Insurance: Mathematics and Economics 64 (2015) 67–76

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

Modeling the number of insureds' cars using queuing theory

Jean-Philippe Boucher *, Guillaume Couture-Piché

Quantact / Département de mathématiques, UQAM. Montréal, Québec, Canada

ARTICLE INFO

Article history:
Received November 2014
Received in revised form
March 2015
Accepted 21 April 2015
Available online 9 May 2015

Keywords:
Queuing theory
Poisson process
Lifetime customer value
Count distribution
Statistical inference

ABSTRACT

In this paper, we propose to model the number of insured cars per household. We use queuing theory to construct a new model that needs 4 different parameters: one that describes the rate of addition of new cars on the insurance contract, a second one that models the rate of removal of insured vehicles, a third parameter that models the cancellation rate of the insurance policy, and finally a parameter that describes the rate of renewal. Statistical inference techniques allow us to estimate each parameter of the model, even in the case where there is censorship of data. We also propose to generalize this new queuing process by adding some explanatory variables into each parameter of the model. This allows us to determine which policyholder's profiles are more likely to add or remove vehicles from their insurance policy, to cancel their contract or to renew annually. The estimated parameters help us to analyze the insurance portfolio in detail because the queuing theory model allows us to compute various kinds of useful statistics for insurers, such as the expected number of cars insured or the customer lifetime value that calculates the discounted future profits of an insured. Using car insurance data, a numerical illustration based on a portfolio from a Canadian insurance company is included to support this discussion.

1. Introduction

In recent years, a value that has gained interest in actuarial science is the *lifetime customer value*, see Guillén et al. (2012), Guelman et al. (2015) or Farris et al. (2010) for a general overview. This is a term from the marketing field that allows targeting of long-term clients. In insurance, it is a value that an insurer can assign to each insured, calculated by discounting all the future profits of the insured. Its application to the field of insurance is particularly recent (see Verhoef and Donkers, 2001), possibly because of the complexity of the actuarial field. To calculate the lifetime customer value of each client, as well as other useful statistics, we propose to model the number of insured cars per household. We use queuing theory (see Gross et al., 2008 for an overview) to construct a new model that needs 4 different parameters: one that describes the rate of addition of new cars on the insurance contract, a second one that models the rate of removal of insured vehicles, a third parameter that models the cancellation rate of the insurance policy, and finally a parameter that describes the rate of renewal. Using regression methods, we also identify insured profiles that are more interesting for insurers. A numerical illustration taken from a portfolio of a car insurance data from a Canadian insurance company is included to support this discussion.

In the second part of the paper, we propose the first approach to model the number of insured vehicles. This includes a process to model the arrival of new vehicles and another process to model the removal of insured vehicles from an existing insurance contract. A third process to model the renewal process will also be added to the model. In Section 3, we generalize the model of Section 2 to include a process that models cancellations during the contract. In Section 4, we propose a method to estimate the parameters of the models, for complete and for censored data. Covariates representing the characteristics of each household will then be added into each parameter of the process. In Section 5, we apply the model and calculate several useful statistics, such as the expected number of insured vehicles or the lifetime customer value. Possible generalizations of the model will be discussed in Section 6, while Section 7 concludes the paper.

1.1. Definition of terms

The term **household** is used to designate a single customer, or an insured. This household can include several members (or drivers) and several cars grouped under one annual **insurance contract**, which can be **renewed** each year. The contract represents the document that binds the insurer with the insured household. In this paper, we focus on the number of cars that the contract

* Corresponding author.
E-mail address: boucher.jean-philippe@uqam.ca (J.-P. Boucher).

http://dx.doi.org/10.1016/j.insmatheco.2015.04.008
0167-6687/

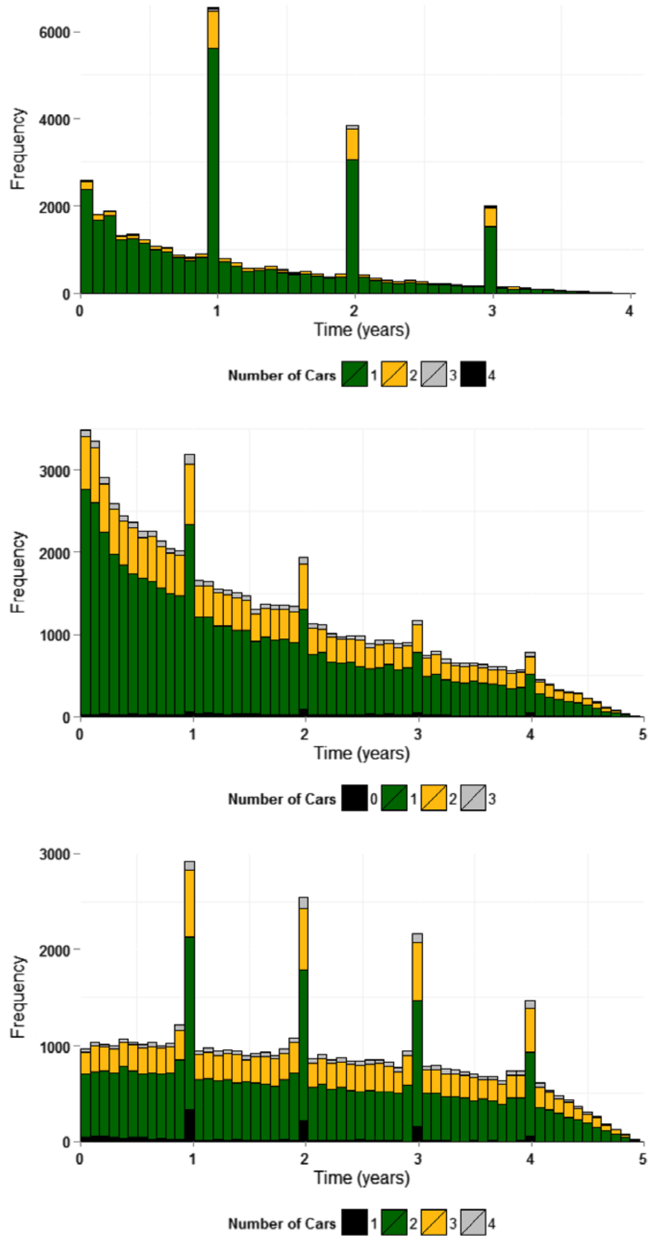

Fig. 1.1. Life of an insurance policy, time prior to the addition of a car and time prior to the removal of a car. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

covers and that are owned by the same household. By extension, **added cars** and **removed cars** from the insurance contracts are also analyzed. Finally, at any time during the insurance coverage, a household can decide to cancel its contract, meaning that all the insured cars are also canceled accordingly. We call this event a **breach of contract** or a **cancellation**.

1.2. Data used and notations

The use of queuing theory is based on the different waiting times before a change in the number of insured cars. We base our research on empirical analyses that come from a Canadian car insurance database. This database contains general insurance information on each of the 322,174 households for the period of 2003–2007. Note that because of the small length of the analyzed time period, the data are censored. For each household, we have information on each of its insured cars. We also have information about new or broken contracts, contract renewal, added or removed cars. Section 4.3 analyzes the database in detail, and describes the insurance data more precisely, particularly the characteristics of each of these policies.

A graphical analysis of waiting times involved in the modeling is presented. First, in Fig. 1.1, the distribution of the life of an insurance policy (in years) is shown. By the life of an insurance policy, we mean the time between the effective date of a new contract and the date of a non-renewal, between the effective date of a new contract and the date of a cancellation. In the first graph of this Figure, to avoid possible bias, we only used policies that were issued and canceled during the 2003–2007 period. This sample allows us to better understand the data. From the figure, we can see that there is a shock at each contract renewal date. Aside from that shock, we can also observe a decreasing exponential trend in the data. The color code shows that most departures happen when there is only one insured car on the insurance policy. In Fig. 1.1, the time before the addition of a car on an existing insurance contract is also shown. For this graph, all the data were used. The last graph of Fig. 1.1 uses again all data, and shows the time before a removal of a car on a policy, which remains in force despite the removal of a car. Again, we can see an exponential trend, shocks at each contract renewal date, even if these shocks are less important than the ones observed in the first graphs. The major purpose of our project is thus to create a mathematical model that will be able to approximate those observations.

Let $N(t)$ be a random variable representing the number of elements in a queuing system at time t. In our case, the number of elements is the number of insured cars owned by a specific household. The probability function of the number of insured cars will be expressed as $\Pr\{N(t) = i\} = p_i(t)$. The probability generating function (PGF) will be expressed as $P_{N(t)}(z, t) = \sum_{i=0}^{\infty} p_i(t) \times z^i$, and its partial derivative with respect to t by:

$$\frac{\partial P_{N(t)}(z, t)}{\partial t}(z, t) = \sum_{i=0}^{\infty} \frac{dp_i(t)}{dt} \times z^i. \tag{1.1}$$

Finally, the conditional probabilities will be represented and denoted as $p_{(j,i)}(s, t) = \Pr\{N(t) = i|N(s) = j\}$.

2. Modeling the number of vehicles

In this section, we introduce how queuing theory, based on Newell (1982), can be used to model the number of insured cars. We first introduce the Poisson process to model the arrival of a new vehicle, and we add another process to model the removal of cars from the contract. Fewer details will be given in this part of the paper because interpreting the result requires only basic knowledge of queuing theory. Nonetheless, this introduction to queuing theory allows us to explain some tools that will be used in complex models, such as the one developed in Section 3.

2.1. Addition and removal of vehicles

In a pure birth process, also called the Poisson process and illustrated in Fig. 2.1, there is only one component of arrival, defined by a parameter λ. The Chapman–Kolmogorov equation is defined in our context by:

$$p_i(t) = \sum_{j=0}^{\infty} p_j(s) p_{(j,i)}(s, t),$$

for $s < t$. We interpret this equation by the fact that a probability can be defined by the sum of all the different paths for a short time period. These equations thus require us to find the conditional