**Hotel Overbooking Strategy Using Cancellation Data**

Joseph Despres

Rishabh Sareen

October 4, 2021

**Abstract** Overbooking has many applications such as airlines, cruise ships and concert halls. This project constructs a profitable overbooking strategy using publicly available hotel cancellation data. The objective is to optimize the number of reservations given the uncertainty of a guest’s arrival. Bookings are maximized when the expected profit the additional reservation no longer justifies the expected value of the overbooking penalty. Using this strategy, the number of empty rooms can be reduced by half, if a hotel is willing to pay modest overbooking penalties.

Work done in partial fulfillment of the requirements of Michigan State University MTH 843; advised by Dr. Peiru Wu, Michigan State University.

**Table of Contents** [This is an example (including pages numbers) to give you some ideas about a report structure.]

[Nomenclature 1](#_Toc36765862)

[Introduction 2](#_Toc36765863)

[Data](#_Toc36765864) 3

[Analysis 4](#_Toc36765864)

[Methods (Procedure)](#_Toc36765866) 6

[Results and Discussion 8](#_Toc36765867)

[Conclusions 12](#_Toc36765868)

[References 14](#_Toc36765871)

[Appendix A 15](#_Toc36765872)

[Appendix B 18](#_Toc36765873)

# Nomenclature (*In order of its appearance*)

# A Number of Guests arriving to a reservation

# b Number of bookings made by the hotel

p Probability

Profit

L Loss function of penalty paid when arrivals exceed capacity

E Expectation operator

Estimated probability

# 

# Introduction

Overselling or overbooking is the practice of committing to sell more of a product or service ability to fulfill. Airlines, for instance often oversell the seats and offer compensation in the event of excess passengers (14 CFR 250.2B). Manufactures often employ similar algorithms to predict deferred or cancelled orders (Leahy 1) which permits them to accept more orders than can be processes. This is a delicate task due to high costs incurred in the event of overselling.

If the probability of a guest arriving is not 1 and a hotel has sufficient capacity, the probability of a hotel having empty rooms is almost surely 1. This is where the opportunity to overbook presents itself. Done properly, overbooking offers the opportunity to gain additional revenue without providing additional products or services. A successful hotel overbooking strategy can be modified, repurposed, and increase profits in a wide variety of industries.

The goal of this project is to construct a profitable overbooking strategy. Hotel booking data available in the public domain is sufficient to estimate the uncertainty of a guest arrival. This strategy is constructed by making some assumptions about overbooking penalties, setting an arbitrary capacity, and quantifying the uncertainty of arrival. Then this strategy is tested with a simulation that resamples bookings until a predetermined capacity is reached.

The Data section contains a description of the hotel cancellation data used. The Analysis section explains the underlying mathematics The Methods section contains the process of developing the overbooking strategy and the validation simulation. The results section contains a discussion the results and a discussion of the applications and limitations. The conclusion contains a summary of the results.

# Data

# 

# Analysis

# Any guest making a hotel reservation is not certain to arrive. This uncertainty is defined as the probability of arrival *p*. A hotel will make many reservations and ask how many arrivals to expect given number of bookings *b* and *p*. The probability of at least *A* guests arriving follows a binomial distribution has a cumulative density function

This function gives the probability of *A* or fewer guests arriving given the number of bookings, and probability of each booking. With this distribution, the total profit a hotel can expect while booking no more than capacity is

Where denotes profit per room. Multiply profit, number of bookings, and sum probability of each arrival to get expected profit.

Booking over capacity allows fewer empty rooms, at the cost of risking insufficient rooms. Therefore, the costs of having more guests arrive than rooms available should be carefully considered. This strategy assumes that the cost of overbooking is providing each overbook additional accommodations at a cost of four times profit . Other industries are certain to have different loss functions, which can be absorbed into this strategy without difficulty. Now relax the assumption that bookings b is less than or equal to capacity n and obtain

# This model gives the profits for a given number of bookings, accounting for the possibility of making more reservations than can be accepted. Then determine the maximum number of bookings by finding the amount of booing’s where or . Determine the number of bookings which maximize profit by solving this equation for *b*. Use an algorithm iterating over integer values of *b* starting at capacity and stopping when

See the source code in Table A.2 in Appendix.

In the hotel booking data (Antonio 1), the average arrival rate of a reservation is 64%. Assuming a hotel has a 100-room capacity, the cost of overbooking *L* is four times profit, and every guest has the same probability of arrival, maximize profit by booking 129 guests. The binomial distribution assumes that each *p* is the same for every trial. Assuming all customers have the same *p*, exposes the hotel to the inevitable event of booking too many customers highly likely to arrive and booking too few customers unlikely to arrive. The profitability of this strategy will depend on the ability to accurately estimate *p*. Therefore, this strategy aims to estimate \phat

The regression results in Table A.1 in Appendix A shows each guest has a unique probability of arrival. To estimate the probability of each guest’s arrival use logistic regression.

Fit the logistic regression to 80% of the data to obtain a for each booking.

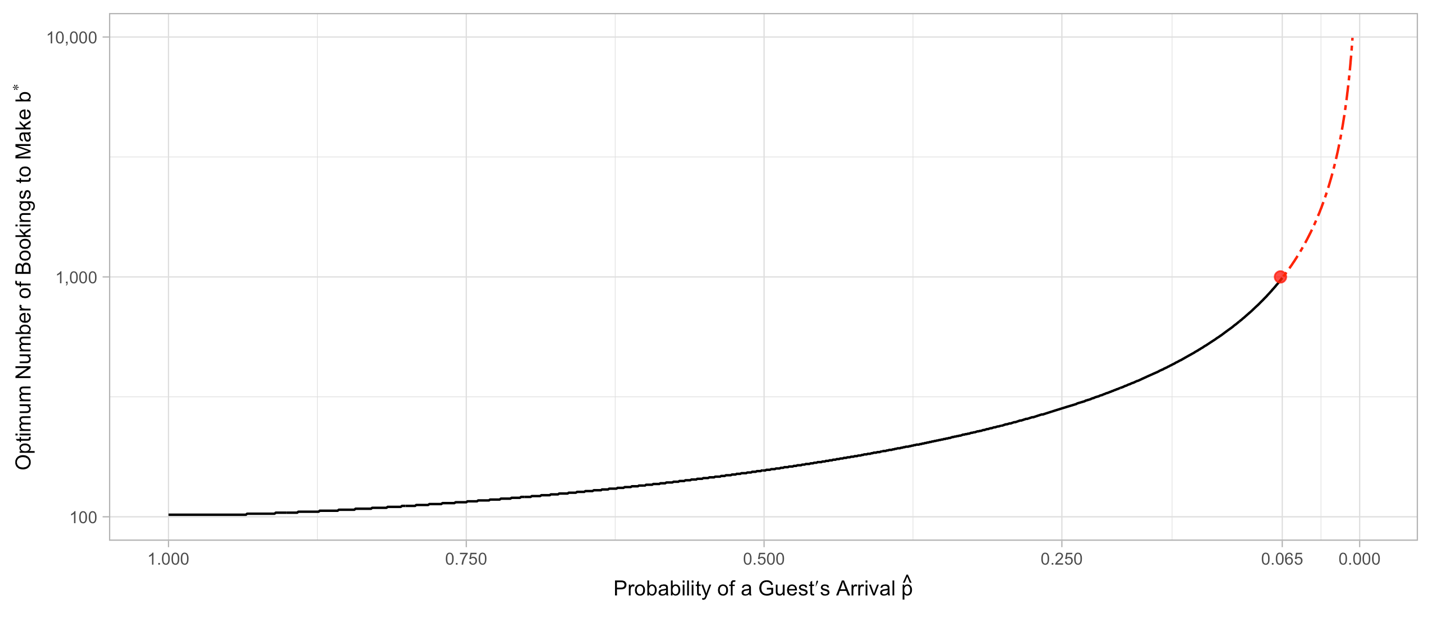
# 

# Methods

Begin constructing this overbooking strategy by estimating the probability a given guest will arrive. Split the booking data into a training dataset and a testing data set. Rather than randomly selecting data, split the data by the first 80% chronologically. Then fit and validate a predictive model. A logistic regression model estimates the probability arrival to be

These are the variables a booking agent would have available at the time of a booking. Therefore, ­ can be computed in real time. The estimated coefficients can be found in Appendix A Table A.1. These regression coefficients are all statistically significant. This supports the claim that the probability of arrival is not the same for each guest. Groups on average tend to have a higher arrival rate than individual short-term guests. The more adults in a single booking, the lower the expected arrival rate. Children seem to increase the odds of arrival until accounting for factors such as group, lead time, and stay length. Guests not making a deposit has a far higher arrival rate than those making non-refundable deposits. Closer inspection shows that non-refundable deposits are asked only from guests requesting flexible booking or had a history of cancelling reservations.

Use to obtain the estimated fraction of capacity of the booking. Let be the optimal number of bookings *b* where given . Figure 1 shows that goes to infinity as the probability of arrival approaches 0. In a commercial setting, this should be restricted to avoid the possibility of a catastrophic loss. This strategy assumes restricts bookings to no more than 10 times capacity.



**Figure 1** Relationship between estimated probability of arrival and optimum bookings.

aFind the expected fraction of capacity of a specific booking by taking the inverse of

.

Assign each booking an expected fraction of capacity. Then accept bookings until the total fraction of capacity exceeds 0.99.

Randomly sample bookings on each date until For the details of the sampling function see the source code in Appendix A Table A3.

# Results and Discussion

# Conclusions

# References [This section must start on a new page. Please stay with uniform format (e.g., MLA format) for the references list below.]

“14 CFR 250.2B -- Carriers to Request Volunteers for Denied Boarding.” *14 CFR 250.2b -- Carriers to Request Volunteers for Denied Boarding.*, https://www.ecfr.gov/current/title-14/chapter-II/subchapter-A/part-250/section-250.2b.

Leahy, John. “Too Many Orders? Yes, Says Consultant. No, Says Ex-Super-Salesman.” *Leeham News and Analysis*, 28 Jan. 2019, https://leehamnews.com/2019/01/28/too-many-orders-yes-says-consultant-no-says-ex-super-salesman/.

# Appendix A

Table

Description automatically generated

This overbookiGraphical user interface, text, application, email

Description automatically generatedng strategy is tested under the conditions that a hotel may make as many bookings as they are willing to accept.

**Graphical user interface, text, application, email

Description automatically generated**

Table A.3 Bootstrap resampling until capacity is reached

# Appendix B [Each appendix section must start on a new page in case you have more than one appendix section.]