Pricing strategy optimization considering customer sensitivity with Monte Carlo simulations

Roberto Maestre   Elena Krasheninnikova

BBVA Data & Analytics

December 11, 2016

Overview

Overview 1

1. Introduction
   - Pricing context
   - Pricing strategy optimization
   - Customer sensitivity
2. Mathematical definition of the problem
3. Simulation, distributions and business constraints
   - Simulation
   - Distributions and business constraints
4. Experiments
   - Data
   - Experiments
5. Results
6. Conclusions

Introduction

The purpose of this research is to provide a simple framework to calculate the Pareto frontier of several pricing strategies through Random Optimization driven by a probabilistic model, maximum decay points and business constraints.

This methodology yields expected retentions and revenues and allows for the testing of new prices.

Pricing context

Generalized Lineal Model (GLM) is the standard in the pricing industry.

$$\varphi = g(\mu) = \mathbf{X}\beta \tag{1}$$

where $g(\mu) \equiv \mathbb{E}(Y)$, g is the link function, $\mathbf{X}$ is the model matrix and β is the vector of unknown parameters, with $\mathbf{Y}$ being independent and distributed as some exponential family distribution.

The car insurance sector commonly deals with two GLM models (frequency-severity method). If $g(\mu_c)$ represents the model for the number of claims and $g(\mu_s)$ the model for the severity of claims, and assuming independence between distributions, the final pricing model is defined as:

$$\varphi = g(\mu_c) * g(\mu_s) \tag{2}$$

where $\varphi \in \mathbb{R}$.

Pricing strategy optimization

As the grade of acceptance of a given price can be represented as a probability function, a GLM with *logit* link function is able to represent the probability of renewal for a given price. The next Equation represents the *logit* function:

$$\eta = \ln\left(\frac{\pi}{1-\pi}\right) = \mathbf{X}\boldsymbol{\beta} \tag{3}$$

Knowing the GLM model parameters, we can obtain the probability of renewal directly from the above equation:

$$\pi = \frac{\mathrm{e}^{\mathbf{X}\boldsymbol{\beta}}}{1+\mathrm{e}^{\mathbf{X}\boldsymbol{\beta}}} = \frac{1}{1+\mathrm{e}^{-\mathbf{X}\boldsymbol{\beta}}} \tag{4}$$

Customer sensitivity (1/2)

The sensitivity of each customer can be defined as the decay ratio ($\frac{\partial \pi}{v}$, $\frac{\partial^2 \pi}{v^2}$, $\frac{\partial^3 \pi}{v^3}, ...$) of the customer's probability for a given range of prices, where v is the variable related to the renewal ratio in the model matrix. We define the sensitivity function of one customer as:

$$\frac{\partial^k \pi}{v^k} = \frac{\partial^k}{v^k}\left[\frac{1}{1 + e^{-(v\boldsymbol{\beta}_v + \mathbf{Y}\boldsymbol{\beta}_y)}}\right] \tag{5}$$

Where k is the order of the partial equation, v is the renewal ratio and y is the rest of the variables included into the model matrix (excluded v).

Customer sensitivity (2/2)

We use the $min\{\psi''\}$ to measure how the rate of change in the probability π is itself changing.

$$\psi'' = \frac{\partial^2 \pi}{\partial v^2} = -\frac{{\boldsymbol{\beta}_v}^2 (e^{v\boldsymbol{\beta}_v + \mathbf{Y}\boldsymbol{\beta}_y} - 1) e^{v\boldsymbol{\beta}_v + \mathbf{Y}\boldsymbol{\beta}_y}}{\left(e^{v\boldsymbol{\beta}_v + \mathbf{Y}\boldsymbol{\beta}_y} + 1\right)^3} \tag{6}$$

Therefore the point where the rate of change of the probability π is minimum is targeted as a Maximum Decay Point (MDP). Later new prices are simulated close to these MDPs.

Overview 2

Mathematical definition of the problem (1/2)

With previous equations we define a set of n customers as:

$$\mathcal{D} = \{(p_i, min\{\psi''_i\}, \mathrm{C}_i, Cost_i) | i = 1, ..., n\} \quad (7)$$

where $p_i = \varphi_i + \varphi_i * v_i$, $\underset{1\times2}{\mathrm{C}_i}, \forall_i \mathrm{C}_{i0} < \mathrm{C}_{i1}$ and $Cost_i$ represents the current customer cost.

Therefore, the pricing strategy optimization is defined as:

$$\arg\max_{\mathbb{v}}(f_1(\mathbb{v}), f_2(\mathbb{v})) \quad (8)$$

where $\mathbb{V} = \bigtimes_{i=1}^{n} [C_{i0}, C_{i1}]$ represents the set of combinations of possible renewal ratios, $\mathbb{v} \subset \mathbb{V}$.

Mathematical definition of the problem (2/2)

The revenue function is represented by $f_1 \in \mathbb{R}$ and $f_2 \in [0, 1]$ represents the retention function. These two functions are defined as:

$$f_1(\mathbb{v}) = \sum_{i \in \mathbb{v}} \mathcal{B}(\pi_i(v_i))p_i - Cost_i \tag{9}$$

where $\mathcal{B}$ represents the Bernoulli distribution.

$$f_2(\mathbb{v}) = \frac{1}{n} \sum_{i \in \mathbb{v}} \pi_i(v_i) \tag{10}$$

where $\forall_i p_i \in [\mathrm{C}_{i0}, \mathrm{C}_{i1}]$ for both equations. As long as π and φ are known, we need to find the best set of $\mathbb{v}$ that maximizes Equation 8.

Overview 3

Simulation

Since the main goal is multi-objective (Equations 9, 10) we can apply the family of Monte Carlo algorithms. Monte Carlo methods (Random Optimization, Metaheuristics, etc.) can solve the problem of finding an optimized price $\mathbb{v}$ that maximizes revenue for several retentions.

We choose Random Optimization for its simplicity and its interesting properties from the engineering point of view.

Distributions and business constraints (1/2)

Two scenarios are possible in an insurance pricing application:

- Renewal prices from optimization have been applied to the customers in the past.
- Renewal prices from business rules have been applied.

In the former, the distributions of renewal prices can be used to simulate new prices, in the latter, we propose the usage of the MDPs as a direct input for simulations. For Normal distributions:

$$\mathbb{v} \sim \mathcal{N}(\mu = min\{\psi''_i\}, \sigma = K) \tag{11}$$

where K is a constant that can be chosen to reduce or increase the simulation variability.

Distributions and business constraints (2/2)

In most applications of practical value we should include business constraints in order to include the global strategy as a part of the optimization.

The Triangular distribution $\mathcal{T}$ could be used to test prices biased to the MDPs along the whole constraints range as shown here:

$$\mathbb{v} \sim \begin{cases} \mathcal{T}(l = m = \mathrm{C}_{i0}, u = \mathrm{C}_{i1}), & \text{if } min\{\psi''_i\} < \mathrm{C}_{i0} \\ \mathcal{T}(l = \mathrm{C}_{i0}, m = u = \mathrm{C}_{i1}), & \text{if } min\{\psi''_i\} > \mathrm{C}_{i1} \\ \mathcal{T}(l = \mathrm{C}_{i0}, m = min\{\psi''_i\}, u = \mathrm{C}_{i1}), & \text{otherwise} \end{cases} \tag{12}$$

where m is the mode and l, u are the lower and the upper values.

Overview 4

Data

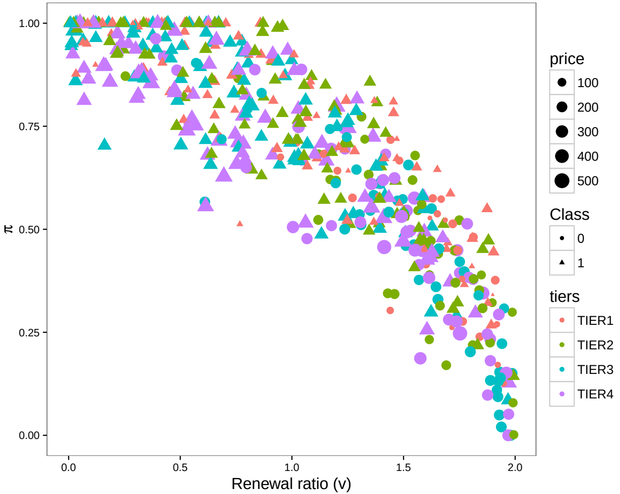

Figure 1: Synthetic data set with *price*, *tier* and *probability*

Experiments (1/2)

Table 1: Simulation scenarios

Scenario	$\mathbb{E}[f_1]$	$\sigma[f_1]$	$\mathbb{E}[f_2]$	$\sigma[f_2]$
baseline	**63719**		**0.7828**	
(I) $\mathcal{U}(\mathrm{C}_{i0}, \mathrm{C}_{i1})$	43580	2323	0.7433	1.22
(II) $\mathcal{N}(min\{\psi''_i\}, 0.2)$	60650	1962	0.7657	0.54
(III) $\mathcal{N}(min\{\psi''_i\}, 0.5)$	49470	2280	0.7162	1.14
(IV) $\mathcal{T}(\mathrm{C}_{i0}, min\{\psi''_i\}, \mathrm{C}_{i1})$	**63580**	1812	**0.7820**	0.11

where, in scenario (I): $\mathrm{C}_{i0} = 0.0, \mathrm{C}_{i1} = 1.9$ and in scenario (IV): $\mathrm{C}_{i0} = min\{\psi''_i\} - 0.1, \mathrm{C}_{i1} = min\{\psi''_i\} + 0.1$.

Experiments (2/2)

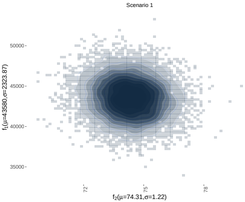

Figure 2: $\mathcal{U}(\mathrm{C}_{i0}, \mathrm{C}_{i1})$

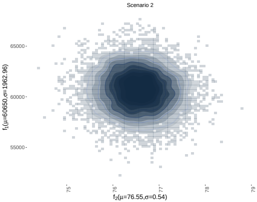

Figure 3: $\mathcal{N}(min\{\psi''_i\}, 0.2)$

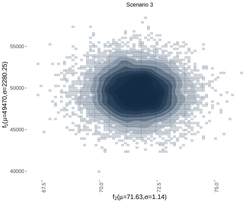

Figure 4: $\mathcal{N}(min\{\psi''_i\}, 0.5)$

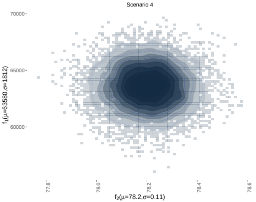

Figure 5: $\mathcal{T}(\mathrm{C}_{i0}, min\{\psi''_i\}, \mathrm{C}_{i1})$

Overview 5

Results (1/2)

Simulating several scenarios $\mathbb{v}$ also has the advantage of having similar expected revenues but with different retention points in some. This concept is related to the Pareto frontier of solutions.

As a consequence, if $\mathbb{v}^k$ represents the k-best scenario in expected revenue terms, i.e. $f_1(\mathbb{v}^1) \sim f_1(\mathbb{v}^2)$, the analyst could choose the scenario with higher probability taking into account the degree of similarity in f_1. Regardless of the final decision of each analyst, several $\mathbb{v}$ can be chosen as optimized pricing strategy for the whole portfolio of customers.

Results (2/2)

The next experiment runs 100000 simulations in order to achieve the best theoretical $\mathbb{E}[f_1]$ for a given $\mathbb{E}[f_2]$. The very positive effect of parallelization is remarkable.

Table 2: Parallel computation performance (in seconds)

Number of threads	User time	System time	Elapsed time
1	359.436	1.944	9.444.352
4	357.584	2.208	2.543.240
8	362.004	5.520	1.771.168
12	375.436	2.608	1.719.724
16	390.460	6.472	1.773.876
20	417.084	4.224	1.822.883

Overview 6

Conclusions

1. The main components such us pricing model, renewal model, business constrains and Maximum Decay Point are used to finding a scenario in which the revenue and the retention are maximized. Because the price sensitivity (and therefore the MDPs) is calculated from the renewal model, we need to ensure a robust model.
2. Several probabilistic distributions such as Normal ($\mathcal{N}$), Triangular ($\mathcal{T}$), Uniform ($\mathcal{U}$), etc. are used to generate renewal prices in order to test new scenarios.
3. We identify the Maximum Decay Point as the highest valuable point. At this point the ratio of change of price sensitivity drops significantly in relation between revenue and retention.
4. By means of Random Optimization, we provide a Pareto Frontier of optimal pricing strategies.

Contacts

Roberto Maestre,

✉ roberto.maestre@bbvadata.com

Elena Krasheninnikova,

✉ elena.krasheninnikova@bbvadata.com