Chapter 3
Overbooking

3.1 Introduction

Early on, many airlines adopted the policy of not penalizing booked customers for canceling reservations at any time before departure. Some would not even penalize those that did not show up for booked flights. In essence, an airline ticket was “like money” since it could be used at full face value for a future flight or redeemed for cash at any future date. In the 1960s, no-shows were becoming a problem for airlines who found that flights that were fully booked were departing with many empty seats. In response, the airlines began to overbook as a means of hedging against no-shows. If a flight had more passengers show up than there were seats available, then the airlines would bump some passengers. The bumped passengers would be re-booked on a later flight. In addition, bumped passengers would be given other compensation, often a meal at the airport and a discount certificate applicable to future travel. The cost to the airline of bumping a passenger is called the denied boarding cost. The denied boarding cost would include the cost of putting a bumped passenger on another flight to her destination, the cost of any direct compensation to the bumped passenger, the cost of the meals or lodging that the airline provides to each bumped passenger, and the cost of “ill will” incurred by bumping the passenger. These costs can be different for each flight. For example, a passenger bumped from the last flight of the day will be provided with a hotel room at the airline’s expense.

While unpopular with passengers, overbooking was effective at increasing load factors and revenues. This raised the issue of determining the right booking limit for a flight. When overbooking is allowed, the booking limit can exceed the capacity on the flight, allowing the airline to book more passengers than the capacity. If the booking limit is set too low, there will be lots of empty seats. On the other hand, if the booking limit is set too high, the benefits of filling the aircraft would be overwhelmed by the denied board costs paid. Determining the optimal booking

G. Gallego, H. Topaloglu, *Revenue Management and Pricing Analytics*,
International Series in Operations Research & Management Science 279,
https://doi.org/10.1007/978-1-4939-9606-3_3

limit was one of the first revenue management problem to be successfully analyzed utilizing the methods of operations research.

Airlines significantly changed their overbooking policies over the years. For example, airlines instituted auctions as a mechanism for identifying people who would be willing to forego their seat on a flight in return for compensation in the form of a future flight discount. This practice proved to be popular with passengers and dramatically reduced the number of involuntary denied boardings. Airlines now also sell non-refundable or partially refundable tickets, particularly at lower costs. Both of these developments have implications for the analysis of overbooking policies.

In this chapter, we study a variety of overbooking models. These models can be viewed as the extensions of the models considered throughout the book to deal with overbooking. In Sect. 3.2, we begin with a static, overbooking model with a single fare class. We characterize the optimal booking limit. In Sect. 3.3, we move on to an overbooking model with multiple fare classes over a single flight leg and characterize the structure of the optimal policy. In Sect. 3.4, we conclude the chapter with overbooking models over a network of flight legs.

3.2 Overbooking for a Single Fare Class

Suppose that a flight has capacity c, the unconstrained demand at a single fare $p > 0$ is D. We assume that passengers who do not show up are given a full refund, and that the unit cost for denied boarding is θ. Let b be the booking limit. Then $N := \min(D, b)$ is the number of bookings. The goal is to find a booking limit that maximizes the expected profit, which is given by the difference between the expected revenue from sold seats and the expected cost of denied boardings.

Let $Z(N) := Z(\min(D, b))$ denote the number of passengers that show up for the flight. We assume that each passenger shows up for the flight with probability q independent of everyone else, so $Z(N)$ is a conditional binomial random variable with parameters (N, q). We can express the expected profit as a function of the booking limit as

$$R(b) := p\,\mathbb{E}[Z(\min(D, b))] - \theta\,\mathbb{E}[Z(\min(D, b)) - c]^+]. \tag{3.1}$$

An optimal booking limit, say b^*, is the largest maximizer of $R(b)$. The next proposition provides a formula for $R(b + 1) - R(b)$ that shows that this quantity is always non-negative when $\theta < p$, and in this case, there is no need to impose a booking limit. When $\theta > p$, the quantity can change only from positive to negative, so b^* is the smallest b such that $R(b + 1) - R(b) < 0$.

Proposition 3.1

$$R(b+1) - R(b) = \mathbb{P}\{D \geq b+1\}\, q\, (p - \theta\, \mathbb{P}\{Z(b) \geq c\}).$$

$$b^* = \min \left\{ b \geq 0 : \mathbb{P}\{Z(b) \geq c\} > \frac{p}{\theta} \right\}. \tag{3.2}$$

Formula (3.2) has some resemblance to Littlewood's rule derived in Chap. 1.

The optimal booking limit given above may yield high booking limits and result in large numbers of denied boardings. In fact, assuming that the demand quantities are large enough that we always have $\min\{D, b^*\} = b^*$, the formula for the optimal booking limit given above implies that the fraction of flights with shows exceeding capacity is roughly p/θ. When $\theta = 2\,p$, roughly half of the flights will have denied boardings. This observation motivates the adoption of frequency-based policies by many airlines, where airlines set a target frequency, say f, for the fraction of booked passengers that would be denied boarding. Under this policy, the airlines would set a booking limit as the largest integer b such that

$$\frac{\mathbb{E}\{[Z(\min(D, b)) - c]^+\}}{\mathbb{E}\{Z(\min(D, b))\}} \leq f.$$

In many cases, airlines use hybrid policies, where they calculate the booking limit that maximizes the expected profit and the booking limit that limits the fraction of passengers that are denied boarding, and use the smallest of the two booking limits.

3.3 Overbooking for Multiple Fare Classes

In this section, we present a model for a single flight with multiple fare classes and overbooking. There are n fare classes indexed by $1, \ldots, n$. We assume that the fare classes are ordered such that $p_1 \geq p_2 \geq \ldots \geq p_n$ and the demand from different fare classes arrive sequentially in the low-before-high order. Throughout this section, we make a number of simplifying assumptions to obtain a tractable model. First, we ignore the cancellations and assume that there are only no-shows. Furthermore, the no-show probability for all customers is the same and the no-show decisions of the different customers are independent of each other. The probability that a customer shows up for the flight does not depend on when she booked the ticket. Finally, the refunds and the denied-service costs are the same for all customers. These assumptions imply that the number of no-shows and the cost of no-shows are only a function of the total number of reservations on hand. As a result, we need to retain only a single state variable that keeps track of the total number of reservations, which helps keep the dynamic programming formulation tractable.

Among our assumptions, the most restrictive ones are perhaps the assumptions that the no-show probability, the refunds, and the denied-service costs are the same for all customers. In practice, cancellation options and penalties are often linked to a particular class, so no-show rates and costs can vary significantly from one class to the next. In certain cases, reservations from groups may be canceled simultaneously, which makes the assumption of independent show-up decisions somewhat unrealistic. There seems to be reasonable empirical evidence to support the assumption that the show-up probabilities of the customers do not depend on when they made their reservations.

As in previous section, the capacity on the flight is c. Each reservation shows up with probability q. We use the random variable $Z(y)$ to capture the number of passengers that show up for the flight given that we have y reservations just before the departure time. Thus, $Z(y)$ is binomially distributed with parameters (y, q). We ignore cancellations and assume that we do not give any refunds to the passengers who do not show up, but we will shortly discuss how to relax both of these assumptions. The cost of denying boarding to a reservation is θ. We use the random variable D_j to capture the demand from fare class j. Our goal is to find a policy to decide how much demand to accept from each class to maximize the total expected profit, where the total expected profit is given by the difference between the revenue from the accepted bookings and the penalty cost from the denied boardings.

For any j, we let $V_j(y)$ to denote the optimal total expected revenue that can be obtained from classes $j, \ldots, 1$, given that we have y reservations on hand at the beginning of stage j. Notice that instead of remaining capacity, we use the number of reservations on hand as the state variable. At the beginning of stage j, we observe the demand from fare class j. Knowing the number of reservations, we decide how many new requests to accept. After all of the n stages, a portion of the reservations show up. If the number of reservations that show up exceed the capacity available, then we incur the denied boarding cost. The sequence of events that we use here is different from the one in Chap. 1, where we first choose the booking limit, then accept as much demand as the booking limit allows. It turns out that both of these sequence of events give rise to the same policy, and our goal is to demonstrate an alternative dynamic programming formulation for the multiple class revenue management problems. Using u to denote the portion of the demand that we accept from a fare class and following the sequence of events that we just described, the dynamic programming formulation of the problem is given by

$$V_j(y) = \mathbb{E}\left\{ \max_{0 \le u \le D_j} p_j\, u + V_{j-1}(y+u) \right\}, \tag{3.3}$$

where we charge the denied boarding cost of the reservations that we cannot accommodate on the flight through the boundary condition

$$V_0(y) = -\theta\, \mathbb{E}\{[Z(y) - c]^+\}. \tag{3.4}$$

In this section, we will show that the optimal policy has the following structure. At each stage j, there exists a booking limit b_j^* such that it is optimal to bring the total number of accepted reservations as close as possible to b_j^* after making the decisions for class j. In the sequence of events for our dynamic program, we observe the demand from fare class j first, then decide what portion of this demand to accept. However, the structure of the optimal policy is such that we bring the total number of accepted reservations as close as possible to some fixed number b_j^* after making the decisions for fare class j. Thus, at the beginning of fare class j, we can set the booking limit to b_j^* before we even observe the demand from fare class j. This implies that assuming that we observe the demand before we decide what portion to accept or setting a booking limit before we observe the demand result in identical policies.

3.3.1 Optimal Booking Limits

Assume that the value functions $\{V_j(\cdot) : j = 1, \ldots, n\}$ computed through the dynamic program in (3.3) are concave. This implies that $\Delta V_j(z) := V_j(z) - V_j(z - 1)$ is decreasing in z for all j.

Under this assumption, we show that the optimal policy can be characterized by a booking limit b_j^* for each stage j, such that it is optimal to bring the total number of reservations as close as possible to b_j^* after making the decisions for class j. Once we show this result, we will verify that concavity of the value function.

Theorem 3.2 *Assume that $V_{j-1}(\cdot)$ is concave, and let b_j^* be the maximizer of the concave function $p_j z + V_{j-1}(z)$ over* $[0, \infty]$. *Then,*

$$b_j^* = \min\{z \geq 0 : p_j + \Delta V_{j-1}(z + 1) < 0\}.$$

In this case, setting

$$u^*(y) = \begin{cases} 0 & \textit{if } b_j^* < y \\ b_j^* - y & \textit{if } y \leq b_j^* \leq y + D_j \\ D_j & \textit{if } b_j^* > y + D_j. \end{cases} \tag{3.5}$$

solves problem (3.3).

The next result confirms the concavity of the value functions.

Theorem 3.3 *The value functions $\{V_j(\cdot) : j = 1, \ldots, n\}$ computed through the dynamic program in* (3.3) *are concave.*

3.3.2 Class-Dependent No-Show Refunds

In the dynamic program in (3.3), we assume that if a passenger does not show up, then we do not give any refund and the probability of showing up for all passengers is the same. In practice, there are different restrictions that come along with different classes. As a result, passengers with tickets for different classes get different refunds when they do not show up and the probability of showing up is different for different classes. Allowing different show-up probabilities for different fare classes is difficult, because this extension requires using a high-dimensional state variable that keeps track of the reservations for each fare class separately. However, we can incorporate no-show refunds without too much difficulty and these no-show refunds could be different for different classes.

Assume that customers of class j who do not show up at the departure time of the flight are given a refund of h_j that is strictly less than the revenue p_j. We continue using all of the assumptions in our earlier model. Since whether a customer does not show up is completely independent of all other decisions and events in the system, we can charge the expected refund at the time the reservation is accepted, instead of the time of service. Thus, if we accept a reservation from a customer of class j, it yields an expected revenue of $p_j-(1-q)\,h_j$. In this case, we can use $p_j-(1-q)\,h_j$ in place of p_j in our earlier dynamic program.

3.3.3 Incorporating Cancellations

We can incorporate cancellations into our model, as long as the cancellation probabilities for the different fare classes are the same. We use ρ to denote the probability that a customer cancels her reservations at any stage. Given that we have y reservations on hand, we use $Z'(y)$ to denote the number of reservations that we still have on hand after observing the cancellations at the current stage. Thus, $Z'(y)$ is binomially distributed with parameters $(y, 1-\rho)$. In this case, the dynamic programming formulation of the problem is given by

$$V_j(y) = \mathbb{E}\left\{ \max_{0 \le u \le D_j} p_j\, u + V_{j-1}(Z'(y+u)) \right\},$$

with the same boundary condition as in (3.4).

Using an induction argument that is very similar to the one used earlier in this section, we can show that the value functions $\{V_j(\cdot) : j = 1, \ldots, n\}$ are concave. In this case, the optimal policy can be characterized by one booking limit b_j^* for each class j such that it is optimal to bring the number of reservations on hand as close as possible to b_j^* after making the decisions for fare class j. The optimal booking limit b_j^* for class j is the maximizer of the function $p_j\, y + V_{j-1}(Z'(y))$ over the interval $[0, \infty]$. Therefore, b_j^* can be computed as

$$b_j^* = \min\{y \ge 0 : p_j + \Delta V_{j-1}(Z'(y+1)) < 0\}.$$

3.4 Overbooking over a Flight Network

In this section, we give a dynamic programming formulation for the network model with overbooking. Following this formulation, we provide a deterministic linear programming approximation that is an upper bound on the optimal total expected revenue. Furthermore, this linear program can be used to extract control policies. Throughout this section, we use the independent demand model and adopt a discrete time formulation. There are m resources in the network indexed by $M := \{1, \dots, m\}$. We denote the vector of initial capacities by $c = (c_1, \dots, c_m) \in \mathcal{Z}^m$. There are T time periods in the selling horizon. We count the time periods backwards. In particular, time period T corresponds to the beginning of the selling horizon, whereas time period 1 is the last time period in the selling horizon. Time period 0 corresponds to the departure time of the flights. We use a single index to capture the ODF's. The set of ODF's in $N := \{1, \dots, n\}$. At time period t, we have a request for ODF j with probability λ_{tj}. The fare for ODF j is p_j. If we deny boarding to customer with a ticket for ODF j, then we incur a penalty of θ_j. Let $a_{ij} = 1$ if ODF j uses resource i, and $a_{ij} = 0$ if ODF j does not use resource i. We allow both cancellations and no-shows. The probability that a reservation for ODF j is retained from time period t to $t-1$ is q_{tj}. In other words, if we have a reservation for ODF j at time period t, this reservation cancels by time period $t-1$ with probability $1-q_{tj}$. Notice that q_{t1} is the probability that a reservation for ODF j is retained from period 1 to period 0, which corresponds to the show probability of a customer with a reservation for ODF j. The cancellation and no-show behavior of each customer is independent of the others. Furthermore, the cancellation decisions at different time periods are independent. Given that we have x_j reservations for itinerary j at time period t, we use $S_{tj}(x_j)$ to denote the number of reservations that we retain from time period t to $t-1$. Due to our assumptions, $S_{tj}(x_j)$ has a binomial distribution with parameters (x_j, q_{tj}). We use the vector $S_t(x) = (S_{tj}(x_j))_{j\in N}$ to capture the vector of retained reservations.

For any time to go t, we use (t, x) to represent the state of the system, where $x = (x_1, \dots, x_n)$ captures the number of reservations on hand for each ODF. To capture the decisions at any time period, we use the vector $u = (u_1, \dots, u_n)$, where $u_j = 1$ if accept a request for ODF j, and $u_j = 0$ otherwise. In this case, using $e_j \in \Re^n_+$ to denote the unit vector with a one in the j-th component, the dynamic programming formulation of the overbooking problem over a network is given by

$$V(t,x) = \max_{u\in\{0,1\}^n} \left\{ \sum_{j\in N} \lambda_{tj}\, u_j \left\{ p_j + \mathbb{E}\{V(t-1, S_t(x+e_j))\} \right\} \right.$$
$$\left. + \left\{ 1 - \sum_{j\in N} \lambda_{tj}\, u_j \right\} \mathbb{E}\{V(t, S_t(x))\} \right\},$$

where the expectations involve the random variables $S_t(x + e_j)$ and $S_t(x)$. Notice that the capacities of the resources do not play a role in the dynamic program above. Since we are allowed to overbook, the number of accepted reservations can exceed the available capacities on the resources. Thus, the capacities come into play when we compute the cost of denying boarding to the passengers that cannot be accommodated on the flights in the boundary condition of the dynamic program. For the boundary condition, we assume that the airline solves an optimization problem to decide which passengers should be allowed boarding so that the total penalty of denied boardings is minimized. (Our boundary condition is perhaps a bit optimistic in the sense that it would be difficult to solve a centralized optimization problem to decide which customers should be denied boarding.) Using the decision variable y_j to capture the number of reservations for ODF j that we deny booking, the boundary condition of our dynamic program is given by

$$\begin{aligned} V(0, x) = -\min \quad & \sum_{j \in N} \theta_j \, y_j \\ \text{s.t.} \quad & \sum_{j \in N} a_{ij} \, [x_j - y_j] \leq c_i \quad \forall i \in M \\ & y_j \leq x_j \;\; \forall j \in N \\ & y_j \in \mathcal{Z}_+ \;\; \forall j \in N. \end{aligned} \tag{3.6}$$

The objective function above minimizes the total cost of denied reservations. The first constraint ensures that the reservations that remain after denied boardings can be accommodated on the flights. The second constraint ensures that the number of denied bookings cannot exceed the number of reservations for each ODF.

Solving the dynamic program above is difficult because the state variable is a high-dimensional vector. Next, we give a tractable linear programming approximation that can be used to obtain an upper bound on the optimal total expected profit.

3.4.1 Linear Programming-Based Upper Bound on $V(T, 0)$

Since we start with no reservations on hand, the optimal total expected profit in our overbooking problem is given by $V(T, 0)$. We give a linear programming approximation that can be used to obtain an upper bound on $V(T, 0)$. We observe that a reservation booked at time period t is retained until the departure time with probability $Q_{tj} := q_{tj} \times q_{t-1,j} \times \ldots \times q_{1,j}$. Using the decision variable w_{tj} to capture the expected number of accepted reservations for ODF j at time period t and y_j to capture the number of reservations for ODF j that we deny boarding, we consider the linear program

$$\bar{V}(T,0) := \max \quad \sum_{t=1}^{T} \sum_{j \in N} p_j\, w_{tj} - \sum_{j \in N} \theta_j\, y_j \tag{3.7}$$

$$\text{s.t.} \quad \sum_{t=1}^{T} \sum_{j \in N} a_{ij}\, Q_{tj}\, w_{tj} - \sum_{j \in N} a_{ij}\, y_j \le c_i \quad \forall\, i \in M$$

$$\sum_{t=1}^{T} Q_{tj}\, w_{tj} - y_j \ge 0 \qquad \forall\, j \in N$$

$$w_{tj} \le \lambda_{tj} \qquad \forall\, t = 1, \ldots, T,\ \ j \in N$$

$$w_{tj},\ y_j \ge 0 \qquad \forall\, t = 1, \ldots, T,\ \ j \in N.$$

The objective function above accounts for the total expected profit, which is the difference between the revenue from the accepted reservations and the penalty cost of denied boardings. The expected number of accepted reservations for ODF j at time period t is w_{tj}. These reservations are retained until the end of the selling horizon with probability Q_{tj}. Therefore, $\sum_{t=1}^{T} Q_{tj}\, w_{tj}$ is the total expected number of reservations for ODF j retained until the departure time, which implies that $\sum_{t=1}^{T} \sum_{j \in N} a_{ij}\, Q_{tj}\, w_{tj}$ corresponds to the total expected capacity consumption of resource i by all the reservations that have been accepted over the selling horizon. On the other hand, $\sum_{j \in N} a_{ij}\, y_j$ gives the capacity of resource i released by the denied boardings. So, the first constraint ensures that the expected capacity consumption for each resource, after considering the capacity released by denied boardings, cannot exceed the capacity of the resource. The second constraint ensures that the number of denied boardings for passengers with a ticket for ODF j does not exceed the expected number of accepted reservations for ODF j. The third constraint is a demand constraint, ensuring that the expected number accepted reservations for ODF j at time period t does not exceed the expected demand for the same ODF at the same time period.

In the next theorem, we show that the optimal objective value of problem (3.7) is an upper bound on the optimal total expected profit. In contrast to our earlier linear programming-based upper bounds, our proof technique here does not use the Jensen's inequality, because owing to cancellations and no-shows, the random quantities would not appear on the right-hand side of a similar perfect hindsight linear program.

Theorem 3.4 $V(T,0) \le \bar{V}(T,0)$.

In the description of the dynamic program and the upper bound (3.7), we kept the fare p_j as time invariant. This was convenient to keep the exposition manageable. In practice, however, consumers who book in period t and either cancel or do not show may get a partial or full refund. Thus, it is more convenient to model the term $p_j w_{tj}$ as $p_{tj} w_{tj}$ where p_{tj} is the net revenue per booking after discounting refunds. As an example, if a consumer obtains a refund $r_{tj} < p_j$ if he cancels a time t booking for product j, then the average revenue per booking is $p_{tj} = p_j - r_{tj}(1 - Q_{tj})$. If $r_{tj} = p_j$, then $p_{tj} = p_j Q_{tj}$ models the case of full refunds, while the case $r_{tj} = 0$, results in $p_{tj} = p_j$ for the case with no refunds. We wrote the dynamic program and the upper bound as if $r_{tj} = 0$, but it is possible to modify both formulations to account for partial refunds so the upper bound remains valid. If we further set $p_{tj} \leftarrow p_{tj}/Q_{tj}$ for all t and all j, then we can modify the objective function in (3.7) to read

$$\sum_{t=1}^{T} \sum_{j \in N} p_{tj}\, Q_{tj}\, w_{tj} - \sum_{j \in N} \theta_j\, y_j.$$

This will be the objective function for (3.7) we will work from now on. In this version, the quantity p_{tj} is the net revenue per surviving booking, whereas $p_{tj} Q_{tj}$ is the net revenue per booking.

3.4.2 *Book-and-Bump Strategy*

A book-and-bump strategy occurs when an airline books passengers at a fare, say p_{tj}, and later bump them if needed by compensating them at level $\theta_j < p_{tj}$. To see how this may happen, suppose first that θ_j is sufficiently high so that $y_j^* = 0$ is optimal in (3.7), and suppose there is a period, say t, such that $w_{tj}^* < \lambda_{tj}$. Suppose now that we reduce θ_j so that now $\theta_j < p_{tj}$. We claim that it is now optimal to accept all requests in period t for product j. Indeed, suppose that we accept $\delta = \lambda_{tj} - w_{tj}^*$ additional requests of product j in period t, this brings additional profits $p_{tj} Q_{tj} \delta$, but we have to pay θ_j for each one of the $Q_{tj}\delta$ units we expect to survive. Thus, the change in profits is equal to $[p_{tj} - \theta_j]Q_{tj}\delta > 0$, showing that if $\theta_j < p_{tj}$, then it is optimal to set $w_{tj}^* = \lambda_{tj}$, and that this may involve booking some consumers with the idea of later bumping them later at a profit. On the other hand, if $p_{tj} < \theta_j$ for all t, then we claim that it is optimal to set $y_j^* = 0$. To see this, suppose for a contradiction that $y_j^* > 0$, so there must be a period t such that $w_{tj}^* > 0$. Reducing w_{tj}^* by ϵ and decreasing y_j^* by ϵ reduces revenues by $p_{tj} Q_{tj} \epsilon$ and costs by $\theta_j Q_{tj} \epsilon$ for a net savings of $-[p_{tj} - \theta_j]Q_{tj}\,\epsilon > 0$, contradicting the optimality of $y_j^* > 0$.

3.4.3 *Upper Bound for High Overbooking Penalties*

A book-and-bump strategy is unfair, unpopular, and illegal. Consequently, most airlines would plan their overbooking models by setting unit overbooking cost $\theta_j > p_{tj}$ for all t and for all j. In this case, $y_j^* = 0$ for all j. This means that in solving the linear program (3.7), we do not overbook beyond adjusting for the expected number of cancellations and no-shows. This implies that we can reformulate the problem ignoring the y_j variables keeping in mind that in the stochastic version of the problem we pay an overbooking cost θ_j for each unit of product j that is overbooked. The updated LP is

$$\begin{aligned} \bar{V}(T,0) := \max \quad & \sum_{t=1}^{T} \sum_{j \in N} p_{tj}\, Q_{tj}\, w_{tj} \\ \text{s.t.} \quad & \sum_{t=1}^{T} \sum_{j \in N} a_{ij}\, Q_{tj}\, w_{tj} \le c_i \quad \forall i \in M \\ & 0 \le w_{tj} \le \lambda_{tj} \qquad \forall t = 1,\ldots,T,\ j \in N \end{aligned} \tag{3.8}$$

This LP is essentially of the same form as the one for the model without cancellations, so there should be hope for heuristics based on its solution. Indeed, if we make the transformation $x_{tj} = Q_{tj} w_{tj}$, the LP (3.8) is as a time-variant version of the model without cancellations except that $x_{tj} \le \lambda_{tj} Q_{tj}$. Notice that $\lambda_{tj} Q_{tj}$ represents the net demand for product j at time t after filtering the demand that cancels or does not show.

3.4.4 Heuristics Based on the Linear Program

We can derive the probabilistic acceptance control (PAC) heuristic from the linear programming upper bound (3.8) exactly as we did in the previous chapter. Let $\{w^*_{tj} : t = 1,\ldots,T,\ j \in N\}$ be the optimal solution to problem (3.8). In period t, a request for product j arrives with probability λ_{tj}, and the PAC heuristic accepts it with w^*_{tj}/λ_{tj} and rejects it with probability $1 - w^*_{tj}/\lambda_{tj}$. Consider now a system where the capacities and the arrival rates λ_{tj} are scaled by an integer factor b, so now the number of arrivals for product j in period t is a binomial with parameters b and λ_{tj}. The PAC heuristic would filter the arrivals by the factor w^*_{tj}/λ_{tj} so in the scaled model, the number of requests accepted by the PAC heuristic is binomial with parameter b and w^*_{tj}. For the upper bound, the expected demand is $b\lambda_{tj}$, and the solution is bw^*_{tj}. Let $\bar{V}^b(T,0)$, $V^b(T,0)$, and $V^b_h(T,0)$ denote the upper bound, the optimal expected revenue, and the expected revenue of the PAC heuristic for the scaled system. Clearly, $V^b(T,0) = b\bar{V}(T,0)$ as bw^*_{tj} is an optimal solution to the scaled LP. We will now show that the PAC heuristic is asymptotically optimal.

Theorem 3.5

$$\lim_{b\to\infty} \frac{V^b_h(T,0)}{V^b(T,0)} \ge \lim_{b\to\infty} \frac{V^b_h(T,0)}{\bar{V}^b(T,0)} \to 1,$$

We can also use a bid-price heuristic based on the solution to the dual problem:

$$\min_{z \geq 0} \left\{ c'z + \sum_{t=1}^{T} \sum_{j \in N} \lambda_{tj} \, Q_{tj} (p_{tj} - \sum_{i \in M} a_{ij} \, z_i)^+ \right\}.$$

The heuristic accepts a request at time t if $p_{tj} \geq \sum_{i \in M} a_{ij} z_i^*$. The heuristic is not asymptotically optimal, but as in the case without cancellations and no-shows, it performs very well if the system is resolved frequently during the sales horizon for moderately large problems as those found in practice.

3.4.5 Other Approximation Strategies

We demonstrated that the linear programming approach that we had developed for network revenue management without overbooking naturally extends to the case where overbooking is allowed. Unfortunately, other approaches that we developed for network revenue management without overbooking do not easily extend to the case where overbooking is allowed. In the chapter on network revenue management problems with independent demand, we discussed two ways of decomposing the dynamic programming formulation of the network revenue management problem by the resources. The first approach exploited the deterministic linear program, and the second approach used Lagrangian relaxation. Under overbooking, even if we can decompose the problem by the resources, the problem that takes place over each resource is intractable because solving the single-resource revenue management problem requires a high-dimensional state variable that keeps track of the numbers of reservations for each ODF. There is some work that is based on decomposing the network overbooking problem by the resources and approximating the single-resource revenue management problems. This work is discussed in the bibliographical remarks at the chapter. In Table 3.1, we compare the bid-price policy derived from the linear program in (3.7) with such a decomposition approach. There are four test problems in this table, encoded by the pair (q, ρ), where q is the probability that an accepted request shows up and ρ is the ratio between the total expected demand for the capacities and the total capacity. In particular, we have $Q_{tj} = q$ for all $t = 1, \ldots, T$ and $j \in N$ and $\rho = q \sum_{i \in M} \sum_{t=1}^{T} \sum_{j \in N} a_{ij} \, \lambda_{tj} / \sum_{i \in M} c_i$. In all of the test problems, the airline network has a hub and spoke structure. There is one hub and four spokes. There is a flight leg from each spoke to the hub and a flight leg from the hub to each spoke. There is a high-fare and a low-fare ODF connecting each origin-destination pair. The fare of a high-fare ODF is eight times the fare of the corresponding low-fare ODF. The arrival process for the requests is set up such that the requests for the low-fare ODF's tend to arrive earlier, whereas the requests for the high-fare ODF's tend to arrive later. The first column in the table shows the upper bound on the optimal total expected profit given by the optimal objective value of problem (3.7), whereas the second and third columns show the total expected revenues obtained by the bid-price heuristic and the decomposition approach. The

Table 3.1 Performance of the bid-price heuristic and the decomposition approach

Problem (q, ρ)	Upper bound	Bid-price heuristic	Decomp. approach	Gaps	
				Bid-price	Decomp.
(0.9, 1.2)	\$30,754	\$29,286	\$29,514	4.77%	4.03%
(0.9, 1.6)	\$31,744	\$30,324	\$30,841	4.47%	2.84%
(0.95, 1.2)	\$28,983	\$27,386	\$27,676	5.51%	4.51%
(0.95, 1.6)	\$23,995	\$22,720	\$22,983	5.31%	4.22%

last two columns show the percent gap between the total expected revenues of the policies and the upper bound on the optimal total expected profit. The results indicate that the decomposition approach yields noticeable improvements over the bid-price heuristic, especially when the capacities are tight.

In the same chapter, we also discussed approximate dynamic programming methods to approximate the value functions. These approaches do not readily extend to the overbooking setting either. In particular, if overbooking is allowed, then the linear program that we used to calibrate the value function approximations includes one constraint for each possible state of the system and the right side of this constraint involves the bumping cost associated with each state. The presence of this constraint makes overbooking problems intractable.

3.5 End of Chapter Problems

1. Consider a flight with 100 seats and a passenger fare of \$130. The denied boarding cost is \$390 per denied boarding, and the no-show rate is 0.16 (assuming a binomial no-show model). Demand for this flight is extremely high; in fact, for any booking limit $b < 200$, bookings will always hit the booking limit.
   (a) Assume that only the passengers who show up for the flight pay the fare of \$130; others are fully refunded. What is the optimal booking limit in this case? What is the corresponding expected net profit? How much does the airline gain from overbooking in this case? (That is, compute the expected revenue under the assumption that the airline does not overbook at all and compare to the overbooking case.)
   (b) For this part, assume that all passengers pay the fare of \$130 at the time of the reservation, regardless of whether or not they show up for the flight. Determine the optimal booking limit, the corresponding expected profit, and the gain from overbooking. (Hint: For this problem, you will need to modify the profit function to account for the fact everyone pays the fare and derive an expression that needs to be satisfied by the optimal booking limit.)

Table 3.2 Complementary cumulative distribution function of $Y(n)$

n	$\mathbb{P}\{Y(n) \geq 50\}$
50	0.0002957647
51	0.0025139996
52	0.0109987484
53	0.0330590953
54	0.0769040347
55	0.1479328364
56	0.2455974389
57	0.3627949618
58	0.4880498145
59	0.6091295054
60	0.7162850318

2. Consider the following overbooking problem. We first choose a booking limit b. After this, a random demand D occurs. For each unit of demand that we accept, we generate a revenue of $\$p$. Assume that every accepted booking request will show up for the flight. The capacity of the plane is c.

   If the number of customers that show up at the departure time exceeds the capacity of the plane, then we offer every customer a voucher worth $\$f$ for use on future flights. The customers who accept the vouchers will voluntarily give up their reservations. We assume that each customer independently declines the voucher, and thus keeps his/her existing reservation with probability $\beta \in (0, 1)$.

   After offering the vouchers, if the number of remaining customers still exceeds the capacity c of the plane, then we begin an involuntary denied boarding process. For each booking that cannot be accommodated on the plane, we incur a penalty cost of $\$\theta$.

   (a) Let $R(b)$ denote the expected profit under the booking limit b. Provide an expression for $R(b)$.
   (b) Assuming that $p - f(1-\beta) < \theta\beta$, determine the integer-valued optimal booking limit. Your answer should only involve probabilities that can be computed by simple table lookups and the problem parameters given above.
   (c) Suppose that $c = 50$, $p = 100$, $f = 200$, $\theta = 300$, and $\beta = 0.85$. Let $Y(n)$ denote a binomial random variable with parameters n and 0.85. Table 3.2 gives the value of $\mathbb{P}\{Y(n) \geq 50\}$ for different values of n. Using this table and the formula from Part (b), determine the optimal booking limit in this case.

3. Consider the model in Sect. 3.3. Assume that the fares satisfy $p_1 \leq p_2 \leq \ldots \leq p_n$. Show that the optimal booking limits $b_1^*, \ldots, b_n^*$ satisfy $b_1^* \leq b_2^* \leq \ldots \leq b_n^*$.
4. We are purchasing a certain product over the time periods $1, 2, \ldots, T$. The demand for the product occurs at the end of these T time periods, say time period $T + 1$. The price of the product fluctuates randomly over the time periods $1, 2, \ldots, T$ and we need to decide how many units of product we should purchase at each time period.

We use the random variable P_t to denote the price of the product at time period t. We use the random variable D to denote the demand for the product, which occurs at time period $T + 1$. For each unit of demand that we cannot satisfy, we incur a shortage cost of \$$\theta$. We are interested in minimizing the total expected cost, which is the sum of the product purchasing cost and the shortage cost.

(a) Formulate the problem as a dynamic program. Clearly give your state and decision variables, and write down the boundary condition at the end of T time periods.
(b) By using backward induction over the time periods, show that the value function is convex.
(c) Assume that the price can take only three different values, a high, a medium, and a low value. Show that in order to be able to make the optimal purchasing decision at each time period, we only need to store $3T$ values. That is, we only need to store three values for each time period. Clearly indicate how each one of these $3T$ values should be computed.

5. Consider a single-flight overbooking problem without any cancellations, but with no-shows. The customers arrive over the time periods $1, 2, \ldots, T$. There are n possible price levels indexed by $1, 2, \ldots, n$. If we sell a ticket at price level j, then we generate a revenue of \$$p_j$. With probability λ_{jt}, a customer that is interested in price level j arrives into the system at time period t. We need to decide whether to accept or reject each customer request. For simplicity, assume that $\sum_{j=1}^{n} \lambda_{jt} = 1$ so that there is always one customer arrival at each time period.

   At the departure time of the flight, which we assume to happen at time period $T+1$, each reservation shows up with probability q. A no-show with a reservation at price level j is given a refund of \$$h_j$. The capacity of the flight is c and for each customer that we cannot board on the flight, we incur a cost of \$$\theta$.

   (a) Let x_{jt} be the number of accepted reservations that we have on hand for price level j at the beginning of time period t. Using the n-dimensional vector $x_t = (x_{1t}, x_{2t}, \ldots, x_{nt})$ as the state variable, formulate a dynamic program that maximizes the expected profit. In your dynamic program, make sure to charge the no-show refunds at the departure time of the flight. (Hint: Let e_j be the j-th unit vector in $\Re^n$. If you accept a request for price level j at time period t, then your state changes from x_t to $x_t + e_j$.)
   (b) We can charge the expected refund cost at the time of accepting a customer request. This amounts to assuming that the revenue associated with price level j is $p_j - (1 - q)\,h_j$. Since we charge the expected refund cost at the time of accepting a customer request and each reservation shows up with the same probability q, we now need to keep track of only the total number of accepted requests.

   Let z_t be the total number of accepted reservations that we have on hand at the beginning of time period t. Using the scalar z_t as your state variable, formulate a dynamic program that maximizes the total expected profit.

(c) Denote the value function in Part a as $V_t(x_t)$ and the value function in Part b as $J_t(z_t)$. Use backward induction over the time periods to show that $V_t(x_t) = J_t(\sum_{j=1}^n x_{jt}) - \sum_{j=1}^n (1-q)\, h_j\, x_{jt}$. (Hint: Recall that if $B_1(n_1, q)$ is a binomial random variable with parameters n_1 and q, and $B_2(n_2, q)$ is a binomial random variable with parameters n_2 and q that is independent of $B_1(n_1, q)$, then $B_1(n_1, q) + B_2(n_2, q)$ is a binomial random variable with parameters $n_1 + n_2$ and q.)

6. Is the PAC heuristic asymptotically optimal if the condition $\theta_j > p_{tj}$ for all t and for all j fails to hold?

3.6 Bibliographical Notes

Simon (1968) proposes auctions as a possible way to handle involuntary denied boardings. Rothstein (1971) gives one of the first systematic treatments of the overbooking problem, where dynamic programming is used to develop an overbooking policy for American Airlines. Chatwin (1998, 1999) give a dynamic programming formulation of the overbooking problem with a single class and characterize the structure of the optimal policy. Lautenbacher and Stidham (1999) study overbooking problems with multiple fare classes over a single resource. The cancellation model in this paper assumes that there can be at most one cancellation at each time period, and the probability of having a cancellation increases as the number of reservations on hand increases. In contrast, we use a binomial cancellation model, which allows multiple cancellations at each time period.

Kleywegt (2001) and Dai et al. (2019) give deterministic approximations to overbooking problems to extract heuristic control policies, some of which have asymptotic optimality guarantees. Karaesmen and van Ryzin (2004a) consider an overbooking model with substitutable flights, where the passengers bumped from one flight can be accommodated on the next one. Karaesmen and van Ryzin (2004b) study various decomposition strategies for the overbooking problem over a flight network. Erdelyi and Topaloglu (2010) leverage the linear programming approximation given in this chapter to decompose the network overbooking problem by the resources. Solving the single-resource overbooking problems is still difficult when the cancellation and no-show probabilities are class-specific. The authors use approximations to the single-resource overbooking problems. The numerical example in Sect. 3.4.5 is taken from this paper. Erdelyi and Topaloglu (2009) use a separable approximation to the bumping cost. In this case, they show that the dynamic programming formulation of the network overbooking problem decomposes by the ODFs, and the single-ODF overbooking problem turns out to be completely tractable. Aydin et al. (2013) present an overbooking model over a single resource. Their cancellation model is similar to the one in our network overbooking model in this chapter, in the sense that the number of cancellations at each time period is binomially distributed. Kunnumkal and Topaloglu (2011b) use stochastic approximation methods to compute bid prices for overbooking over a network of flight legs. Kunnumkal et al. (2012) give a randomized version of the linear program in (3.7) to capture the randomness in the show-up decisions more accurately.

Appendix

Proof of Proposition 3.1 We can write $Z(N)$ as $Z(N) = \sum_{i=1}^{N} X_i$, where the X_i's are independent Bernoulli random variables with probability q. Clearly $Z(\min(D, b+1)) - Z(\min(D, b)) = X_{b+1} \times \mathbf{1}(D \geq b+1)$, where $\mathbf{1}(\cdot)$ is the indicator function. Consequently, we get

$$\mathbb{E}\{Z(\min(D, b+1)) - Z(\min(D, b))\} = q\,\mathbb{P}\{D \geq b+1\}, \tag{3.9}$$

Similarly, note that we always have $Z(\min(D, b+1)) \geq Z(\min(D, b))$. Furthermore, $Z(\min(D, b+1))$ and $Z(\min(D, b))$ can differ by at most 1. Thus, if $Z(\min(D, b)) < c$, then we have $Z(\min\{b+1, D\}) \leq c$. On the other hand, if $Z(\min(D, b)) \geq c$, then $Z(\min\{b+1, D\}) \geq c$. In this case, we obtain

$$\begin{aligned}
&[Z(\min(D, b+1)) - c]^+ - [Z(\min(D, b)) - c]^+ \\
&\qquad = \begin{cases} Z(\min(D, b+1)) - Z(\min(D, b)) & \text{if } Z(\min(D, b)) \geq c \\ 0 & \text{otherwise} \end{cases} \\
&\qquad = \begin{cases} X_{b+1} & \text{if } D \geq b+1 \text{ and } Z(\min(D, b)) \geq c \\ 0 & \text{otherwise} \end{cases} \\
&\qquad = \begin{cases} X_{b+1} & \text{if } D \geq b+1 \text{ and } Z(b) \geq c \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

Using the chain of equalities above, we get

$$\begin{aligned}
&\mathbb{E}\{[Z(\min(D, b+1)) - c]^+ - [Z(\min(D, b)) - c]^+\} \\
&\qquad\qquad = q\,\mathbb{P}\{D \geq b+1\}\,\mathbb{P}\{Z(b) \geq c\}.
\end{aligned} \tag{3.10}$$

Using (3.9) and (3.10) in (3.1), we obtain

$$R(b+1) - R(b) = \mathbb{P}\{D \geq b+1\}\, q\,(p - \theta\,\mathbb{P}\{Z(b) \geq c\}),$$

from which the formula for b^* follows. □

Proof of Theorem 3.2 In (3.3), we need to solve the problem

$$\max_{0 \le u \le D_j} \Big\{ p_j\, u + V_{j-1}(y+u) \Big\}.$$

We define a new decision variable z such that $z = y + u$. Since y is the number of reservations just before making the decisions for class j and u is the number of reservations we accept from class j, the decision variable z can be interpreted as the number of reservations after making the decisions for fare class j. After the change of variables, the problem is equivalent to

$$\max_{y \le z \le y + D_j} \Big\{ p_j\, z + V_{j-1}(z) \Big\} - p_j\, y. \tag{3.11}$$

Since the last term $p_j\, y$ does not affect the optimal solution, we can concentrate on the following problem

$$\max_{y \le z \le y + D_j} \Big\{ p_j\, z + V_{j-1}(z) \Big\}. \tag{3.12}$$

Since $V_{j-1}(\cdot)$ is concave, the objective function of problem (3.12) above is concave. Thus, the problem above maximizes a concave function subject to the constraint that the decision variable lies in the interval $[y, y + D_j]$.

Let b_j^* be the maximizer of the concave function $p_j\, z + V_{j-1}(z)$ over $[0, \infty]$. The maximizer can be computed as

$$b_j^* = \min\{z \ge 0 : p_j\,(z+1) + V_{j-1}(z+1) \le p_j\, z + V_{j-1}(z)\}.$$

which yields the desired result.

We can characterize the optimal solution to the constrained problem above depending on whether b_j^* is in the interval $[y, y + D_j]$ or lies to the left or the right side of this interval. In particular, using z^* to denote the solution to problem (3.12), we have

$$z^* = \begin{cases} y & \text{if } b_j^* < y \\ b_j^* & \text{if } y \le b_j^* \le y + D_j \\ y + D_j & \text{if } b_j^* > y + D_j. \end{cases} \tag{3.13}$$

We show the three cases above, along with the maximizer b_j^* of the function $p_j\, z + V_{j-1}(z)$ and the interval $[y, y + D_j]$ in Fig. 3.1. If $b_j^* < y$, then the number of reservations we have y is already larger than the optimal booking limit b_j^*. Thus, the only way to get as close as possible to b_j^* after making the decisions for class j is not to accept any reservations from class j. In other words, we keep the number of reservations on hand at y. This situation corresponds to the first case above. If $b_j^* < y$, then it is optimal to set $z^* = y$. If $y \le b_j^* \le y + D_j$, then $b_j^* - y \le D_j$. So, we can accept $b_j^* - y$ reservations from class j to bring the number of reservations on

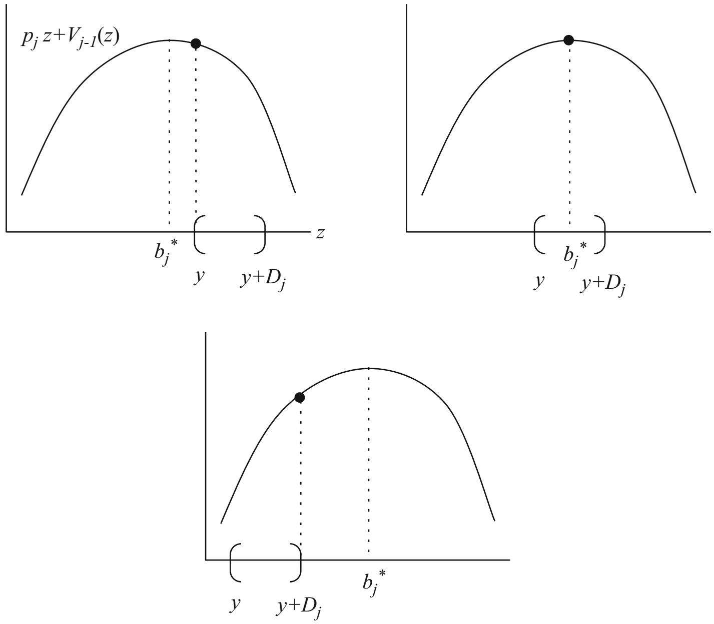

Fig. 3.1 Optimal decision for class j

hand to exactly b_j^* after making the decisions for class j. This situation corresponds to the second case above. If $y \le b_j^* \le y + D_j$, then it is optimal to set $z^* = b_j^*$. Lastly, if $b_j^* > y + D_j$, then $D_j < b_j^* - y$. Thus, the only way to get as close as possible to b_j^* after making the decisions for class j is to accept all of the demand from class j, in which case, the number of reservations that we have after making the decisions for class j goes up to $y + D_j$. This situation corresponds to the third case above. If $b_j^* > y + D_j$, then it is optimal to set $z^* = y + D_j$. Noting the change of variables $z = y + u$ and using (3.13), as a function of y, an optimal solution to problem (3.3) is given by the expression in the theorem. □

Proof of Theorem 3.3 We show the result by using induction over the classes in reverse order. Since $Z(y)$ is a binomial random variable with parameters (y, q), we can write $Z(y) = \sum_{i=1}^{y} X_i$, where $X_1, X_2, \ldots$ are independent Bernoulli random variables with parameter q. In this case, we have

$$[Z(y+1) - c]^+ - [Z(y) - c]^+ = \begin{cases} Z(y+1) - Z(y) & \text{if } Z(y) \ge c \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} X_{y+1} & \text{if } Z(y) \ge c \\ 0 & \text{otherwise,} \end{cases}$$

which implies that $\mathbb{E}\{[Z(y+1)-c]^+ - [Z(y)-c]^+\} = q\,\mathbb{P}\{Z(y) \geq c\}$. Since $Z(y)$ is a binomial random variable with parameters (y, q), $\mathbb{P}\{Z(y) \geq c\}$ is increasing in y. Therefore, $\mathbb{E}\{[Z(y+1)-c]^+ - [Z(y)-c]^+\}$ is increasing in y. In this case, $\mathbb{E}\{[Z(y)-c]^+\}$ is convex in y, which implies that $V_0(y) = -\theta\,\mathbb{E}\{[Z(y)-c]^+\}$ is concave in y, as desired. This discussion establishes the base case for the induction argument. Next, we assume that the value function $V_{j-1}(\cdot)$ is concave and show that $V_j(\cdot)$ is also concave.

Assume that $V_{j-1}(\cdot)$ is concave. By using the same change of variables used to obtain problem (3.11), we can write the dynamic program in (3.3) as

$$V_j(y) = \mathbb{E}\left\{ \max_{y \leq z \leq y+D_j} [p_j\, z + V_{j-1}(z)] \right\} - p_j\, y.$$

We define

$$W_j(y, D_j) = \max_{y \leq z \leq y+D_j} [p_j\, z + V_{j-1}(z)], \tag{3.14}$$

so that $V_j(y) = \mathbb{E}\{W_j(y, D_j)\} - p_j\, y$. If we can show that $W_j(y, D_j)$ is concave in y, then $\mathbb{E}\{W_j(y, D_j)\}$ is concave in y as well, in which case, it follows that $V_j(y) = \mathbb{E}\{W_j(y, D_j)\} - p_j\, y$ is concave, which is the result we are after. Thus, we proceed to showing that $W_j(y, D_j)$ is concave in y.

By the induction assumption $V_{j-1}(\cdot)$ is concave. We let b_j^* be the maximizer of the concave function $p_j\, z + V_{j-1}(z)$ over the interval $[0, \infty]$. Since $V_{j-1}(\cdot)$ is concave, the discussion that we used to obtain the three cases in (3.13) still holds. In this case, letting z^* be the optimal solution to problem (3.14), z^* is still given by the three cases in (3.13). Noting that the optimal objective function of problem (3.14) is $W_j(y, D_j)$, we have

$$W_j(y, D_j) = \begin{cases} p_j\, y + V_{j-1}(y) & \text{if } b_j^* < y \\ p_j\, b_j^* + V_{j-1}(b_j) & \text{if } y \leq b_j^* \leq y + D_j \\ p_j\, (y + D_j) + V_{j-1}(y + D_j) & \text{if } b_j^* > y + D_j \end{cases}$$

$$= \begin{cases} p_j\, y + V_{j-1}(y) & \text{if } b_j^* < y \\ p_j\, b_j^* + V_{j-1}(b_j^*) & \text{if } b_j^* - D_j \leq y \leq b_j^* \\ p_j\, (y + D_j) + V_{j-1}(y + D_j) & \text{if } y < b_j^* - D_j. \end{cases}$$

We plot the function $p_j\, y + V_{j-1}(y)$ as a function of y on the left side of Fig. 3.2. Notice that the maximizer of this function over $[0, \infty]$ is b_j^*. We plot the function $W_j(y, D_j)$ on the right side of Fig. 3.2. Notice that the functions $p_j\, y + V_{j-1}(y)$ and $W_j(y, D_j)$ are identical for y in $[b_j^*, \infty]$. For y in the interval $[b_j^* - D_j, b_j^*]$, the function $W_j(y, D_j)$ takes the constant value $b_j^* + V_{j-1}(b_j^*)$, which is the maximum value of $p_j\, y + V_{j-1}(y)$. Lastly, for y in the interval $[0, b_j^* - D_j]$, the

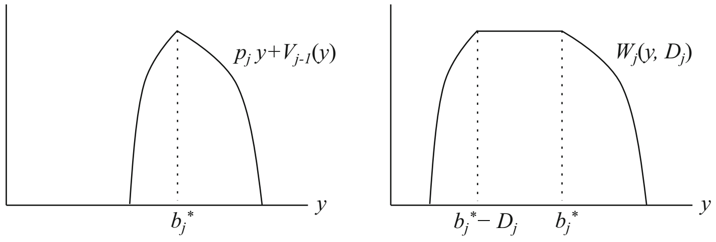

Fig. 3.2 Concavity of the value function for class j

function $W_j(y, D_j)$ takes the value of $p_j\ (y+D_j)+V_{j-1}(y+D_j)$. In other words, over the last interval, the function $W_j(y, D_j)$ is a shifted version of the function $p_j\ (y + D_j) + V_{j-1}(y + D_j)$. Thus, intuitively speaking, the function $W_j(y, D_j)$ is obtained by "cutting" the function $p_j\ y + V_{j-1}(y)$ in half at the point $y = b_j^*$, "shifting" the left portion of the function D_j units to the left, and "filling in" the middle with the constant value $b_j^* + V_{j-1}(b_j^*)$. Since $b_j^* + V_{j-1}(b_j^*)$ is the maximum value of the function $p_j\ y + V_{j-1}(y)$, it follows that $W_j(y, D_j)$ is concave, which is the desired result. □

Proof of Theorem 3.4 We let $D_{tj} = 1$ if there is a demand for ODF j at time period t, otherwise $D_{tj} = 0$. In this case, D_{tj} is a Bernoulli random variable with parameter λ_{tj} so that $\mathbb{E}\{D_{tj}\} = \lambda_{tj}$. We let the random variable W_{tj}^* be the number of accepted bookings for ODF j at time period t under the optimal policy and the random variable X_{tj}^* be the number of bookings for ODF j accepted at time period t that survive until the departure time. Thus, X_{tj}^* is a binomial random variable with parameters (W_{tj}^*, Q_{tj}). Thus, we have $\mathbb{E}\{X_{tj}^*\} = Q_{tj}\,\mathbb{E}\{W_{tj}^*\}$. Lastly, we let the random variable Y_j^* be the number of denied bookings for ODF j under the optimal policy. Under the optimal policy, we have the inequalities

$$\sum_{t=1}^{T} \sum_{j\in N} a_{ij}\, X_{tj}^* - \sum_{j\in N} a_{ij}\, Y_j^* \le c_i \quad \forall\, i \in M$$

$$Y_j^* \le \sum_{t=1}^{T} \sum_{j\in N} X_{jt}^* \quad \forall\, j \in N$$

$$W_{tj}^* \le D_{tj} \quad \forall\, t = 1, \ldots, T,\ j \in N.$$

The first inequality states that the capacity consumption of each resource, after accounting for the denied boardings, does not exceed the available capacity of the resource. The second inequality states that the number of denied boardings for each ODF cannot exceed the accepted bookings for the ODF. The third inequality states that the number of accepted bookings for each ODF at each time period cannot exceed the demand for the ODF. Taking expectations on both sides of the inequalities above and noting that $\mathbb{E}\{X_{tj}^*\} = Q_{tj}\,\mathbb{E}\{W_{tj}^*\}$, the inequalities above imply that setting $w_{tj} = \mathbb{E}\{W_{tj}^*\}$ and $z_j = \mathbb{E}\{Y_j^*\}$ for all $t = 1, \ldots, T$, $j \in N$ provides a feasible solution to problem (3.7). The total profit from the optimal policy is $\sum_{t\in T}\sum_{j\in N} p_j\, W_{tj}^* - \sum_{j\in N}\theta_j\, Y_j^*$, in which case, taking expectations, the total expected profit from the optimal policy is $V(T,0) = \sum_{t\in T}\sum_{j\in N} p_j\,\mathbb{E}\{W_{tj}^*\} - \sum_{j\in N}\theta_j\,\mathbb{E}\{Y_j^*\}$. Thus, setting $w_{tj} = \mathbb{E}\{W_{tj}^*\}$ and $y_j = \mathbb{E}\{Y_j^*\}$ for all $t = 1, \ldots, T$, $j \in N$ provides a feasible solution to problem (3.7) and the objective value provided by this solution is equal to $V(T,0)$. In this case, it follows that the optimal objective value of problem (3.7) is at least $V(T,0)$, so we obtain $\bar{V}(T,0) \geq V(T,0)$.

□

Proof of Theorem 3.5 Since the number of requests that arrive for product j in period t is a Bernoulli random variable with success probability λ_{tj}, the number admitted by the PAC heuristic is a thinned Bernoulli with probability w_{tj}^*. From this number, a fraction Q_{tj} will survive, so the number of bookings for period t that survive is also thinned Bernoulli with probability $Q_{tj}w_{tj}^*$. This shows that the expected revenues associated with the PAC heuristic, aggregating over all products, is equal to $\sum_{t=1}^{T}\sum_{j\in N} p_{tj}\,Q_{tj}\,w_{tj}^* = \bar{V}(T,0)$, where the equality uses the fact that $\{w_{tj}^* : t = 1, \ldots, T,\ j \in N\}$ is an optimal solution to problem (3.8).

Now, we consider the expected cost $\mathbb{E}[V(0,X)]$, where X is the vector of reservations on hand at the end of the horizon and $V(0,x)$ is the optimal objective value of problem (3.6). Clearly $X_j = \sum_{t=1}^{T} X_{tj}$, where X_{tj} is Bernoulli random variable with mean $Q_{tj}w_{tj}$. Since the X_{tj}'s are independent over t, it follows that X_j has mean $\sum_{t=1}^{T} Q_{tj}w_{tj}^*$ and variance $\sum_{t=1}^{T} Q_{tj}w_{tj}^*(1 - Q_{tj}w_{tj}^*) \leq \sum_{t=1}^{T} Q_{tj}w_{tj}^*$.

A feasible solution to program $V(0,X)$ in (3.6) is to pay the overbooking fee θ_j for each unit of product j booking in excess of the mean, yielding the feasible solution $y = \{y_j : j \in N\}$ with $y_j = (X_j - \mathbb{E}[X_j])^+$. Consequently, it follows that

$$\mathbb{E}[V(0,X)] \geq -\sum_{j\in N}\theta_j\,\mathbb{E}(X_j - \mathbb{E}[X_j])^+ \geq -\frac{1}{2}\sum_{j\in N}\theta_j\sqrt{\sum_{t=1}^{T} Q_{tj}w_{tj}^*},$$

where we have used the fact that for any random variable with finite second moment $\mathbb{E}[(X-\mathbb{E}[X])^+] \leq 0.5\sqrt{\mathrm{Var}[X]}$. Thus, a lower bound on the expected revenue from the PAC heuristic is given by

$$V_h(T,0) = \bar{V}(T,0) + \mathbb{E}[V(0,X)] \geq \bar{V}(T,0) - \frac{1}{2}\sum_{j\in N}\theta_j\sqrt{\sum_{t=1}^{T} Q_{tj}w_{tj}^*}.$$

Clearly bw_{tj}^* is the solution to the linear program scaled by a factor b, so $\bar{V}^b(T,0) = b\bar{V}(T,0) \geq V^b(T,0) \geq V_h^b(T,0)$. From the bound on $\mathbb{E}[V(0,X)]$ we see that

$$V_h^b(T,0) = \bar{V}^b(T,0) - \mathbb{E}[V^b(0,X)] \geq \bar{V}^b(T,0) - \frac{1}{2}\sum_{j\in N}\theta_j\sqrt{b\sum_{t=1}^{T} Q_{tj}w_{tj}^*}.$$

Dividing by $\bar{V}^b(T,0)$ and letting $b \to \infty$, we find that

$$\lim_{b\to\infty}\frac{V_h^b(T,0)}{V^b(T,0)} \geq \lim_{b\to\infty}\frac{V_h^b(T,0)}{\bar{V}^b(T,0)} \to 1,$$

completing the proof. □