Data Is (Still) BIG

Jaqui Liebenberg
Client Solutions Actuary
RGA Reinsurance Company of South Africa Limited

Very big in fact! I recently attended the International Congress of Actuaries in Washington, D.C., and given the number of presentations on predictive modelling and big data, it was quite clear that the clever use of data to enhance our risk assessment capabilities is quite a hot topic in actuarial circles. This, at a time when accounting standards are under the spotlight, the European Union's Solvency II (and the Solvency Asset Management project locally) is front of mind, healthcare reform, tax changes and the legislative landscape are all in 'draft' territory and hence under actuarial scrutiny and subject to debate.

While big data and predictive analytics are current buzzwords within the insurance industry, they have been around for a while, and perhaps as an industry we have been relatively slow to catch up. All sorts of industries and companies can and do benefit from the power of using historical and current data to create predictions about future events or behaviours. Mobile phone companies use it to predict which of their customers are likely to buy a new data plan or cancel an existing data plan. Financial services companies can use it to detect fraudulent transactions or the likelihood of default on a loan or mortgage. The predictive power of data can help online advertising companies predict which advert we are likely to click on, and hence they can carefully select those they present to us, sneaky as it is! Its use even extends into the public services sector, where past data can be harnessed to predict the likelihood of criminal offenders becoming repeat

offenders, information that can then be used as a guide to appropriate prison terms.

Within life insurance, predictive analytics can be very powerful and its applications multifold. As Eric Siegel wrote in his book on the subject, predictive analytics is "The power to predict who will click, buy, lie or die", and all four of these elements and more can be extended to and explored within the context of life insurance.

Who will (and who should) click

Effective client segmentation in life insurance is no longer a 'nice-to-have' – it has become a necessity. The average American consumer was exposed to 500 advertisements each day in 1970. It has been estimated that this number grew to 5,000 per day in the early 1990s and that today people are bombarded by close to 30,000 marketing messages per day.[1] The result of this 'advertising clutter', as it has been dubbed, is that recall rates have fallen sharply. In 1965, consumers recalled 34% of the commercials they had seen. By the early 1990s this had dropped to 8%, and in 2007, in a survey asking consumers to recall recent commercials they had seen on a given day, consumers could barely name two commercials.[2]

Given the competitive environment we are in, this malfunction of the traditional marketing model is far from ideal for three main reasons. First, no company wants to annoy its most valuable asset: its customer base. Second, margins are consistently being driven

1 Jay Walker-Smith, Yankelovich Consumer Research.
2 Empire Research Group.

lower, so cost and resource efficiency are increasingly fundamental to success. Third, a company can only 'touch' a customer so many times, so it needs to make the most of each opportunity. Predictive analytics presents an opportunity to fix this situation by using data about the customer to ensure that as an industry we decide who will and who should click by developing the right offer for the right customer and delivering it to them in the right way.

Who will die (or at least how soon)

One of the most powerful and common uses of predictive modelling in life insurance is predicting those lives to whom we are comfortable offering life insurance, with or without extensive underwriting. RGA worked on a predictive modelling exercise with a bancassurer in Asia for this very purpose. The client was fortunate to have a large customer pool but was struggling to penetrate the life insurance market. A key finding that emerged as a result of our research was that the invasiveness of and time involved in undergoing medical underwriting were the main barriers to life insurance sales. So RGA was brought in to assist the client in developing a model to predict underwriting decisions on the bank's existing customer base and thereby identify which of its customers the

bancassurer could successfully offer guaranteed-issue cover to without significantly increasing its risk. The model was used to rank the risks from best to worst and split them into 10 buckets of equal size (deciles). The in-sample model results showed that in the top 3 deciles of model output, only 0.5% of the lives were substandard. On validation, using a portion of the bank's data set aside for the purpose of testing the model performance under real business conditions, 0.6% of the lives in the top 3 deciles of model output proved to be substandard.

The outcome: Traditionally, these lives would all have been underwritten, but the model results suggested that the bank could offer these and similar lives guaranteed-issue cover. This would assist the bank in achieving its goals of deeper market penetration and significantly reduced acquisitions costs.

At RGA, our data scientists (a cooler name than actuary, I have been told) have been involved in developing many predictive analytics models for various purposes for clients across the globe, and we will be happy to assist your organisation in its data modelling needs. As Arthur C. Nielsen, Sr., founder of ACNielsen, quite aptly wrote, "The price of light is less than the cost of darkness." ∎