

CAS PREDICTIVE MODELING CASE STUDY DETAILS



Case Overview

You head up the pricing function for Alpha Beta Gamma (ABG) Insurance, a regional personal lines carrier interested in expanding their business by increasing their penetration in the renter's insurance market. In the past few years, ABG began selling a renter's insurance product covering dormitory dwellings to college students. The CEO has asked to see an outline of a newly developed pricing strategy for current and new Dormitory Insurance business.

Company Background

As ABG has been selling the Dormitory policy, the company was able to capture claim and policy data related to the various coverages of the Dormitory policy. When the company initially launched the program, a very basic rating plan was launched, and it charged every customer the same premium. Now, as ABG collected more data points, the firm is in a position to update its pricing algorithm, in order to charge a more statistically accurate premium for each customer.

For current (or in-force) customers (or policyholders), ABG has information which will be provided to you related to prior claims (loss history) as well as certain additional (demographic) information that is collected each year via an underwriting questionnaire.

The information provided should be used to help you identify and segment policyholders according to their risk level. The dataset provided contains a complete list of information available (or variables) for each potential policyholder. Below is a sample of the types of information provided:

- Student's 'year' of study (freshman, sophomore, etc.)
- Dorm location (on/off campus)
- Distance to campus
- Greek (fraternity/sorority) versus not
- Etc.

Your goal is to develop a predictive model which uses both claim history and other demographic data provided to determine the actuarially-sound price for each policyholder.

Relevant Industry Information

Predictive Models

Predictive models optimize the use of available data to enhance underwriting and pricing decisions, predict customer behavior, and improve risk selection. Within the U.S., Generalized Linear Models are widely utilized and accepted as the modeling method of choice for insurance pricing. These pricing models can then be further supplemented with underwriting models utilizing advanced machine learning algorithms, such as Random Forest or Gradient Boosting Machines.

For the consumer, the (often) automated pricing and underwriting decisions from these models provide a consistent and objective outcome with a quicker turnaround time, thus shortening the time to policy issuance and improving the overall customer experience. U.S. Property and Casualty (P&C) insurers are progressively

expanding their use of predictive models, both in pricing and underwriting.

Insurance Coverage Provided:

Renters insurance generally covers the following:

Personal Property Protection

This coverage safeguards policyholders' valuables such as furniture, electronics and clothing. It covers the loss of belongings if they're stolen or damaged. The basic policy offered by ABG for dormitories covers up to \$10,000 of personal property.

Renters insurance might also include:

Additional Living Expenses

If a covered loss makes the dorm (or home away from home) uninhabitable, Renters Insurance (and the Dormitory Insurance offered by ABG) will help pay for increased living costs. This might include things such as a temporary place to live or other costs incurred while living in the temporary location (such as reasonable meal expense).

Liability Protection

Liability Insurance helps protect policyholders against financial uncertainty that can arise if someone sues for damages after being injured on the property. The liability protection offered by ABG's dormitory policy is limited to \$500,000.

Guest Medical Protection

Guest medical coverage, included in most residential insurance policies, provides protection if someone incurs medical expenses from an injury suffered at the insured property, regardless of who is at fault. The guest medical protection offered by ABG's dormitory policy is limited to \$150,000.

Exposures generally not covered by renters' or standard homeowners' policies:

- Acts of terrorism and biological or nuclear events
- Pandemics
- School property
- Flood, earthquakes (offered as separate policies or as endorsements in certain areas)
- Intentional acts

Insurance Product Structure:

For purposes of this exercise, the rate will be defined as the actuarially determined price per unit of exposure for a given risk. The rate will usually include an amount to cover expected loss, the expenses associated with the policy, and some expected underwriting profit.

For the Dormitory Insurance product, all four of the above-mentioned coverages are provided for each policyholder, and each coverage will have a stand-alone rating algorithm. The simple rating scheme will begin with a "base rate", which is the price charged for an "average" policyholder.

Underwriters have assigned three rating tiers to each individual policy based on underwriting expertise: Preferred, Standard, and Non-Standard. Base rates will vary with each tier. Also, within each rating tier, there may be additional factors that can be applied to the base rate to increase or decrease the base rate within that tier. For example, whether the insured property is sprinklered is correlated with the property's loss potential.

Data

The dataset provided includes 10,000 policyholders with varying degrees of risk. The dataset is on a coverage level for each policyholder, and the claim history is provided for each policyholder and coverage. In addition to the claim history, there is additional information provided about each policy, including information about the dorm (unit safety features, location, etc.) and information about the policyholder (year of study, major, etc.).

Deliverables

Your department's goal is to develop a method to determine the price for each insured. You will need to utilize the Claim History, the Underwriting Tiers, and other information provided to create a rating algorithm that can rate each policyholder.

Your primary focus should be on the risk segmentation described above. However, you may also wish to address:

- Are there any underwriting guidelines which you feel are important? This could include things like whether to decline coverage given certain policyholder characteristics.
- For continuous variables, how do you recommend implementing these variables into your pricing algorithm?
- Do you have thoughts or suggestions about other rating/pricing considerations such as overhead costs, advertising, claims processing, etc. How are these considerations different for new business vs renewal business?
- Are there any other relevant items?

Your presentation should cover the following:

- 1. Data & Inputs:**
 - a. Talk about whether the data is appropriate for the task at hand.
 - b. Are there additional sources of information which would improve your predictions?
 - c. What transformation(s) did you make to ensure the data was prepared for modeling? For example, did you need to cleanse the data?
- 2. Modeling Methods:**
 - a. Provide a description of the various statistical modeling techniques you decided to use. Discuss any assumptions you needed to make. If you used more than one statistical technique, explain your rationale for the final model you selected.
 - b. How did you decide which variables to include in your model? Did you decide to include or exclude a variable on something other than statistical grounds? Why?
 - c. Discuss the extent to which you feel your model is appropriate for new policyholders. Can you point to anything you might observe in future data that would cause you to reconsider your model or its parameters?
 - d. What evaluation graphs did you use to validate your modeled result?
- 3. Additional Considerations:**
 - a. How does customer satisfaction drive renewal probability? What is the cost of new versus renewal business?
 - b. Have you considered other factors that will impact the final rate/price charged? For example, will you offer additional credits/debits within each rating tier based on insured characteristics (that is further refine customer segments)?

- c. How would you consider the expense component of the rate (i.e., overhead, marketing, claims adjusting, etc.)?
- d. Without additional information in subsequent years, what is your rating strategy over time?

4. Results & Insights:

- a. You will need to rationalize not only the technical aspects of your model, but also recommend how to utilize the output of the model.
- b. Based on your model, what do you ultimately expect your insured population and the profitability of the Renters book of business looks like over the next 1-3 years?
- c. Will you develop any type of pro forma or business plan recommendation for the CEO based on the outcome of your predictive model?
- d. Your presentation should be targeted towards your CEO but provide enough detail to validate your work at a high enough level to engage your audience.

Evaluation Criteria

Your project will be evaluated based on the following criteria:

Presentation Skills

This is an important skill not just for actuaries, but for anyone who will be working in a business setting. The participants will be graded on general presentation skills, clarity of presentation, and demonstration of technical knowledge of the topic.

Technical Analysis

The participants should demonstrate proper usage of the data, think critically about the problem at hand, and make reasonable additional assumptions as needed. The analysis should be clear and well-documented, and sound from an actuarial perspective. Assumptions should be explained and be reasonable.

Use of Tools

Participants may have a number of options for tools at their disposal. Each participant should choose an appropriate tool for the task at hand and justify their selection. They should demonstrate appropriate usage of the capabilities of their chosen tools.

Business Knowledge

In addition to demonstrating an understanding of the techniques used in the analysis, the participants should understand the impact of their analysis from the perspective of the business. The participants should make business recommendations based on their analysis. Additional recommendations not directly rooted in the technical analysis may further demonstrate business acumen. The case study lists additional considerations that should be discussed; these considerations should be touched on at the very least.

Suggested Resources

CAS Student Central Webinars

- Predictive Analytics and Your Actuarial Career – April 2018
- Predictive Analytics and Data Science for Aspiring Actuaries – November 2015

Online articles:

- [An Introduction to Statistical Learning, with Application in R](#)
- [Introduction to Predictive Modeling Using GLMs](#)

- [CAS GLM Monograph](#)

Suggested Software Platform:

- R
- Python
- SAS
- SPSS
- MatLab
- Excel

Glossary of Insurance Terms

Claim – A formal request by a policyholder to an insurance company for compensation for a covered loss or policy event. The insurance company validates the claim and, once approved, issues payment to the insured or an approved interested party on behalf of the insured

Coverage – The amount or type of risk or liability protection that is provided for an individual or entity by way of insurance services

Covered Loss – Types of accidents, losses, or situations for which an insurance policy will pay compensation

Customer – Current or potential insured/policyholder

Demographic - A type of information related to identification. Examples of common demographics include age, gender, race, location, etc. Not all demographic information may legally be used in rating.

In-force – Active policy

Insured/Policyholder – Party/parties covered by an insurance contract

Personal Lines – Personal lines insurance includes property and casualty insurance products that protect individuals from losses they couldn't afford to cover on their own. These types of insurance lines make it possible to do things such as drive a car and own a home without risking financial ruin.

Policy – Legal contract containing binding terms of insurance coverage

Risk – The hazard or chance of loss and/or the degree of probability of such loss

Underwriting – The process of evaluating the risk of insuring a home, car, driver, etc. to determine if it's acceptable for the insurance company to take the chance on providing insurance (or to help determine the appropriate class, rate, discounts/surcharges, etc. that would support the acceptability of the risk)

Variable – A quantity or function that may assume any given value or set of values