NEW MODELS FOR SINGLE LEG AIRLINE REVENUE MANAGEMENT WITH OVERBOOKING, NO-SHOWS, AND CANCELLATIONS

by

NURŞEN AYDIN

Submitted to the Graduate School of Engineering and Natural Sciences
in partial fulfillment of
the requirements for the degree of
Master of Science

Sabancı University
Summer 2009

NEW MODELS FOR SINGLE LEG AIRLINE REVENUE MANAGEMENT WITH OVERBOOKING, NO-SHOWS AND CANCELLATIONS

APPROVED BY

Assoc. Prof. Dr. Ş. İlker Birbil ....
(Thesis Supervisor)

Assist. Prof. Dr. Nilay Noyan ....
(Thesis Co-supervisor)

Assist. Prof. Dr. Güvenç Şahin ....

Assist. Prof. Dr. Fatma Tevhide Altekin ....

Assist. Prof. Dr. Dilek Tüzün Aksu ....

DATE OF APPROVAL: ....

to my family

Acknowledgments

I would like to express my sincere gratitude to my thesis advisor, Associate Professor Ş. İlker Birbil for his invaluable support, encouragement, supervision and useful suggestions throughout this research. His moral support and continuous guidance enabled me to complete my work successfully.

I would like to thank my thesis co-advisor, Assistant Professor Nilay Noyan whose enthusiasm has guided me through all my research. Her inspiration and sincerity always motivates me. I would also like to thank Associate Professor Hans Frenk for his guidance, sincerity and crucial contribution, which made him a backbone of this research and of this thesis. Without my advisors' guidance and support, completion of this thesis would be impossible. Their motivation and encouragement will always guide me throughout my professional career.

I am grateful to all my friends from Sabancı University for being the surrogate family during the two years I stayed there. I would like to express my heartiest thanks to Taner, Figen, Birol, Belma, Burak, Serkan, İbrahim, Özlem and Merve. I am also grateful to my colleague and friend, Gamze Belen for her trust, motivation and endless friendship. I especially want to thank Mehmet Karaca for his invaluable support, understanding and help with the simulations.

I would like to thank TÜBITAK for providing me the financial support.

Lastly, I would like to thank my parents and sister for their concern, caring, endless love and support they have provided throughout my life. Without them, I would not succeed.

NEW MODELS FOR SINGLE LEG AIRLINE REVENUE MANAGEMENT WITH OVERBOOKING, NO-SHOWS, AND CANCELLATIONS

Nurşen Aydın

Industrial Engineering, Master of Science Thesis, 2009

Thesis Supervisors: Assoc. Prof. Dr. Ş. İlker Birbil
Assist. Prof. Dr. Nilay Noyan

Keywords: single-leg, airline revenue management, seat allocation, overbooking.

Abstract

Airline revenue management (ARM) problem focuses on finding a seat allocation policy, which results in the maximum profit. Overbooking has been receiving significant attention in ARM over the years, since a major loss in revenue results from cancellations and no-shows. Basically, overbooking problem aims at maximizing the profit by minimizing the number of vacant seats. However, this problem is difficult to handle due to the demand and cancellation uncertainties and the size of the problem. In this study, we propose new models for the static and the dynamic overbooking problems.

Due to the complex analytical form of the overbooking problem, in the static case we introduce models that give upper and lower bounds on the optimal expected profit. In the dynamic case, however, we propose a new dynamic programming model, which is based on two streams of arrivals; one for booking and the other one is for cancellation. This approach allows us to come up with a computationally tractable model. We also present numerical results to show the effectiveness of our models.

KAPASİTE ÜSTÜ REZERVASYONU, GELMEYENLERİ VE İPTALLERİ İÇEREN TEK BACAK HAVAYOLLARI GELİR YÖNETİMİ İÇİN YENİ MODELLER

Nurşen Aydın

Endüstri Mühendisliği, Yüksek Lisans Tezi, 2009

Tez Danışmanları: Doç. Dr. Ş. İlker Birbil
Yrd. Doç. Dr. Nilay Noyan

Anahtar Kelimeler: tek bacak, havayolları gelir yönetimi, kapasite dağılımı, kapasite üstü rezervasyon.

Özet

Havayolları gelir yönetimi problemi maksimum karı sağlayacak kapasite dağıtım politikasını bulmaya odaklanmıştır. Gelir kaybının en büyük kısmı iptaller ve gelmeyen yolculardan kaynaklandığı için kapasite üstü rezervasyon stratejisi havayolları gelir yönetiminde senelerdir önemli bir yer tutmaktadır. Kapasite üstü rezervasyonu içeren havayolları gelir yönetimi problemi temel olarak boş koltuk saysını emküçükleyerek en büyük geliri elde etmeyi amaçlamaktadır. Ancak, talep-iptal belirsizlikleri ve problem büyüklüğü gibi nedenlerle bu problemi çözmek oldukça zordur. Bu çalışmada, statik ve dinamik kapasite üstü rezervasyon problemleri için çeşitli modeller önerilmiştir.

Kapasite üstü rezervasyon probleminin karmaşık analitik yapısından dolayı, statik problemde en iyi beklenen kar için alt ve üst sınırları belirleyen yeni modeller sunulmuştur. Dinamik problemde ise rezervasyon ve iptaller için iki akış temelli dinamik programlama modeli ortaya konulmuştur. Bu yaklaşım bize çözülebilir bir model sağlar. Modellerin etkinliğini gösteren sayısal örnekler de çalışmada sunulmuştur.

Table of Contents

Abstract **vi**

Özet **vii**

1 INTRODUCTION **1**
1.1 Motivation . . . 3
1.2 Contributions . . . 5
1.3 Outline . . . 5

2 LITERATURE REVIEW **6**
2.1 General Single-Leg ARM without Overbooking . . . 7
2.1.1 General Static Single-Leg Problem . . . 7
2.1.2 General Dynamic Single-Leg Problem . . . 8
2.2 Single-Leg ARM with Overbooking . . . 9
2.2.1 Static Overbooking Problem . . . 9
2.2.2 Dynamic Overbooking Problem . . . 10
2.3 Robust Single Leg Problem . . . 12

3 NEW MATHEMATICAL MODELS **13**
3.1 Problem Statement . . . 13
3.2 Notation . . . 14
3.3 Static Overbooking Problem . . . 23
3.3.1 Model With No Booking Limit For Each Fare Class . . . 23
3.3.2 Model With Booking Limit For Each Fare Class . . . 27
3.4 Dynamic Overbooking Problem . . . 30

4 SOLUTION APPROACHES **36**
4.1 Static Models . . . 36
4.2 Dynamic Model . . . 38

5 COMPUTATIONAL RESULTS **40**
5.1 Static Models . . . 40
5.2 Dynamic Models . . . 46

6 CONCLUSION AND FUTURE WORK **52**

Bibliography **54**

List of Figures

5.1 Truncated probability distributions for different fare classes 42
5.2 The relative difference between the objective function values of problem P_I^{UB} and problem P_I^{LB} with respect to β 43
5.3 The relative difference between the objective function values of problem P_I^{UB} and problem P_I^{LB} with respect to α 44
5.4 Overbooking 45
5.5 Overbooking amount for different values of show-up and overbooking capacity 46
5.6 The change of adopted Dirichlet distribution parameters over time . . . 47
5.7 An example of the change of cancellation probabilities over time and the number of reserved seats 48
5.8 $J_t(n)$ versus n for different t values 49
5.9 The relative difference between the objective function values of the dynamic model and static upper bounding problem P_I^{UB} for varying β . . 50
5.10 The relative difference between the objective function values of the dynamic model and static upper bounding problem P_I^{UB} for varying s parameter 51

List of Tables

5.1 Parameters used in the truncated Poisson distribution 42
5.2 Parameters used in the simulation for β 43
5.3 Parameters used in the simulation for α 43
5.4 Parameters used in the simulation of static problem P_I^{LB} 44
5.5 Parameters used in the simulation of dynamic model 47
5.6 Parameters used in the simulation of dynamic and static models 50
5.7 Parameters used in the simulation of dynamic and static models 51

CHAPTER 1

INTRODUCTION

Airline revenue management (ARM) has been one of the most successful application areas of operations research. With the Airline Deregulation Act in 1979, US Civil Aviation Board (CAB) loosened control of airline prices and new low-cost and charter airlines entered the market. Therefore, airline companies began to explore the ways of competing effectively, and different approaches in revenue management (RM) evolved. As world aviation markets were increasingly deregulated, airline companies began to adopt ARM. Following the success at the airlines, RM has been applied by a number of other industries, including hotels, car rental agencies, rail transportation, and cruise lines. Development and investment in RM continues in many of these industries today. We refer the reader to Talluri and Van Ryzin [34] and Phillips [27] for a comprehensive overview on revenue management.

ARM is the strategy of managing the available capacity among different fare classes over time in order to maximize the revenue. It is basically concerned with the demand-management decisions. In particular, it is related to setting and updating the availability of the fare classes. ARM decisions are executed at three levels. Strategic level includes identification of customer segments and establishment of products targeted at those segments. The key point is to make a distinction between leisure and business customers. Tactical level requires determining and updating reservation limits on the seat capacity of a particular fare class. It is the key point of the process. Such tactical decisions include capacity allocation and overbooking related decisions. Lastly, booking control level determines which bookings to accept and which to reject. It is a function of the reservation system. Such operational level decisions may depend on the time and the characteristics of the request, the segment of customer, or the combinations of all of these.

The techniques applied in ARM have been expanded with the increasing competition among the airlines. Capacity allocation and overbooking are two main strategies of ARM. While capacity allocation is important for determining the number of seats reserved for booking for different fare classes, overbooking plays a critical role in setting the total bookings that should be accepted for a fare class in the face of uncertain no-shows and cancellations.

Capacity allocation is the problem of determining how many seats are to be booked for different fare classes. In this case, the critical decision is optimal control of the seat inventory. If tickets are sold on a first-come first-serve basis, the flight capacity is going to fill up with low fare customers and the airline loses the potential high fare passengers. By imposing booking limits on fare classes, potential revenue losses can be avoided. However, inefficient booking limits may result in unsold seats. Therefore, determining how many seats is crucial and we refer to this problem as the capacity allocation (seat allocation) problem.

Typically, the booking limits are determined at the beginning of the booking process based on demand estimates. The basic capacity allocation problem is concerned with the allocation of the limited seats to the demand that occurs over time before the flight departure. A straightforward mathematical model of the seat allocation problem is,

$$\begin{aligned} \max \quad & \sum_{i=1}^{m} r_i min\{x_i, d_i\} \\ \text{s.t} \quad & \sum_{i=1}^{m} x_i \leq C, \\ & x \in Z_+^m, \end{aligned} \tag{1.1}$$

where x_i denotes the number of reserved seats for fare class i, $1 \leq i \leq m$ and d_i is the realized demand for that class. Here, r_i is the price of fare class i seat, and we assume without loss of generality that $r_1 > r_2 > ... > r_m$. The objective is to find the optimal allocation of the seat capacity C such that the revenue is maximized. It is clear that an optimal allocation policy is given as follows. By considering the demand for each fare class, we reserve all the seats for the higher priced classes as long as the capacity is still available. However, this simple model ignores the demand uncertainty. Since the demand realization cannot be known in advance, we may estimate its distribution. Let $D_i(\omega)$ be the realization of demand d_i. Then, the total revenue is given by $\sum_{i=1}^{m} r_i min\{x_i, D_i(\omega)\}$ and the expected revenue equals $\sum_{i=1}^{m} r_i E[min\{x_i, D_i\}]$. Consequently, the seat allocation problem for random demand is given by

$$
\begin{aligned}
\max \quad & \sum_{i=1}^{m} r_i E[min\{x_i, d_i\}] \\
\text{s.t} \quad & \sum_{i=1}^{m} x_i \leq C, \\
& x \in Z_+^{\cdot}
\end{aligned}
\tag{1.2}
$$

This problem is first formulated by Wollmer [37] and it is a standard separable problem which can be solved by dynamic programming (see also [6]).

1.1 Motivation

In ARM problem, while the revenue consideration tends to protect seats for potential high-fare passengers, it is important to utilize as many seats as possible. Solving the capacity allocation problem effectively does not guarantee the maximum possible revenue in practice due to the fact that passengers with reservations may cancel or not show up at the departure time. If airlines strictly stick to the physical flight capacity, they will experience revenue loss due to the cancellations and no-shows. Therefore, many airlines try to compensate for such losses by a strategy called overbooking. Overbooking is the policy of accepting more reservations than the physical capacity to protect themselves against unanticipated cancellations or no-shows. Without overbooking, many of the cancellations and each one of the no-shows would result in an empty seat. As a result, the airline would not only lose potential revenue, it would also maintain and support huge amounts of useless capacity. Current statistics on major airlines depict that the average no-show rate is around 8% [2]. Smith et al. [32] have estimated that 50% of American Airlines reservations resulted in either a cancelation or a no-show in 1992. More recently, US Airways have reported that there would have been a loss of $1 billion revenue in 2007 if the airline had not been overbooked.

Considering the overbooking issue poses an additional challenge, since the airline company incurs a penalty if there is not an available seat for an overbooked passenger. Basically, taking into consideration the overbooking requires decisions on the number of physically available seats on a flight leg that are allowed to be oversold given that the accepted requests may not show up at the departure time. Therefore, the optimal overbooking level should balance lost revenue due to empty seats with the penalties and loss of customer goodwill when the firm faces more demand than the physical capacity. When a flight is oversold, the airline will rebook some customers on a later flight which is called bumping. If the flight is much later, the bumped passengers are provided with a meal and overnight accommodation when necessary. In addition, a penalty charge is paid to each bumped passenger.

There are mainly four overbooking policies:

1. *Deterministic Heuristic*: It calculates booking limit with respect to the ratio of capacity and show-up rate.

2. *Risk-Based Policy*: It estimates the cost of denied service and then weighs that cost against the expected revenue to determine booking levels that maximize the expected profit.

3. *Service-Level Policy*: It determines the booking limits according to a specific target such as aiming no more than one denied service for every 500 shows.

4. *Hybrid Policy*: It joins risk-based policy and service-level policy. Risk-based limits are calculated according to service-level constraints.

Although companies differ in their overbooking policies, they mainly aim to maximize revenue by setting booking limits, which balance the expected cost of denied service with the potential additional contribution from more sales. It is clear that how many seats to offer to different fare classes on a flight depends on how much the airline is willing to overbook; thus, capacity allocation and overbooking decisions are interrelated. Therefore, the problem of determining optimal booking limits for multiple fare classes is the combination of overbooking and capacity allocation problem, and it is extremely difficult to solve in general. Since 1950, various solution methodologies have been developed to solve this problem. After now, we refer to capacity allocation with overbooking as overbooking problem. Although overbooking has been studied over the years, the proposed models generally simplify the problem. Therefore, they do not represent the actual overbooking problem in real world. These simplifications can be classified as follows: To deal with the uncertainty some models ignore the cancellation or no-show penalties [10, 13, 17]. On the other hand, some studies use historical data to model randomness. However, they can only be applied to a specific overbooking problem [31,35,36]. To reduce the size of the problem some models consider one or two fare classes and they generally focus on only overbooking costs, instead of the revenue contribution from different fare classes [1, 4, 29, 33, 36]. Furthermore, in dynamic overbooking problem it is generally considered that arrival and cancellation probabilities depend on the number of reservations [9, 33]. However, one may argue that no arrival and arrival probabilities should be independent of the number of the reserved seats.

1.2 Contributions

In this research, we firstly review and discuss the overbooking problem in detail. This review allows us to position our work in the literature. Then, we propose new mathematical models for static and dynamic single-leg problems. Our models consider overbooking, no-shows and cancellations. Considering the class dependency and demand uncertainty, the proposed approaches provides us realistic ARM policies. In the static case, several models that examine the overbooking problem from different angles are introduced. Since the static overbooking problem has a very complex structure, we present new models, which give upper and lower bounds on the optimal expected profit. On the other hand, in the dynamic case we propose a new dynamic programming model which is based on two streams of arrivals: one for booking and the other one for cancellation. We relax the unrealistic assumption used in the literature and allow cancellation and no-show probabilities depend on the total number of already booked seats.

1.3 Outline

The outline of the thesis is as follows. Chapter 2 gives a literature review of the airline revenue management problem with a particular emphasis on overbooking. This literature review is followed in Chapter 3 by the introduction of the static and dynamic models. Solution approaches to the proposed models are given in Chapter 4. Chapter 5 presents the computational study. We end this thesis with Chapter 6, which contains the conclusion and the planned future work.

CHAPTER 2

LITERATURE REVIEW

Despite decades of research and practice, the airline revenue management is still a challenging research area. One stream of papers on ARM develop leg-based models, whereas the other stream of papers focuses on network-based models. While the leg-based models aim at optimizing the passenger mix on a single-leg flight, network models find similar optimal decisions when booking requests for multiple legs are considered simultaneously. Although many practical overbooking problems observed in the airline industry are network based, single leg problems still play an important role. Because network based airline problems generally require solving many single leg problems and some small airline companies have special one-hub networks with single legs and so they only need to solve single leg problems.

Before reviewing the relevant literature on the single-leg airline revenue management, we list some common terminology [34]:

- **Flight leg:** A nonstop flight.
- **Standby:** A passenger without a reservation who wants to get on a flight and waits at the airport for last minute openings.
- **No-show:** A passenger who has booked a ticket on an airline flight but does not show up for the departure.
- **Go-show:** A passenger without a reservation who has been assigned to an empty seat right before the departure.
- **Overbooking pad:** The number of seats to be overbooked on a given flight.
- **Authorization level:** The maximum number of bookings allowed for each fare class.

2.1 General Single-Leg ARM without Overbooking

General single-leg problem without overbooking can be defined as the allocation of limited seats to the demand. This problem is also known as the capacity allocation problem. Most of the research on ARM focuses on the capacity allocation problem. This problem can be examined using two models; static and dynamic. Static models focus on setting booking limits on different fare classes at the beginning of the booking period. It is assumed that reservation requests come in sequentially in order of increasing fare level. For instance, low fare booking requests come in before high fare booking requests. This means that booking control policies can be based on the total demand for each fare class and they do not need to consider the actual arrival process. On the other hand, dynamic models monitor the actual demand over the booking period and decide whether to accept a reservation request at its arrival time. There is no assumption on the arrival order of the booking requests.

2.1.1 General Static Single-Leg Problem

The static capacity allocation problem is firstly addressed by Littlewood [24]. He has formulated a two fare class model to determine the booking limits. The idea behind his model is closing the low fare class when the revenue from selling another low fare class seat is less than the expected revenue of selling the same seat at a higher fare. In other words, he determines a protection level for high fare class.

Belobaba [5] extends Littlewood's model to a multi-class problem and introduces the method of expected marginal seat revenue (EMSR) for the general approach. However, this method can generate optimal booking limits only for the two fare class problem. Curry [12], Wollmer [37], and Brumelle and McGill [8] work on EMSR method and obtain optimal policies for the multi-class static problem. Furthermore, Curry proposes an approach to deal with multi flight legs, when the capacities are not shared among different origin-destinations. He models the problem with continuous time dynamic programming formulation with a recursive equation for the optimal value functions. Wollmer presents a method to find a seat allocation policy by establishing a critical value for each fare class as EMSR method does. Different than Belobaba's method, he uses discrete demand distributions. Brumelle and McGill formulate a model which is capable of handling both continuous and discrete demand, and they show that under certain conditions the optimal allocation can be found by equating the marginal revenues in the various fare classes. All of these proposed static models result in optimal policies under the assumption that reservation requests arrive in the order of increasing fare class prices. However, Robinson [28] makes no assumption on the order of the demand arrival and he obtains close approximations to the optimal policy by using Monte Carlo integration.

On the other hand, Van Ryzin and McGill [38] develop a simple adaptive approach to find protection levels for the multi class problem, which does not need any demand forecasting. The method adjusts protection levels by using historical observations. They show that under reasonable regularity conditions, the algorithm converges to the optimal protection levels. Since this method does not need any demand forecasting, it is a way to remove all of the difficulties involving demand uncertainty. However, to get a good approximation of the protection levels, the updating procedure requires a large sequence of flights to obtain sufficient historical data.

2.1.2 General Dynamic Single-Leg Problem

Brumelle and McGill [8] demonstrate that under the low-to-high fare arrival assumption, a static solution method is optimal as long as no change is realized in the probability distributions of demand. However, dynamic solution methods do not assume a specific arrival order of the booking requests and they do not determine a booking policy at the start of the reservation period like the static methods do. Instead, they observe the state of the system over time and decide whether to accept a particular request when it arrives. In this case, a booking policy based on the total demand for each fare class is not optimal, and dynamic programming methods are required.

Lee and Hersh [22] propose a discrete-time dynamic programming model where the demand for each fare class is modeled by a nonhomogeneous Poisson process; they formulate the problem as a Markov decision process (MDP). In this model, the reservation period is divided into decision periods. These decision periods are chosen sufficiently small to allow only one arrival within a period and the state of the system changes when a decision period ends. In each period, a reservation request is accepted if its fare is higher than expected marginal revenue of the seat. The model determines the optimal capacity level for each fare class. Liang [23] reformulates and solves the model in Lee and Hersh in continuous-time. On the other hand, Kleywegt and Papastavrou [18] present a different approach. They show that the problem can also be formulated as a dynamic and stochastic knapsack problem (DSKP). Their model considers cost of unused capacity and penalty for rejection different than the other existing models.

Lautenbacher and Stidham [21] combine the dynamic and static approaches under a common MDP formulation which yields both models as special cases. The proposed model allows passenger arrival throughout the booking period. They first develop a dynamic model and then make the necessary adjustments for the static model.

2.2 Single-Leg ARM with Overbooking

By the end of 1950s, no-shows were becoming a major problem since customers were allowed to cancel or became no-show without penalty. In 1961, CAB estimated a no-show rate of 1 out of 10 passengers [34]. To deal with the economical consequences, airlines have been allowed to overbook. There is a huge literature on overbooking problem and we focus on the most related work. Again, the types of models within this field can be classified into static models and dynamic models.

2.2.1 Static Overbooking Problem

Before the Littelewood's research [24] on capacity allocation problem, almost all research on ARM focusses on determining overbooking pad without considering contribution of fare classes. These models either take a cost-based approach that controls the expected number of denied passengers or a sevice level-based approach that controls the probability of denied boarding within limits set by airline regulators. The first scientific work on overbooking appeared in 1958 by Beckman [4]. Beckman works on the single-leg one fare class cost-based problem. He presents a simple static overbooking model, which determines the total overbooking by balancing the lost revenue due to empty seats with the cost of denied boardings. A more implementable model was published by Thompson [36], which entirely ignores the probability distribution of demand and requires only data of cancellation proportions out of any fixed number of reservations. He tries to determine the risk of overselling by allowing extra bookings in two fare classes problem or, alternatively, determine the extra bookings in two fare classes so that the overselling risk equals some pre-assigned value. His work has been examined and refined by Taylor [35], and Rothstein and Stone [31]. They have developed models for certain airline problems by using British European Airways data. Therefore, they face problems in estimating the parameters for more general models. Rothstein [30] also presents a study on the history of overbooking in the airline industry.

Later work on the static overbooking problem was published by Bodily and Pfeifer [7]. They give optimal decision rules for overbooking in single fare class problem. They have developed a generic decision rule, then adapted it to specific models. As in the model proposed by Beckman, they trade off between the number of unutilized seats and the number of denied customers. The decision rule they have developed maximizes the expected payoff including the cost of both spoilage and oversales. These rules provide insights to the user. Probabilities for spoilage is assessed subjectively at each decision point by using the number of reserved seats and previous experience.

Chi [10] formulates multi class static overbooking problem as a dynamic programming model. Given the flight capacity, fares and the distributions of demand, he derives the maximum number of seats allowed for the lowest open fare class. In this model, fare classes constitute the stages and he assumes that the demand for the lowest fare occurs first and the booking for a class starts, if all the bookings are made for the lower classes. In addition, to simplify the model he assumes that reservations can be canceled without any penalty. He proposed an approximate dynamic programming algorithm as a solution method which provides near-optimal solutions. Coughlan [11] has also studied the overbooking problem in the multi-class case and tried to determine overbooking levels for each fare class. In this model, the empty seats are considered right before the plane departure and go-shows are allocated to empty seats with the same ticket price. To simplify the model, Coughlan [11] assumes that the demand and the number of bookings in each fare class are independent and normally distributed. However, in literature it is common to assume that demand follows a Poisson distribution. Furthermore, they assume that the minimum of demand and the number of bookings are also independently normally distributed. In addition, he assumes that the number of go-shows in any class is independent of the number of show-ups in that class. This assumption may not be valid in practice, since the number of show-ups limits the number of go-shows. As a solution method, he proposes direct search methods, Hooke and Jeeves [16] and Nelder and Mead [26]. He tests his model only for three fare classes and states that the solution procedures employed do not guarantee optimality.

2.2.2 Dynamic Overbooking Problem

Several researchers have addressed dynamic overbooking models for a single-leg flight. Generally, dynamic overbooking problem is modeled as a MDP. Rothstein [29], Alstrup et al. [1] and Subramanian et al. [33] are three examples that use Markov decision processes. Rothstein [29] has formulated the one fare-class overbooking problem and has constructed a general model for determining overbooking policies. The number of reservations is the state space of the system and the system changes the state space according to time-dependent transition probabilities. In order to simplify the model, he assumes that cancellation probabilities are independent of the number of reserved seats. However, this assumption may not be realistic. On the other hand, Alstrup et al. [1] have developed a dynamic-programming approach to solve an overbooking model with two fare classes by extending [29]. The objective is to determine the optimal allocation of seats such that the difference between the maximal obtainable gain and the actual gain is maximized. Different than Rothstein [29], they also consider the cost of transferring higher fare-class passengers to lower fare-class passengers (downgrading). As a solution method, two dimensional stochastic dynamic programming has been used. However, the dynamic programming treatment of overbooking grows exponentially in size and becomes computationally burdensome for real-world problems.

Subramanian et al. [33] formulate the multi class overbooking problem as a discrete-time MDP without making any assumption on the arrival pattern. They extend the model of Lee and Hersh [22] that does not permit overbooking. The state space is the number of reserved seats in each fare class. Although, they propose a model which has class dependent cancellations and no-shows, their model can be applied only to small size problems. Due to the computational intractability, they try to reduce the size of the state space into one dimension. In the proposed model, it is assumed that only an arrival, a cancellation or a null event can be realized at each stage. Furthermore, no arrival, cancellation and arrival probabilities are assumed to depend on the number of reserved seats. However, one may argue that no arrival and arrival probabilities should be independent of the number of reserved seats. In our model, while cancellation probabilities depend on the number of bookings, arrival and no arrival probabilities are independent of the current bookings. Chatwin [9] formulates the problem as a birth-and-death process and proposes two models. While model he ignores refunds and no-show penalties in the first, in the second model, he allows that the refunds and the fares may vary over time. The state of the system is the number of reservations on hand. He assumes that customers cancel their reservations independently according to an exponential distribution with a common rate and the number of reservation requests is dependent on the number of current bookings as Subramanian et al. [33]. As mentioned before, it may be preferable to relax the latter assumption that the demand depends on the number of reserved seats.

Karaesmen and Van Ryzin [17] examine the problem differently. Their model also permits that classes can substitute for one another. They formulate the problem as a two-period problem. In the first period, reservations are accepted and in the second period, cancellations are realized and reserved seats allocated to the classes. The problem is then to decide how many additional reservation requests to accept. In the service period, after the cancellations and no-shows all remaining customers are either assigned to a class seat or are denied. The assignment to a class is modeled as a network flow problem. In this formulation, they assume that the service provider decides allocation with perfect knowledge of the number of survivals in each class. As a solution method, they propose a stochastic gradient algorithm.

All of these proposed models try to find optimal booking limits for fare classes, assuming the aggregate overbooking limit is pre-determined. However, Feng and Xiao [13] treat overbooking upper bound as a decision variable and derive its optimal level. This model is closely related to the model present by Chatwin [9] but they do not consider cancellations. In addition, they extend it by taking fare dependent no-show rates and refunds. It is assumed that the overbooking penalty is an increasing and uniformly convex function of the number oversold seats. This uniform convexity means that extra overbooking decreases revenues. They show that the expected revenue increases with the overbooking level up to a certain point and then remains the same.

2.3 Robust Single Leg Problem

In seat allocation and overbooking models, probability distributions are used to model uncertainty in demand and cancellations. These probabilities are usually estimated by analyzing the historical data, and hence, they are prone to inaccuracies. Robust model takes into account the inaccurate estimate of the probability distributions. Recent research in ARM discusses the availability of information. Adaptive methods are used to find optimal booking limits with limited information. In each iteration booking limits are updated with respect to historical observations [19, 20]. Ball and Queyranne [3] use online algorithms to solve robust problem and present closed-form optimal booking limits. Lan et al. [20] formulate the robust problem by assuming that demand in each fare class lies in a given interval. Birbil et al. [6] present the robust version of classical static and dynamic single leg problem which considers the inaccuracies associated with estimated probability distributions of the demand for different fare classes.

CHAPTER 3

NEW MATHEMATICAL MODELS

The overbooking problem has been studied in the literature since the 1950's. In the classical single-leg problem, information needed for the state of the system is the number of seats still available. However, with cancellations, no-shows and overbookings, we will need to monitor how many seats are already booked for each fare class. This makes the problem more complex and difficult to solve. Therefore, some of the proposed models have examined the problem by simplifying some of the components. Others present heuristics or approximation algorithms as a solution method. In this study, we propose new mathematical models for the static and dynamic single-leg problems that cover overbooking, no-shows and cancellations. The proposed approaches seek to achieve a better ARM policy by considering class dependency and using realistic probability distributions to optimize overbooking. Usually demand and cancellations are random variables and we do not know in advance their realizations. In order to deal with this uncertainty, we need to know the probability distributions of these demands and cancellations which are consistent with the system.

3.1 Problem Statement

Early studies on the overbooking problem propose a cost or service level based approach [34]. They either try to balance overbooking cost with the cost of empty seats or limit the probability of denied service [4, 31, 35, 36]. Later studies focus on finding optimal booking limits by maximizing revenue. Generally, overbooking models aim to find optimal booking limits in each fare class by reducing the revenue losses resulting from no-shows or cancellations.

The overbooking problem can be defined as follows. Consider a flight with a known seat capacity C. The airline operator may overbook passengers with a corresponding penalty s up to an overbooking limit C'. If d_i is the demand for fare class i it is assumed in the our basic model that passengers can reserve one of the m fare classes with probability $p_i, 1 \leq i \leq m$. In the other models the demand assumptions are more complex. r_i is the price of a fare class i seat, $1 \leq i \leq m$, and without loss of generality $r_1 > r_2 > ... > r_m$. In our model, passengers who already have a reservation may cancel at any time or become a no-show customer at the departure time of the flight. In case of cancellation, the airline company refunds those passengers with a percentage $\alpha_i, 1 \leq i \leq m$ of the corresponding fare class i ticket price. While in static models cancelation and no-shows are considered together, dynamic models due to their dynamic nature relax this assumption. The objective is now to determine the optimal allocation of overbooking capacity of each fare class in terms of expected revenue maximization.

3.2 Notation

Before explaining the proposed models of this system, we introduce some notation. We first need to define a Bernoulli selection type random variable to model demand and list therefore the following convention. If $\mathbf{X}$ and $\mathbf{Y}$ are random variables, then $\mathbf{X} =^d \mathbf{Y}$ means that the cumulative distribution functions of $\mathbf{X}$ and $\mathbf{Y}$ are the same. Also, if $\mathbf{X}$ denotes the nonnegative integer random size of a population, then the random variable $\mathbf{B}(p, \mathbf{X})$ denotes the total number within the population of size $\mathbf{X}$ having a certain property under the condition that each member in the population has this property with probability p independent of each other. Hence the random variable $\mathbf{B}(p, \mathbf{X})$ is given by

$$\mathbf{B}(p, \mathbf{X}) := \begin{cases} \sum_{i=1}^{\mathbf{X}} \mathbf{1}_{\{\mathbf{U}_i \leq p\}} & \text{if } \mathbf{X} \geq 1 \\ 0 & \text{if } \mathbf{X} = 0 \end{cases} \tag{3.1}$$

with $\mathbf{U}_n, n \in \mathbb{N}$ a sequence of independent standard uniformly distributed random variables and the random variable $\mathbf{X}$ is independent of the sequence $\mathbf{U}_n, n \in \mathbb{N}$.

An alternative way to look at this random variable is the value of a simple increasing random walk with probability $1 - p$ of staying in the same state and p of moving one upwards evaluated at the random time $\mathbf{X}$. Clearly for every $j \leq k$

$$\mathbb{P}(\mathbf{B}(p, \mathbf{X}) = j | \mathbf{X} = k) = \binom{k}{j} p^j (1-p)^{k-j} \tag{3.2}$$

and so with $q_k := \mathbb{P}(\mathbf{X} = k)$ it follows that

$$\mathbb{P}(\mathbf{B}(p, \mathbf{X}) = j) = \sum\nolimits_{k=j}^{\infty} \binom{k}{j} p^j (1-p)^{k-j} q_k = p^j \sum\nolimits_{k=j}^{\infty} \binom{k}{j} (1-p)^{k-j} q_k \tag{3.3}$$

In general, this distribution might be difficult to compute unless $\mathbf{X}$ has a particular distribution function. Also by relation (3.1) we simply obtain

$$\mathbb{E}(\mathbf{B}(p, \mathbf{X})) = p\mathbb{E}(\mathbf{X}). \tag{3.4}$$

The following result is well known. For completeness we give its proof.

Lemma 1 *The probability generating function of the random variable* $\mathbf{B}(p, \mathbf{X})$ *is given by*

$$\mathbb{E}(z^{\mathbf{B}(p,\mathbf{X})}) = \mathbb{E}\left((1-p+pz)^{\mathbf{X}}\right)$$

and $\mathbf{B}(q, \mathbf{B}(p, \mathbf{X})) =^d \mathbf{B}(pq, \mathbf{X})$ *for any* $0 \leq p, q \leq 1$

Proof. To compute the generating function of the random variable $\mathbf{B}(p, \mathbf{X})$ we observe by the conditional expectation formula that

$$\mathbb{E}\left(z^{\mathbf{B}(p,\mathbf{X})}\right) = \sum\nolimits_{k=0}^{\infty} \mathbb{E}(z^{\mathbf{B}(p,\mathbf{X})} | \mathbf{X} = k) \mathbb{P}(\mathbf{X} = k). \tag{3.5}$$

By relation (3.1) we obtain for $k \geq 1$

$$\mathbb{E}(z^{\mathbf{B}(p,\mathbf{X})} | \mathbf{X} = k) = \mathbb{E}\left(z^{\mathbf{B}(p,k)}\right) = (1-p+pz)^k.$$

This shows by relation (3.5)

$$\mathbb{E}(z^{\mathbf{B}(p,\mathbf{X})}) = \mathbb{E}\left((1-p+pz)^{\mathbf{X}}\right) \tag{3.6}$$

and the first result is proved. To show the second result we observe by the first part that

$$\mathbb{E}(z^{\mathbf{B}(q,\mathbf{B}(p,\mathbf{X}))}) = \mathbb{E}\left((1-q+qz)^{\mathbf{B}(p,\mathbf{X})}\right) = \mathbb{E}((1-p+z^*p)^{\mathbf{X}})$$

with $z^* = 1-q+qz$. Since $1-p+z^*p = 1-pq+zpq$ we obtain

$$\mathbb{E}(z^{\mathbf{B}(q,\mathbf{B}(p,\mathbf{X}))}) = \mathbb{E}(z^{\mathbf{B}(pq,\mathbf{X})})$$

and the second result is proved. □

In the remainder of this thesis we need the following class of functions.

Definition 1 *A function $f : \mathbb{Z}_+ \mapsto \mathbb{Z}$ is discrete concave if and only if the differences $n \mapsto f(n+1) - f(n)$ are decreasing. It is called discrete convex if the function $-f$ is discrete concave. If the function f is both discrete concave and discrete convex it is said to be discrete linear.*

We will now collect some results related to discrete concavity which are needed in the next sections. The first result is also listed in [25]. For completeness we give a proof.

Lemma 2 *If the function $\mu : \mathbb{Z}_+ \mapsto [0, 1]$ is a discrete concave function and the function $f : \mathbb{Z}_+ \mapsto \mathbb{R}$ is a nonincreasing discrete concave function then the function $g : \mathbb{Z}_+ \mapsto \mathbb{R}$ given by*

$$g(n) = \mu(n)f(n-1) + (1-\mu(n))f(n)$$

is a nonincreasing discrete concave function.

Proof. We need to show that the function $n \mapsto g(n) - g(n+1)$ is increasing. Introducing the increasing sequence $d(n) = f(n) - f(n+1) \geq 0$ it follows that

$$g(n) - g(n+1) = \mu(n)d(n-1) + (1-\mu(n+1))d(n)$$

Clearly $g(n) - g(n+1) \geq 0$ and so g is nonincreasing. It follows now after some computations that the difference $p(n) := (g(n) - g(n+1)) - (g(n+1) - g(n+2))$ equals

$$p(n) = a_1(n) + a_2(n) + a_3(n)$$

with

$$a_1(n) = d(n)(\mu(n+2) - 2\mu(n+1) + \mu(n)) \leq 0$$

and

$$a_2(n) = (d(n-1) - d(n))\mu(n) \leq 0$$

and

$$a_3(n) = (1 - \mu(n+2))(d(n) - d(n+1)) \leq 0$$

Hence the function p is nonpositive and this means that the function $n \mapsto g(n+1)-g(n)$ is decreasing. $\square$

The next result is also proved in [25].

Lemma 3 *If the function* $f : \mathbb{Z}_+ \mapsto \mathbb{R}$ *is discrete concave, then the function* $h : \mathbb{Z}_+ \mapsto \mathbb{R}$ *given by*

$$h(n) = \begin{cases} \max\{r + f(n+1), f(n)\} & \textit{if } n \in \mathbb{N} \\ f(0) & \textit{if } n = 0 \end{cases}$$

is also discrete concave.

Proof. The function $f : Z_+ \to R$ is discrete concave, so this function is increasing on $\{0, ..., n^*\}$ and decreasing on $\{n^*, n^* + 1, ..., \}$ with

$$n^* = arg\ max f(p) : p \in Z_+.$$

Hence it follows for every $n \in \{0, ..., n^* - 1\}$ that

$$f_1(n) := \begin{cases} \max\{f(n+1), f(n)\} = f(n+1) & \text{if } 1 \leq n \leq n^* - 1 \\ f(0) & \text{if } n = 0 \end{cases}$$

and for every $n \in \{n^*, n^* + 1, ..., \}$ that

$$f_1(n) = max\{f(n+1), f(n)\} = f(n)$$

Since f is discrete concave this implies for x belonging to $\{1, ..., n^* - 1\}$ that

$$\begin{aligned} f_1(x-1) - f_1(x) &= f(x) - f(x+1) \\ &\leq f(x+1) - f(x+2) \\ &= f_1(x) - f_1(x+1) \end{aligned}$$

Also for $x = n^*$

$$\begin{aligned} f_1(n^* - 1) - f_1(n^*) &= 0 \\ &\leq f(n^*) - f(n^* + 1) \\ &= f_1(n^*) - f_1(n^* + 1) \end{aligned}$$

and for $x \in \{n^* + 1, ..., \}$

$$\begin{aligned} f_1(x-1) - f_1(x) &= f(x-1) - f(x) \\ &\leq f(x) - f(x+1) \\ &= f_1(x) - f_1(x+1) \end{aligned}$$

and hence we have shown the function $f_1 : Z_+ \to R$ is discrete concave. Clearly for all $n \in \{1, 2, ...\}$

$$\begin{aligned} g(n) &= max\{r + f(n+1), f(n)\} \\ &= max_{p \in \{n+1,n\}}\{f(p) + r(p-n)\} \\ &= max_{p \in \{n+1,n\}}\{f(p) + rp\} - rn \end{aligned} \tag{3.7}$$

Since the function is discrete concave also the function $p \mapsto f(p)+rp$ is discrete concave and this implies that the function

$$n \mapsto \begin{cases} \max_{p \in \{n+1,n\}}\{f(p) + rp\} & \text{if } n \in N \\ f(0) & \text{if } n = 0 \end{cases}$$

is discrete concave. Applying now relation 3.7 yields the desired result. □

In the next lemma we will derive an important property of expectations of discrete concave functions of the random variable $\mathbf{B}(p, n)$).

Lemma 4 *If the function $f : \mathbb{Z}_+ \mapsto \mathbb{R}$ is discrete concave, then the function $n \mapsto \mathbb{E}f(\mathbf{B}(p, n))$ is also discrete concave.*

Proof. We need to show that

$$n \mapsto \mathbb{E}f(\mathbf{B}(p, n+1)) - \mathbb{E}f(\mathbf{B}(p, n))$$

is decreasing. By the definition of $\mathbf{B}(p, n+1)$ in relation (3.1) and the conditional expectation formula we obtain that

$$\begin{aligned}
\mathbb{E}f(\mathbf{B}(p,n+1)) - \mathbb{E}f(\mathbf{B}(p,n)) &= p\mathbb{E}(f(\mathbf{B}(p,n+1)) - f(\mathbf{B}(p,n))|\mathbf{U}_{n+1} \leq p) \\
&= p(\mathbb{E}(f(1+\mathbf{B}(p,n)) - f(\mathbf{B}(p,n))|\mathbf{U}_{n+1} \leq p) \\
&= p\mathbb{E}(f(1+\mathbf{B}(p,n)) - f(\mathbf{B}(p,n))).
\end{aligned} \tag{3.8}$$

Since $\mathbf{B}(p,n+1) \geq \mathbf{B}(p,n)$ and f discrete concave we obtain that

$$n \mapsto f(1+\mathbf{B}(p,n)) - f(\mathbf{B}(p,n))$$

is decreasing and by relation (3.8) the result follows. □

Usually the demand for fare classes is a random variable $\mathbf{D}$ and we do not know in advance its realization. Let n be the number of reserved seats and $\mathbf{D}$ the random demand for seats. Consequently, the total number of reserved seats on the selected flight is equal to the random variable

$$\mathbf{N}(n) := \min(n, \mathbf{D}) \tag{3.9}$$

We will now investigate the behavior of the expectation of any function of the random $\mathbf{N}(n)$. If $f : \mathbb{Z}_+ \mapsto \mathbb{R}$ is such a function and we consider $n \mapsto \mathbb{E}(f(\mathbf{N}(n)))$ then clearly

$$\mathbb{E}(f(\mathbf{N}(n))) = \mathbb{P}(\mathbf{D} \geq n)f(n) + \sum\nolimits_{k=0}^{n-1} f(k)\mathbb{P}(\mathbf{D} = k)$$

This shows for every $n \in \mathbb{Z}_+$ that

$$\mathbb{E}(f(\mathbf{N}(n+1))) - \mathbb{E}(f(\mathbf{N}(n))) = \mathbb{P}(\mathbf{D} \geq n+1)f(n+1) + \mathbb{P}(\mathbf{D} = n)f(n) - \mathbb{P}(\mathbf{D} \geq n)f(n). \tag{3.10}$$

Since

$$\mathbb{P}(\mathbf{D} \geq n) = \mathbb{P}(\mathbf{D} \geq n+1) + \mathbb{P}(\mathbf{D} = n) \tag{3.11}$$

we obtain by relation (3.10) that

$$\begin{aligned}
\mathbb{E}(f(\mathbf{N}(n+1))) - \mathbb{E}(f(\mathbf{N}(n))) &= \mathbb{P}(\mathbf{D} \geq n)(f(n+1) - f(n)) - \mathbb{P}(\mathbf{D} = n)(f(n+1) - f(n)) \\
&= (\mathbb{P}(\mathbf{D} \geq n) - \mathbb{P}(\mathbf{D} = n))(f(n+1) - f(n))
\end{aligned}$$

Using again relation (3.11) this implies

$$\mathbb{E}(f(\mathbf{N}(n+1))) - \mathbb{E}(f(\mathbf{N}(n))) = \mathbb{P}(\mathbf{D} \geq n+1)(f(n+1) - f(n)) \tag{3.12}$$

Now we show a rather surprising result needed in the next subsection.

Lemma 5 *If $f : \mathbb{Z}_+ \mapsto \mathbb{R}$ is a discrete concave function and the optimization problem*

$$\max\{f(n) : n \geq C\}$$

has a finite optimal solution n_{opt}, then this is also an optimal solution of the problem

$$\max\{\mathbb{E}f(\mathbf{N}(n)) : n \geq C\}.$$

Proof. By the discrete concavity of f implying discrete unimodality we obtain for every $n \geq n_{opt}$ that

$$f(n+1) \leq f(n) \tag{3.13}$$

and for every $n < n_{opt}$

$$f(n+1) \geq f(n). \tag{3.14}$$

This shows by relations (3.12), (3.13) and (3.14) that for every $n \geq n_{opt}$

$$\mathbb{E}(f(\mathbf{N}(n+1)) \leq \mathbb{E}f(\mathbf{N}(n))$$

and for every $n < n_{opt}$ that

$$\mathbb{E}(f(\mathbf{N}(n+1)) \geq \mathbb{E}(f(\mathbf{N}(n)).$$

Hence n_{opt} is also an optimal solution of the optimization problem $\max\{\mathbb{E}f(\mathbf{N}(n)) : n \geq C\}$. □

If the random variable $\mathbf{D}$ has a bounded support then it does not hold that every optimal solution of the optimization problem $\max\{\mathbb{E}f(\mathbf{N}(n)) : n \geq C\}$ is also an optimal solution of the optimization problem $\max\{f(n) : n \geq C\}$. As an example we take $C = 2$, $f(n) = -(n-4)^2$ and $\mathbb{P}(\mathbf{D} = 3) = 1$. In this case the unique optimal solution of optimization problem $\max\{f(n) : n \geq C\}$ is given by $n = 4$. The set of optimal solution of optimization problem $\max\{\mathbb{E}f(\mathbf{N}(n)) : n \geq C\} = \max\{f(\min\{n, 3\}) : n \geq C\}$ is given by $n \geq 3$ and so this set contains elements which are not an optimal solution of the first problem. However, if the support of the random varible $\mathbf{D}$ is given by the set $\mathbb{Z}_+$, then one can also show the reverse implication of Lemma 5 as shown in the next result.

Lemma 6 *If the support of the random variable* $\mathbf{D}$ *is given by* $\mathbb{Z}_+$ *and the optimization problem*

$$\max\{f(n) : n \geq C\}$$

with f *a discrete concave function has a finite optimal solution, then the optimal solution sets of the optimization problems* $\max\{f(n) : n \geq C\}$ *and* $\max\{\mathbb{E}f(\mathbf{N}(n)) : n \geq C\}$ *coincide.*

Proof. By Lemma 5 we only need to show that there exists an optimal solution of the optimization problem $\max\{\mathbb{E}f(\mathbf{N}(n)) : n \geq C\}$ and that an optimal solution of this problem is also an optimal solution of $\max\{f(n) : n \geq C\}$. Since by assumption the optimization problem $\max\{f(n) : n \geq C\}$ has a finite optimal solution it follows by Lemma 5 that this is also an optimal solution of the optimization problem $\max\{\mathbb{E}f(\mathbf{N}(n)) : n \geq C\}$. Consider now an optimal solution n_{opt} of $\max\{\mathbb{E}f(\mathbf{N}(n)) : n \geq C\}$. By relation (3.12) we obtain for every $n > n_{opt}$ that

$$\begin{aligned} 0 &\geq \mathbb{E}f(\mathbf{N}(n)) - \mathbb{E}f(\mathbf{N}(n_{opt})) \\ &= \textstyle\sum_{k=n_{opt}}^{n-1} \mathbb{E}f(\mathbf{N}(k+1)) - \mathbb{E}f(\mathbf{N}(k)) \\ &= \textstyle\sum_{k=n_{opt}}^{n-1} \mathbb{P}(\mathbf{D} \geq k+1)(f(k+1) - f(k)) \end{aligned} \tag{3.15}$$

Since the set of optimal solutions of the optimization problem $\max\{f(n) : n \geq C\}$ is nonempty, there exists a minimal element n_s of this set. For this minimal element it follows using f is discrete concave that for every $k < n_s$

$$f(k+1) - f(k) \geq f(n_s) - f(n_s - 1) > 0$$

If $n_s > n_{opt}$ and applying $\mathbb{P}(\mathbf{D} \geq k+1) > 0$ for every k this yields by relation (3.15) that

$$0 \geq \mathbb{E}f(\mathbf{N}(n_s)) - \mathbb{E}f(\mathbf{N}(n_{opt})) > 0$$

and hence it must follow that $n_s \leq n_{opt}$. Without loss of generality, we may assume that

$n_s < n_{opt}$. Since by lemma 5 the value n_s is also an optimal solution of $\max\{\mathbb{E}f(\mathbf{N}(n)) : n \geq C\}$ we obtain

$$\begin{aligned} 0 &= \mathbb{E}f(\mathbf{N}(n_{opt})) - \mathbb{E}f(\mathbf{N}(n_s)) \\ &= \sum_{k=n_{st}}^{n_{opt}-1} \mathbb{P}(\mathbf{D} \geq k+1)(f(k+1) - f(k)) \end{aligned} \tag{3.16}$$

By the discrete concavity of the function f and n_s an optimal solution of $\max\{f(n) : n \geq C\}$ it follows that

$$f(k+1) - f(k) \leq 0$$

for every $k \geq n_s$. This shows by relation (3.16) that

$$f(k+1) - f(k) = 0$$

for $n_s \leq k \leq n_{opt} - 1$ and so $f(n_{opt}) = f(n_s)$. Hence n_{opt} is also an optimal solution of $\max\{f(n) : n \geq C\}$ and the result is proved. □

We will now discuss a generalization of the selection process. If our selection process yields the possible properties $E_i, 0 \leq i \leq m$ to reflect the demand for each fare class, then we introduce for $\mathbf{X}$ a nonnegative integer random size of the population, the $m+1$ dimensional random vector

$$(\mathbf{B}_0(\mathbf{p}, \mathbf{X}), ..., \mathbf{B}_m(\mathbf{p}, \mathbf{X}))$$

with $\mathbf{p} = (p_0, ..., p_m), p_i > 0, \sum_{i=0}^{m} p_i = 1$ and

$$\mathbf{B}_k(\mathbf{p}, \mathbf{X}) = \begin{cases} \sum_{i=1}^{\mathbf{X}} \mathbf{1}_{\{\xi_i = E_k\}} & \text{if } \mathbf{X} \geq 1 \\ 0 & \text{if } \mathbf{X} = 0 \end{cases}$$

where $\xi_i, i \in \mathbb{N}$ a sequence of independent and identically distributed random variables having state space $\{E_0, ..., E_m\}$. The random variable $\mathbf{B}_k(\mathbf{p}, \mathbf{X})$ denotes the total number of members within the population having property E_k if we assume that independent of each other each member within the population has property E_j with probability $p_j, 0 \leq j \leq m$. If we want to know the total number of the population $\mathbf{X}$ having one of the different properties $E_1, ..., E_m$, then in each trial we have probability p_0 that a member does not have either of these properties and so the total number of the population having either one of the different properties $E_1, ..., E_m$ has a binomial distribution with parameter $1 - p_0 = \sum_{k=1}^m p_k$ of success.

Lemma 7 *It follows that* $\sum_{k=1}^m \mathbf{B}_k(\mathbf{p}, \mathbf{X}) =^d \mathbf{B}(1 - p_0, \mathbf{X})$.

Proof. By definition

$$\begin{aligned} \sum_{k=1}^m \mathbf{B}_k(\mathbf{p}, \mathbf{X}) &= \sum_{k=1}^m \sum_{i=1}^{\mathbf{X}} \mathbf{1}_{\{\xi_i = E_k\}} \\ &= \sum_{i=1}^{\mathbf{X}} \sum_{k=1}^m \mathbf{1}_{\{\xi_i = E_k\}} \\ &= \sum_{i=1}^{\mathbf{X}} \mathbf{1}_{\{\cup_{k=1}^m \{\xi_i = E_k\}\}} \end{aligned}$$

Since

$$\mathbb{P}(\cup_{k=1}^m \{\xi_i = E_k\}) = \sum_{k=1}^m p_k = 1 - p_0$$

and hence

$$\sum_{i=1}^{\mathbf{X}} \mathbf{1}_{\{\cup_{k=1}^m \{\xi_i = E_k\}\}} =^d \sum_{i=1}^{\mathbf{X}} \mathbf{1}_{\{\mathbf{U}_i \leq 1 - p_0\}}$$

we obtain

$$\sum_{k=1}^m \mathbf{B}(\mathbf{p}_k, \mathbf{X}) =^d \mathbf{B}(\sum_{k=1}^m p_k, \mathbf{X}).$$

□

3.3 Static Overbooking Problem

In this section, we introduce our static overbooking models. Static models focus on determining booking limits in each fare class at the beginning of a reservation period. In this section, we first start with a basic model capturing this behavior. In a second static model, we introduce demand assumptions at the fare class level.

3.3.1 Model With No Booking Limit For Each Fare Class

For simplicity, we start with the most basic model capturing overbooking in the reservation process of seats for an airplane. In this simple model, we only try to determine the optimal size of the overbooking of a C seat capacity flight with m different fare classes. Let y be the total size of the overbooking. Hence the total number of seats to be allocated is given by $n := C + y \geq C$. If $\mathbf{D}$ denotes the total random demand for seats in this airplane, then clearly $\mathbf{N}(n) := \min(n, \mathbf{D})$ is the random number of reserved seats before departure. To model the revenue we assume in this static model that each reserved seat is a fare class i seat with probability $p_i, 1 \leq i \leq m$. Clearly these probabilities satisfy $\sum_{i=1}^{m} p_i = 1$. This shows that $\mathbf{B}(p_i, \mathbf{N}(n))$ is the random number of reserved fare class i seats before departure and the associated random revenue before departure is therefore given by $r_i \mathbf{B}(p_i, \mathbf{N}(n))$. Since with probability $1 - \beta_i$ each passenger having a reserved seat will not show up it follows by Lemma 1 that the total random number of no shows within fare class i equals $\mathbf{B}((1-\beta_i)p_i, \mathbf{N}(n))$. Hence with α_i denoting the fraction of the price refunded for a fare class i ticket the total random revenue generated by fare class i customers at the departure time of the airplane is given by

$$r_i \mathbf{B}(p_i, \mathbf{N}(n)) - \alpha_i r_i \mathbf{B}((1-\beta_i)p_i, \mathbf{N}(n)).$$

By relation (3.4) this implies that $(r_i p_i - \alpha_i(1-\beta_i) r_i p_i)\mathbb{E}(\mathbf{N}(n))$ is the total expected revenue of fare class i customers and so the total expected revenue over all fare classes equals

$$\theta_0 \mathbb{E}(\mathbf{N}(n)). \tag{3.17}$$

with

$$\theta_0 := \sum_{i=1}^{m} r_i p_i (1 - \alpha_i(1-\beta_i))$$

To model the penalty cost of overbooking, we first observe applying again Lemma 1 that the total number of arriving fare class i customers before departure is given by $\mathbf{B}(\beta_i p_i, \mathbf{N}(n))$ with β_i denoting the probability that a reserved fare class i customer will show up. Adding up all the different arriving fare class i customers the total number of overbooked seats equals

$$\max\{\sum_{i=1}^{m} \mathbf{B}(\beta_i p_i, \mathbf{N}(n)) - C, 0\} \tag{3.18}$$

By Lemma 7 we know that

$$\max\{\sum_{i=1}^{m} \mathbf{B}(\beta_i p_i, \mathbf{N}(n)) - C, 0\} =^d \max\left\{\mathbf{B}\left(\sum_{i=1}^{m} \beta_i p_i, \mathbf{N}(n)\right) - C, 0\right\}. \tag{3.19}$$

Hence by relations (3.18) and (3.19) and s the penalty costs of an overbooking the total expected overbooking costs are given by

$$s\mathbb{E}\left(\max\left\{\mathbf{B}\left(\sum_{i=1}^{m}\beta_i p_i, \mathbf{N}(n)\right) - C, 0\right\}\right). \tag{3.20}$$

Adding the terms in (3.17) and (3.20), we finally obtain the expected revenue as

$$g(n) := \theta_0\mathbb{E}(\mathbf{N}(n)) - s\mathbb{E}\left(\max\left\{\mathbf{B}\left(\sum_{i=1}^{m}\beta_i p_i, \mathbf{N}(n)\right) - C, 0\right\}\right). \tag{3.21}$$

Now, the optimal overbooking limit is a solution of the optimization problem

$$\max\{g(n) : n \geq C, n \in \mathbb{Z}_+\} \tag{P_T}$$

To analyse this optimization problem, we first rewrite the objective function g. Since for any $0 \leq p \leq 1$,

$$\max\{\mathbf{B}(p, \mathbf{N}(n)) - C, 0\} + \min\{\mathbf{B}(p\mathbf{N}(n)) - C, 0\} = \mathbf{B}(p, \mathbf{N}(n)) - C,$$

we obtain

$$-s\mathbb{E}(\max\{\mathbf{B}(p, \mathbf{N}(n)) - C, 0\}) = -sp\mathbb{E}(\mathbf{N}(n)) + sC + s\mathbb{E}(\min\{\mathbf{B}(p, \mathbf{N}(n)) - C, 0\}).$$

Hence, the objective function g in relation (3.21) can be written as

$$g(n) = \mathbb{E}f(\mathbf{N}(n)) \tag{3.22}$$

with the function $f : \mathbb{Z}_+ \mapsto \mathbb{R}$ given by

$$f(x) := sC + (\theta_0 - s\sum_{i=1}^{m}\beta_i p_i)x + s\mathbb{E}\left(\min\{\mathbf{B}\left(\sum_{i=1}^{m}\beta_i p_i, x\right) - C, 0\}\right) \tag{3.23}$$

To analyze the global behavior of this function we consider the following cases;

1. $\theta_0 - s\sum_{i=1}^{m}\beta_i p_i \geq 0$. To analyse this case we first observe using $\mathbf{B}(p, n+1) \geq \mathbf{B}(p, n)$ that the function

$$x \mapsto \mathbb{E}\left(\min\{\mathbf{B}\left(\sum_{i=1}^{m}\beta_i p_i, x\right) - C, 0\}\right)$$

is increasing. This shows by relation (3.23) that the function f is increasing.

Hence by the the monotonicity of $n \mapsto \mathbf{N}(n)$ the function $n \mapsto \mathbb{E}f(\mathbf{N}(n))$ is increasing and an optimal solution of our booking problem is to set $n = \infty$. An intuitive interpretation of this result is as follows. Since $(1 - \beta_i)$ is the no-show rate and α_i is the refund percentage, the expected net revenue per customer given that this is a fare class i customer equals

$$r_i - r_i\alpha_i(1 - \beta_i) - s\beta_i$$

Hence with p_i denoting the probability that a customer is a fare class i customer the expected revenue per customer is given by

$$\sum_{i=1}^{m} p_i(r_i - r_i\alpha_i(1 - \beta_i) - s\beta_i) = \theta - s\sum_{i=1}^{m} \beta_i p_i.$$

Hence for $\theta - s\sum_{i=1}^{m} \beta_i p_i \geq 0$, it is always profitable despite the overbooking costs to accept all demand. This means that the overbooking limit should be set to infinity. Clearly, this is a pathological case and will probably never happen in practice. A more reasonable assumption is given by the following

2. $\theta_0 - s\sum_{i=1}^{m} \beta_i p_i < 0$. To analyze this case we observe by Lemma 4 that the function

$$x \mapsto \mathbb{E}\left(\min\{\mathbf{B}\left(\sum_{i=1}^{m} \beta_i p_i, x\right) - C, 0\}\right)$$

is a discrete concave function. Hence by relation (3.23) the function f is a discrete concave function. Since $\lim_{x\uparrow\infty} f(x) = -\infty$ this shows that the optimization problem

$$\max\{f(n) : n \geq C\}$$

is easy to solve and there exist a finite optimal solution $n_{opt} \geq C$. Applying now Lemma 5 yields that n_{opt} is also a solution of the optimization problem P_T. A surprising consequence of this result is that the total booking limit does not depend on the cumulative distribution function (cdf) of the total demand $\mathbf{D}$. To compute this optimal solution we first need to evaluate the function

$$n \mapsto f(n+1) - f(n)$$

for every $n \geq C$ with f given by relation (3.23). Introducing

$$\theta_1 := s \sum_{i=1}^{m} \beta_i p_i$$

it follows by relation (3.23) and (3.8) that

$$f(n+1) - f(n) = \theta_0 - \theta_1 + \theta_1 \mathbb{E} f_0 \left(\mathbf{B} \left(\sum_{i=1}^{m} \beta_i p_i, n \right) \right) \tag{3.24}$$

with

$$f_0(x) = \min\{x - C + 1, 0\} - \min\{x - C, 0\} = \begin{cases} 1 & \text{if } x \leq C - 1 \\ 0 & \text{otherwise} \end{cases}$$

This shows for every $n \geq C$ that

$$f(n+1) - f(n) = \theta_0 - \theta_1 + \theta_1 \mathbb{P} \left(\mathbf{B} \left(\sum_{i=1}^{m} \beta_i p_i, n \right) \leq C - 1 \right). \tag{3.25}$$

By our assumption we know that $0 < \theta_0 \theta_1^{-1} < 1$ and this implies by relation (3.25)

$$f(n+1) - f(n) \leq 0 \Leftrightarrow \mathbb{P} \left(\mathbf{B} \left(\sum_{i=1}^{m} \beta_i p_i, n \right) \leq C - 1 \right) \leq 1 - \theta_0 \theta_1^{-1}.$$

Using the discrete concavity of the function f, an optimal solution of our optimization problem is therefore given by

$$n_{opt} = \inf \left\{ n \geq C : \mathbb{P} \left(\mathbf{B} \left(\sum_{i=1}^{m} \beta_i p_i, n \right) \leq C - 1 \right) \leq 1 - \theta_0 \theta_1^{-1} \right\}.$$

3.3.2 Model With Booking Limit For Each Fare Class

In problem P_T, we only consider the total demand and we do not distinguish different fare classes. We extend this model by considering overbooking and no-shows at the class level and propose two models, which provide us with the lower and upper bounds of the optimal expected revenue. In these models, we try to determine the optimal size of the reserved fare class i seats, $1 \leq i \leq m$. Therefore, we do not have the binomial relationship between overall demand and the demand for individual fare classes as in the problem P_T.

Let $\mathbf{D}_i$ denote the random demand for fare class $i, 1 \leq i \leq m$ and n_i be the number of reserved seats in fare class i such that $\sum_{i=1}^{m} n_i \geq C$. The random variable $\mathbf{N}_i(n_i) = \min\{n_i, \mathbf{D}_i\}$ denotes the number of customers having a reserved seat in fare class i before the departure time of the plane. Since with probability $(1-\beta_i)$ a customer having a fare class i seat will not show up, the number of no-shows within fare-class i is given by $\mathbf{B}((1-\beta_i), \mathbf{N}_i(n_i))$, while the number of occupied fare class i seats at the departure of the plane is given by $\mathbf{B}(\beta_i, \mathbf{N}_i(n_i))$. Since the total number of overbookings is given by $\max\{\sum_{i=1}^{m} \mathbf{B}(\beta_i, \mathbf{N}_i(n_i)) - C, 0\}$, the random overbooking cost is given by $s \max\{\sum_{i=1}^{m} \mathbf{B}(\beta_i, \mathbf{N}_i(n_i)) - C, 0\}$. Hence for any feasible vector $\mathbf{n} = (n_1, ..., n_m)$ the random revenue is

$$\sum_{i=1}^{m} r_i \mathbf{N}_i(n_i) - \sum_{i=1}^{m} \alpha_i r_i \mathbf{B}((1-\beta_i), \mathbf{N}_i(n_i)) - s \max\{\sum_{i=1}^{m} \mathbf{B}(\beta_i, \mathbf{N}_i(n_i)) - C, 0\}. \tag{3.26}$$

Applying now relation (3.4) and relation (3.26), we obtain that the total expected revenue of a given feasible booking vector $\mathbf{n} = (n_1, ..., n_m)$ as follows

$$q(n) = \sum_{i=1}^{m} (1 - \alpha_i(1-\beta_i)) r_i \mathbb{E}(\mathbf{N}_i(n_i)) - s\mathbb{E}\left(\max\{\sum_{i=1}^{m} \mathbf{B}(\beta_i, \mathbf{N}_i(n_i)) - C, 0\}\right).$$

Hence the associated optimization problem becomes

$$\max\{q(n) : \sum_{i=1}^{m} n_i \geq C, n_i \in \mathbb{Z}_+\}. \tag{P_I}$$

This problem is very difficult to solve. However, we can find upper and lower bounds on the optimal objective function. Let y_i be the actual allocation of physical capacity to fare class $i, 1 \leq i \leq m$ and C' denote the overbooking capacity. Then, we allocate at most C' seats, to different fare classes in such a way that the objective function, including the revenue and the penalty costs representing the inability to keep the occupied fare class i seats at departure below a the target value y_i, is maximized.

Lemma 8 *It follows that*

$$\mathbb{E}(\max\{\sum_{i=1}^{m} \mathbf{B}(\beta_i, \mathbf{N}_i(n_i)) - C, 0\}) \leq \sum_{i=1}^{m} \mathbb{E}(\max\{\mathbf{B}(\beta_i, \mathbf{N}_i(n_i)) - y_i, 0\})$$

for any vector $\mathbf{y}$ *satisfying* $\sum_{i=1}^{m} y_i = C$.

Proof. Clearly the function $f(x) = \max\{x, 0\}$ satisfies the subadditivity property given by

$$f(y_1 + y_2) \leq f(y_1) + f(y_2).$$

By this subadditivity property and $\sum_{1=1}^{n} y_i = C, y_i \in \mathbb{Z}_+$ it follows that

$$\begin{aligned} f\left(\sum_{i=1}^{m} \mathbf{B}(\beta_i, \mathbf{N}_i(n_i)) - C\right) &= f(\sum_{i=1}^{m} \mathbf{B}(\beta_i, \mathbf{N}_i(n_i)) - y_i) \\ &\leq \sum_{i=1}^{m} f(\mathbf{B}(\beta_i, \mathbf{N}_i(n_i)) - y_i). \end{aligned}$$

Hence we obtain for any $\sum_{i=1}^{m} y_i = C, y_i \in \mathbb{Z}_+$ that

$$\mathbb{E}\left(f\left(\sum_{i=1}^{m} \mathbf{B}(\beta_i, \mathbf{N}_i(n_i)) - C\right)\right) \leq \sum_{i=1}^{m} \mathbb{E}(f(\mathbf{B}(\beta_i, \mathbf{N}_i(n_i)) - y_i)),$$

and so the result is verified. □

It follows by the above lemma that

$$q(\mathbf{n}) \geq \sum_{i=1}^{m} (1 - \alpha_i(1 - \beta_i)) p_i \mathbb{E}(\mathbf{N}_i(n_i)) - s \sum_{i=1}^{n} \mathbb{E}\left(\max\{\mathbf{B}(\beta_i, \mathbf{N}_i(n_i)) - y_i, 0\}\right).$$

Hence, to obtain a lower bound on the optimal objective value we could solve the following separable problem:

$$\begin{aligned} \max \quad & \sum_{i=1}^{m} (1 - \alpha_i(1 - \beta_i)) r_i \mathbb{E}(\mathbf{N}_i(n_i)) - s \sum_{i=1}^{m} \mathbb{E}\left(\max\{\mathbf{B}(\beta_i, \mathbf{N}_i(n_i)) - y_i, 0\}\right) \\ \text{s.t} \quad & \sum_{i=1}^{m} n_i \geq C, \\ & \sum_{i=1}^{m} n_i \leq C', \qquad\qquad (P_I^{LB}) \\ & \sum_{i=1}^{n} y_i = C, \\ & n_i \in \mathbb{Z}_+, y_i \in \mathbb{Z}_+. \end{aligned}$$

In this study, we determine C' by using deterministic heuristic (1). By solving the problem **??** we obtain a lower bound on the expected revenue. Since this model is a separable problem, it can be decomposed into separate problems for each fare class. In this case, dynamic programming can be utilized. The problem P_I^{LB} can be solved by dynamic programming with two state space, where the fare classes correspond to stages, and the physical airline and overbooking capacities are the state spaces.

By using Jensen's inequality we can find an upper bound on the expected revenue. We can define the overbooking penalty part of the objective function as

$$\mathbb{E}\left(f\left(\sum_{i=1}^{m}\mathbf{B}(\beta,\mathbf{N}_i(x_i))-C\right)\right).$$

Then, we observe that

$$\begin{aligned}\mathbb{E}\left(f\left(\sum_{i=1}^{m}\mathbf{B}(\beta,\mathbf{N}_i(x_i))-C\right)\right) &\geq f(\mathbb{E}(\sum_{i=1}^{m}\mathbf{B}(\beta,\mathbf{N}_i(x_i))-C)) \\ &= f(\beta\sum_{i=1}^{m}\mathbb{E}(\mathbf{N}_i(x_i))-C),\end{aligned}$$

and so the function

$$\mathbf{x}\mapsto(1-\alpha(1-\beta))\sum_{i=1}^{m}r_i\mathbb{E}(\mathbf{N}_i(x_i))-f(\beta\sum_{i=1}^{m}\mathbb{E}(\mathbf{N}_i(x_i))-C)$$

yields an upper bound on the objective function. Hence, to obtain an upper bound on the optimal objective value we could solve

$$\begin{aligned}\max\quad & \sum_{i=1}^{m}(1-\alpha_i(1-\beta_i))r_i\mathbb{E}(\mathbf{N}_i(n_i))-s\,max(\sum_{i=1}^{m}\beta_i\mathbb{E}\left(\mathbf{N}_i(n_i)\right)-C,0)\\ \text{s.t}\quad & \sum_{i=1}^{m}n_i\geq C,\\ & \sum_{i=1}^{m}n_i\leq C', \qquad\qquad (P_I^{UB})\\ & n_i\in\mathbb{Z}_+.\end{aligned}$$

Again this problem is a separable problem and it can be solved by dynamic programming, where the fare classes and overbooking capacity of the airplane correspond to the stages and the state space, respectively.

3.4 Dynamic Overbooking Problem

In this section, we introduce our discrete-time dynamic model for the overbooking problem. Dynamic models decide whether to accept or reject a particular reservation request at its arrival time. They relax the static assumption on the arrival order of fare classes. Booking requests for each fare class can arrive throughout the reservation period and arrivals and cancellations are modeled as time dependent processes.

Consider a flight with m fare classes and a known capacity C. Booking requests for each fare class arrive according to a time-dependent process. The ticket sales period is partitioned into periods $1, 2, ..., T$, where T corresponds to the flight departure time and stage 1 corresponds to the opening of the flight for reservations. Customers who have already booked may cancel their tickets up to the departure time of the flight. At the time of the cancellation, customers are refunded an amount which is a class dependent percentage of the ticket price r_i. In addition, customers can execute no-shows with probability $(1-\beta_i)$ right before the departure of the plane, which is also class dependent; we assume that no-shows are not refunded. To prevent the empty seats arising from the cancellations and no-shows, customers can be overbooked with overbooking penalty s. At each period up to departure, we assume that two streams of events msy occur: a cancellation and an arrival. We assume that the cancellation request occurs before the arrival process. The state variable is the number of reserved seats in each fare class $\mathbf{n} = (n_1, ..., n_m)$. Let p_{it} denote the probability of an arrival in fare class i in period t and p_{0t} denote the null event in period t. Similarly, $q_{it}(\mathbf{n})$ is the probability of a cancellation in fare class i in period t and $q_{0t}(\mathbf{n}) = 1 - \sum_{i=1}^{m} q_{it}(\mathbf{n})$ is the probability of no-cancellation. Clearly, the demand probability is independent but the cancellation probability depends on the number of total reserved seats. To obtain a more realistic model, we assume that cancellation probabilities are nondecreasing and concave functions of $\mathbf{n}$. At each stage, upon arrival we decide to accept or reject a customer's request.

Next we introduce the random variables $\xi_t \in R^2$, $1 \leq t \leq T$. The first component of ξ_t represents the cancellation and the second component represents the arrival. Below we list the four possible cases for ξ_t

$\xi_t = (0, r_i) \Leftrightarrow$ no cancellation and a fare class i arrival in period t.

$\xi_t = (k_j, r_i) \Leftrightarrow$ a fare class j cancellation and a fare class i arrival in period t.

$\xi_t = (k_j, r_0) \Leftrightarrow$ a fare class j cancellation and no arrival in period t.

$\xi_t = (0, r_0) \Leftrightarrow$ no cancellation and no arrival in period t.

Observe that we set $i = 0$ and $r_0 = 0$ for the null event, and $p_{0t} = 1 - \sum_{i=1}^{m} p_{it}$. We assume that the random variables $\xi_t, t = 1, ..., T$ are independent.

As a function of the state in period t, let $R_t(\mathbf{n})$ be the optimal random revenue from t up to T before an event occurs in period t and the optimal value function J_t be given by

$$J_t(\mathbf{n}) = E(R_t(\mathbf{n})),$$

where $E(R_t(\mathbf{n})|\xi_t)$ denotes the conditional expectation on the information ξ_t. Then, we obtain the following relations by the principle of optimality in dynamic programming;

$E(R_t(\mathbf{n})|\xi_t = (k_0, r_i)) = max\{r_i + J_{t+1}(\mathbf{n} + e_i), J_{t+1}(\mathbf{n})\}$

$E(R_t(\mathbf{n})|\xi_t = (k_j, r_i)) = max\{r_i - k_j + J_{t+1}(\mathbf{n} + e_i - e_j), -k_j + J_{t+1}(\mathbf{n} - e_j)\}$

$E(R_t(\mathbf{n})|\xi_t = (k_0, r_0)) = J_{t+1}(\mathbf{n})$

$E(R_t(\mathbf{n})|\xi_t = (k_j, r_0)) = -k_j + J_{t+1}(\mathbf{n} - e_j)$

Here e_i is the unit vector which denotes the arrival in fare class i and k_j is the refund paid to customer in class j. We set $j = 0$ and $k_0 = 0$ for no-cancellation. By the definition of a conditional expectation it follows that

$$\begin{aligned} J_t(\mathbf{n}) &= E(R_t(\mathbf{n})) \\ &= \sum_{j=0}^{m} \sum_{i=0}^{m} E(R_t(\mathbf{n})|\xi_t = (k_j, r_i)) p_{it} q_{jt}(\mathbf{n}). \end{aligned} \tag{3.27}$$

Then,

$$\begin{aligned} J_t(\mathbf{n}) = & \sum_{j=1}^{m} \sum_{i=0}^{m} p_{it}\, q_{jt}(\mathbf{n})\, max\{r_i - k_j + J_{t+1}(\mathbf{n} - e_j + e_i), -k_j + J_{t+1}(\mathbf{n} - e_j)\} \\ & + \sum_{i=0}^{m} p_{it}\, q_{0t}(\mathbf{n})\, max\{r_i + J_{t+1}(\mathbf{n} + e_i), J_{t+1}(\mathbf{n})\}. \end{aligned} \tag{3.28}$$

The overbooking penalty constitutes the boundary condition which is

$$J_T(\mathbf{n}) = -sE(max\{\sum_{i=1}^{m} B(\beta_i, n_i) - C, 0\}).$$

Due to the multi-dimensional state space, this problem is very difficult to solve. By reducing the state space into one dimension, it can be simplified. In this new problem, states are the total number of reserved seats denoted by n. Therefore, the cancellation probability $q_t(n)$ does not depend on the fare class and we assume that the refund, denoted by k, is the same for all classes. Again in each period, it is assumed that two streams of events occur. First a cancellation be realized and then the an arrival may occur. Let p_{it} denote the probability of an arrival in fare class i, $q_t(n)$ be the probability of a cancellation, p_{0t} and $q_{0t}(n)$ denote the no arrival and no cancellation in period t, respectively. In this model, cancellation probabilities are linearly dependent on n, $q_t(n) = \omega_t n$. ω_t be the time dependent cancellation parameter. As in the previous model, we introduce the random variables $\xi_t \in R^2$, $1 \leq t \leq T$. Its first component denotes the cancellation and the second component represents the arrival. Then the four cases are given by

$\xi_t = (0, r_i) \Leftrightarrow$ no cancellation and a fare class i arrival in period t.

$\xi_t = (1, r_i) \Leftrightarrow$ a cancellation and a fare class i arrival in period t.

$\xi_t = (1, r_0) \Leftrightarrow$ a cancellation and no arrival in period t.

$\xi_t = (0, r_0) \Leftrightarrow$ no cancellation and no arrival in period t.

It is also assumed that the random variables ξ_t, $1 \leq t \leq T$, are independent. As a function of the state in period t, let $R_t(n)$ be the optimal random revenue from t up to T, before an event occurs in period t. Then, the expected optimal value function J_t is given by $J_t(0) = E(R_t(0))$ while the number of reservations at the beginning of period t is 0. Clearly, $J_t(0) = E(E[R_t(0)|\xi_t])$ and by the principle of dynamic programming, the conditional expectations are given by

$E(R_t(n)|\xi_t = (0, r_i)) = max\{r_i + J_{t+1}(n+1), J_{t+1}(n)\}$,

$E(R_t(n)|\xi_t = (1, r_i)) = max\{r_i - k + J_{t+1}(n), -k + J_{t+1}(n-1)\}$,

$E(R_t(n)|\xi_t = (0, r_0)) = J_{t+1}(n)$,

$E(R_t(n)|\xi_t = (1, r_0) = -k + J_{t+1}(n-1)$.

Now using the definition of a conditional expectation it follows that

$$\begin{aligned} J_t(n) &= E(R_t(n)) \\ &= \sum_{i=0}^{m} E(R_t(n)|\xi_t = (1, r_i))p_{it}q_t(n) \\ &+ \sum_{i=0}^{m} E(R_t(n)|\xi_t = (0, r_i))p_{it}q_{0t}(n). \end{aligned} \tag{3.29}$$

Then, the recursive relation becomes

$$\begin{aligned} J_t(n) = \sum_{i=0}^{m} \, p_{it} \, q_t(n) \, max\{r_i - k + J_{t+1}(n), -k + J_{t+1}(n-1)\} + \\ \sum_{i=0}^{m} \, p_{it} \, q_{0t}(n) \, max\{r_i + J_{t+1}(n+1), J_{t+1}(n)\}, \end{aligned} \tag{3.30}$$

Introducing the function

$$g_{i,t+1}(n) \mapsto \begin{cases} \max\{r_i + J_{t+1}(n+1), J_{t+1}(n)\}, & \text{if } n \in N \\ J_{t+1}(0), & \text{if } n = 0 \end{cases}$$

and using $q_{0t}(n) = 1 - q_t(n) \geq 0$ we obtain

$$J_t(n) + k\sum_{i=0}^{m} p_{it}q_t(n) = \sum_{i=0}^{m} p_{it}(q_t(n)g_{i,t+1}(n-1) + (1 - q_t(n))g_{i,t+1}(n)). \qquad (3.31)$$

By Lemma 3 it follows in case $n \mapsto J_{t+1}(n)$ is a nonincreasing discrete concave function that also $n \mapsto g_{it+1}(n)$ is a nonincreasing discrete concave function. Since $n \mapsto q_t(n)$ is discrete linear on $\{0, ..., C'\}$ we conclude from Lemma 2 that the function

$$n \mapsto (q_t(n)g_{i,t+1}(n-1) + (1 - q_t(n))g_{i,t+1}(n))$$

is a nonincreasing discrete concave function. Applying now relation (3.31) it follows that the function

$$n \mapsto J_t(n) + k\sum_{i=0}^{m} p_{it}qt(n)$$

is a nonincreasing discrete concave function. Finally by the linearity of the cancellation probabilities we obtain

$$n \mapsto k\sum_{i=0}^{m} p_{it}qt(n)$$

is an increasing discrete linear function on $\{0, ..., C'\}$. This depicts that the function $n \mapsto J_t(n)$ is a nonincreasing discrete concave function on $\{0, ..., C'\}$. Hence this shows by the induction that the function $n \mapsto J_t(n)$ is a nonincreasing discrete concave function once we have verified that $n \mapsto J_T(n), n = 0, ..., C'$ is discerete concave. The boundary condition of dynamic programming model is given by

$$J_T(n) = -s\mathbb{E}(max\{B(\beta, n) - C, 0\})$$

To analyze the boundary condition, we first rewrite $J_T(n)$. Since for any $0 \leq \beta \leq 1$

$$\max\{B(\beta, n) - C, 0\} + \min\{B(\beta, n) - C, 0\} = B(\beta, n) - C$$

We obtain

$$-s\mathbb{E}(max\{B(\beta,n)-C,0\}) = -s\beta n + sC + s\mathbb{E}(min\{B(\beta,n)-C,0\}).$$

We observe by Lemma 4 that the function

$$n \mapsto \mathbb{E}\left(\min\{\mathbf{B}\left(\beta,n\right)-C,0\}\right)$$

is a nonincreasing discrete concave function. Therefore, $J_T(n)$ is also nonincreasing discrete concave function.

CHAPTER 4

SOLUTION APPROACHES

4.1 Static Models

Static models determine the booking limit for each fare class at the beginning of the reservation period and they assume that reservations for different fare classes happen sequentially. In this way, the reservation period is divided into intervals during which all arriving passengers request the same fare. A natural solution approach for such a problem is dynamic programming where fare classes constitute the stages [5,12,24,37]. Without cancellations and overbookings, the state of the system is the number of seats still available. However, with cancellations and overbookings, the information required to characterize the state is the number of reservations in each fare class.

Problem P_I^{LB} has two dimensional state space, which corresponds to the overbooking capacity and the physical capacity. Introduce for every $p \leq m$, $c \in \{0, ..., C\}$ and $x \in \{0, ..., C'\}$ the value $R_p^l(x, c)$ as the maximal expected revenue for fare classes p up to m, then we have

$$R_p^l(x,c) = max\left\{\sum\nolimits_{i=p}^{m} S(n_i,y_i) | \sum_{i=p}^{m} n_i \leq x, \sum_{i=p}^{m} y_i \leq c, n_i, y_i \in \mathbb{Z}, i = p, ..., m\right\},$$

where

$$S(n_i, y_i) = (1 - \alpha_i(1 - \beta_i))r_i\mathbb{E}(\mathbf{N}_i(n_i)) - s\mathbb{E}\left(\max\{\mathbf{B}(\beta_i, \mathbf{N}_i(n_i)) - y_i, 0\}\right).$$

By the optimality principle of Bellman it now follows for every $p+1 \leq m$, $c \in \{0, ..., C\}$ and $x \in \{0, ..., C'\}$ that

$$R_p^l(x, c) = \max_{\substack{0 \leq n_p \leq x \\ 0 \leq y_p \leq c}} \left\{R_{p+1}^l(x - n_p, c - y_p) + S(n_p, y_p)\right\},$$

where $R_m^l(x,c) = (1-\alpha_m(1-\beta_m))r_m\mathbb{E}(\mathbf{N}_m(n_m)) - v\mathbb{E}\left(\max\{\mathbf{B}(\beta_m, \mathbf{N}_m(n_m)) - y_m, 0\}\right), c \in \{0, ..., C\}$ and $x \in \{0, ..., C'\}$.

Consequently, we can recursively compute the optimal objective value $R_1(C', C)$ and corresponding pseudocode is given in Algorithm 1. The computational complexity of the algorithm is of the order of $O(mCC')$.

Algorithm 1: Solving problem (**??**)

```
Input: C, C', m, r, s, β, α
for i = m to 1 do
    if i = m then
        for y_i = 0 to C do
            for n_i = 0 to C' do
                Compute S(n_i, y_i)
                R^l_m(n_i, y_i) = S(n_i, y_i)
    else
        for c = 0 to C do
            for x = 0 to C' do
                Set Z(x, c) = 0
                for y_i = 0 to c do
                    for n_i = 0 to x do
                        Compute S(n_i, y_i)
                        Z(n_i, y_i) = S(n_i, y_i) + R^l_{i+1}(x - n_i, c - y_i)
                R^l_i(x, c) = max(Z)
obj = max(R^l_1)
```

Problem P_I^{UB} is also a standard separable problem and can be solved by dynamic programming, where the overbooking capacity corresponds to the state space. Let us define for every $p \leq m$ and $x \in \{0, ..., C'\}$ the function $F_p(x) = \sum_{i=p}^{m}(1 - \alpha_i(1 - \beta_i))r_i\mathbb{E}(\mathbf{N}_i(n_i))$ for the revenue and $\phi_p(x) = s\ max(\sum_{i=p}^{m} \beta_i \mathbb{E}\left(\mathbf{N}_i(n_i)\right) - C, 0)$ for the overbooking cost. Then, $R_p^u(x)$ denotes the maximal expected revenue for fare classes p up to m and it is given by

$$R_p^u(x) = \max\left\{F_p(x) + \phi_p(x) | \sum_{i=p}^{m} n_i \leq x, n_i \in \mathbb{Z}, i = p, ..., m\right\}.$$

By the optimality principle of Bellman, it now follows for every $p+1 \leq m$ and $x \in \{0, ..., C'\}$ that

$$R_p^u(x) = \max_{0 \leq n_p \leq x} \left\{ F_{p+1}(x - n_p) + \phi_p(x) + (1 - \alpha_p(1 - \beta_p)) r_p \mathbb{E}(\mathbf{N}_p(n_p)) \right\},$$

where $R_m^u(x) = (1 - \alpha_m(1 - \beta_m)) r_m \mathbb{E}(\mathbf{N}_m(n_m)) - s\mathbb{E}\left(\max\{\mathbf{B}(\beta_m, \mathbf{N}_m(n_m)) - C, 0\}\right)$ and $x \in \{0, ..., C'\}$.

We can then recursively compute the optimal objective value $R_1(C')$. The pseudocode is given in Algorithm 2, which shows that the complexity of the algorithm is of the order of $O(mC')$.

Algorithm 2: Solving problem (3.3.2)

Input: $C, C', m, r, s, \beta, \alpha$
for $i = m$ **to** 1 **do**
    if $i = m$ **then**
        for $x = 0$ **to** C' **do**
            Compute $F_m(x)$ and $\phi_m(x)$
            $R_m^u(x) = F_m(x) + \phi_m(x)$
    else
        for $x = C'$ **to** 0 **do**
            Z(x)=0
            for $n_i = 0$ **to** x **do**
                Compute $\phi_i(x)$
                $Z(n_i) = (1 - \alpha_i(1 - \beta_i)) r_i \mathbb{E}(\mathbf{N}_i(n_i)) + F_{i+1}(x - n_i) + \phi_i(x)$
            $R_i^u(x) = max(Z)$
obj = $\max(R_1^u)$

4.2 Dynamic Model

The proposed dynamic model decides whether to accept or reject a booking request according to its arrival time and the state of the system. The demand and cancellation for each fare class are modeled as time dependent processes. The primary solution technique is dynamic programming, where the stages correspond to the remaining time until departure. The state of the system is the total number of reserved seats.

A backward recursive solution requires an overall computational complexity of $O(mTC')$ and the pseudocode is given in Algorithm 3.

Algorithm 3: Solving the dynamic overbooking problem

Input: $C, C', T, m, r, s, \beta, k, p_{mt}, q_t$
for $t = T$ **to** 1 **do**
  if $t = T$ **then**
    for $n = C'$ **to** C **do**
      Compute $J_T(n)$
  else
    for $n = C'$ **to** 0 **do**
      if $n = 0$ **then**
        $J_t(n) = \sum_{i=0}^{m} p_{it} \;\; max\{r_i + J_{t+1}(n+1), J_{t+1}(n)\}$
      else
        Compute $J_t(n)$

CHAPTER 5

COMPUTATIONAL RESULTS

We evaluate the performance of the proposed models by performing several experiments using simulation and present the results in two sections. In the first section, we compare the static upper bounding problem P_I^{UB} and the lower bounding problem P_I^{LB} for different show-up probabilities and refund percentages. Then, we analyze the static problem P_I^{LB} with respect to overbooking parameters, since P_I^{LB} gives the overbooking amount for each fare class.

In the second section, additional simulation experiments are conducted to see the differences between the static and dynamic modeling approaches. In this way, we can measure the effectiveness of the dynamic model. To give a lower bound on the gap between the static and dynamic models, we compare the solution of the dynamic model with the solution of the static problem P_I^{UB} because problem P_I^{UB} yields an upper bound for the maximum expected revenue that can be obtained by the static models. The simulation experiments depict that the dynamic model performs consistently better than the static problem P_I^{UB}. Also, in any of the cases, where the variability of the show-up probability is high and the overbooking penalty range is wide, the improvement is more significant. In all our simulation experiments we have used MATLAB 7.0 on a personal computer with 1.6 GHz Intel Celeron M processor and 1015 MB of RAM.

5.1 Static Models

We have implemented the models given in Section 3.1. In the first part, we compare the two proposed models. To simulate the models, we need to provide a probability vector $p_i \in \Re^{K+1}$, show-up probability β_i, and refund percentages $\alpha_i, 1 \leq i \leq m$. Then, we use the proposed models to find the upper and lower bounds on the optimum objective value.

Computation of overbooking levels depends on many parameters. One of them is show-up probability, $\beta_i, 1 \leq i \leq m$. Different class customers request different service and get refunds. Therefore, in real world applications show up probability are class dependent. Generally, it is higher in low fare classes because they are not flexible (changeable) and do not have any refunds [11]. In our model, we scale the show-up probabilities by taking into account this observation. On the other hand, refunds are calculated by using refund percentage parameter, $\alpha_i \geq 0, i = 1, ..., m$. We assume without loss of generality that $\alpha_1 > \alpha_2 > ... > \alpha_m = 0$ to reflect the higher refunds for relatively more expensive (flexible) fare class seats and no refunds for the cheapest class. Another important parameter in overbooking calculations is the cost of denial. Although overbooking aims to minimize the number of empty seats, it has the risk of customer denial because of the insufficient flight capacity. Bumping cost may include relatively intangible elements, such as loss of reputation, as well as any direct compensation. For an overbooking strategy to make sense, the revenue gain from boarding passengers must outweigh the loss from bumping, including all penalties and ill-will cost that might be incurred. Denied boardings realize at the cabin level. Therefore, calculation of the overbooking cost cannot be made exactly. However, in reality bumping cost depends on the flight properties. For example, Turkish Airlines (THY) arranges substitute transportation to get the denied passenger to her final destination and overbooking cost depends on the flight length not fare class price [39]. In the literature, a weighted average fare is used as a overbooking cost [11]. In our model, overbooking cost, s, is estimated by taking the weighted average of fares with respect to β.

In our simulation experiments, the probability vectors are generated by using truncated Poisson distribution with parameters $\lambda_i > 0, i = 1, ..., m$ and K [6]. K shows the maximum possible total demand for a fare class. As a result, total demand for a fare class i is concentrated on $\{0, ..., K\}$. In each run λ_i values are uniformly generated from the intervals $[\kappa_i, \nu_i]$, respectively and sorted in ascending order ($\lambda_1 < \lambda_2 < ... < \lambda_m$) to show the higher demand for relatively cheaper fare class seats. The parameters and their values are given in Table 5.1. An example of the probability vectors obtained by using truncated Poisson distribution is given in Figure 5.1.

Table 5.1: Parameters used in the truncated Poisson distribution

Parameters	Values
$[K, C, C', m]$	[100,100,120,4]
$(\kappa_1, \kappa_2, \kappa_3, \kappa_4)$	(2,20,30,40)
$(\nu_1, \nu_2, \nu_3, \nu_4)$	(12,40,60,80)

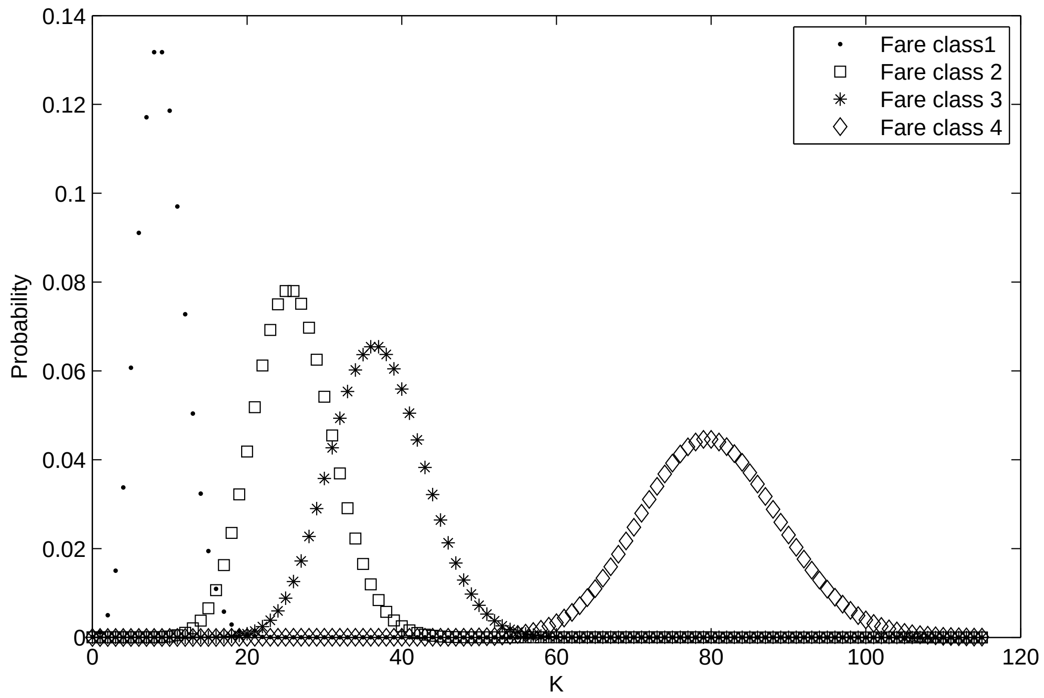


Figure 5.1: Truncated probability distributions for different fare classes

After generating the overbooking parameters and probability vectors, we simulate the models using these parameters. First, the problems P_I^{UB} and P_I^{LB} are tested with respect to β changes. To make a fair comparison between two models, we use the same β for all fare classes. We make 20 simulation runs for each β values and in each run we provide the probability vectors $p_i \in \Re^{K+1}, 1 \leq i \leq m$. As our statistics, we store the mean value in each run. Table 5.2 gives the parameters used in this simulation. With these parameters, the running time of algorithm for P_I^{LB} is approximately 20 sec. and the running time of algorithm for P_I^{UB} is around 0,01 sec.

Figure 5.2 depicts the relative difference between two models over β changes. The expected revenues of both models decrease as β decreases since we use fixed overbooking capacity. This means that as show-up probability reduces, it is needed to make more reservations otherwise it results in revenue loss due to no-shows and related refunds. In addition, the relative differences decrease with β since two models only differ in their overbooking cost calculation. Therefore, as β goes to zero, the results of these models become similar.

Table 5.2: Parameters used in the simulation for β

Parameters	Values
$[K, C, C', m]$	[100,100,120,4]
$(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$	(0.4, 0.3, 0.2, 0)
(r_1, r_2, r_3, r_4)	(160,135,115,95)

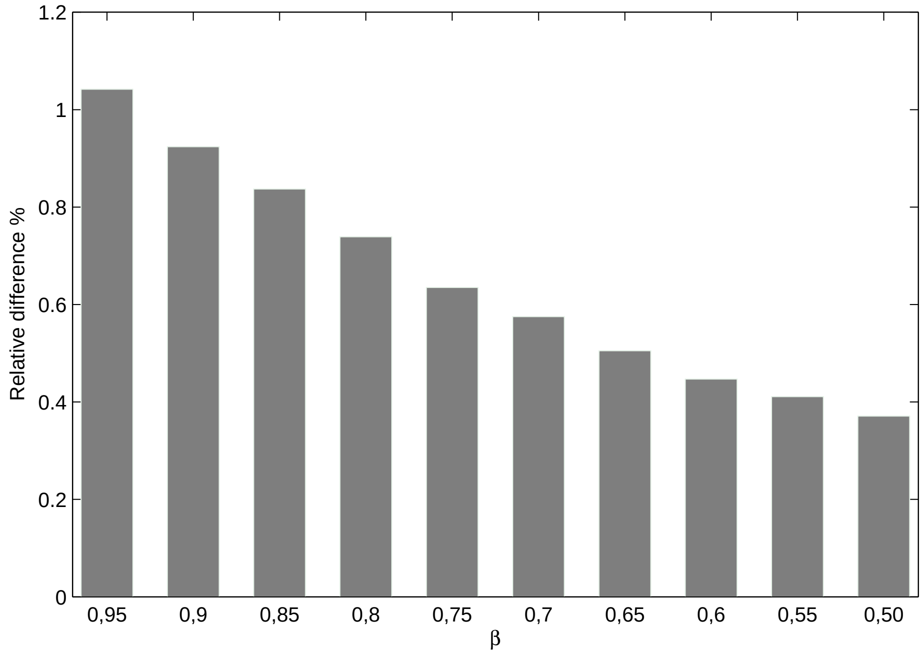


Figure 5.2: The relative difference between the objective function values of problem P_I^{UB} and problem P_I^{LB} with respect to β

Then, we test these problems with respect to α changes. We use the parameters in Table 5.3. Again to make a fair comparison, we use the same α value for all fare classes. The results are given in the Figure 5.3. As it is seen in the Figure 5.3, as α increases, the relative difference increases since the refunds of no-shows increases.

Table 5.3: Parameters used in the simulation for α

Parameters	Values
$[K, C, C', m]$	[100,100,120,4]
$(\beta_1, \beta_2, \beta_3, \beta_4)$	(0.95,0.85,0.80,0.95)
(r_1, r_2, r_3, r_4)	(160,135,115,95)

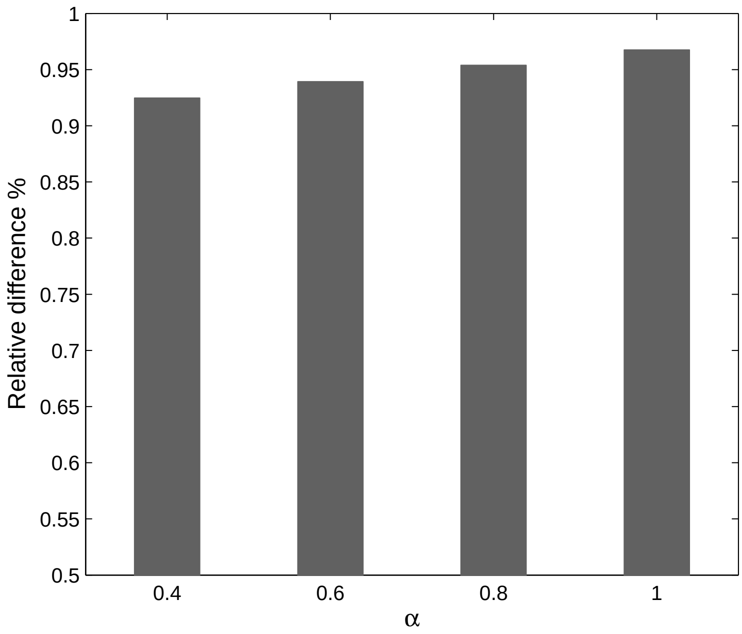


Figure 5.3: The relative difference between the objective function values of problem P_I^{UB} and problem P_I^{LB} with respect to α

We also conduct sensitivity analysis for the static problem P_I^{LB} with respect to β since the model gives the overbooking amounts in each fare class. β_i values are different for different fare classes. Therefore, we only change β_1 value to observe the effect of altering the show-up probability of one class to others. In the first simulation, we use fixed overbooking capacity. The parameters that we use for this simulation are given in Table 5.4. Again, we make 20 simulation runs and in each run we generate the probability vectors $p_i \in \Re^{K+1}, 1 \leq i \leq m$. As shown in Figure 5.4 the overbooking amount in fare class 1 increases as β_1 decreases and overbooking amount in other classes decrease because of the fixed overbooking capacity.

Table 5.4: Parameters used in the simulation of static problem P_I^{LB}

Parameters	Values
$[K, C, C', m]$	[100,100,120,4]
$(\beta_2, \beta_3, \beta_4)$	(0.85,0.80,0.95)
$(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$	(0.4, 0.3, 0.2, 0)
(r_1, r_2, r_3, r_4)	(160,135,115,95)

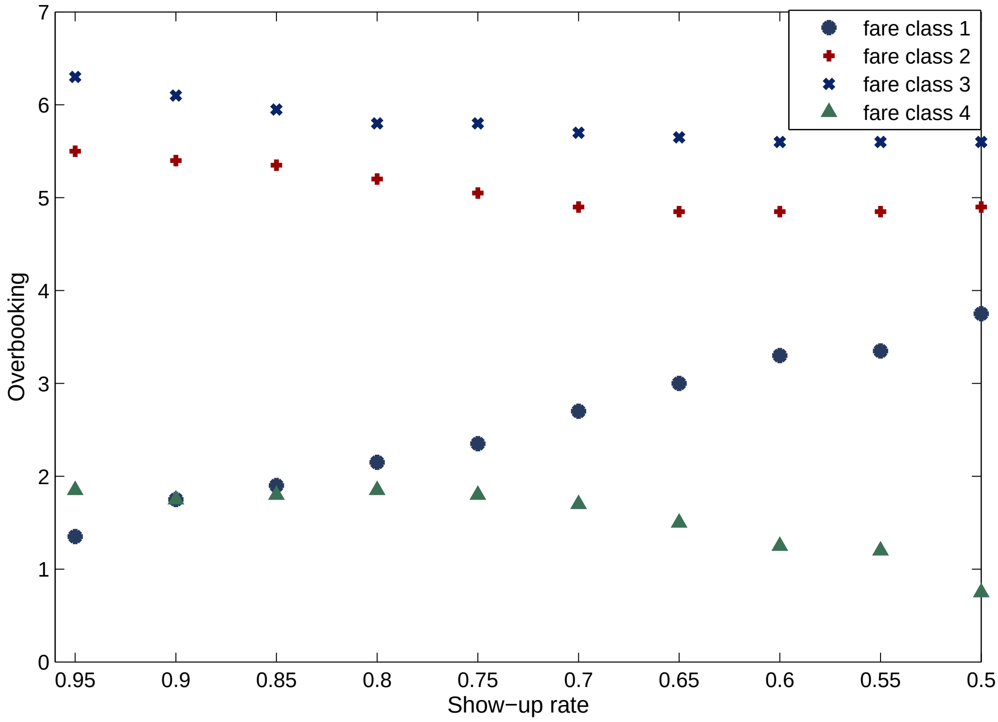
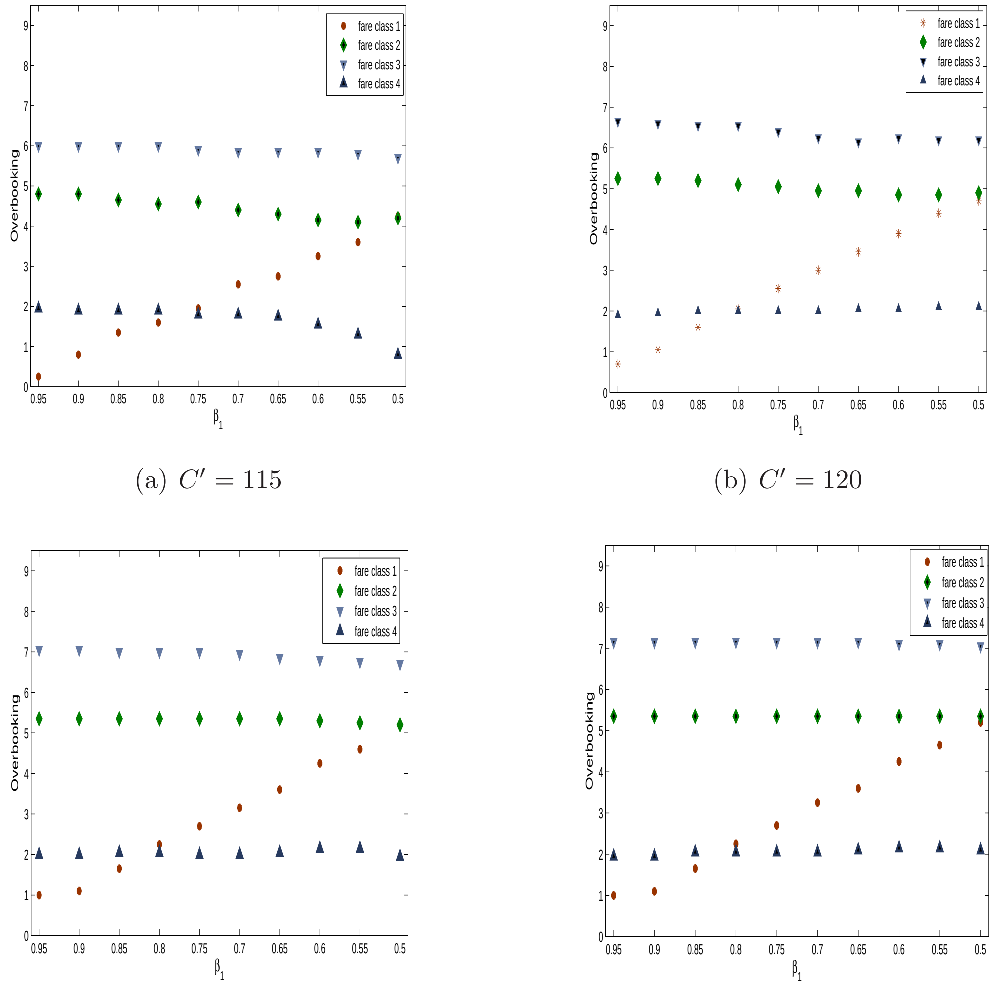


Figure 5.4: Overbooking

Then, we test the static problem P_I^{LB} according to β_1 and overbooking capacity changes. The parameters used in the simulation are the same as in Table 5.4. The results are given in the Figure 5.5. In this simulation, we compare overbooking changes in fare classes with respect to β_1 and overbooking capacity C'. We make 20 simulations for each C' value. Figure 5.5(a) depicts that while overbooking amount in fare class 1 increases, it decreases in other fare classes, mostly in fare class 4 due to its low fare and high show-up rate. However, when we allow to make more reservations by increasing the overbooking capacity, overbooking amount in the fare class 4 does not change due to its high show-up rate and overbooking amount in fare class 2 and 3 increase but stay stable with respect to β_1 changes. Static problem P_I^{LB} is concave in C'. Therefore, even if we increase C', after some point overbooking amount in each fare class does not change.



(a) $C' = 115$ (b) $C' = 120$

(c) $C' = 125$ (d) $C' = 130$

Figure 5.5: Overbooking amount for different values of show-up and overbooking capacity

5.2 Dynamic Models

In this section, we conduct simulation experiments to compare the static problem P_I^{UB} and the dynamic model (3.30). The motivation of these simulations is to show the effect of having more information, as one has more information in the dynamic model than the static model.

In the dynamic model (3.30), for each period t up to departure of the plane we consider a cancellation and an arrival process. In order to provide the arrival probability vector p_t of period t, Dirichlet distribution with parameters $\gamma_i(t), 0 \leq i \leq m$ is used [6]. Dirichlet distribution has been used to describe the distribution of purchase probabilities for a population of individuals buying one and only one brand of a particular product as described by Goodhardt [15]. In the dynamic model, Dirichlet distribution allows us to provide arrival probabilities at each period t for each fare class. It is reasonable to predict that as the departure time T approaches, the request for cheaper fare classes reduce, whereas requests for the more expensive fare classes increase. To achieve this, we adjust the adopted Dirichlet distribution parameters monotonically. Figure 5.6 shows the change of these variables over time and Table 5.5 gives the values of the parameters that we use. With these parameters, the running time of the dynamic programming algorithm is around 0,30 sec.

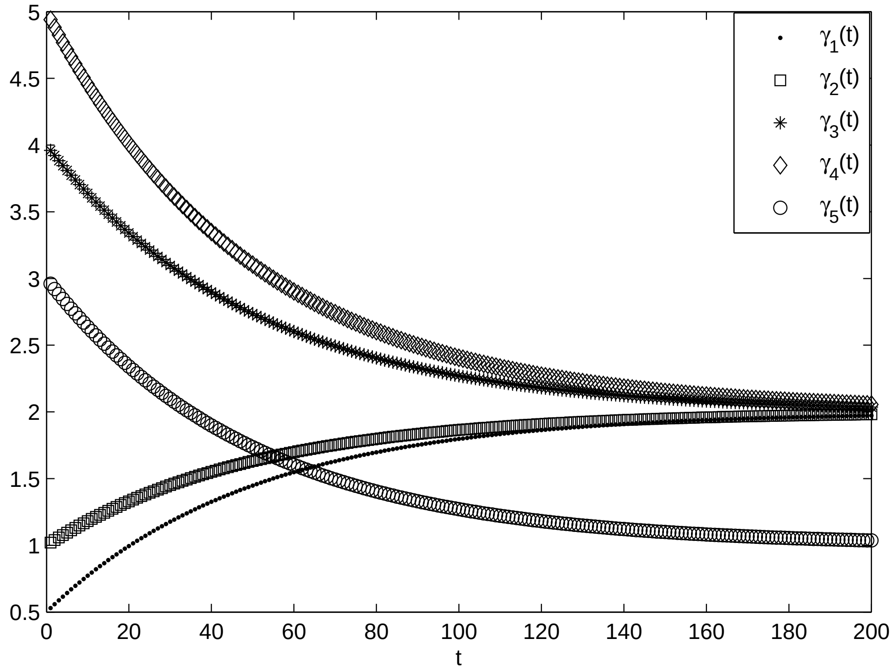


Figure 5.6: The change of adopted Dirichlet distribution parameters over time

Table 5.5: Parameters used in the simulation of dynamic model

Parameters	Values
$[T, C, C', m, \beta, k]$	[200,100,120,4,0.80,30]
(r_1, r_2, r_3, r_4)	(160,135,115,95)
$(\bar{v}_0, \bar{v}, v_0, v_1, v_2, v_3)$	(1,2,3,0.5,1,4,5)

In our model, we assume that booking requests are independent of the number of seats already reserved, whereas cancellation and no-show probabilities depend on the total number of booked seats. This means that the higher the number of reserved seats, the higher the probability of cancellation [33]. In addition, we consider cancellations and no-shows at class independent rates. For class-based cancellations it is observed that cancellation intensities of all fare classes are different. While cancellation amount of the cheaper fare classes are decreasing in the remaining time before departure of the plane, it is increasing for the more expensive fare classes in the remaining time before departure. Therefore, when estimating a probability of cancellation in period t for all m fare classes, we adjust ω_t to reflect intensities of all fare classes. Figure 5.7 shows cancellation probabilities for different values of n over time. Cancellation probability function $q_t(n)$ is linearly dependent on n.

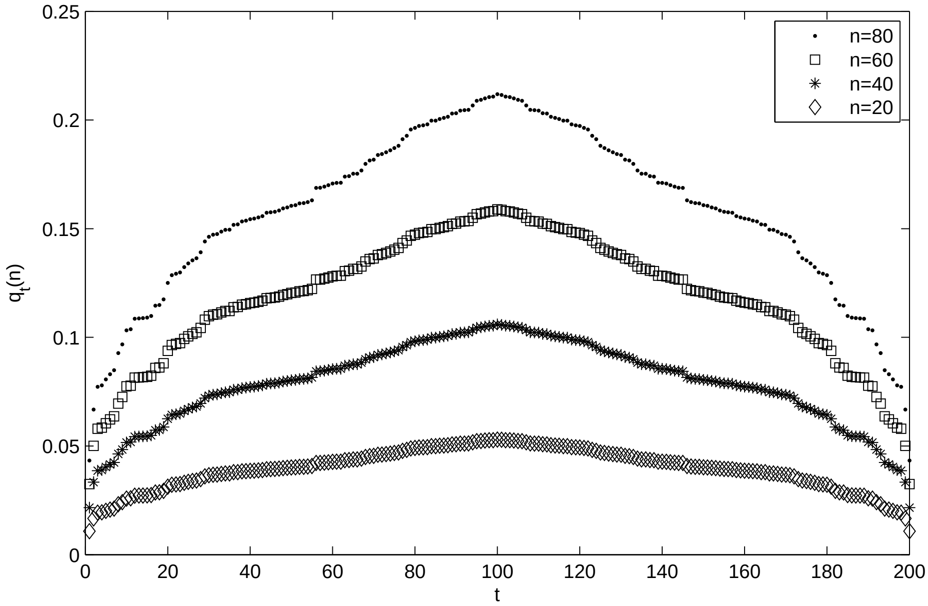


Figure 5.7: An example of the change of cancellation probabilities over time and the number of reserved seats

After generating the parameters and probability vectors, we have firstly implemented a dynamic programming algorithm to solve (3.30). As it is shown in the Figure 5.8, $J_t(n)$ is nonincreasing in n.

Then we make 20 simulation runs to compare the solution of static and dynamic models. In each run, we first provide for $1 \leq t \leq T$ the arrival probability vector and cancellation probability vector. Then, we compute the expected optimal revenue for the dynamic model (3.30). To be able to compare static and dynamic models, we need to compute the demand probabilities $p_{il} = P(D_i = l), 1 \leq l \leq T$, by using the arrival probabilities $p_t, 1 \leq t \leq T$. If ζ_i denotes the revenue generated by a random arrival in period t, we may assume that it may take $m+1$ different values $r_0, r_1, ..., r_m$ and its discrete density is given as $P(\zeta_t = r_i) = p_{it}, 0 \leq i \leq m, 1 \leq t \leq T$. Then we have

$$D_i = \sum_{i=1}^{T} \mathbf{1}_{\{\zeta_t \leq r_i\}}$$

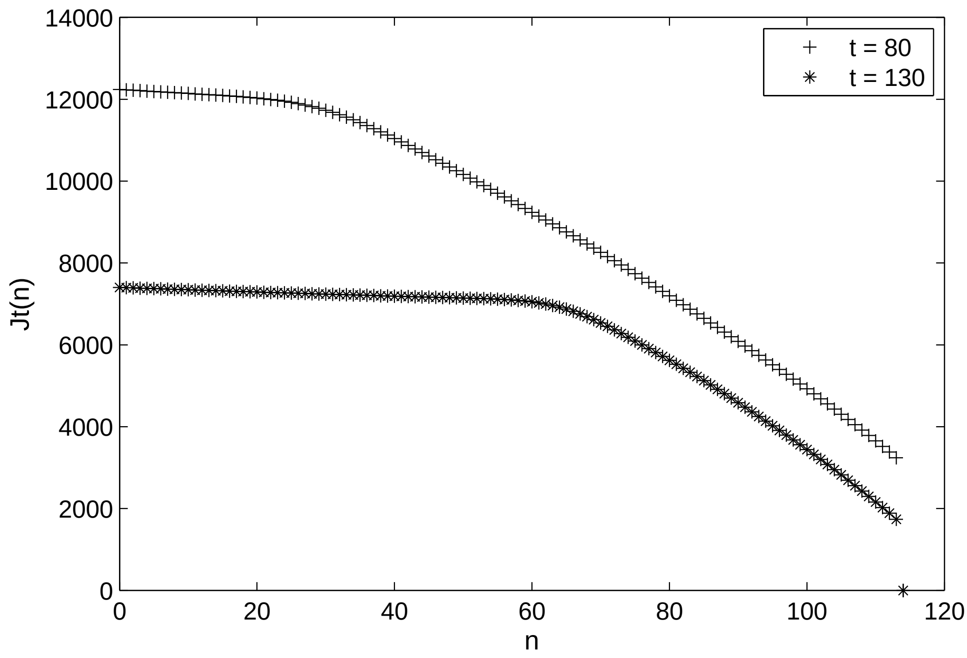


Figure 5.8: $J_t(n)$ versus n for different t values

Since these random variables are assumed to be independent, the Bernoulli random variables $\mathbf{1}_{\{\zeta_t \leq r_i\}}, 1 \leq t \leq T$, are also independent. It is observed that for every $\alpha \in (0, 2\pi)$ the discrete Fourier transform

$$P(\alpha) = E(exp(i\alpha(\sum_{i=1}^{T} \mathbf{1}_{\{\zeta_t \leq r_i\}}))) = \prod_{t=1}^{T} E(exp(i\alpha \mathbf{1}_{\{\zeta_t \leq r_i\}})).$$

Consequently,

$$E(exp(i\alpha \mathbf{1}_{\{\zeta_t \leq r_i\}})) = p_{it} exp(i\alpha) + (1 - p_{it}) = 1 - p_{it}(1 - exp(i\alpha))$$

and as a result, we obtain

$$P(\alpha) = \prod_{t=1}^{T} (1 - p_{it}(1 - exp(i\alpha))).$$

It is known that

$$p_{it} = \frac{1}{T+1} \sum_{n=0}^{T} P(\frac{2\pi n}{T+1}) exp(\frac{-2\pi ink}{T+1})$$

We can easily obtain the probabilities p_{ik} by using the FFT algorithm [14]. After generating the probabilities, we can compute the expected optimal revenue for the static upper bounding problem P_I^{UB}. Dynamic model and static problem P_I^{UB} are firstly compared with respect to different show-up rates. The parameters we use are given in Table 5.6. We conduct 20 simulation runs for different β values. To make a fair comparison between the static and dynamic models, we use the same β values and fixed refund amount. Figure 5.9 shows our results as a stacked bar plot. Each bar plot represents the relative difference in percentages between the revenue obtained with dynamic model and the revenue obtained with the static model. Figure 5.9 depicts that relative difference increases as show-up rates decreases. This means that dynamic model gives better results for the systems where randomness is high.

Table 5.6: Parameters used in the simulation of dynamic and static models

Parameters	Values
$[T, C, C', m, k]$	[200,100,120,4,30]
(r_1, r_2, r_3, r_4)	(160,135,115,95)

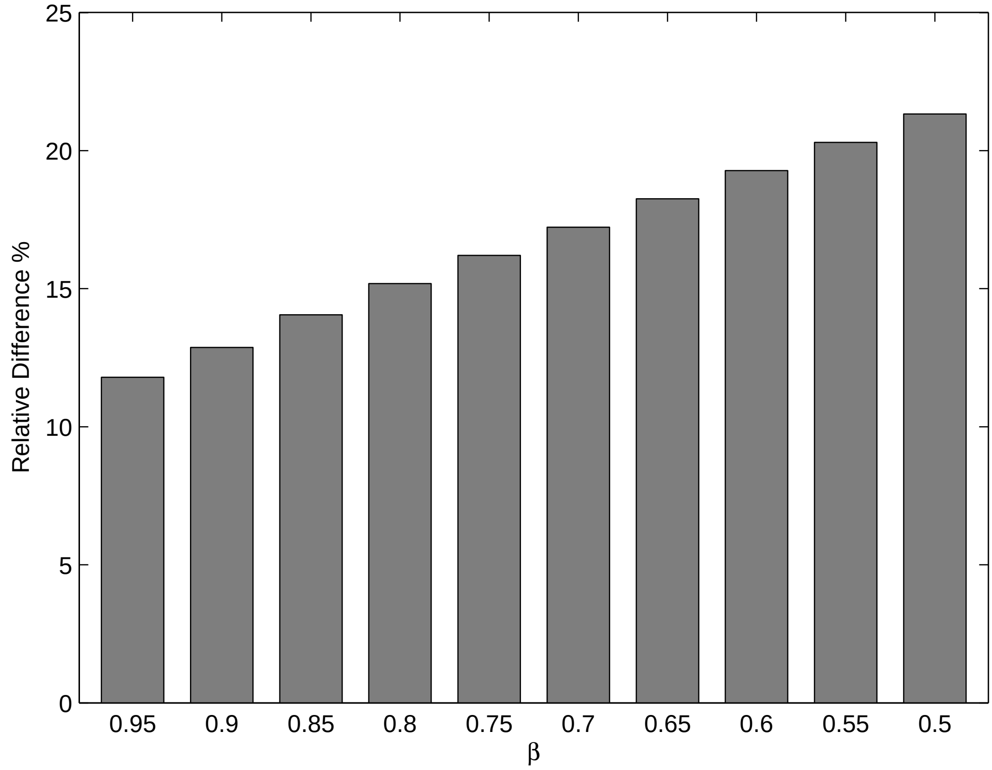


Figure 5.9: The relative difference between the objective function values of the dynamic model and static upper bounding problem P_I^{UB} for varying β

Then we compare these models with respect to overbooking penalty s changes. The parameters are given in Table 5.7. We make 20 simulation runs for different s values. Figure 5.10 depicts that relative difference slightly decreases as s increase. This is reasonable since as s increases overbooking amount decreases; therefore gap due to overbooking decreases.

Table 5.7: Parameters used in the simulation of dynamic and static models

Parameters	Values
$[T, C, C', m, \beta, k]$	[200,100,120,4,0.80,30]
(r_1, r_2, r_3, r_4)	(160,135,115,95)

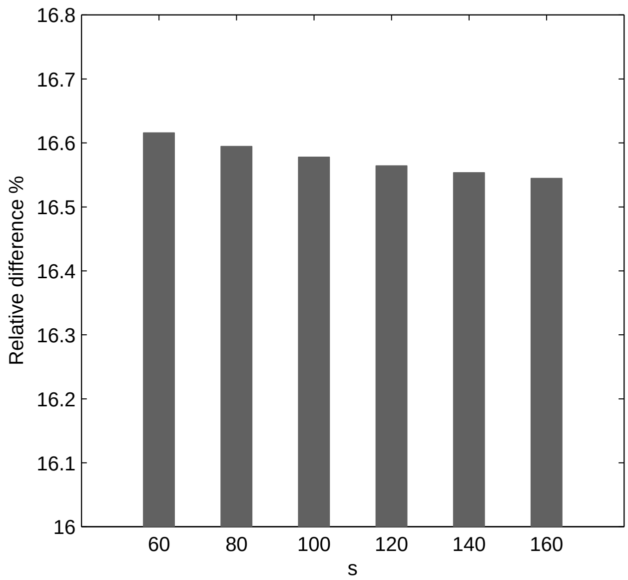


Figure 5.10: The relative difference between the objective function values of the dynamic model and static upper bounding problem P_I^{UB} for varying s parameter

CHAPTER 6

CONCLUSION AND FUTURE WORK

In this study, we consider the single-leg ARM problem with overbooking, no-shows, and cancellations and introduce new static and dynamic models. These models are different than the proposed models in the literature in terms of the objective function and modeling approaches. In the literature, generally overbooking problem is simplified by ignoring cancelation penalty or reducing problem size. In all cases we provide support for the models with computational results.

For the static overbooking problem, we first introduce an overbooking problem that considers total demand. Then we extend it to a model with a booking limit for each fare class. However, due to its complex analytical form we propose two models which provide the upper and lower bounds on the optimal expected revenue. In this way, we can also show the possible worst and best cases of static overbooking problem. In the computational studies we demonstrate that the relative differences between these models are very low. Therefore, they can give a close approximation of the optimal expected revenue.

For the dynamic overbooking problem, we propose a new model at which each period up to departure of the plane, either a cancellation, an arrival or both of them can be realized. In this way, we independently handle cancelation and arrival processes. On the other hand, our proposed model differs from the existing studies in the literature by its assumptions. We assume that while arrivals are independent of the number of reservations, cancellations depend on it, which is reasonable. Experimental studies demonstrate that the relative difference between static and dynamic models is comparatively high even if we compare the dynamic model with the static upper bound model and it increases as the show-up probability decreases.

We can extend this research in several directions and study these extensions on the model P_T. One of the extensions might be introducing a service level constraint. In this way, we can provide the limit on overbooking risk. In addition, we can set a certain target level with respect to the expected earnings and minimize the probability of overbooking.

Furthermore, we can extend the model P_T by considering revenue obtained from the empty seats before departure. These seats can be allocated to the passengers without reservations who wait at the airport.

Bibliography

[1] Alstrup, J., Boas, S., Madsen, O.B.G. and Vidal, R.V.V., Booking Policy for Flights with Two Types of Passengers, EJOR, 27, 274-288, 1986.

[2] Bailey, J., Bumped Fliers and No Plan B, The New York Times, May 30, 2007.

[3] Ball, M.O. and Queyranne, M., Toward Robust Revenue Management: Competitive Analysis of Online Booking, Working paper, University of Maryland, 2006.

[4] Beckman, J.M., Decision and team problems in airline reservations, Econometrica 26, 134145, 1958.

[5] Belobaba, P.P., Application of a Probabilistic Decision Model to Airline Seat Inventory Control, Operations Research, 37(2), 183-197, 1989.

[6] Birbil, Ş.İ., Frenk, J.B.G., Gromicho, J.A.S. and Zhang, S., The Role of Robust Optimization in Single-leg Airline Revenue Management, Management Science, 55(1), 148-163, 2009.

[7] Bodily, S.E. and Pfeifer, P.E., Overbooking decision rules, Omega, 20, 129133, 1992.

[8] Brumelle, S.L. and McGill, J.I., Airline Seat Allocation with Multiple Nested Fare Classes, Operations Research, 41, 127-137, 1993.

[9] Chatwin, R.E., Continuous-Time Airline Overbooking with Time Dependent Fares and Refunds, Transportation Science, 33, 182-191, 1999.

[10] Chi, Z., Airline Yield Management in a Dynamic Network Environment, PhD Thesis, MIT, 1995.

[11] Coughlan, J., Airline overbooking in the multi-class case, Journal of the Operational Research Society, 50, 1098-1103, 1999.

[12] Curry, R.E., Optimal Airline Seat Allocations with Fare Classes Nested by Origins and Destinations, Transportation Science, 24(3), 193-204, 1990.

[13] Feng, Y. and Xiao, B., A continuous-time seat control model for single-leg flights with no-shows and optimal overbooking upper bound, European Journal of Operational Research, 174(2), 1298-1316, 2006.

[14] Golub, G.H. and van Loan, C.F., Matrix Computations, 3rd Edition, The Johns Hopkins University Press, Baltimore, 1996.

[15] Goodhardt, G., Ehrenberg, A. and Chatfield, C., The Dirichlet: A Comprehensive Model of Buying Behavior, Journal of the Royal Statistical Society, 147(5), 621-655, 1984.

[16] Hooke, R. and Jeeves, T.A., Direct search solution of numerical and statistical problems, Journal of the Association for Computing Machinery, 8, 212-229, 1961.

[17] Karaesmen, I. and van Ryzin, G., Overbooking with Substitutable Inventory Classes, Operations Research, 52, 83-104, 2004.

[18] Kleywegt, A.J. and Papastavrou, J.D., The Dynamic and Stochastic Knapsack Problem, Operations Research, 46, 17-35, 1998.

[19] Kunnumkal, S. and Topaloglu, H., A stochastic approximation method for the revenue management problem on a single flight leg with discrete demand distributions. Working paper, School of Operations Research and Industrial Engineering, Cornell University, 2007.

[20] Lan, Y., Gao, H., Ball, M., and Karaesmen, I., Revenue Management with Limited Information. Working paper, University of Maryland, Robert H. Smith School of Business, 2007.

[21] Lautenbacher, C.J. and Stidham, S., The Underlying Markov Decision Process in the Single-Leg Airline Yield Management Problem, Transportation Science, 33(2), 1999.

[22] Lee, T.C. and Hersh, M., A Model for Dynamic Airline Seat Inventory Control with Multiple Seat Bookings, Transportation Science, 27, 252-265, 1993.

[23] Liang, Y., Solution to the Continuous Time Dynamic Yield Management Model, Transportation Science, 33, 117-123, 1999.

[24] Littlewood, K., Forecasting and Control of Passengers, 12th AGIFORS Symposium Proceedings, 95-128, 1972.

[25] Lippman, S.A. and Stidman, S., Individual versus social optimization in exponential congestion systems, Operations Research, 25(2), 233247, 1977.

[26] Nelder, J.A. and Mead, R., A simplex method for function minimization, Computer Journal, 7, 308-313, 1965.

[27] Phillips, R.L., Pricing and Revenue Optimization, Stanford University Press, Stanford, 2005.

[28] Robinson, L.W., Optimal and Approximate Control Policies for Airline Booking with Sequential Nonmonotonic Fare Classes, Operations Research, 43(2), 252-263, 1995.

[29] Rothstein, M., An Airline Overbooking Model, Transportation Science, 5, 180-192, 1971.

[30] Rothstein, M., O.R. and the airline overbooking problem, Operations Research, 33, 237248, 1985.

[31] Rothstein, M. and Stone, A.W., Passenger booking levels, AGIFORS Symposium Proceeding 7, Noordwijk, The Netherlands, 1967.

[32] Smith, B., Leimkuhler, J. and Darrow, R., Yield Management at American Airlines, Interfaces, 22(1), 8-31, 1992.

[33] Subramanian, J., Stidham, S. and Lautenbacher, C., Airline Yield Management with Overbooking, Cancellations, and No-Shows, Transportation Science, 33 (2), 147-167, 1999.

[34] Talluri, K. and van Ryzin, G., The Theory and Practice of Revenue Management, Kluwer Academic Publishers, Boston, 2004.

[35] Taylor, C.J., The determination of passenger booking levels, AGIFORS Symposium Proceedings, Fregene, Italy, (2), 1962.

[36] Thompson, H.R., Statistical problems in airline reservation control, Operational Research Quarterly, 12(3), 167185, 1961.

[37] Wollmer, R.D., An Airline Seat Management Model for a Single Leg Route when Lower Fare Classes Book First, Operations Research, 40(1), 26-37, 1992.

[38] van Ryzin, G.J. and McGill, J., Revenue Management without Forecasting or Optimization: An Adaptive Algorithm for Determining Airline Seat Protection Levels, Management Science, 46, 760-775, 2000.

[39] Turkish Airlines (THY), http://www.thy.com.tr, March, 2009.