Managing overbooking in hotels: A probabilistic model using Poisson distribution

Gopi Nath Vajpai
qvaipai@gmail.com
Renaissance College of Hotel Management and Catering Technology, Ramnagar, Uttarakhand

ABSTRACT

Yield Management is commonly practiced in all hotels and other industries where advance booking is required. Various revenue and non-revenue tools of yield management have been practiced and applied. One important tool for yield management is using overbooking in hotels and airlines industries. For management at operational level, a sound understanding of overbooking will allow maximizing revenue.

The real challenge arises when faced with the question regarding how many overbooking is optimum. If we are overly cautious and overbook very less, it would mean the loss of revenue for the hotel. On the other hand, if we overbook to a large extent it could result in confirmed reservations being walked (turned away). This may lead to unsatisfied guests and a poor brand image for the hotel.

Hence a balanced approach is required which will help us optimize the overbooking level. In most hotels, this is achieved by adjusting overbooking level based on the experience of Front Office manager.

This paper utilizes Poisson distribution as a tool to model the probability of no-shows in a hotel using historical data and hence optimize the number of overbooking to be accepted in a given time period.

Keywords: Overbooking, Revenue Management, Room Reservation, Poisson distribution.

1. INTRODUCTION

When examining historical data it has been observed that even during peak seasons hotels are not able to attain 100% occupancy. This is due to the fact that many of the customers make last-minute cancellations, some of them do not turn up on the day of arrival due to various reasons (no-show) while many guests check-out before their scheduled departure date (called understand or early departure). These things significantly hamper the revenue of the hotel as the unsold rooms could have been sold to other guests to increase hotel profitability.

Consider a hotel having 200 rooms with 5% no-show or last minute cancellations. During peak season 10 rooms go unoccupied despite the hotel being fully booked. If the average tariff for the hotel was 80 US Dollars. It would mean the hotel could have earned 800 additional Dollars by selling the unused rooms. Which means a loss of 24000 Dollars in a month.

Business-wise this is a huge loss of revenue and inability of the manager to cash in on the opportunity.

To overcome this problem, it has become a common practice to take more bookings than the number of total rooms in the hotel. This is done in order to accommodate for no-shows, late cancellations, and understands. This technique is in use in many Airlines, Hotels, and Restaurants.

In most cases, overbooking is viewed by the general public with skepticism (Bardi, 2006).

When a hotel overbooks a sold-out date, it is taking a calculated risk that more guests will understay, cancel or no-show than the number of room by which the hotel has overbooked (Vallen & Vallen, 2009).

The amount of overbooking taken is decided based on historic data of the number of no-shows, late cancellations and understays as well as the experience of the Front Office/Reservation manager.
It is noteworthy that there are other revenue management techniques also available to Front Office managers. However, our research paper will only look at overbooking technique and how we can effectively use it to accurately predicted overbooking level to maximize hotel revenue.

**Suitability of Poisson distribution for Modelling**

Poisson distribution is a discrete probability distribution frequently used to model the number of events happening in a given time interval if the average number of times the event occurs in a given time period is known. Since the number of no-shows, late cancellations and understays (NULs) occur in whole numbers (0, 1, 2…) and they are non-negative, Poisson distribution can easily be applied to accurately arrive at probabilities of different level of NULs happening on a given day. These probabilities will, in turn, help us to find the optimum level of overbookings to be accepted for the hotel for a given day.

Generally Poisson model fits if the data in consideration adheres to following assumptions:

- Events (NULs) occur independently of each other i.e. the happening or not happening of one event does not affect other events.
- The probability of an occurrence of a single event over a small interval is approximately proportional to the size of the interval.
- The events occur in the whole number, i.e. they can take values 0, 1, 2… and so on.

**2. LITERATURE REVIEW**

Overbooking has been widely practiced and advised. Vallen & Vallen (2009) mentioned that the fault for overbooking is not the hotel industry’s alone. Tour operators, convention housing committees, and individual guests are all to blame. Each the hotel included attempts to maximize its own position at the risk of overbooking. They further argued that there is no perfect solution to the problem of overbooking. As long as hotels overbook to compensate for no-show and last minute changes in occupancy, there will be walk guests.

A stochastic cancellation model was initially developed by Yechiali and Varda (1978) which was a model to overcome hotel overbooking problem. This model was restricted to a single day or a given period of time which was sold as a single entity. The model also considered compensation to customers whose reservation could not be honored.

The interrelatedness of yield management, overbooking and pricing in general and many types of problems thereof have been discussed by Weatherford and Bodily (1992). They have proposed the term perishable-asset revenue management (PARM) to denote the field that combines the area of yield management, overbooking, and pricing for perishable assets. The discussion was formulated in a generalized manner to which can be implemented for a variety of asset classes.

An early model for hotel overbooking was also developed by Toh and Decay (2002) which shows the derivation and also a simple to a final formula which can be used easily.

From the point of view of Operation Research Karaesmen and Ryzin (2004) have discussed an overbooking problem with multiple reservation and inventory classes, in which the multiple inventory classes may be used as substitutes to satisfy the demand of a given reservation class. The problem is to jointly determine overbooking levels for the reservation classes, taking into account the substitution options.

Hwang and Wen (2009) in their study tried to investigate the effect of hotel overbooking and compensation practices on customer’s perception of fairness and loyalty.

A study on the effect of cash-based compensation and voucher-based compensation was also done by Lee and Noone (2010). Additionally, it also gave an insight into the face that overcompensation does not significantly influence customer repatronage intentions.

In yet another attempt Ling & Dong (2015) attempted to discuss pricing and overbooking strategies of a hotel in context of cooperation with multiple third-party websites. The emphasis was on analyzing how these strategies influence the cooperation process.

More recently, Tse and Poon (2016) have also developed a conceptual model to maximize revenue in a restaurant using reservation data. For different days of the week, they observed the demand for the restaurant seat. By modeling cancellations, no-shows and walk-ins in the restaurant they arrived at the optimum overbooking to be accepted.

Other attempts have also been made to more accurately predict the optimum level of overbooking. Some of these are based on probability and many others use simulation approach for the same.

**3. OBJECTIVES**

By using this model following objectives will be achieved

- Find the probability of a required number of NULs happening based on historical data
- Based on above probability optimize the number of overbooking to be accepted
4. DEFINITIONS AND NOTATIONS

No-Show: An event wherein a confirmed reservation does not turn up at the hotel on the day of his check-in.

Understay/Early Departure: It means customers checking-out before their scheduled departure date e.g. a customer who is booked for 3 nights checking-out after 2 nights.

Late Cancellation: Last minute cancellation done by customers generally few hours before arrival.

NUL(s): For the purpose of ease of writing, in this research paper we will henceforth be referring NUL to mean either of the event/events: a No-show, Understay or Late Cancellation. Practically every NUL will increase the available rooms in the hotel by 1 for that day.

\( \lambda \) (Lambda): The Poisson parameter, here representing the average number of events(NULs) happening per unit time(1 day in our case).

Walking: The event when a confirmed reservation cannot be accommodated due to unavailability of rooms. In other words, denying a room to a customer even if he has a confirmed reservation.

5. ASSUMPTIONS

For the simplicity of modeling data some assumptions are necessary which have been listed below:

- The number of NULs occur at a constant rate for a given time frame which generally is one season/month.
- All the rooms under consideration are of same category and tariff.
- Each NUL is independent of other NUL(s) i.e. these events happen independently of each other.
- The time period under study is taken in such a manner that the general occupancy levels and NULs follow similar pattern within that time-frame.

6. METHODOLOGY

We will first assume a hotel where the NULs are occurring at a rate of \( \lambda \) per day. This is the average rate of NULs happening and the value of \( \lambda \) here can be calculated from the past records of the hotel for the time period say a month or quarter.

Our next step will be to find the probability of a particular number of NULs occurring on a given day. This can be calculated by using Poisson distribution.

We can denote the random variable X following a Poisson distribution as below

\[ X \sim \text{Poi}(\lambda) \]

here the random variable, X denotes the number of NULs happening in a time period.

To calculate the probability of \( x \) number of NULs occurring for each value of \( x \) we can use the formula:

\[ P(x=X) = e^{-\lambda} \frac{\lambda^x}{x!} \]

Where,

\[ e = \text{Euler’s constant (approximately 2.7182)} \]

\[ x! = \text{factorial of } x \]

\( x \) here can take values 0, 1, 2 and so on.

By using above formula we will arrive at the probabilities of the different number of NULs occurring on a given day. A cumulative probability can also be arrived at by adding all the probability values starting from 0 up to \( x \).

Our next step will be to find out what is the probability the hotel will be successful in accommodating a particular number of overbookings say \( k \). This will be possible only if minimum \( k \) NUL’s happen on that day.

Therefore we will need to find the probability of minimum \( k \) NULs happening.

Since all the probabilities add up to 1. We can easily find the probability of minimum \( k \) NULs happening by subtracting from 1 the probability of \( k-1 \) NULs happening.

By using above method we can plot the probability of success for different values of Overbooking (denoted by OB) for different mean values of \( \lambda \).
Table – 1 Probability of Success(No Walking) for each level of $\lambda$.

<table>
<thead>
<tr>
<th>OB</th>
<th>$\lambda=3$</th>
<th>$\lambda=4$</th>
<th>$\lambda=5$</th>
<th>$\lambda=6$</th>
<th>$\lambda=7$</th>
<th>$\lambda=8$</th>
<th>$\lambda=9$</th>
<th>$\lambda=10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.9502129</td>
<td>0.9816844</td>
<td>0.9932621</td>
<td>0.9990881</td>
<td>0.9996645</td>
<td>0.9998766</td>
<td>0.9999546</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.8008517</td>
<td>0.9084218</td>
<td>0.9595723</td>
<td>0.9826487</td>
<td>0.9927049</td>
<td>0.9969808</td>
<td>0.9987659</td>
<td>0.9995006</td>
</tr>
<tr>
<td>3</td>
<td>0.5768099</td>
<td>0.7618967</td>
<td>0.8753480</td>
<td>0.9380312</td>
<td>0.9703638</td>
<td>0.9862460</td>
<td>0.9937678</td>
<td>0.9972306</td>
</tr>
<tr>
<td>4</td>
<td>0.3527681</td>
<td>0.5665299</td>
<td>0.7349741</td>
<td>0.8487961</td>
<td>0.9182346</td>
<td>0.9576199</td>
<td>0.9787735</td>
<td>0.9896639</td>
</tr>
<tr>
<td>5</td>
<td>0.1847368</td>
<td>0.3711631</td>
<td>0.5595067</td>
<td>0.7149435</td>
<td>0.8270084</td>
<td>0.9003676</td>
<td>0.9450364</td>
<td>0.9707473</td>
</tr>
<tr>
<td>6</td>
<td>0.0839179</td>
<td>0.2148696</td>
<td>0.3840393</td>
<td>0.5543204</td>
<td>0.6992917</td>
<td>0.8087639</td>
<td>0.8843095</td>
<td>0.9329140</td>
</tr>
<tr>
<td>7</td>
<td>0.0330885</td>
<td>0.1106740</td>
<td>0.2378165</td>
<td>0.3936972</td>
<td>0.5502889</td>
<td>0.6866257</td>
<td>0.7932192</td>
<td>0.8698586</td>
</tr>
<tr>
<td>8</td>
<td>0.0119045</td>
<td>0.0511336</td>
<td>0.1333717</td>
<td>0.2560202</td>
<td>0.4012862</td>
<td>0.5470392</td>
<td>0.6761030</td>
<td>0.7797794</td>
</tr>
<tr>
<td>9</td>
<td>0.0038030</td>
<td>0.0213634</td>
<td>0.0680936</td>
<td>0.1527625</td>
<td>0.2709087</td>
<td>0.4074527</td>
<td>0.5443474</td>
<td>0.6671803</td>
</tr>
<tr>
<td>10</td>
<td>0.0011025</td>
<td>0.0081322</td>
<td>0.0318281</td>
<td>0.0839240</td>
<td>0.1695041</td>
<td>0.2833757</td>
<td>0.4125918</td>
<td>0.5420703</td>
</tr>
<tr>
<td>11</td>
<td>0.0000714</td>
<td>0.0009152</td>
<td>0.0054531</td>
<td>0.0200920</td>
<td>0.0533496</td>
<td>0.1119240</td>
<td>0.1969916</td>
<td>0.3032239</td>
</tr>
<tr>
<td>12</td>
<td>0.0000161</td>
<td>0.0002737</td>
<td>0.00020189</td>
<td>0.0088275</td>
<td>0.0269998</td>
<td>0.0637972</td>
<td>0.1242266</td>
<td>0.2084435</td>
</tr>
<tr>
<td>13</td>
<td>0.0000034</td>
<td>0.0000763</td>
<td>0.00006980</td>
<td>0.0036285</td>
<td>0.0128114</td>
<td>0.0341807</td>
<td>0.0738508</td>
<td>0.1355356</td>
</tr>
<tr>
<td>14</td>
<td>0.0000007</td>
<td>0.0000199</td>
<td>0.00002263</td>
<td>0.0014004</td>
<td>0.0057172</td>
<td>0.0172570</td>
<td>0.0414663</td>
<td>0.0834585</td>
</tr>
<tr>
<td>15</td>
<td>0.0000001</td>
<td>0.0000049</td>
<td>0.0000690</td>
<td>0.0005901</td>
<td>0.0024066</td>
<td>0.0082310</td>
<td>0.0220357</td>
<td>0.0487404</td>
</tr>
</tbody>
</table>

The graph below shows the probability of success (plotted on Y-axis) against a number of overbooking accepted by the hotel (shown on X-axis) for different values of $\lambda$.

![Chart -1 Probability Chart for Accepting Overbookings](image)

**Interpretation for example**

Each colored line in this graph represents different mean values of $\lambda$. We can see that a hotel experiencing NULs ($\lambda$) at the rate of 10 per day (extreme right curve on the graph) can accept up to 7 overbookings with roughly 85% surety of success (no customer walked).

Similarly, a hotel having average NULs of 5 for a particular season can accept 3 overbookings with about 90% surety of success.

**7. CONCLUSION**

We can see that Poisson distribution can help in accurately predict future arrivals for a hotel if the data for the past period is available. However for better managing the yield of any hotel these predictions need to be clubbed with other revenue management strategies such as accepting walk-in guests in case of unsold rooms, last minute discounts and other similar strategies. One important factor that is a cost for compensating a walked guest may be different in different hotels and may vary on a case to case basis. This factor has not been considered in the present model and if included can give an overall better strategy for the hotel.

Further, this model can be assembled into a Microsoft-excel spreadsheet using historic data or developed into a software programme to assist hotel owners and managers to optimize their overbooking levels and hence maximize hotel profitability.
Although these derivations can help in managing the yield of the hotel in a better way, there are some limitations which may be addressed in future researches. Some of these are pointed below:

- Although using Poisson distribution gives a better estimation of the probability, however the variance (spread) of historic data has not been considered in this model. More spread data will be riskier and revenue managers will need to be more cautious when dealing with same.

- Presently this model does not consider the magnitude of revenue loss which must be paid as compensation for all walked bookings. Therefore it can be made more robust by examining the exact loss to the hotel for each walked booking and thus maximizing hotel revenue.

- Although all NULs are assumed to be independent, however in actual operations few reservations could be dependent such as a guest having 2 room reservations for the same date.

8. REFERENCES


