistically independent of demands in other categories. It is widely believed that this assumption is not satisfied in practice.

**Indexing:** The process of assigning individual ODF categories to virtual nesting buckets.

**Interspersed arrivals:** Characteristic of an arrivals process in which booking requests in different booking classes do not arrive in any particular order. (Compare with sequential booking classes.)

**Itinerary:** For purposes of this paper, an itinerary is a trip from an origin to a destination across one or more airline networks. A complete specification of an itinerary includes departure and arrival times, flight numbers, and booking classes. The term is used ambiguously to include both one-way and round-trip travel. That is, used in the first way, a round-trip involves two itineraries and, in the second way, one itinerary.

**Leg:** See flight leg.

**Leg based control:** An older, but still common, method of reservations control and revenue management in which limits are set at the flight leg level on the number of passengers flying in each booking class. Such systems are unable to properly control multileg traffic, although virtual nesting provides a partial solution.

**Littlewood's rule:** This simple two-fare allocation rule was proposed by Littlewood (1972). Given average high fare f1, average discount fare f2, random full fare demand Y, and s seats remaining, Littlewood's rule stipulates that a discount seat should be sold as long as the discount fare equals or exceeds the expected marginal return from a full fare booking of the last remaining seat; that is, discount demand should be satisfied as long as f2 f1 Pr(Y . s). This is essentially equivalent to the classic optimal stocking rule for single period stochastic inventory (newsvendor) problems.

**Load factor:** The ratio of seats filled on a flight to the total number of seats available.

**Low-before-high fares:** (Also called monotonic fares or sequential fares.) The sequential booking class assumption is often augmented by the additional assumption that booking requests arrive in strict fare sequence, generally from lowest to highest as flight departure approaches. The existence of low standby fares violates this assumption.

**Minimum acceptable fare (MAF):** See bid-price.

**Monotonic fares:** See low-before-high fares.

**Multileg:** A section of an itinerary or network involving more than one leg.

**Multiple booking:** See batch booking.

**Nested booking:** In fully nested (also called serially nested) booking systems, seats that are available for sale to a particular booking class are also available to bookings in any higher fare booking class, but not the reverse. Thus, a booking limit L for a discount booking class defines an upper bound on bookings in that class and any lower valued classes and a corresponding protection level of (C 2 L) for all higher classes; where C is the total capacity of the pool of seats shared by all classes. This should be contrasted with the older distinct bucket approach to booking control. See, also, parallel nesting.

**Network effects:** A booking on any leg in the airline network may block booking of any itinerary that includes that leg. Subsequent interactions of the blocked itinerary with other legs in the network can, in a similar fashion, propagate across the full network.

**Newsvendor problem:** The problem of choosing the quantity of a perishable item to stock (e.g., newspapers) given known cost, selling price, and salvage values, and subject to uncertain future demand. (Also called the newsboy or single period stocking problem.) This classic problem is essentially equivalent to the simple two-fare seat allocation problem with sequential arrivals.

**No-shows:** Booked passengers who fail to show up at the time of flight departure, thus allowing no time for their seat to be booked through normal reservations processes. No-shows are particularly common among full fare passengers whose tickets are fully refundable in the event of cancellation or no-show.

**ODF control (O-D problem):** Origin-destination fare control. An approach to revenue management that accounts for all possible passenger itineraries between origins and destinations in the airline network, at all fare levels. See network effects.

**Opportunity cost:** See displacement cost.

**Optimal booking limits:** This term is often used to refer to exact booking limits for the single leg seat inventory control under assumptions 1 through 6 in Section 4.1. They are only optimal within the context of that basic model. At present, there are no truly optimal booking limits for the full ODF revenue management problem, and likely never will be.

**Overbooking:** The practice of ticketing seats beyond the capacity of an aircraft to allow for the probability of no-shows.

**Oversold:** An ambiguous term sometimes used when more passengers show up for a flight than there are seats available. Such situations must be resolved with denied boardings.

**Parallel nesting:** See nested booking. This is an approach to booking that is intermediate between simple distinct bucket control and full nesting. A number of lower fare classes are assigned to distinct buckets, but these buckets are nested in one or more higher fare classes. This approach reduces the revenue potential of the combined fare classes, but may facilitate control. Perishable asset revenue management

**Protected seats:** Seats that are restricted to bookings in one or more fare classes. In fully nested booking systems, seats are protected for bookings in a fare class or any higher fare class.

**Protection levels:** The total number of protected seats for a booking class. In fully nested booking systems the protection level for a fare class applies to that class and all higher fare classes.

**RCS:** Reservations Control System.

**Recapture:** The booking of a passenger who is unable to obtain a reservation for a particular flight or set of flights with an airline onto alternative flights with the same airline.

**Reservation system controls:** The internal logic used by the reservation system for controlling the availability of seats. This logic is usually difficult to change and is often a significant constraint when implementing a yield management system. See controllable booking classes.

**Restrictions:** Sets of requirements that are applied to discount fare classes to differentiate them as fare products and discourage diversion. Examples are fourteen-day advance booking requirements, cancellation penalties, Saturday night stay over, and midweek departure requirements. Also referred to as booking fences.

**Revenue management:** The practice of controlling the availability and/or pricing of travel seats in different booking classes with the goal of maximizing expected revenues or profits. This term has largely replaced the original term yield management.

**Rules:** See restrictions.

**Seamless availability:** A capability of reservation and information systems that allows for direct transmission of availability requests from ticket agents to airlines. With this capability, airlines may be able to provide unrestricted origin-destination control of their seat inventory.

**Seat allocation:** See seat inventory control.

**Seat inventory control:** The component of a revenue management system that controls the availability of seats for different booking classes.

**Segment:** One or more flight legs covered by a single flight number. Thus, if a flight originates at airport A, makes an intermediate stop at B, and terminates at C; the possible flight segments are AB, BC, and ABC.

**Segment closed indicator (SCI):** A flag in reservations control systems that indicates that a booking class is closed to bookings over a particular segment. The same booking class may be open for bookings over other segments of the same flight. This allows for O-D control at the segment level.

**Segment control:** A level of itinerary seat inventory control that accounts for the revenue value of flight segments, but does not account for itineraries that involve other flight segments. In the case of a two leg flight A to B to C, segment control would permit closing the AB segment to a discount booking class but leaving the ABC segment open for the same class. This system fails to account for the possibly high revenue value of a booking that includes, for example, the segment AB in its itinerary but switches to a different flight at B.

**Sequential booking classes:** The assumption that requests for bookings in particular classes are not interleaved; for example, all B-class requests arrive before any Y-class requests. This assumption is rarely satisfied in practice; however, it is close enough to permit significant revenue gains from methods based on the assumption. Also, early booking restrictions on many discount booking classes ensure a degree of compliance.

**Sequential fares:** See low-before-high fares.

**Serial nesting:** See nested booking.

**Show-ups:** Passengers who appear for boarding at the time of flight departure. The number of showups is (final bookings 1 go-shows 1 standbys 2 no-shows).

**Single-leg control:** See leg based control.

**Space control:** See seat inventory control.

**Spill:** Unsatisfied demand that occurs because a capacity or booking limit has been reached. See censorship of demand data.

**Spill formula:** A formula or algorithm that estimates the amount of spill that has occurred on past flights.

**Spoilage:** Seats that travel empty despite the presence of sufficient demand to fill them. This will occur, for example, if discount booking classes are closed too early, and full fare demands do not fill the remaining seats. This should be distinguished from excess capacity-seats that are empty because of insufficient total demand.

**Standby fares:** Some airlines will sell last minute discount seats to certain categories of travelers (e.g., youth or military service personnel) who are willing to wait for a flight that would otherwise depart with empty seats.

**Static models:** Models that set current seat protection policies without consideration of the possibility of adjustments to the protection levels later in the booking process. (Compare with dynamic models.)

**Structural solution:** A solution to an optimization problem in the form of specifications (frequently equations) that reveal the pattern of behavior of optimal solutions. These are important because they lead to a deeper understanding of the nature of optimal solutions and can lead to development of efficient solution algorithms.

**Threshold curves:** Threshold curves are functions that return time-dependent booking limits for overbooking or seat inventory control.

**Unconstrained demand:** An estimate of the demand for a past flight or fare class that has been corrected for censorship.

**Upgrade:** This term is used in two ways. First, it refers to an offer to a passenger to fly in a higher service class without additional charge (e.g., in exchange for frequent flyer points, or to avoid a denied boarding). Second, it refers to a decision by a customer to book in a higher fare class than originally intended when he or she is advised that no seats are available at their preferred fare.

**Virtual nesting/virtual classes:** This is one approach to incorporating origin-destination information into leg or segment based control systems. Multiple ODFs are grouped into virtual buckets on the basis of similar revenue characteristics (e.g., comparable total fare values, or similar total bid prices). The virtual buckets may easily contain a mixture of traditional fare classes. The buckets are then nested and assigned to traditional booking classes for control in a leg based reservation system.

**Yield management:** The early term used for what is now more commonly called revenue management.

# APPENDIX B

# DATA AND RESULTS FOR DIFFERENT SCENARIOS

## B.1 DATA FOR DIFFERENT SCENARIOS

### B.1.1 BASE PROBLEM

Table B.1: Fares settings for the base problem

| *OD* Number | Origin-Destination | Fare Class 3 | Fare Class 2 | Fare Class 1 |
|---|---|---|---|---|
| 1 | A-B | 75 | 125 | 250 |
| 2 | A-C | 130 | 370 | 400 |
| 3 | A-D | 200 | 320 | 460 |
| 4 | B-C | 100 | 150 | 330 |
| 5 | B-D | 160 | 200 | 420 |
| 6 | C-D | 80 | 110 | 235 |

Table B.2: Demand settings for the base problem

| | Fare Class 3 | | | | Fare Class 2 | | | | Fare Class 1 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Itinerary* | $p_j$ | $\delta_j$ | $E_j$ | $SD_j$ | $p_j$ | $\delta_j$ | $E_j$ | $SD_j$ | $p_j$ | $\delta_j$ | $E_j$ | $SD_j$ |
| $AB$ | 80 | 1.6 | 50 | 9.01 | 80 | 2 | 40 | 7.75 | 3 | 0.1 | 30 | 18.17 |
| $AC$ | 80 | 2 | 40 | 7.75 | 50 | 2 | 25 | 6.12 | 2 | 0.1 | 20 | 14.83 |
| $AD$ | 60 | 2 | 30 | 6.71 | 72 | 3 | 24 | 5.66 | 2 | 0.1 | 20 | 14.83 |
| $BC$ | 60 | 2 | 30 | 6.71 | 40 | 2 | 20 | 5.48 | 2 | 0.1 | 20 | 14.83 |
| $BD$ | 60 | 2 | 30 | 6.71 | 60 | 3 | 20 | 5.16 | 6 | 0.3 | 20 | 9.31 |
| $CD$ | 80 | 1.6 | 50 | 9.01 | 80 | 2 | 40 | 7.75 | 6 | 0.2 | 30 | 13.42 |

Table B.3: Request arrival settings for the base problem

| | Fare Class 3 | | Fare Class 2 | | Fare Class 1 | |
|---|---|---|---|---|---|---|
| Itinerary | $\alpha$ | $\beta$ | $\alpha$ | $\beta$ | $\alpha$ | $\beta$ |
| 1-6 | 5 | 6 | 2 | 5 | 2 | 13 |

### B.1.2 LOW-BEFORE-HIGH ARRIVAL PATTERN

Table B.4: Request arrival settings for the case with low-before-high arrival pattern

| | Fare Class 3 | | Fare Class 2 | | Fare Class 1 | |
|---|---|---|---|---|---|---|
| *Itinerary* | $\alpha$ | $\beta$ | $\alpha$ | $\beta$ | $\alpha$ | $\beta$ |
| $1-6$ | 13 | 2 | 13 | 13 | 2 | 13 |

### B.1.3 INCREASED LOW-FARE DEMAND VARIANCE

Table B.5: Demand settings for the case with increased low-fare demand variance

| | Fare Class 3 | | | | Fare Class 2 | | | | Fare Class 1 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Itinerary* | $p_j$ | $\delta_j$ | $E_j$ | $SD_j$ | $p_j$ | $\delta_j$ | $E_j$ | $SD_j$ | $p_j$ | $\delta_j$ | $E_j$ | $SD_j$ |
| *AB* | 20 | 0.4 | 50 | 13.23 | 20 | 0.5 | 40 | 10.95 | 3 | 0.1 | 30 | 18.17 |
| *AC* | 20 | 0.5 | 40 | 10.95 | 5 | 0.2 | 25 | 12.25 | 2 | 0.1 | 20 | 14.83 |
| *AD* | 15 | 0.5 | 30 | 9.49 | 18 | 0.75 | 24 | 7.48 | 2 | 0.1 | 20 | 14.83 |
| *BC* | 15 | 0.5 | 30 | 9.49 | 10 | 0.5 | 20 | 7.75 | 2 | 0.1 | 20 | 14.83 |
| *BD* | 15 | 0.5 | 30 | 9.49 | 15 | 0.75 | 20 | 6.83 | 6 | 0.3 | 20 | 9.31 |
| *CD* | 20 | 0.4 | 50 | 13.23 | 20 | 0.5 | 40 | 10.95 | 6 | 0.2 | 30 | 13.42 |

### B.1.4 SMALLER DIFFERENCES BETWEEN FARES

Table B.6: Fare settings for case with smaller fare spreads

| *OD* Number | Origin-Destination | Fare Class 3 | Fare Class 2 | Fare Class 1 |
|---|---|---|---|---|
| 1 | A-B | 75 | 125 | 175 |
| 2 | A-C | 130 | 170 | 220 |
| 3 | A-D | 200 | 320 | 440 |
| 4 | B-C | 100 | 150 | 210 |
| 5 | B-D | 160 | 200 | 250 |
| 6 | C-D | 80 | 110 | 160 |

### B.1.5 REALISTIC COEFFICIENTS OF VARIATION AND CLOSE FARES

Table B.7: Fare settings for the case with realistic variations and close fares

| *OD* Number | Origin-Destination | Fare Class 3 | Fare Class 2 | Fare Class 1 |
|---|---|---|---|---|
| 1 | A-B | 75 | 125 | 175 |
| 2 | A-C | 130 | 170 | 220 |
| 3 | A-D | 200 | 320 | 460 |
| 4 | B-C | 100 | 150 | 210 |
| 5 | B-D | 180 | 210 | 250 |
| 6 | C-D | 80 | 110 | 160 |

Table B.8: Demand settings for the case with realistic variations and close fares

| | Fare Class 3 | | | | Fare Class 2 | | | | Fare Class 1 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Itinerary* | $p_j$ | $\delta_j$ | $E_j$ | $SD_j$ | $p_j$ | $\delta_j$ | $E_j$ | $SD_j$ | $p_j$ | $\delta_j$ | $E_j$ | $SD_j$ |
| *AB* | 80 | 1.6 | 50 | 9.01 | 80 | 2 | 40 | 7.75 | 3 | 0.1 | 30 | 18.17 |
| *AC* | 80 | 2 | 40 | 7.75 | 50 | 2 | 25 | 6.12 | 2 | 0.1 | 20 | 14.83 |
| *AD* | 60 | 2 | 30 | 6.71 | 72 | 3 | 24 | 5.66 | 2 | 0.1 | 20 | 14.83 |
| *BC* | 60 | 2 | 30 | 6.71 | 40 | 2 | 20 | 5.48 | 2 | 0.1 | 20 | 14.83 |
| *BD* | 60 | 2 | 30 | 6.71 | 60 | 3 | 20 | 5.16 | 6 | 0.3 | 20 | 9.31 |
| *CD* | 80 | 1.6 | 50 | 9.01 | 80 | 2 | 40 | 7.75 | 6 | 0.2 | 30 | 13.42 |

Table B.9: Request arrival settings for the case with realistic variations and close fares

| | Fare Class 3 | | Fare Class 2 | | Fare Class 1 | |
|---|---|---|---|---|---|---|
| *Itinerary* | $\alpha$ | $\beta$ | $\alpha$ | $\beta$ | $\alpha$ | $\beta$ |
| 1 – 6 | 2 | 2 | 2 | 2 | 2 | 2 |

## B.2 RESULTS OF THE OPTIMIZATION MODELS

### B.2.1 BASE PROBLEM

Table B.10: Optimal allocations of the mathematical models for the base problem

| *ODF* | | | | | *EMVLP* ($\theta$) | | *SLP-RM* ($L$)(x1000) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *itinerary* | *class* | *DLP* | *EMR* | *SLP* | 0.001 | 0.005 | 60 | 65 | 70 | 75 | 80 |
| *AB* | 3 | 41 | 42 | 44 | 44 | 42 | 43 | 44 | 46 | 44 | 44 |
| | 2 | 40 | 40 | 40 | 41 | 38 | 40 | 43 | 41 | 42 | 40 |
| | 1 | 30 | 40 | 40 | 39 | 25 | 32 | 30 | 33 | 36 | 36 |
| *AC* | 3 | 0 | 0 | 0 | 1 | 18 | 18 | 7 | 7 | 0 | 0 |
| | 2 | 25 | 18 | 20 | 20 | 18 | 20 | 21 | 20 | 20 | 20 |
| | 1 | 20 | 22 | 20 | 19 | 10 | 14 | 16 | 17 | 20 | 20 |
| *AD* | 3 | 0 | 0 | 0 | 0 | 23 | 5 | 7 | 0 | 0 | 0 |
| | 2 | 24 | 21 | 20 | 21 | 18 | 20 | 21 | 22 | 22 | 24 |
| | 1 | 20 | 17 | 16 | 15 | 8 | 8 | 12 | 14 | 16 | 16 |
| *BC* | 3 | 30 | 23 | 24 | 25 | 24 | 24 | 24 | 23 | 24 | 23 |
| | 2 | 20 | 19 | 16 | 19 | 18 | 20 | 19 | 20 | 20 | 20 |
| | 1 | 20 | 27 | 28 | 24 | 13 | 19 | 17 | 21 | 24 | 25 |
| *BD* | 3 | 1 | 15 | 20 | 20 | 21 | 20 | 20 | 20 | 18 | 12 |
| | 2 | 20 | 16 | 16 | 16 | 15 | 16 | 20 | 16 | 16 | 16 |
| | 1 | 20 | 22 | 20 | 20 | 14 | 17 | 17 | 20 | 20 | 24 |
| *CD* | 3 | 45 | 38 | 40 | 39 | 40 | 40 | 40 | 40 | 36 | 38 |
| | 2 | 40 | 36 | 36 | 36 | 35 | 36 | 36 | 36 | 36 | 36 |
| | 1 | 30 | 35 | 32 | 33 | 26 | 38 | 27 | 32 | 36 | 36 |

### B.2.2 INCREASED LOW FARE DEMAND VARIANCE

Table B.11: Optimal allocations of the mathematical models for increased low-fare demand variance case

| *ODF* | | | | | *EMVLP* ($\theta$) | | *SLP-RM* ($L$)(x1000) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *itinerary* | *class* | *DLP* | *EMR* | *SLP* | 0.001 | 0.005 | 60 | 65 | 70 | 75 | 80 |
| *AB* | 3 | 41 | 41 | 40 | 41 | 41 | 44 | 44 | 42 | 40 | 40 |
| | 2 | 40 | 41 | 40 | 40 | 38 | 38 | 43 | 44 | 40 | 40 |
| | 1 | 30 | 41 | 40 | 38 | 26 | 29 | 35 | 36 | 41 | 41 |
| *AC* | 3 | 0 | 0 | 0 | 3 | 26 | 10 | 0 | 4 | 0 | 0 |
| | 2 | 25 | 15 | 16 | 15 | 14 | 16 | 16 | 16 | 15 | 16 |
| | 1 | 20 | 23 | 20 | 19 | 10 | 14 | 20 | 20 | 21 | 24 |
| *AD* | 3 | 0 | 0 | 8 | 9 | 18 | 16 | 5 | 1 | 4 | 0 |
| | 2 | 24 | 21 | 20 | 20 | 18 | 20 | 22 | 20 | 20 | 20 |
| | 1 | 20 | 18 | 16 | 15 | 9 | 13 | 16 | 16 | 20 | 20 |
| *BC* | 3 | 30 | 22 | 24 | 23 | 24 | 22 | 24 | 23 | 20 | 20 |
| | 2 | 20 | 19 | 16 | 28 | 17 | 18 | 20 | 20 | 20 | 18 |
| | 1 | 20 | 28 | 28 | 25 | 14 | 19 | 24 | 24 | 28 | 28 |
| *BD* | 3 | 1 | 17 | 16 | 18 | 21 | 20 | 17 | 19 | 16 | 17 |
| | 2 | 20 | 15 | 16 | 15 | 15 | 15 | 16 | 16 | 14 | 16 |
| | 1 | 20 | 22 | 20 | 20 | 14 | 18 | 20 | 20 | 22 | 22 |
| *CD* | 3 | 45 | 35 | 32 | 35 | 41 | 34 | 36 | 35 | 33 | 33 |
| | 2 | 40 | 36 | 36 | 35 | 36 | 36 | 36 | 36 | 36 | 36 |
| | 1 | 30 | 36 | 36 | 33 | 28 | 29 | 32 | 36 | 36 | 36 |

### B.2.3 SMALLER DIFFERENCES BETWEEN FARES

Table B.12: Optimal allocations of the mathematical models for smaller spread among fares case

| *ODF* | | | | | *EMVLP* ($\theta$) | | *SLP-RM* ($L$)(x1000) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *itinerary* | *class* | *DLP* | *EMR* | *SLP* | 0.001 | 0.005 | 60 | 65 | 70 | 75 | 80 |
| *AB* | 3 | 41 | 45 | 44 | 44 | 42 | 48 | 45 | 44 | 44 | 44 |
| | 2 | 40 | 41 | 40 | 40 | 38 | 41 | 42 | 44 | 44 | 40 |
| | 1 | 30 | 36 | 32 | 33 | 25 | 23 | 28 | 33 | 36 | 32 |
| *AC* | 3 | 0 | 4 | 8 | 9 | 23 | 21 | 11 | 2 | 0 | 8 |
| | 2 | 25 | 20 | 20 | 20 | 19 | 20 | 20 | 21 | 20 | 20 |
| | 1 | 20 | 14 | 12 | 12 | 8 | 4 | 12 | 16 | 16 | 12 |
| *AD* | 3 | 0 | 0 | 8 | 6 | 18 | 15 | 6 | 1 | 0 | 8 |
| | 2 | 24 | 22 | 20 | 21 | 19 | 20 | 23 | 24 | 20 | 20 |
| | 1 | 20 | 18 | 16 | 15 | 8 | 8 | 12 | 16 | 20 | 16 |
| *BC* | 3 | 30 | 25 | 24 | 25 | 24 | 24 | 25 | 28 | 24 | 24 |
| | 2 | 20 | 20 | 20 | 19 | 18 | 18 | 20 | 20 | 20 | 20 |
| | 1 | 20 | 22 | 20 | 20 | 13 | 17 | 17 | 20 | 25 | 20 |
| *BD* | 3 | 1 | 21 | 20 | 21 | 22 | 22 | 22 | 20 | 20 | 20 |
| | 2 | 20 | 17 | 16 | 16 | 15 | 16 | 16 | 16 | 16 | 16 |
| | 1 | 20 | 17 | 16 | 16 | 13 | 15 | 16 | 16 | 19 | 16 |
| *CD* | 3 | 45 | 38 | 40 | 39 | 42 | 40 | 40 | 40 | 40 | 40 |
| | 2 | 40 | 37 | 36 | 37 | 37 | 36 | 36 | 36 | 36 | 36 |
| | 1 | 30 | 30 | 28 | 29 | 26 | 28 | 28 | 31 | 29 | 28 |

### B.2.4 REALISTIC VARIATIONS AND CLOSE FARES

Table B.13: Optimal allocations of the mathematical models for realistic variations and close fares case

| *ODF* | | | | | *EMVLP* ($\theta$) | | *SLP-RM* ($L$)(x1000) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *itinerary* | *class* | *DLP* | *EMR* | *SLP* | 0.001 | 0.005 | 60 | 65 | 70 | 75 | 80 |
| *AB* | 3 | 50 | 43 | 40 | 42 | 43 | 44 | 44 | 44 | 40 | 40 |
| | 2 | 40 | 42 | 40 | 41 | 38 | 39 | 40 | 40 | 40 | 40 |
| | 1 | 30 | 34 | 32 | 33 | 30 | 32 | 32 | 36 | 32 | 32 |
| *AC* | 3 | 0 | 4 | 16 | 12 | 25 | 16 | 12 | 1 | 16 | 16 |
| | 2 | 25 | 16 | 16 | 15 | 14 | 16 | 16 | 16 | 16 | 16 |
| | 1 | 20 | 18 | 16 | 17 | 16 | 16 | 16 | 20 | 16 | 16 |
| *AD* | 3 | 11 | 20 | 20 | 20 | 21 | 20 | 20 | 20 | 20 | 20 |
| | 2 | 24 | 22 | 20 | 21 | 18 | 19 | 20 | 24 | 20 | 20 |
| | 1 | 20 | 21 | 20 | 19 | 15 | 18 | 19 | 20 | 20 | 20 |
| *BC* | 3 | 30 | 22 | 24 | 22 | 22 | 23 | 20 | 20 | 24 | 24 |
| | 2 | 20 | 19 | 16 | 18 | 17 | 16 | 20 | 20 | 16 | 16 |
| | 1 | 20 | 22 | 20 | 21 | 18 | 20 | 20 | 24 | 20 | 20 |
| *BD* | 3 | 10 | 21 | 20 | 21 | 22 | 24 | 20 | 20 | 20 | 20 |
| | 2 | 20 | 16 | 16 | 16 | 15 | 16 | 16 | 16 | 16 | 16 |
| | 1 | 20 | 19 | 16 | 18 | 17 | 16 | 20 | 20 | 16 | 16 |
| *CD* | 3 | 45 | 35 | 40 | 38 | 44 | 43 | 36 | 33 | 40 | 40 |
| | 2 | 40 | 35 | 36 | 36 | 38 | 32 | 36 | 35 | 36 | 36 |
| | 1 | 30 | 31 | 32 | 31 | 30 | 32 | 32 | 32 | 32 | 32 |

# APPENDIX C

# MATLAB PROGRAMMING CODES

In this section MATLAB Programming codes are given. The brief definitions of ten different Matlab .m files are given below:

Mainpart.m: Main part for optimization models and used for test purposes.

Input.m: Initial data for legs, seats and demands.

Demandpart.m: This .m file is used for calculating probabilities of demands for all predetermined integer values.

SLP-RM.m: File for *SLP-RM* procedure.

PMP-RC.m: File for *PMP-RC* model.

RRS.m: File for *RRS* procedure.

Partitioned.m: Simulation model for partitioned control policy.

Nested.m: Simulation model for nested control policy.

BidPrice.m: Simulation model for bid price control policy.

Bayesian.m: Simulation model for Bayesian update.

Mainpart.m

```
% This .m file is the mainpart for optimization models and used for ↙
test
% purposes. By setting test and den values and changing parameters ↙
of the
% models, performances of the models are tested.

tic;      % Used for calculating CPU time

for test=1:1

    for den=1:1

    Input                    % Go to SLP_input .m file and get the ↙
input data
    Demandpart               % Go to SLP_demand .m file and solve the ↙
problem
    revden(den,:)=rev;       % Get objective function values
    bidden(den,:)=bid;       % Get bid prices
    seatden(den,:)=seat;     % Get seat allocations
    %probden(den,:)=prob;     % Get probabilily of revenue being↙
lower than
                             % the target level(used only for ↙
proposed models)

    end

    revtest(:,:,test)=revden;        % Write objective function↙
values to a 3-d matrix
    bidtest(:,:,test)=bidden;        % Write bid prices to a 3-d↙
matrix
    seattest(:,:,test)=seatden;      % Write seat allocations to a 3- ↙
d matrix
    %probtest(:,:,test)=probden;      % Write probabilities to a 3-d↙
matrix(used only
                                     % for proposed models)

end

toc;                 % Used for calculating CPU time

clear test den        % clear variables test and den
```

```
% This .m file is used for describing initial values for legs, seats ↙
and demands.

%*********NETWORK STRUCTURE DATA************

C=200;                          % Upper bound of x variables
CAPACITY=[200 200 200];         % Capacity of the flights
D=18;                           % Number of ODFs
NL=3;                           % Number of legs
NF=3;                           % Number of fare classes
K=200;                          % Number of demand segments
Leg=[1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0;0 0 0 1 1 1 1 1 1 1 1 1 1 1↙
1 0 0 0;...
    0 0 0 0 0 0 1 1 1 0 0 0 1 1 1 1 1 1]; % ODFs in legs


%************DEMAND AND FARE DATA******************

M=100000;        % Big number (used only for proposed models)
L=65000;         % Target level (used only for proposed models)
Rho=0;           % Limit for probability being lower than L (used ↙
only for proposed models)
NOD=200;         % Number of demand realizations (used only for ↙
proposed models)
theta=0.005;     % Theta value for EMVLP model

beT=[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0]; % Used for bayesian↙
update purposes

%***************Base Problem*******************
FF=[75 125 250 130 170 400 200 320 460 100 150 330 160 200 420 80 ↙
110 235]; % Price of tickets
DDODF=[80 1.6;80 2;3 0.1;80 2;50 2;2 0.1;60 2;72 3;2 0.1;60 2;...
    40 2;2 0.1;60 2;60 3;6 0.3;80 1.6;80 2;6 0.2];↙
% Demand parameters
Inbeta=[5 6;2 5;2 13];↙
% Beta distribution parameters
%
%
% %**********Changed Arrival Process************
% FF=[75 125 250 130 170 400 200 320 460 100 150 330 160 200 420 80 ↙
110 235]; % Price of tickets
% DDODF=[80 1.6;80 2;3 0.1;80 2;50 2;2 0.1;60 2;72 3;2 0.1;60 2;...
%     40 2;2 0.1;60 2;60 3;6 0.3;80 1.6;80 2;6 0.2];↙
% Demand parameters
%  Inbeta=[13 2;13 13;2 13];↙
% Beta distribution parameters
```

```
%
%
%*******Increased Low-Fare Demand Variance******
% FF=[75 125 250 130 170 400 200 320 460 100 150 330 160 200 420 80 ↙
110 235]; % Price of tickets
% DDODF=[20 0.4;20 0.5;3 0.1;20 0.5;5  0.2;2 0.1;15 0.5;18 0.75;2↙
0.1;15 0.5;...
%      10 0.5;2 0.1;15 0.5;15 0.75;6 0.3;20 0.4;20 0.5;6 0.2];↙
% Demand parameters
% Inbeta=[13 2;13 13;2 13];↙
% Beta distribution parameters
%
%
% %*********** Changed Fares 1*****************
% FF=[75 125 175 130 170 220 200 320 440 100 150 210 160 200 250 80 ↙
110 160]; % Price of tickets
% DDODF=[80 1.6;80 2;3 0.1;80 2;50 2;2 0.1;60 2;72 3;2 0.1;60 2;...
%      40 2;2 0.1;60 2;60 3;6 0.3;80 1.6;80 2;6 0.2];↙
% Demand parameters
% Inbeta=[13 2;13 13;2 13];↙
% Beta distribution parameters
%
%
% %*********** Changed Fares 2*****************
% FF=[75 125 175 130 170 220 230 340 460 100 150 210 180 210 250 80 ↙
110 160]; % Price of tickets
% DDODF=[20 0.4;20 0.5;30 1;20 0.5;5 0.2;20 1;15 0.5;18 0.75;20 1;15↙
0.5;...
%      10 0.5;20 1;15 0.5;15 0.75;60 3;20 0.4;20 0.5;60 2];↙
% Demand parameters
% Inbeta=[2 2;2 2;2 2];↙
% Beta distribution parameters
% CAPACITY=[220 220 220];     % Capacity of the flights


for i=1:D
    Expdem(1,i)=DDODF(i,1)/DDODF(i,2);       % Calculate expected↙
demand
end
clear i;
```

```
% This .m file is used for calculating probabilities of demands for ↙
all predetermined integer values.


%*****************Calculate Pr(Dj>=i) and Pr(Dj<i) ↙


pdfno=(0:C-1);

for i=1:D
    ProbDist(:,i)=1-nbincdf(pdfno,DDODF(i,1),(DDODF(i,2)+beT(1,i))/ ↙
(DDODF(i,2)+1));
end

ProbDist=1-ProbDist;

for i=1:K
    ProbDistcdf(i,:)=ProbDist(i*C/K,:);     % Group demands
end

ProbDistrcdf=1-ProbDistcdf;

%***********************Calculate Pr(Dj=i) ↙


pdfno=(0:C);
for i=1:D
    ProbDistpdf(:,i)=nbinpdf(pdfno,DDODF(i,1),DDODF(i,2)/(DDODF(i,2) ↙
+1));
    ProbDistpdf(C,i)=1-nbincdf(C-1,DDODF(i,1),DDODF(i,2)/(DDODF(i,2) ↙
+1));
end

clear pdfno i ProbDist
SLPRM          % Go to mathematical model
```

```
% In this .m file, first of all  average of v(a) values is minimized ↙
and
% then the results of this model are used to maximize expected ↙
revenue.

%%%%%%%%%%%%%FIRST MODEL%%%%%%%%%%%%%%

% VARIABLES

xodf=sdpvar(K,D,'full');
v=sdpvar(NOD,1);


%CONSTRAINTS

for l=1:NOD

% b values for 0-1 constraints
bin1(l,1)=-L;
bin2(l,1)=M+L;

    for rn=1:D
        RDM(l,rn) = random('nbin',DDODF(rn,1),DDODF(rn,2)/(DDODF(rn, ↙
2)+1),1,1);
    end

        for j=1:D
            RDMj=min(max(RDM(l,j)),C);
            for i=1:K
                if RDMj>(i-1)*C/K
                    MA(i,j)=1;
                else
                    MA(K,j)=0;
                    break
                end
            end
        end

        MAT(:,:,l)=MA;


    bc1(l,1)=-M*v(l,1)-sum(MA.*xodf)*FF';
    bc2(l,1)=M*v(l,1)+sum(MA.*xodf)*FF';

    clear MA;

end
```

```
clear l rn RDMj j i;

cap=sum(xodf*Leg')';                % Capacity constraint

for i=1:NL
    bcap(i,1)=CAPACITY(1,i);        % Right hand side of capacity
constraint
end

clear i;

% Constraint set

F=set(cap <= bcap)+set(bc1 <= bin1)+set(bc2 <= bin2)+set(0 <= xodf
<= C/K)+set(0<= v <=1);

% OBJECTIVE FUNCTION

pav=sum(v)/NOD;

% SOLVE MODEL

solvesdp(F,pav,sdpsettings('solver','glpk'));

% RESULTS


for l=1:NOD

   if double(v(l,1))>0
       vi(l,1)=1;
   else
       vi(l,1)=0;
   end

end
clear l;

prob=sum(vi)/NOD;

%%%%%%%%%%%%%SECOND MODEL%%%%%%%%%%%%%%%%%

% Clear constraints and variables

clear F xodf cap bc1 bc2 v;
```

```
% VARIABLES

xodf=sdpvar(K,D,'full');

%CONSTRAINTS

for l=1:NOD

    MA=MAT(:,:,l);

    bc1(l,1)=-M*vi(l,1)-sum(MA.*xodf)*FF';
    bc2(l,1)=M*vi(l,1)+sum(MA.*xodf)*FF';

    clear MA;

end
clear l;

cap=sum(xodf*Leg')';                    % Capacity constraint

% Constraint set

F=set(cap <= bcap)+set(bc1 <= bin1)+set(bc2 <= bin2)+set(0 <= xodf ↙
<= C/K);

% OBJECTIVE FUNCTION

obj=-sum(xodf*FF')+sum((ProbDistcdf.*xodf)*FF');

% SOLVE MODEL

solvesdp(F,obj,sdpsettings('solver','glpk'));

% RESULTS

bid=dual(F(1))';                                        % Bid price

% Round results

for i=1:D
    for j=1:K
        dxodf(j,i)=round(double(xodf(j,i)));
    end
end
clear i j;

seat=sum(dxodf);                                        % Seat↙
```

```
allocation

rev=sum(dxodf*FF')-sum((ProbDistcdf.*dxodf)*FF');       % Revenue

clear bc1 bc2 bcap bin1 bin2 cap dxodf obj pav vi xodf F MA MAT↙
ProbDist ProbDistcdf ProbDistpdf;
```

```
% This .m file minimize SLP problem with an additional constraint ↙
sum(pa.va)<rho.
% Instead of using pa values, average of va's are used in the model.

% VARIABLES

xodf=sdpvar(K,D,'full');
v=binvar(NOD,1);


%CONSTRAINTS

for l=1:NOD

% b values for 0-1 constraints
bin1(l,1)=-L;
bin2(l,1)=M+L;

    for rn=1:D
        RDM(l,rn) = random('nbin',DDODF(rn,1),DDODF(rn,2)/(DDODF(rn, ↙
2)+1),1,1);
    end

        for j=1:D
            RDMj=min(max(RDM(l,j)),C);
            for i=1:K
                if RDMj>(i-1)*C/K
                    MA(i,j)=1;
                else
                    MA(K,j)=0;
                    break
                end
            end
        end

        MAT(:,:,l)=MA;


    bc1(l,1)=-M*v(l,1)-sum(MA.*xodf)*FF';
    bc2(l,1)=M*v(l,1)+sum(MA.*xodf)*FF';

    clear MA;

end

clear l rn RDMj j i;
```

```
cap=sum(xodf*Leg')';         % Capacity constraint

for i=1:NL
    bcap(i,1)=CAPACITY(1,i);            % Right hand side of ↙
capacity constraints
end
clear i;

pav=sum(v)/NOD;              % Constraint for Pr(R<L)

%Constraint Set

F=set(cap <= bcap)+set(bc1 <= bin1)+set(bc2 <= bin2)+set(0 <= xodf ↙
<= C/K)+set(pav <= Rho);

% OBJECTIVE FUNCTION

obj=-sum(xodf*FF')+sum((ProbDistcdf.*xodf)*FF');

% SOLVE MODEL

solvesdp(F,obj,sdpsettings('solver','glpk'));

% RESULTS

rev=-double(obj);                    % Revenue

prob=double(pav);                    % Pr(R<L)

seat=round(sum(double(xodf)));       % Seat allocation

bid=[0 0 0];                         % Integer prograaming, no bid↙
price

clear bc1 bc2 bcap beT bin1 bin2 cap obj pav v xodf MAT MA↙
ProbDistpdf ProbDistcdf ProbDist;
```

```
% RRS Procedure

A1=0;
B1=0;
A2=0;
B2=0;
A3=0;
B3=0;
A4=0;
B4=0;

for s=1:300                                   % Number of realized demands

    % VARIABLES

    x_rlp=sdpvar(1,D,'full');

    % CONSTRAINTS

    i=1:1:NL;
    b_rlp(i,1)=CAPACITY(1,i);          % RHS of the capacity constraint
    clear i

    for rn=1:D
        RDM_rlp(1,rn) = random('nbin',DDODF(rn,1),(DDODF(rn,2)+beT↙
(1,rn))/(DDODF(rn,2)+1),1,1);   % Generate random demand
    end
    clear rn

    cap_rlp=(x_rlp*Leg')';                % Capacity constraint

    % Constraint Set

    F=set(cap_rlp <= b_rlp)+set(0 <= x_rlp <=RDM_rlp);

    % OBJECTIVE FUNCTION

    obj_rlp(s,1)=-x_rlp*FF';

    % SOLVE MODEL

    solvesdp(F,obj_rlp(s,1),sdpsettings('solver','glpk'));

    % OUTPUT

    bid_rlp(s,:)=dual(F(1))';
```

```
    if -double(obj_rlp(s,1))<75000;
        A1=A1+1;
        objm1(A1,1)=-double(obj_rlp(s,1));
        for i=1:NL
            bidm1(A1,i)=bid_rlp(s,i);
        end
    else
        B1=B1+1;
        objm2(B1,1)=-double(obj_rlp(s,1));
        for i=1:NL
            bidm2(B1,i)=bid_rlp(s,i);
        end
    end

    if -double(obj_rlp(s,1))<80000;
        A2=A2+1;
        objm3(A2,1)=-double(obj_rlp(s,1));
        for i=1:NL
            bidm3(A2,i)=bid_rlp(s,i);
        end
    else
        B2=B2+1;
        objm4(B2,1)=-double(obj_rlp(s,1));
        for i=1:NL
            bidm4(B2,i)=bid_rlp(s,i);
        end
    end

    if -double(obj_rlp(s,1))<85000;
        A3=A3+1;
        objm5(A3,1)=-double(obj_rlp(s,1));
        for i=1:NL
            bidm5(A3,i)=bid_rlp(s,i);
        end
    else
        B3=B3+1;
        objm6(B3,1)=-double(obj_rlp(s,1));
        for i=1:NL
            bidm6(B3,i)=bid_rlp(s,i);
        end
    end

    if -double(obj_rlp(s,1))<90000;
        A4=A4+1;
        objm7(A4,1)=-double(obj_rlp(s,1));
        for i=1:NL
```

```
            bidm7(A4,i)=bid_rlp(s,i);
        end
    else
        B4=B4+1;
        objm8(B4,1)=-double(obj_rlp(s,1));
        for i=1:NL
            bidm8(B4,i)=bid_rlp(s,i);
        end
    end

clear b_rlp cap_rlp x_rlp RDM_rlp obj_rlp F i;

end

    bidp(1,:)=mean(bidm1);
    bidp(2,:)=mean(bidm2);
    bidp(3,:)=mean(bidm3);
    bidp(4,:)=mean(bidm4);
    bidp(5,:)=mean(bidm5);
    bidp(6,:)=mean(bidm6);
    bidp(7,:)=mean(bidm7);
    bidp(8,:)=mean(bidm8);

    obj(1,:)=mean(objm1);
    obj(2,:)=mean(objm2);
    obj(3,:)=mean(objm3);
    obj(4,:)=mean(objm4);
    obj(5,:)=mean(objm5);
    obj(6,:)=mean(objm6);
    obj(7,:)=mean(objm7);
    obj(8,:)=mean(objm8);

bid=bidp(2,:)';
rev=obj(2,:);
seat=[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0];
```

```
% Simulation model for partitioned booking limits

tic;

    %%%%%%%%%%%%%%INPUT PART%%%%%%%%%%%%%%

    %----Static input part

    Input           % Get input data from optimization part

    notp=3600;      % Number of total periods (3600 periods=24 ↙
hours*150 days)

    IND=DDODF;      % Duplicate demand parameters

    Allocations;    % Get booking limits from allocations (for DLP, ↙
EMR, SLP and EMVLP models)

    INCAP=CAPACITY; % Initial Capacity

    %----Dynamic input part

    TM=24;           % Total number of models will be solved in the ↙
simulation

    PM=0;           % Total number of proposed models will be solved ↙
in the simulations

    Update=0;       % 3 if bayesian update is used, 0 otherwise

    upn=3;          % Number of updates for one replication

    trep=5000;        % Replication number

    %%%%%%%%%%%%BOOKING LIMITS%%%%%%%%%%%%%%%%

    if Update==3    % Get booking limits of proposed models
        TM=1;
        clear BL
    else
        for i=1:PM
             BL(i+(TM-PM),:)=seatden(i,:);
        end
    clear i;
    end

   %%%GENERATE ARRIVAL RATE USING BETA DISTRIBUTION%%%
```

```
    for i=1:3
        for j=1:notp
            arr_rate(j,i)=betapdf(j/notp,Inbeta(i,1),Inbeta(i,2)) ↙
/notp;
        end
    end
    clear i j;

    %%%%%%%%%%%%%REPLICATIONS%%%%%%%%%%%%%

for rep=1:trep              % Replication number

    rep                     % Show replication number on screen

    if Update==3
        up=1;               % Intialize update number
        Mainpart            % Go to Mainpart and solve the model
        BL=seat;            % Get booking limit
    end

    %%%%%GENERATE INTERARRIVAL TIMES%%%%%%%%%

    % Generate random gamma expected demand for all ODFs

    for i=1:6
        for j=1:3
            Aodf(i,j)=gamrnd(DDODF((i-1)*3+j,1),1/DDODF((i-1)*3+j, ↙
2));
        end
    end
    clear i j

    % Generate Lambda(t)=Beta(t).Aodf

    for k=1:notp
        for i=1:6
            for j=1:3
                arr_nhpp(k,(i-1)*3+j)=arr_rate(k,j)*Aodf(i,j);
            end
        end
    end
    clear i j k

    % Generate interarrival times with interpolation
    inter_arrival=[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0];
```

```
    for k=1:18
        y=0;
        m=0;
        for cnt=1:1000
            rnd=unifrnd(0,1,1);
            y =y-log(rnd);
            if y>Aodf(ceil(k/3),k-floor((k-1)/3)*3);
                break;
            end
            slam=0;
            for i=1:notp
                slam=slam+arr_nhpp(notp-i+1,k);
                if slam>y
                    m=m+1;
                    inter_arrival(m,k)=(i-1)+(y-slam+arr_nhpp(notp- ↙
i+1,k))/arr_nhpp(notp-i+1,k);
                    break
                end
            end
        end
    end
    clear k y m slam

    sizeint=size(inter_arrival);        % Size of the inter_arrival ↙
matrix
    rsize=sizeint(1,1)*sizeint(1,2);    % Number of cells in ↙
inter_arrival matrix

    % Set null cells to a big number

    for k=1:18
        for n=1:sizeint(1,1)
            if inter_arrival(n,k)==0
                inter_arrival(n,k)=100000;
            end
        end
    end
    clear k n

    % Initialize counters

    for i=1:TM
        ER(i)=0;                                                % Total ↙
revenue
        CAP(i,:)=CAPACITY;                                      % ↙
Capacity
        AR(i,:)=[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0];          % ↙
```

```
Accepted Requests
        reject(i)=0;                                        %↙
Rejected Requests
        acount(1,:)=[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0];  % Total↙
number of requests
    end
    clear i

    %%%%%%%%%%SIMULATION PART%%%%%%%%%%

    intarr=inter_arrival;       % Copy inter_arrival matrix

    for r=1:rsize

        %----Find the type and time of the arrival

        minint=min(intarr);
        m_minint=min(minint);       % Find minimum number (first↙
arrival time) in the matrix

        if m_minint>notp
            break                   % If this value is greater than↙
the departure time of the flight, stop.
        end

        for k=1:18
            if minint(1,k)==m_minint
                atype=k;
                atime=m_minint;
                for n=1:sizeint(1,1)
                    if intarr(n,k)==m_minint
                        intarr(n,k)=100000;     % Find the ODF of↙
the arrival and set cell to 10000
                        break
                        break
                    end
                end
            end
        end
        clear k n

        %----Go to bayesian update

        if Update==3
            Bayesian;
        end
```

```
        %----Decision making

        for ms=1:TM
            if BL(ms,atype)-AR(ms,atype)>0              % If total↙
accepted is lower than booking limit and the capacity is available ↙
then accept
                if min(CAP(ms)-Leg(:,atype)')>=0;
                    ER(ms)=ER(ms)+FF(atype);
                    CAP(ms,:)=CAP(ms,:)-Leg(:,atype)';
                    AR(ms,atype)=AR(ms,atype)+1;
                else
                    reject(ms)=reject(ms)+1;
                end
            end

        end
        clear ms

        acount(1,atype)=acount(1,atype)+1;       % Count requests

    end

    for ms=1:TM
        LF(rep,ms)=(sum(INCAP)-sum(CAP(ms,:)))/sum(INCAP);     % Load↙
factor
        Rev(rep,ms)=ER(ms);↙
% Revenue
    end
    clear ms

    clear inter_arrival

end

LoadFactor=mean(LF);     % Average of load factor
Average_Rev=mean(Rev);  % Average of revenue

toc;
```

```
% Simulation model for nested booking limits

tic;

    %%%%%%%INPUT PART%%%%%%%%%

    %----Static input part

    Input           % Get input data from optimization part

    notp=150;       % Number of total periods (3600 periods=24 ↙
hours*150 days)

    IND=DDODF;      % Duplicate demand parameters

    Allocations ;   % Get booking limits from allocations (for DLP, ↙
EMR, SLP and EMVLP models)

    INCAP=CAPACITY; % Initial Capacity

    %----Dynamic input part

    TM=10;          % Total number of models will be solved in the ↙
simulation

    PM=5;           % Total number of proposed models will be solved ↙
in the simulation

    Update=2;       % 2 if bayesian update is used, 0 otherwise

    upn=3;          % Number of updates for one replication

    trep=500;       % Replication number

    %%%%%%%%BOOKING LIMITS%%%%%%%%%%

    if Update==2    % Get booking limits of proposed models
        TM=1;
        clear BL
    else
        for i=1:PM
             BL(i+(TM-PM),:)=seatden(i,:);
        end
    clear i;
    end

   %%%GENERATE ARRIVAL RATE USING BETA DISTRIBUTION%%%
```

```
    for i=1:3
        for j=1:notp
            arr_rate(j,i)=betapdf(j/notp,Inbeta(i,1),Inbeta(i,2)) ↙
/notp;
        end
    end
    clear i j;

    %%%%%%%%%%%%%REPLICATIONS%%%%%%%%%%%%

for rep=1:trep             % Replication number

    rep                    % Show replication number on screen

    if Update==2
        up=1;              % Intialize update number
        Mainpart           % Go to Mainpart and solve the model
        BL=seat;           % Get booking limit
    end

    %%%%%%%%%GENERATE INTERARRIVAL TIMES%%%%%%%%%%%

    % Generate random gamma expected demand for all ODFs

    for i=1:6
        for j=1:3
            Aodf(i,j)=gamrnd(DDODF((i-1)*3+j,1),1/DDODF((i-1)*3+j, ↙
2));
        end
    end
    clear i j

    % Generate Lambda(t)=Beta(t).Aodf

    for k=1:notp
        for i=1:6
            for j=1:3
                arr_nhpp(k,(i-1)*3+j)=arr_rate(k,j)*Aodf(i,j);
            end
        end
    end
    clear i j k

    % Generate interarrival times with interpolation
    inter_arrival=[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0];
```

```
    for k=1:18
        y=0;
        m=0;
        for cnt=1:1000
            rnd=unifrnd(0,1,1);
            y =y-log(rnd);
            if y>Aodf(ceil(k/3),k-floor((k-1)/3)*3);
                break;
            end
            slam=0;
            for i=1:notp
                slam=slam+arr_nhpp(notp-i+1,k);
                if slam>y
                    m=m+1;
                    inter_arrival(m,k)=(i-1)+(y-slam+arr_nhpp(notp- ↙
i+1,k))/arr_nhpp(notp-i+1,k);
                    break
                end
            end
        end
    end
    clear k y m slam

    sizeint=size(inter_arrival);        % Size of the inter_arrival ↙
matrix
    rsize=sizeint(1,1)*sizeint(1,2);    % Number of cells in ↙
inter_arrival matrix

    % Set null cells to a big number

    for k=1:18
        for n=1:sizeint(1,1)
            if inter_arrival(n,k)==0
                inter_arrival(n,k)=100000;
            end
        end
    end
    clear k n

    % Initialize counters

    for i=1:TM
        ER(i)=0;                                               % Total ↙
revenue
        CAP(i,:)=CAPACITY;                                     % ↙
Capacity
        AR(i,:)=[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0];         % ↙
```

```
Accepted Requests
        reject(i)=0;                                               %↙
Rejected Requests
        acount(1,:)=[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0];  % Total↙
number of requests
    end
    clear i

    %%%%%%%%%%%%SIMULATION PART%%%%%%%%%%%%%%%

    intarr=inter_arrival;        % Copy inter_arrival matrix

    for r=1:rsize

        %----Find the type and time of the arrival

        minint=min(intarr);
        m_minint=min(minint);    % Find minimum number (first arrival↙
time) in the matrix

        if m_minint>notp         % If this value is greater than the↙
departure time of the flight, stop.
            break
        end

        for k=1:18
            if minint(1,k)==m_minint
                atype=k;
                atime=m_minint;
                for n=1:sizeint(1,1)
                    if intarr(n,k)==m_minint
                        intarr(n,k)=100000;     % Find the ODF of the↙
arrival and set cell to 10000
                        break
                        break
                    end
                end
            end
        end
        clear k n

        %----Go to bayesian update

        if Update==2
            Bayesian;
        end
```

```
        %----Nesting and decision making

        for ms=1:TM
            % Nesting Part
            B=[0 0 0];
            for b=1:D
                cb=All_convert(1,b);
                if cb == atype
                    break;
                else
                    B=B+max(BL(ms,cb)-AR(ms,cb),0)*Leg(:,cb)';
                end
            end

            for j=1:NL
                if Leg(j,atype)==1
                    Bmin(1,j)=CAP(ms,j)-B(1,j);
                else
                    Bmin(1,j)=1;
                end
            end

            if Bmin>0      % If booking limit is greater than 0, ↙
accept
                ER(ms)=ER(ms)+FF(atype);
                CAP(ms,:)=CAP(ms,:)-Leg(:,atype)';
                AR(ms,atype)=AR(ms,atype)+1;
            else
                reject(ms)=reject(ms)+1;
            end

        end
        clear ms b cb j B Bmin

        acount(1,atype)=acount(1,atype)+1;       % Count requests

    end

    for ms=1:TM
        LF(rep,ms)=(sum(INCAP)-sum(CAP(ms,:)))/sum(INCAP);     % Load↙
factor
        Rev(rep,ms)=ER(ms);↙
% Revenue
    end
    clear ms

    clear inter_arrival;
```

```
end

LoadFactor=mean(LF);     % Average of load factor
Average_Rev=mean(Rev);   % Average of revenue

toc;
```

```
% Simulation model with bid price policy and bayesian update

tic;

    %%%%%%INPUT PART%%%%%%%%%

    %----Static input part

    Input           % Get input data from optimization part

    notp=3600;      % Number of total periods (3600 periods=24 ↙
hours*150 days)

    IND=DDODF;      % Duplicate demand parameters

    Bids;           % Get bid prices from allocations (for DLP, EMR, ↙
SLP and EMVLP models)

    INCAP=CAPACITY; % Initial Capacity

    %----Dynamic input part

    TM=1;          % Total number of models will be solved in the ↙
simulation

    PM=0;           % Total number of proposed models will be solved ↙
in the simulations

    Update=1;       % 1 if bayesian update is used, 0 otherwise

    upn=3;          % Number of updates for one replication

    trep=500;      % Replication number

    %%%%%%%BID PRICES%%%%%%%%%%%

    if Update==1    % Get bid prices of proposed models
        TM=1;
        clear Bid
    else
        for i=1:PM
             Bid(i+(TM-PM),:)=bidden(i,:);
        end
    clear i;
    end

   %%%GENERATE ARRIVAL RATE USING BETA DISTRIBUTION%%%
```

```
    for i=1:3
        for j=1:notp
            arr_rate(j,i)=betapdf(j/notp,Inbeta(i,1),Inbeta(i,2)) ↙
/notp;
        end
    end
    clear i j;

    %%%%%%%%%%%%%%REPLICATIONS%%%%%%%%%%%%

    if Update==1
        Mainpart            % Go to Mainpart and solve the model
        Inbid=bid';
    end

for rep=1:trep              % Replication number

    rep                     % Show replication number on screen

    if Update==1
        up=1;               % Intialize update number
        Mainpart            % Go to Mainpart and solve the model
        Bid=bid';           % Get bid price
    end

    %%%%%%%%GENERATE INTERARRIVAL TIMES%%%%%%%%%%%

    % Generate random gamma expected demand for all ODFs

    for i=1:6
        for j=1:3
            Aodf(i,j)=gamrnd(DDODF((i-1)*3+j,1),1/DDODF((i-1)*3+j, ↙
2));
        end
    end
    clear i j

    % Generate Lambda(t)=Beta(t).Aodf

    for k=1:notp
        for i=1:6
            for j=1:3
                arr_nhpp(k,(i-1)*3+j)=arr_rate(k,j)*Aodf(i,j);
            end
        end
    end
```

```
    clear i j k

    % Generate interarrival times with interpolation
    inter_arrival=[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0];

    for k=1:18
        y=0;
        m=0;
        for cnt=1:1000
            rnd=unifrnd(0,1,1);
            y =y-log(rnd);
            if y>Aodf(ceil(k/3),k-floor((k-1)/3)*3);
                break;
            end
            slam=0;
            for i=1:notp
                slam=slam+arr_nhpp(notp-i+1,k);
                if slam>y
                    m=m+1;
                    inter_arrival(m,k)=(i-1)+(y-slam+arr_nhpp(notp- ↙
i+1,k))/arr_nhpp(notp-i+1,k);
                    break
                end
            end
        end
    end
    clear k y m slam

    sizeint=size(inter_arrival);          % Size of the inter_arrival ↙
matrix
    rsize=sizeint(1,1)*sizeint(1,2);      % Number of cells in ↙
inter_arrival matrix

    % Set null cells to a big number

    for k=1:18
        for n=1:sizeint(1,1)
            if inter_arrival(n,k)==0
                inter_arrival(n,k)=100000;
            end
        end
    end
    clear k n

    % Inıtialize counters

    for i=1:TM
```

```
        ER(i)=0;                                                   % Total↙
revenue
        CAP(i,:)=CAPACITY;                                         %↙
Capacity
        AR(i,:)=[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0];             %↙
Accepted Requests
        reject(i)=0;                                               %↙
Rejected Requests
        acount(1,:)=[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0];  % Total↙
number of requests
    end
    clear i

    %%%%%%%%%%SIMULATION PART%%%%%%%%%%%%

    intarr=inter_arrival;       % Copy inter_arrival matrix

    for r=1:rsize

        %----Find the type and time of the arrival

        minint=min(intarr);
        m_minint=min(minint);       % Find minimum number (first↙
arrival time) in the matrix


        if min(minint)>notp
            break                   % If this value is greater than↙
the departure time of the flight, stop.
        end


        for k=1:18
            if minint(1,k)==m_minint
                atype=k;
                atime=m_minint;
                for n=1:sizeint(1,1)
                    if intarr(n,k)==m_minint
                        intarr(n,k)=100000;      % Find the ODF of↙
the arrival and set cell to 10000
                        break
                        break
                    end
                end
            end
        end
        clear k n
```

```
        %----Go to bayesian update

        if Update==1
            Bayesian;
        end

        %----Decision making

        for ms=1:TM

            if FF(atype)>=Leg(:,atype)'*Bid(ms,:)'       % If bid↙
price is lower than the fare and the capacity is available then ↙
accept
                if min(CAP(ms,:)-Leg(:,atype)')>=0;
                    ER(ms)=ER(ms)+FF(atype);
                    CAP(ms,:)=CAP(ms,:)-Leg(:,atype)';
                    AR(ms,atype)=AR(ms,atype)+1;
                else
                    reject(ms)=reject(ms)+1;
                end
            end
        end
        clear ms

        acount(1,atype)=acount(1,atype)+1;       % Count requests
    end

    for ms=1:TM
        LF(rep,ms)=(sum(INCAP)-sum(CAP(ms,:)))/sum(INCAP);↙
% Load factor
        Rev(rep,ms)=ER(ms);↙
% Revenue
    end
    clear ms

    clear inter_arrival;

end

LoadFactor=mean(LF);     % Average of load factor
Average_Rev=mean(Rev);   % Average of revenue

toc;
```

```
% Bayesian update

if m_minint>((notp*up)/upn)

        % Calculate percentage of the arrivals to come in the ↙
remainig time
        beT=[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0];

        for i=1:6
            for j=1:3
                for k=1:(notp*up/upn)
                    beT(1,(i-1)*3+j)=beT(1,(i-1)*3+j)+arr_rate ↙
(notp+1-k,j);
                end
            end
        end

        % Update demand data and solve the model again

        for i=1:6
            for j=1:3
                Expdem(1,(i-1)*3+j)=(IND((i-1)*3+j,1)+acount(1,(i- ↙
1)*3+j))/(IND((i-1)*3+j,2)+beT(1,(i-1)*3+j))*(1-beT(1,(i-1)*3+j));
                DDODF((i-1)*3+j,1)=IND((i-1)*3+j,1)+acount(1,(i-1) ↙
*3+j);
            end
        end

        L=L-ER;             % Update target level
        CAPACITY=CAP;       % Update capacity
        Demandpart;         % Solve model
        Bid=bid';           % Get bid price
        BL=seat;            % Get booking limit
        AR(1,:)=[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0];  % Initialize↙
accepted requests
        up=up+1;            % Increase update number
end
```