ATELIER THÉMATIQUE I.A.R.D.
NON-LIFE INSURANCE WORKSHOP

organisé par *Quantact, le Laboratoire de mathématiques actuarielles et financières* du CRM
organized by *Quantact*, the CRM Laboratory of Actuarial and Financial Mathematics

Université du Québec à Montréal (UQAM)
Pavillon Président-Kennedy Building
Room/salle : PK-3605
201, avenue du Président-Kennedy, Montréal
Vendredi, 6 avril 2018, 9h00-16h30
Friday, April 6, 2018, 9 :00-16 :30

Invité(e)s/Speakers

Julien Trufin (Université libre de Bruxelles)
9h00 – 12h00 (*short course*)

Machine learning techniques in non-life insurance ratemaking

In this short course, we will introduce some machine learning techniques in a non-life insurance context. Specifically, we will present several techniques such as regression trees, random forest and gradient boosting. With this talk, we aim at giving a survival kit to non-life actuaries that are more familiar with traditional GLM techniques. We will present some relevant R packages and numerical illustrations will be performed in R.

Quantact

CENTRE DE RECHERCHES MATHÉMATIQUES

Emiliano Valdez (University of Connecticut)
14h00 – 15h00

Joint Modeling of Customer Loyalty and Risk in Personal Insurance
(with Catalina Bolancé, Edward Frees and Montserrat Guillén)

This work connects two strands of research of modeling personal (automobile and homeowners) insurance. One strand involves understanding the joint outcomes of separate personal insurance contracts, e.g., do higher automobile claims suggest more severe homeowner claims? Joint modeling of personal insurance is complicated by the fact that the outcomes typically have a mass at zero, corresponding to no claims, and when there are claims, distributions tend to be right-skewed and long-tailed. Moreover, it is important to account for insured personal characteristics as well as characteristics of the contract and, in the case of auto and homeowners, features of the automobile and the house. A second strand of the literature involves understanding determinants of customer loyalty. For example, we now know that when a customer cancels one insurance contract, he or she is likely to cancel all other contracts soon after.

This paper examines longitudinal data from a major Spanish insurance company that offers automobile and homeowners insurance. The dataset tracks 890,542 clients over five years, many of whom subscribed to both automobile and homeowners insurance (75,536, or approximately 8.5%). To represent this data, we use copula regression to model the joint outcomes of auto and home claims as well as customer loyalty. Including customer loyalty, or duration with the company, is complicated because of the censoring of this time variable as well as the discreteness. Although customers may cancel the contract at any time, cancelation typically occurs at contract renewal, making this variable essentially a discrete outcome. Composite likelihood and generalized method of moments techniques allow us to address the special features of this data structure. Our estimation results provide evidence of interesting relationships among auto claims, home claims and customer loyalty.

Consistent with findings from other studies, we find that intertemporal dependencies are important, e.g., high auto claims from one year signal high auto claims for the following year. Work is ongoing to develop strategies that will allow the insurance manager to identify profitable portfolios through measurement of a customer loyalty index.

Marie-Pier Côté (Université Laval)
15h30 – 16h30

A Bayesian model for multivariate micro-level insurance claims
(with Christian Genest and David A. Stephens)

A Bayesian model for granular insurance claim amounts is proposed. It accounts for the multi-level, multivariate features of individual claims, e.g., multiple claimants for the same event, each of whom may receive benefits under different coverages. To avoid sampling bias induced when relying only on closed files, a multiple imputation procedure exploiting open file data is proposed. For a given claim, the combination of coverages under which payments are made forms a type which is modeled with multinomial regression. The presence of legal and claims expert fees follows a logistic regression, given the type. The strictly positive severities are then modeled with log skewed normal regressions linked by a Student *t* copula. The Bayesian framework yields a predictive distribution for the amounts paid, including parameter risk and process risk, while handling missing covariates and open files. The approach is illustrated with Accident Benefits car insurance claims from a large Canadian company.

Site web: www.quantact.uqam.ca/pages/seminaires