Yield Management

B.S.c. Thesis by

Márta Treszl

Mathematics B.S.c., Mathematical Analyst

Supervisors:

Róbert Fullér

Professor at Åbo Akademi University

Viktória Villányi

Assistant Professor at the Department of Operations Research

Budapest

2012
# Table of Contents

## Preface

- 1

## Chapter 1

1.1 Introduction ................................................................. 3  
1.2 History of Yield Management ........................................ 3  
1.3 Levels of Revenue Management ..................................... 5  
1.4 Overview of a revenue management system ..................... 6  
1.5 Measuring effectiveness ................................................ 7

## Chapter 2: Quantity-based Revenue Management

- 8

2.1 Introduction ................................................................. 8  
2.2 Capacity Allocation .......................................................... 8  
2.2.1 Capacity allocation with independent demand ............... 9  
2.2.1.1 Two-class problem ............................................... 9  
2.2.1.2 Multiple-fare classes .......................................... 12  
2.2.2 Capacity Allocation with dependent demands ............... 15  
2.2.2.1 Two fare classes with dependent demands ................... 16  
2.2.2.2 Multiple-fare classes with dependent demands ............. 17  
2.3 Overbooking ................................................................. 18  
2.3.1 Static Model ............................................................ 18  
2.3.2 Dynamic Booking Limits ............................................. 21

## Chapter 3: Price-based Revenue Management

- 23

3.1 Introduction ................................................................. 23  
3.2 Dynamic Pricing ............................................................ 23
3.2.1 Style- and Seasonal Goods Markdown Pricing .............................................. 23
3.2.2 Discount Airline Pricing ............................................................................. 24
3.2.3 Consumer-Packaged Goods Promotion ...................................................... 24
3.3 Auctions ......................................................................................................... 25
3.3.1 Types of Auctions ....................................................................................... 25

Chapter 4: Forecasting ......................................................................................... 27
4.1 Introduction ...................................................................................................... 27
4.2 Forecasting Methods ....................................................................................... 28
  4.2.1 Ad-Hoc forecasting .................................................................................... 28
    4.2.1.1 M-period Moving Average ................................................................. 28
    4.2.1.2 Exponential smoothing .................................................................... 29

Summary .............................................................................................................. 32

Bibliography ........................................................................................................ 33
Preface

The main reason I chose Yield Management as the topic of my thesis is because I wanted to get an insight to an interesting and practical application of mathematics and I was especially interested in a financial topic. I think I made a good decision, because yield management is used every day by very different companies. From the major ones like airlines to even the smallest one, like a barber shop on the corner. So there is no doubt that it is practical. Its main goal is to maximize the profit with the help of mathematical models. These models are easy to implement but they can make a great difference in a company’s revenue.

We are going to make many assumptions to create simple conditions, and although these assumptions are not necessarily true in real life, they are useful, because with the help of them we can create models, which are easy to implement, and calculating them can be real fast. These are the most important aspects for companies using yield management tactics, as we will see later they have to update the calculated numbers frequently in order to make the best decision. And when we can get a nearly optimal solution to our problem in seconds, it is way more useful than if we got the exact solution, but we had to wait hours for it. In today’s extremely fast world, this is simply not an option. Companies who can react, and adapt the fastest to market changes, are going to stand out, and get an advantage against the others. This is the goal of every businessman, to be the best, among his competitors.

The aim of this thesis is to get an insight to this interesting strategy, yield management. Chapter 1 is a brief introduction to yield management, its history, how it works and how we can measure its effectiveness. In Chapter 2 we discuss quantity based yield management, used mainly in the airline industry. In this chapter we will be dealing with companies, which are not flexible concerning the price, so our variable will be the capacity. First we will see a simple model, and then we move on to more complex ones, which are more related to real life problems. Next, in Chapter 3, we move on to price based yield management where, in contrast to the quantity based methods, the focus will be on the price, and on how we should change it over time to make the most profit. In Chapter 4 we will talk about forecasting, which is a common element of both quantity- and price based yield
management. It is vital to make the best forecasts we can, because without a good forecast our models will not work as well as expected.

At last but not least, I would like to thank my supervisors Róbert Fullér and Viktória Villányi for their help, and also my family and friends for their support.
Chapter 1

1.1 Introduction

Yield management is a strategy used by many different types of companies (mainly airlines) in order to maximize the profit. Yield management is not about how many employees we hire, how much we pay for their work, or what we invest our money in. This strategy maximizes profit from another point of view. It focuses on selling the right product to the right customer for the right price at the right time. It is about allocating the capacity to different fare classes. In contrast to other profit maximizing strategies, yield management is not setting and updating prices, it is setting and updating the availability of fare classes. Companies applying yield management can be very different, but they have to meet 4 conditions.

1.) The seller is selling fixed stock of perishable capacity
2.) Customers book capacity prior to the time of service (e.g. in case of an airline company the time of service is the moment the plane departs)
3.) The seller manages a set of fare classes, each of which has a fixed price (at least in the short run)
4.) The seller can change the availability of fare classes over time

1.2 The history of Yield Management

The idea of yield management originates from airline companies in the United States. These were the first ones to apply this tactic, that’s why I put a little more emphasis on the airline companies during my thesis. Before 1978 the airline industry in the United States did not have much freedom, in fact the fares and the schedules were controlled by the Civil Aeronautics Board (CAB). During this period flying was a luxury and fares were very high.
In 1978 the whole airline industry was shocked and everything changed, when the Congress passed the Airline Deregulation Act. It said that in four years the industry would be completely deregulated. This meant companies were free to decide their fares, and their domestic routes. This was a huge change from a totally restricted industry to complete freedom. The main reason this Act was passed, was to encourage new entrants into the business. One of these newcomers was PeopleExpress. It was a small company with extremely low prices, 70% below the bigger airlines. It came as a shock for major airlines, like American Airlines, because they just simply could not compete with such low prices. They had to cover their costs, and this meant low fares were not an option. This was a very hard time, and something needed to be done. If they were to lower their fares they were going to go bankrupt, but if they kept the higher fares they would lose their passengers. It seemed like there was no solution to this problem, and American Airlines will fail to compete in this new deregulated environment. But Robert Crandall, former CEO of American Airlines, thought it otherwise. In 1985 American announced its “Ultimate Super Saver Fares”. People thought it was a joke, a final attempt to avoid bankruptcy, but it was real. American introduced low fares, just like PeoleExpress, or in some cases even lower. There were only two differences:

1.) If a passenger wanted to purchase an “Ultimate Super Saver” fare he needed to book at least two weeks prior to departure, and stay at his destination over a Saturday night.

2.) The number of seats that could be sold for discount price was restricted. In this way American could save seats for full fare customers who book just days before departure.

With these two changes American Airlines segmented the market between leisure and business travelers. Both segments preferred the major airline’s better service, so eventually PeoleExpress was on the edge of bankruptcy.

This is considered to be the birth of yield management. In time, more and more airline companies started applying it and later even other industries too, like hotels and car rental companies. That’s why today the name revenue management is preferred; it is a more general naming. Yield management is used only by airline companies.
1.3 Levels of Revenue management

Revenue management has three levels: strategic, tactical and booking control. Revenue management strategy is the level, where market segments and the price, set for that particular segment, are determined. This is done mainly annually.

Revenue management tactics is the process of determining and updating the limit of how much unit of the capacity can be sold to a particular fare class. This is done frequently, in most cases daily or weekly. In the following chapters I am going to focus on this second level of revenue management.

Booking control has to be done in real time. When a booking request arrives for a flight, we have to decide to accept it, or maybe refuse it and wait for another one that can increase the profit more. The decision is made with the help of a reservation system. This system has the booking limits for all the fare classes.

Airline companies determine the number of available discount seats with the help of booking limits. These are controls that limit the amount of capacity that can be sold to any particular class at a given point in time. If the demand reaches this limit, the class needs to be closed to any other customer from the same class. There are two kinds of booking limits:

- **Partitioned**: A partitioned booking limit contains separate blocks, one for each class.
- **Nested**: In case of a nested booking limit, the available capacity for different classes overlaps in hierarchical order. It excludes the opportunity that a unit is unavailable for a high class customer, but in the same time it is available for a lower class customer.

After checking the booking limit the reservation system determines if we have enough available capacity. If yes, the request is accepted, if not, it is rejected.
1.4 Overview of a revenue management system

Now let us have a look at the structure of a revenue management system, and the way it works. A Revenue management system calculates and updates the booking limits within a reservation system.

The RM system goes through four steps:

1) **Data Collection:** Companies using revenue management have to store their historical data of customer behavior, prices, demand and other factors in order to make good forecasting and estimation. This is the foundation of all yield management systems. The more precisely we can store our historical data, the more precise forecast we can make.

2) **Forecasting and estimation:** During this step we have to estimate the parameters of our model, and after that we have to make predictions using these parameters. Sometimes companies not only forecast demand, but no-shows, or cancellations too. As data is the basis of forecasting, so is forecasting the basis of optimization. Without a good prediction of demand, no-shows, customer behavior, etc. we cannot optimize the controls.

3) **Optimization:** We have to find the optimal set of controls, which we are going to use. These can be booking limits, prices, discounts, etc.

4) **Control:** The final step is to control the sale of inventory using the previously optimized controls.
1.5 Measuring effectiveness

Before the deregulation act was passed in 1978, airlines used load factor to measure their effectiveness. It is calculated by dividing the number of seats sold on a flight by the total number of seats.

This was perfect when there were regulations, but in the 80s, when the conditions changed, the way how performance was measured had to change too. It became more complex. E.g. in the past, the more passengers simply meant more money, but after the deregulation the types of customers needed to be taken into consideration too. Maybe a flight with full of discount passengers is not as profitable as a flight which has empty seats, but all passengers are full fare customers. Because of these changes airlines not only looked at load factor, but yield also. Yield is revenue per passenger mile. But yield does not take into consideration the capacity of the flight. So we need the combination of these two metrics. This is revenue per available seat mile (RASM), which is revenue from available seats normalized by the distance. So e.g. if we have a plane with 100 seats, flight which is 2000 miles, and our total revenue is $100,000, then RASM=$100,000/ (100 x 2000) =$0.5. How revenue is distributed does not have an effect on RASM.

Similarly to airline companies, other industries can also apply this metric, e.g. in case of hotels we speak about revenue per available room night, in case of car rentals revenue per available rental day. So we can measure and compare performance within markets, and also between different markets. It is important to mention that if a company’s RASM is high, it does not automatically mean that its profit is high too, since RASM ignores costs. It is a metric, completely based on revenue.
Chapter 2

Quantity-based Revenue Management

2.1 Introduction

Quantity based yield management is very useful for airline companies. Due to the fact that a plane has a fixed capacity of seats, the company should find the optimal customer mix if it wants to achieve the highest revenue possible. They have to decide how many seats to sell to discount passengers, and how many to save for future passengers, who are willing to pay the full fare.

The main problem of quantity based revenue management is to find an optimal allocation of our resource among the different fare classes. In contrast with the price based revenue management, quantity base methods focus on our resource. How can we make the most of what we have, without changing the prices too much?

We can achieve this with the help of booking limits, which we already mentioned, and protection levels. Besides booking limits protection levels are also important controls. Protection levels set the number of capacity units to reserve for a particular class. These can also be partitioned or nested, like booking limits.

2.2 Capacity Allocation

The main idea behind capacity allocation is that customers, who are sensitive to price, are willing to lower their needs, and sacrifice comfort and time in order to get a cheaper discount ticket. On the other hand, higher fare customers are not sensitive to price, because they want to travel fast and comfortably, so they are willing to pay the full price in order to receive a service suitable for their needs. With these assumptions we can segment the
market into customer groups, which we call *fare classes*, according to a customer’s level of price sensitivity.

We assume that the different fare classes are distinct, and they require the same resource. We also presume that we have homogeneous units of capacity, and that customers demand only a single unit of it. For now we do not allow no-shows or cancellations, we are going to deal with them later.

### 2.2.1 Capacity allocation with independent demand

First we start with the less complex models, where demands are independent. This means that when we close a fare class, it does not affect the demand for other classes.

**2.2.1.1 Two-class problem:**

This is the simplest case, where we assume that we have a fixed capacity \( C < \infty \) and two classes of customers: one is the *discount class passengers*, paying \( p_d \) for a unit (or seat in case of airlines); the other is the *full fare passengers* paying \( p_f \). Obviously it is easy to see that \( p_d > 0 \) and \( p_f > p_d \). Another important assumption we have to make is that all discount requests come before the full fare requests.

With these assumptions in mind the basic question of the two class capacity allocation problem is the following: How many seats of discount customers should we book, when there is a chance of future full fare demand? In other words: What is the optimal booking limit \( b \), for the discount class, for which we can gain the highest possible revenue? If we find the optimal booking limit, then it is easy to find the optimal protection level too. According to the above definition the protection level \( y \) can be calculated from the capacity and the booking limit: \( y = C - b \).

When calculating the optimal booking limit we want to avoid two scenarios: setting the limit too high or setting it too low. If we set \( b \) too low then we talk about *spoilage*, because we turned away discount passengers for future high fare customers, but the number of high fare customers is smaller than the protection level. In this case we are spoiling our capacity. (E.g. in the airline business our plane is flying with empty seats). On the other hand if we set
too low we could have made more revenue by saving the seats for future high class customers, but we sold it to discount passengers. E.g. we have a plane with 100 seats out of which we sold 80 to discount passengers, and 20 to full fare passengers, but the actual demand for the higher fare class seats was 30. So we turned away 10 better paying customers, for discount customers. This obviously decreased our revenue. This scenario is called *dilution*. Our goal is to balance the risks of these two outcomes to maximize expected revenue.

With the help of a decision tree we can illustrate the different outcomes of increasing the booking level by one unit.

\[
d_f = \text{full fare demand} \\
d_d = \text{discount demand} \\
d_d \text{ and } d_f \text{ are independent random variables} \\
F_f(x) = P(d_f \leq x) \\
F_d(x) = P(d_d \leq x)
\]

The right side of the tree gives us the change in expected revenue. If we take the probability weighted sum of these outcomes we can calculate the *expected change in revenue* caused by the fact that we increased \( b \) by 1.


\[ E(h(b)) = F_d(b) \times 0 + [1 - F_d(b)] \left[ [1 - F_f(C - b)](p_d - p_f) + F_f(C - b)p_d \right] \]

\[ = [1 - F_d(b)] \left[ p_d - [1 - F_f(C - b)p_f] \right] \]

where \( 1 - F_f(C - b) = P(\text{full fare demand} > \text{protection level}) \). This number is small when protection level is high.

From the above equation we can determine whether or not we should increase the booking limit for low fare class. If the value of \( E(h(b)) < 0 \), then increasing the booking limit by one decreases our revenue, but if \( E(h(b)) > 0 \), then we should definitely increase the booking limit, because we have a chance for higher profit.

So if \( p_d < [1 - F_f(C - b)]p_f \) we should not allocate any more seats to discount passengers.

If \( p_d > [1 - F_f(C - b)]p_f \) we should allocate at least one seat for a discount customer.

We can easily see that we get the optimal booking limit \( (b^*) \) if

\[ p_d = [1 - F_f(C - b^*)]p_f. \]

This equation can also be written as

\[ \frac{p_d}{p_f} = 1 - F_f(C - b^*). \]

Note that the optimal protection level \( y^* = C - b^* \). So the above equation is equivalent to

\[ \frac{p_d}{p_f} = 1 - F_f(y^*). \]

This was first described by Kenneth Littlewood in 1972, and is known as Littlewoods’s Rule.

Littlewood’s Rule says that: “to maximize expected revenue, the probability that full fare demand will exceed the protection level should equal to the ratio \( \frac{p_d}{p_f} \) (Robert L. Phillips, Pricing and Revenue Optimization (2005))

An interesting fact is that the optimal protection level only depends on the two fares and the distribution of expected full fare demand. The discount demand or the capacity is not in the equation, so the optimal protection level does not depend on them in the two class model.
2.2.1.2 Multiple-fare classes:

The next model we are going to discuss is the case of not just two, but multiple fare classes. In this model we also have fixed capacity \( C \), but we have \( n \) classes with \( n \) fares for each one. Class one has the highest price and the \( n \)th class has the lowest. \( p_1 > p_2 > \cdots > p_n \).

d\(_i\) represents the demand in class \( i \). These are independent random variables.

\[
f_i(x) = P(\text{demand for class } i \text{ will be } x \geq 0) \]

\[
F_i(x) = P(d_i \leq x)
\]

We assume that bookings arrive in increasing fare order, so class \( n \) customers book first. We still assume that there are no cancellations or no-shows. Our goal is the same as in the case of the two class model: to maximize expected revenue.

The main problem of the multiple fare class capacity allocation is that when a booking period starts we have to set a booking limit for that particular class. \( b_j \) denotes the booking limit for class \( j \). So we have to calculate \( n \) booking limits.

Suppose we are at the start of period \( j \). We know the number of bookings we already accepted:

\[
\sum_{i=j+1}^{n} x_i.
\]

Due to the fact that we do not allow no-shows and cancellations, if we subtract this number from our total capacity, we get the unbooked capacity at the beginning of booking period \( j \), we denote this capacity by \( C_j \).

\[
C_j = \left[ C - \sum_{i=j+1}^{n} x_i \right]
\]

In order to find the optimal booking limit we have to go backwards. Namely from the beginning of the final period. At that time we have \( C_1 \) remaining capacity, and since we have no no-shows or cancellations we allow all of the unbooked capacity to be booked. After all, we do not want to spoil our resources. So the first class booking limit is \( b_1 = C_1 \). Then we go to the next period, class 2. Now we have \( C_2 \) unbooked capacity, and we have two fare
classes: the first and the second. So this case is exactly the same as the two class model discussed previously, so we can use Littlewood’s Rule to determine the optimal booking limit for class 2 ($b_2^*$). $b_2^*$ is the smallest value of $0 \leq b_2 \leq C_2$ for which

$$F_1(C_2 - b_2) < 1 - \frac{p_2}{p_1}.$$ 

When we get to period 3, the main goal is the same: to balance spoilage and dilution. We can also use a decision tree for illustrating the different changes in expected revenue.

The hard part of the multiple fare class problem is that usually we don’t know the values of probabilities $q_1$ and $q_2$, and there’s no easy way to calculate them. So instead of determining the optimal booking limit, we have to calculate a good, but not optimal one. For this we can use the expected marginal seat revenue (EMSR) heuristics. The main idea of these heuristics is to use a bunch of two class problems and applying Littlewood’s rule to calculate the nearly optimal solution. The two types of EMSR we are going to discuss are called version a and version b.

EMSR-a:

This version calculates the booking limits in two steps:

1) It determines the protection levels for the current class relative to each of the higher classes with the help of Littlewood’s rule

2) We get the protection level of the current class by adding up the previously calculated protection levels.
For example, let’s consider the case when we are in class 3. Now EMSR-a calculates protection level for class 3 against class 1 \((y_{31})\) with the help of Littlewood’s Rule

\[
y_{31} = F_1^{-1}\left(\frac{p_1 - p_3}{p_1}\right)
\]

and protection level for class 3 against class 2 \((y_{32})\)

\[
y_{32} = F_2^{-1}\left(\frac{p_2 - p_3}{p_2}\right)
\]

The sum of these two protection levels gives us the protection level for class 3 \((y_3)\)

\[
y_3 = y_{31} + y_{32} =
\]

\[
= F_1^{-1}\left(\frac{p_1 - p_3}{p_1}\right) + F_2^{-1}\left(\frac{p_2 - p_3}{p_2}\right)
\]

The booking limit for class 3 is

\[
b_3 = \left[C_3 - F_1^{-1}\left(\frac{p_1 - p_3}{p_1}\right) - F_2^{-1}\left(\frac{p_2 - p_3}{p_2}\right)\right]
\]

We can use this method for any given number. If \(j \geq 2\) the EMSR-a protection level is for class \(j\):

\[
y_j = \sum_{i=1}^{j-1} F_i^{-1}\left(\frac{p_i - p_j}{p_i}\right)
\]

EMSR-b:

This version of EMSR creates an “artificial class” in every booking period, for which the demand \((\hat{d}_j)\) is the sum of all the future periods,

\[
\hat{d}_j = \sum_{i=1}^{j} d_i
\]

and its fare \((\hat{p}_j)\) is the average of the expected fares from future bookings
\[ \hat{p}_j = \frac{\sum_{i=1}^{j} p_i \cdot d_i}{\sum_{i=1}^{j} d_i} \]

After we determine these demands and fares, we can calculate \( b_j \) relative to the artificial class with the help of Littlewood’s Rule.

\[ y_j = F^{-1}\left( \frac{\hat{p}_j - p_j}{\hat{p}_j} \right) \]

2.2.1.3 Bid Pricing

Bid pricing can be described with a single rule:

“Accept a single-seat booking request if it’s associated fare is greater than or equal to the fare in the lowest open class. Otherwise reject it.” (Robert L. Phillips, Pricing and Revenue Optimization (2005))

We call the minimum acceptable fare a bid price.

The main difference between booking limits and bid prices is that booking limits are revenue based, while bid prices are class based.

Although bid prices are easy to use, they have some disadvantages:

1) With bid pricing we can only handle booking for a single seat
2) It needs to be recalculated every time request happens.

2.2.2 Capacity Allocation with dependent demands

So far we only discussed independent demands among fare classes, but in the real world things are more complex. In reality we cannot just separate customers into different classes; we cannot say that one person is a full fare customer, the other one is a discount customer.
In fact, everybody, even the full fare customers, are in search for the lowest available prices, that suits their needs best. In this case we have to introduce a new definition.

*Buy up* (sell up): Buy up means that when we close a discount class, it affects future full fare demand; more specifically it increases the demand.

So in contrast to the previous models, now the demands are dependant.

We will follow the same structure as in the case of independent demand model, so we start with only two classes.

2.2.2.1 Two fare classes with dependent demands:

The notations are the same as in the independent demand model, but we have to redefine demands.

\[ d_d = \text{total discount demand} \]

\[ d_f = \text{total full fare demand assuming all discount bookings are accepted} \]

Also, we have to introduce a new element \( \alpha \), which is a number that describes the *fraction of discount demand that are willing to book for full price in case they are rejected.* (\( 0 \leq \alpha \leq 1 \))

With the help of this \( \alpha \) we can calculate the actual full fare demand: \( \hat{d}_f = d_f + \alpha [d_d - b] \)

As before, we illustrate the changes in expected revenue with the help of a decision tree.
We get the impact of increasing $b$ by one on expected revenue the same way: by adding up the probability weighted sum of the different outcomes.

$$E[h(b)] = F_d(b) \times 0 + [1 - F_d(b)][(1 - F_f(C - b))(p_d - p_f) + F_f(C - b)(p_d - \alpha p_f)]$$

$$= [1 - F_d(b)][p_d - \alpha p_f - (1 - F_f(C - b))(1 - \alpha)p_f]$$

If $\alpha \geq \frac{p_d}{p_f}$ the term

$$[1 - F_d(b)][p_d - \alpha p_f - (1 - F_f(C - b))(1 - \alpha)p_f]$$

will be less than or equal to 0 for any value of $b$. In this case $b^* = 0$.

If $\alpha < \frac{p_d}{p_f}$, we can find the optimal protection level applying Littlewood’s Rule, but with a modified fare ratio:

$$1 - F_f(y^*) = \frac{p_d - \alpha p_f}{(1 - \alpha)p_f}$$

$$\Downarrow$$

$$1 - F_f(y^*) = \frac{1}{1 - \alpha} \left( \frac{p_d}{p_f} - \alpha \right)$$

Where $\frac{1}{1 - \alpha} \left( \frac{p_d}{p_f} - \alpha \right)$ is the modified fare ratio.

2.2.2.2 Multiple fare classes with dependent demand

When we have multiple fare classes we can apply the modified EMSR heuristics. These heuristics use an additional assumption, namely that buy up is only possible in the next highest class but no higher.

$\alpha_j$ ($j = 2, 3, \ldots, n$) denotes the fraction of class $j$ customers who are willing to buy up in case of rejection, and pay the fare of class $j - 1$.

EMSR-a with buy-up:
\[ y_j = \sum_{i=1}^{j-1} F_i^{-1} \left[ \frac{1}{1 - \alpha_j} \left( \frac{p_i - p_j}{p_i} \right) \right] \]

EMSR-b with buy-up:

\[ y_j = F^{-1} \left[ \frac{1}{1 - \alpha_j} \left( \frac{\hat{p}_j - p_j}{\hat{p}_j} \right) \right] \]

### 2.3 Overbooking

Overbooking is a very important part of quantity based revenue management. So far our main goal was to optimize customer mix, however when we talk about overbooking we want to increase the total volume of sales. Overbooking is when a seller sells more units of capacity than he actually has available. Why would somebody do this?

In the previous models we assumed that there are no no-shows or cancellations, but in real life it happens all the time that somebody changes his mind and cancels his booking, or does not show up at all. Companies would lose huge amounts of revenue if they didn’t overbook. The main challenge is when more customers show up, than the available capacity. The company has to manage the negative effects of denying its service to a customer. A very important rule is that selection must not be discriminatory (except for price discrimination of course).

#### 2.3.1 Static model

First we start with the easier model, where there is no distinction between cancellations and no-shows. Furthermore we assume that no-shows are independent, and they have the same probability. The first thing a company needs to do is determine its objective. It is important because with the objective we know what function we want to maximize.
Overbooking policies:

1) **Deterministic heuristic**: it is the simplest one, it calculates booking limits with the help of consistent historic show rate ($\rho$). This policy says that we should set the booking limit $b$ so that $C = \rho \times b$. It is a simple solution, but in real life it is often used.

2) **Risk-based policies**: Before we talk about risk based policies, we have to introduce a new definition, namely the denied boarding cost. This is the amount of money we have to pay if a passenger, we expected to be a no show, would actually show up. In this case, unfortunately we have to deny our customer to board, because we have a fixed capacity. This is a very risky situation; we have to handle it carefully because we don’t want to have disappointed customers.

The risk based policies focus on setting $b$ by balancing the expected denied service cost with the potential revenue we can make by selling more seats. We denote the total number of passengers, who show up at departure by $s$. If $p$ represents the price, then the total revenue equals to $s \times p$. The denied service cost of one passenger is denoted by $D$, and if $s$ is greater than our capacity ($C$), we have to refuse $s - C$ passengers. So the total denied service cost is $D(s - C)$. We want to maximize our revenue, so the objective function will be the net revenue, which is calculated as follows:

$$R = ps - D(s - C)$$

Let $n$ represent the number of bookings at departure. It is easy to see that $n = \min(d, b)$, where $d$ stands for total demand. If $x$ represents the number of no shows, then $s = n - x = \min(d, b) - x$. From these we can calculate the fraction of passengers who showed up $\rho = \frac{s}{n}$, and then we get the no show rate which is

$$1 - \rho = \frac{x}{n}$$

We assume that the number of no shows does not depend on the number of bookings ($d$).

$$F(d) = P(\text{total number of booking requests} \leq d)$$
\[ G(x) = P(\text{total number of no shows} \leq x) \]

So to calculate the probability that demand is less than \( d \) and no shows are less than \( x \), we can just simply multiply \( F(d) \) and \( G(x) \).

After we clarified every notation and assumption we can maximize our objective function, which is

\[
E[R|b] = p \cdot E[s] - D \cdot E[(s - C)] \\
= p \cdot E[\min(b, d) - x] - D \cdot E[(\min(b, d) - x - c)]
\]

With a decision tree we can illustrate the different outcomes of increasing \( b \) by one, and its effect on the revenue.

As previously, we can add up the probability weighted values of the branches of the tree, and we get the change in expected net revenue.

\[
E[R|b + 1] - E[R|b] = [1 - F(b)]\{G(b - C)(p - D) + [1 - G(b - C)]p\} \\
= [1 - F(b)]\{p - G(b - C)D\}
\]

So the optimal booking limit is the smallest value of \( b \) for which

\[
\frac{p}{D} < G(b - C)
\]
3) **Service-level policies:** As we mentioned above, managers try to avoid denying service, because they do not want to disappoint their customers and lose them. Apart from this, you can’t really measure this in any number. Maybe some customers are more flexible and they are not going to fly with the competitor airline because of one disappointing flight, but maybe they will. And you can never tell which one it is going to be. This is the basis of service level policies.

These methods try to minimize the fraction of booked customers, to whom we will need to deny service, because of overbooking. First we set a number, e.g. 10,000, out of which we only want to allow one denied service. Then we calculate the smallest value of $b$ for which

$$\frac{E[(s|b) - C]}{E[(s|b) - C]} = \frac{1}{10,000}$$

Another variation of service level policies is setting the fraction of customers served as allowed number of denied service customers. Let this fraction be denoted by $q$. Then we want the smallest value of $b$ for which

$$E[(s|b) - C] = qE[\min((s|b), C)]$$

4) **Hybrid policies:** When applying this kind of policy, you calculate both risk based and service level policies, and then you choose the minimum of the two.

2.3.2 Dynamic Booking Limits

So far we did not make any difference between no shows and cancellations. Although, in reality these two are not the same. *No shows* are the customers, who keep their bookings until departure; they just simply do not show up. However *cancellations* do not keep their booking requests, but they cancel it prior to departure, so our booking limit changes over time. That’s why we want to use *dynamic booking limits*. 


Let’s use a dynamic cancellation rate $r(t)$, where $t$ represents the number of remaining days until departure. We denote the total booking at time $t$ by $m(t)$. Then

$$[1 - r(t)]m(t)$$

people will keep their booking requests at time $t$, out of which

$$\rho[1 - r(t)]m(t)$$

customers will actually show up. A commonly used method is to consider this $\rho[1 - r(t)]$ as a show rate that changes over time. Then we can simply apply a risk based policy, discussed earlier.

After the calculation we will get a dynamic booking limit, which we have to update regularly. Usually as time goes by, and departure approaches, the number of cancellations will decrease, and the number of bookings will increase. Airline companies update their booking limits more frequently as $t \to 0$. 
Chapter 3

**Price-based Revenue Management**

### 3.1 Introduction

In this chapter we talk about price based revenue management. These methods are becoming more and more widespread, as companies can react faster to market changes than they used to. In this chapter our variable will be the price, not the capacity. Companies which are flexible regarding their prices will prefer this kind of revenue management, instead of quantity based revenue management systems. One of the most important methods is dynamic pricing.

### 3.2 Dynamic Pricing

#### 3.2.1 Style- and Seasonal-Goods Markdown Pricing

As I mentioned in the beginning of my thesis, yield management applies to perishable assets. This means that our goods will lose their value and we cannot sell them after that, so we cannot make profit. Every company’s goal is to reduce spoilage to the minimum. So before the inventory loses its value, companies will lower prices, in hope of increase in demand. It is always better to sell something for a low price, then to not sell it at all. This is a very simple, but useful use of dynamic pricing to gain more revenue.

Another scenario can be that we do not have enough information about the demand of our clients. So first we need to set all prices high and as time goes by, we can observe which products are the ones for which people are willing to pay the high price. For the rest of the products we can lower our prices. Applying this kind of dynamic pricing is also a good way to learn about customer demand.
3.2.2 Discount Airline Pricing

It is worth mentioning this particular case, because when it comes to airlines we have to set our prices differently. As departure day approaches (the day when our inventory loses its value), in contrast to style goods, prices become higher. There is a simple explanation to this. Airline companies prefer customers who book early, but if customers book tickets early and then see that prices dropped, they are going to be furious. It would lead to high uncertainty if everybody wanted to buy their tickets the day before departure.

3.2.3 Consumer-Packaged Goods Promotion

The main difference in this case is that a promotion is a short term price reduction. Consumer-packaged goods are products that customers buy repeatedly, so they know how the prices change, and they are willing to adapt to these changes. For example if the price of coffee drops, then customers tend to buy a lot of packages, they stockpile it. So we can see a drop in demand after the promotion ends. We have to be careful about the length of the promotion. If it is too long, our product can lose its value. But it is great way if we want extra attention on our product and we want to advertise it.

There are three kinds of promotions, depending on the person who runs it.

- *Trade promotions*: These are run by a manufacturer to retailers. It may not even be passed on to customers
- *Retail/Consumer promotions*: In this case the promotion is run by the retailer to the customers
- The third type, which does not have a general name, is when manufacturers run promotions directly to the customers (e.g. coupons you get in mail)

The goals of the different types of promotions are different too. When we speak of a trade promotion, the manufacturer would like to increase his profit, by selling more from its own product. But in case of retail promotion the retailer would like to increase his profit with the maximization of his overall sell, not with maximizing the selling of just one particular product.
3.3 Auctions

Besides dynamic pricing, auctions are another form of tactics used in price based revenue management. The main difference between the two is that in case of dynamic pricing companies decide about the price of their products, but when it comes to auctions, the customers are the ones offering prices. This is called bid, and companies decide which one to accept. An auction is basically a set of rules, which is called mechanism. In case of auctions, prices can easily follow the changes in the market and customer demand. One of the biggest advantages of this kind of price based revenue management is that we don’t need much information about customer demand and their willingness to buy our product.

3.3.1 Types of auctions

There are a great variety of auctions, but now we just discuss a case, in which we want to sell one indivisible product, and \( N \) customers are offering the bids.

- **Open ascending (English) auction**
  
  The pricing method of open ascending auction is that the seller offers increasing prices to the buyers, and with the show of their hands they can indicate if they are willing to buy the product for that price. This goes on until only one buyer remains. This method is used mainly for selling art in auction houses.

- **Open descending (Dutch) auction**
  
  In case of open descending auction the sellers offer prices, but in contrast to the previous type, the prices are now decreasing. The first customer, who is willing to pay the offered price, wins the product. The Aalsmer and Naaldwijck flower markets are a great example for this type of mechanism.

- **Sealed-bid, first price auction**
  
  The way this auction works is that potential buyers submit sealed bids to the company. Naturally, the buyers don’t know how many customers make bids, and how much they offer. The customer, who submitted the highest bid, wins the product.

- **Sealed-bid, second-price auction (Vickrey auction)**
It works exactly like the previous auction, the only difference is that the customer, who made the highest bid and won the product, does not have to pay the highest bid, but the second highest one.
Chapter 4

Forecasting

4.1 Introduction

Estimation and forecasting methods are vital elements of both the price based revenue management systems and the quantity based revenue management systems. Quantity based systems usually collect historical data and use time series methods to predict future demand. Besides demand booking profile forecasting is also important, it describes how bookings arrive. Cancellation, no show probabilities and buy up factors are also important to calculate.

In case of the price based systems we predict the demand as a function of a variable, e.g. price. So we have to estimate the parameters of this function. Additional predictions like how the demand is affected by the size of the population, behavior changes or a low-inventory may also be important. We can forecast these with the help of historical price-demand relationships.

When we want to estimate something we need to find parameters that best describe a given set of observed data. Estimation is done much less frequently than forecasting, because it is basically the setup of the parameters of a forecasting model.

On the other hand forecasting is about predicting unobserved values, and it is done relatively frequently. Because of its frequent use the forecasting method should be fast and easy to implement. In this chapter we are going to focus on forecasting demand, because for a revenue management system it is the most important information to predict. But there are cases when we have to forecast other quantities, such as cancellation, no shows, and prices or, in case of a hotel, length of stay.
4.2 Forecasting methods

4.2.1 Ad-hoc forecasting

These methods are basically heuristics. They try to average out the noise from the data and predict the trend, seasonality and level.

- **Trend**: A predictable increase or decrease in values over time. Usually modeled with linear functions.
- **Seasonality**: A periodic or repeating pattern in the values over time
- **Level**: Typical or average value of the data.

These methods smooth the data, and with the help of this smoothed series, we can predict future demand.

Next we talk about how this smoothing can be done.

4.2.1.1 $M$-period moving average

This method’s basic assumption is that the most recent observations describe the future better than older data. That’s why it only uses the past $M$ observed data, not all the data.

$t$ represents present time and we want to make predictions for values at time $t + k$. We denote the $k$-period ahead forecast by $\hat{Z}_{t+k}$.

$z_1, \ldots, z_t$ denotes the observed demand so far, $\hat{Z}_{t+1}, \ldots, \hat{Z}_{t+K}$ denotes the forecasts. If we want to forecast just one period ahead a good prediction is to take the average of the past $M$ observations:

$$\hat{Z}_{t+1} = \frac{z_t + z_{t-1} + \cdots + z_{t-M+1}}{M} = \hat{Z}_t + \frac{z_t - z_{t-M}}{M}$$

This is called the simple $M$-period moving-average forecast. With the help of this we can forecast $k$-period ahead:

$$\hat{Z}_{t+k} = \hat{Z}_{t+1}, \quad k = 2, \ldots, K$$
4.2.1.2 Exponential smoothing

These methods have good accuracy; in addition they are simple and fast so they are very popular. Before we go on some notations need to be clarified:

\( A_t \) = the estimate of the level for period \( t \)

\( T_t \) = the estimate of the trend for period \( t \)

\( S_t \) = the estimate of the seasonality for period \( t \).

Now we have a look at three different methods of exponential smoothing.

**Simple exponential smoothing**

This is the simplest method, it has only one parameter, the smoothing constant for the level, denoted by \( \alpha \) (0 < \( \alpha \) < 1). We do not take trend and seasonality into consideration.

\[
\hat{Z}_{t+1} = A_t = \alpha z_t + (1 - \alpha)\hat{Z}_t =
\]

\[
= \alpha z_t + (1 - \alpha)(\alpha z_{t-1} + (1 - \alpha)\hat{Z}_{t-1}) =
\]

\[
= \alpha z_t + \alpha(1 - \alpha)z_{t-1} + (1 - \alpha)^2\hat{Z}_{t-1} =
\]

\[
= \cdots =
\]

\[
= \alpha \sum_{j=0}^{\infty} (1 - \alpha)^j z_{t-j}
\]

\[
\hat{Z}_{t+k} = \hat{Z}_{t+1}, \quad k = 1, \ldots, K
\]
The value of $\alpha$ is determined before the start of the system. Usually it is set between 0.05 and 0.3. Smaller $\alpha$ makes the forecast more stable and larger values make it more sensitive to noise, but also to recent changes in level.

In contrast to the above mentioned $M$-period moving average, this method uses all past data, not just some of them. As you can see from the formula this is the weighted sum of past data, where the weights are decreasing as we go further back in time.

**Exponential smoothing with linear trend**

If we can observe a relatively stable trend through a period of time over and over again (e.g. annually) in our data, it is recommended to use this method. Now we have two parameters: the smoothing factors for level ($0 < \alpha < 1$), and for trend ($0 < \beta < 1$). Besides this $\beta$ should be smaller than or equal to $\alpha$, because we want the smoothing of the trend to be at least as stable as the smoothing of the level.

We can calculate the $k$th period ahead forecast the following way

$$\hat{Z}_{t+k} = A_t + kT_t, \ k = 1, ..., K$$

where

$$A_t = \alpha z_t + (1 - \alpha)(\hat{Z}_t + T_t)$$

is an average level

and

$$T_t = \beta(\hat{Z}_t - \hat{Z}_{t-1}) + (1 - \beta)T_{t-1}$$

is the estimation of the trend.

We calculate the average level once throughout the process. A great advantage of this is that if the trend stops ($T_t = 0$) our forecast would still be close to the average level. The estimated trend needs to be updated in every period.
**Holt-Winter’s method**

We can apply this method if our data has periodic seasonality, next to trend. We denote this period by $L$. $\alpha, \beta,$ and $\gamma$ are parameters that control the smoothing of level, trend and seasonality.

$$0 < \alpha < 1, \quad 0 < \beta < 1, \quad 0 < \gamma < 1$$

We calculate the forecast for $t + k$ with the following formula:

$$Z_{t+k} = (A_t + kT_t)S_{t+k-L}, \quad k = 1, \ldots, K$$

Where

$$A_t = \frac{Z_t}{S_{t-L}} + (1 - \alpha)(\hat{Z}_t + T_t)$$

is the average level,

$$T_t = \beta_t(\hat{Z}_t - \hat{Z}_{t-1}) + (1 - \beta)T_{t-1}$$

is the estimated trend and

$$S_t = \gamma \left(\frac{Z_t}{S_t}\right) + (1 - \gamma)S_{t-L}$$

represents the estimated seasonality.
Summary

In my thesis I gave a little insight into the complex system of revenue management tactics. We saw that the first step is to collect as many historical data as possible, and make the best database we can, because without that the system could not work right. Next we have to predict the uncertain probabilities of future demands, no-shows, cancellations, etc. It is also recommended to learn about customer behavior as much as we can, so we can recognize possible periodicity, and peek seasons. After that we have to decide which type of revenue management we are going to apply, what controls we are going to use. If a company can change its prices greatly over periods of time it is recommended to use price based revenue management. But if a company cannot do this it has to rely on finding the right customer mix. It should make the most of what it has without changing the prices too much. The final step is to control the sale of inventory using the previously optimized controls.

Of course the topic deepens much further, than I could discuss in my thesis, but in the real world the tendency is that companies prefer these simple models, because they are fast, easy to understand and easy to implement. And at the end everything in business depends on just this. How fast can we adapt to the conditions of a rapidly changing world? How fast can we react to the continuous changes of customers’ behavior, who nowadays have more power than ever? With the widespread use of the Internet, customers can learn about the competitor companies’ prices in just a second, so they have more options to find the product that suits their needs best. In an environment like this companies should change their point of view about profit. It can be increased not only by managing the internal market (like employees, wages, etc.), but by managing the external one too (customers). And yield management is the perfect choice, as it deals with this exact problem.
Bibliography


