Simulating hidden demand in dynamic pricing for hospitality industry

Anton Muratov, Atomize AB

November 8, 2018

1 Motivation

Atomize AB is a Gothenburg-based company building an intelligent pricing engine for hospitality industry. We do this in two ways: either providing our customers price recommendations that they would then have to manually accept in our GUI, or in a fully automated fashion, taking real-time control of customers prices for them, given the restrictions they set up.

Dynamic pricing aims to answer a question: what would happen to the total revenue if the current price is changed by $\pm X\%$, and based on the answer, set the correct price. Good dynamic pricing requires a lot of sales data: it is necessary to probe the demand at different price levels against different market conditions.

In hotel industry it is not always possible to obtain perfect historical sales data. Reservation history often is incomplete and price history typically only contains the last observed price of the roomnight, not the whole trajectory. Also, new clients of Atomize often have fixed prices across their inventory for the entire season. This makes training dynamic pricing policies harder in general, as there is very little historical data that is of use. Worse, a big part of demand is *hidden* either by price censoring (when the price is too high, the person that would otherwise have booked the room, does not), or by availability censoring (when the room is booked out, no new bookings are observed), making it very hard to answer the dynamic pricing question on the historical data only.

Benchmarking the dynamic pricing policies can also be hard in practice, as it is typically impossible to conduct a clean A/B test. The simplest benchmark should answer a question: how does implementing the new policy affect the goal function (total revenue in this case) versus the default policy. In the current business model, however, it is hard to set up an A/B test controlling for day-of-week effects, nearby events and seasonality, and practically impossible to also take into account the market situation, competitors prices and availability etc.

2 Project scope

The end result of the project should be a model of hidden demand and unobserved customer behavior addressing these two problems. The model is to be trained on the sales history to simulate the hidden demand (potential reservations) so that given the historical pricing policy (historical prices) the simulated realized demand process (observed reservations) passes a formal distribution equality test against the observed realized demand.

It will be required to define the distribution equality criterion/criteria.

Simulator will then be used for training the price optimizer and benchmarking.

3 Suggested methodology

A non-homogeneous Poisson Process in a general space could be used as a starting point for the model, [1, Ch. 2] gives a good introduction. Additional information on statistics of point processes can be found in [4], [2].

The data analysis and modelling should preferably be done in Python 3, see [3] for a primer.

4 Data

Anonymized data from Atomize client base will be available for studying. Hotels vary in size (10 - 1000 rooms) and length of available history (typically around a year of data is available). For the historical period available, the data includes the history of price changes for the hotel's different roomtypes for the future stay dates up to one year ahead, as well as the change history of every reservation within the period, including the creation timestamp, price at which the room was booked, various additional costs etc. The data in Atomize storage already is purged from any sensitive guest data (we do not store names, locations, IP addresses etc.). Additionally, any variables that would allow identification of the hotel (such as location data, currency, real room names, etc.) will be removed from the study.

References

[1] Sung Nok Chiu, Dietrich Stoyan, Wilfrid S Kendall, Joseph Mecke, *Stochastic geometry and its applications*, John Wiley & Sons, 2013.

[2] Alan Karr, *Point processes and their statistical inference*, Routledge, 2017.

[3] Wes McKinney, *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*, O'Reilly Media, Inc., 2012

[4] Jesper Moller, Rasmus Waagepetersen, *Statistical inference and simulation for spatial point processes*, Chapman and Hall/CRC, 2003.