LISBON
SCHOOL OF
ECONOMICS &
MANAGEMENT
UNIVERSIDADE DE LISBOA

Master in

Actuarial Science

Master's Final Work

Internship Report

Health Insurance Pricing with Generalised Linear Models

Ana Beatriz Marques Cabral Valente

October - 2020

LISBON SCHOOL OF ECONOMICS & MANAGEMENT
UNIVERSIDADE DE LISBOA

Master in

Actuarial Science

Master's Final Work

Internship Report

Health Insurance Pricing with Generalised Linear Models

Ana Beatriz Marques Cabral Valente

Supervision:

João Manuel de Sousa Andrade e Silva

Maria Isabel Teixeira da Silva Pimenta Ribeiro

October - 2020

Acknowledgments

The realization of this internship report was a pleasant challenge and this accomplishment was only possible with all the support I received.

First of all, I would like to express my deepest appreciation to my thesis advisor, professor João Andrade e Silva, who provided valuable and constructive suggestions. His extensive knowledge and commitment were essential for the materialization of this report.

I would also like to acknowledge my colleagues at Allianz for their patient guidance and assistance. Despite all the barriers created by the pandemic situation and the home confinement, your support was fundamental and is greatly cherished. Special gratitude should be given to my company's supervisor, Isabel Ribeiro, for embracing this project by my side with an enthusiastic encouragement present at all times.

Finally, I must express my gratitude for the unconditional love of my family, friends and dogs. Thank you for always being there, your moral support and confort throughout this confinement phase were necessary to maintain the focus and finish this report.

Abstract

Generalized Linear Models (GLMs) are being broadly used in the Non-Life Insurance Pricing. The premium charged by the insurance company is calculated based on a tariff. The most standard procedure to estimate the pure premium is by assuming that the claim counts and claim amounts are independent. From this independence, the claim frequency and severity can be forecasted by distinct GLMs and the Tariff is obtained by combining both models.

The present report gives a brief introduction on the methodology and describes how we prepared the data prior to the GLM application. The models obtained for the Stomatology Treatments and Appointments, one of the many coverages that can be included in a Health Insurance policy, are analyzed in this report. The SAS software was used to construct the datasets and to properly organize the data and R was the software used for the modelling process. Once the models were estimated, the pure premium was calculated and a tariff for the mentioned coverage was constructed.

Finally, we compared the results obtained by modelling the coverage in R with the output obtained by my colleagues, using the software implemented by the company. We conclude that both models are not significantly different, despite having some structural distinctions.

Keywords: Health Insurance, Insurance Pricing, Tariff, Generalized Linear Models, Claim Frequency, Claim Severity

Resumo

Os Modelos Lineares Generalizados (GLMs) são amplamente utilizados na precificação de seguros do ramo Não Vida. O prémio cobrado pela seguradora é calculado com base em uma tarifa. A abordagem clássica para estimar o prémio é feita assumindo a independência entre o número de sinistros e o seu custo. A partir desta independência, a frequência e a severidade dos sinistros são estimados através de GLMs separados e a tarifa é obtida combinando os dois modelos.

O presente relatório fornece uma breve introdução sobre a metodologia e descreve como preparámos os dados antes da aplicação do GLM. Os modelos obtidos para os Tratamentos e Consultas de Estomatologia, uma das muitas coberturas que podem ser incluídas numa apólice de Seguro Saúde, são analisados neste relatório. O software SAS foi utilizado para construir as bases de dados e para organizar adequadamente a informação e o software R foi utilizado para o processo de modelagem. Uma vez estimados os modelos, o prémio puro foi calculado e a tarifa, para a cobertura mencionada, foi construída.

Por fim, comparámos os resultados obtidos em R com as conclusões obtidas pelos meus colegas, utilizando o software implementado pela empresa. Concluímos que ambos os modelos não são significativamente diferentes, apesar de apresentarem algumas distinções estruturais.

Palavras-Chave: Seguro de Saúde, Precificação de Seguros, Tarifa, Modelo Linear Generalizado, Frequência de Sinistros, Severidade de Sinistros

List of Tables

List of Figures

Contents