Article from:

Product Matters!

February 2010 – Issue 76

Why Predictive Modeling For Life Insurance And Annuities?

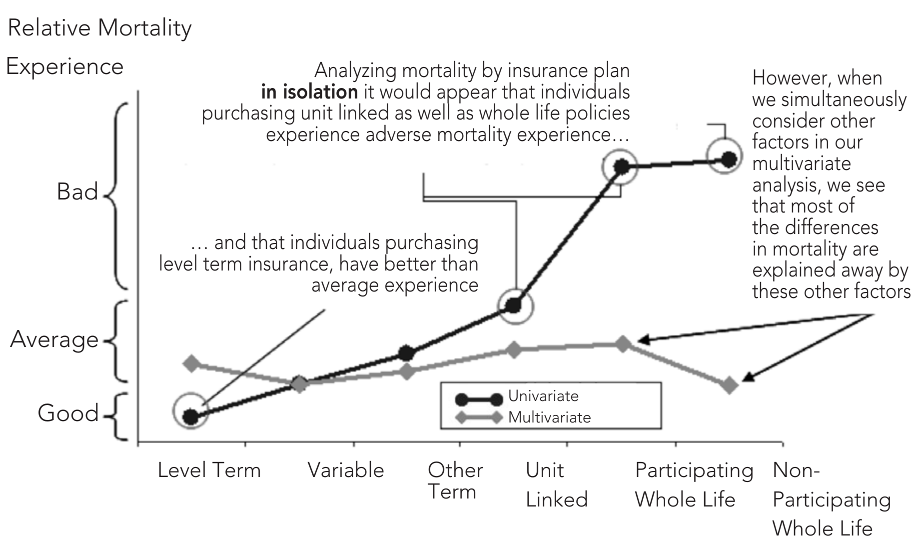
by Andy Staudt

Andy Staudt, FCAS, MAAA is an actuary with Towers Watson in San Francisco. He can be reached at andy.staudt@towerswatson.com.

Mark Twain wrote that "the art of prophecy is very difficult, especially with respect to the future." While this will always be true, it is even more true if we continue to rely on methods that require unnecessary assumptions to model the past, let alone the future. Yet predictive modeling offers an alternative where, by making use of more advanced statistical methods and algorithms, we can avoid making some of these simplifying assumptions. We can then start to both better understand what has happened historically as well as make more educated estimates as to where we are going in the future.

To help understand why predictive modeling should be a necessary (but by no means sufficient) tool for the life actuary, this article identifies examples of various limitations in traditional mortality experience analyses. It also posits various predictive modeling techniques as a means of avoiding these constraints.

FIGURE 1



Multivariate Analysis

Historically, the driving motivation for predictive modeling in many industries has been the desire to simultaneously analyze the effect of different factors on an item of interest—a process known as multivariate analysis. In contrast, most mortality analyses are univariate, where the effect of factors such as age, gender or smoking status on mortality are evaluated and presented in isolation. Although univariate analysis does provide a strong indication of how mortality experience varies, based on an explanatory factor, any combination of these results invariably leads to some redundancy or inadequacy. For example, if in our data set, smokers have higher mortality, but only women are smokers, then the effect of smoking will be captured twice, both by the gender variate as well as by the smoking status variate. Combining these two variates will lead to an overstatement of the mortality for women smokers. One band-aid for this problem works by slicing the data into various buckets (i.e., male/smokers, female/smokers, male/non-smokers, etc.) and evaluating the mortality experience for each bucket separately. This technique certainly solves the problem; however, as we increasingly want to slice along more and more dimensions, the credibility in each bucket decreases rapidly. As such, we turn to multivariate techniques that allow us to model the correlations and interactions among many different variables simultaneously.

Figure 1 shows how a univariate analysis may produce misleading and inaccurate results by ignoring correlation among several explanatory variables. Based on our univariate analysis, we may be inclined to believe there is significant disparity in mortality experience based on type of policy purchased. However, the multivariate result suggests that the type of policy has a much less pointed effect on mortality because variables, such as face amount of insurance purchased and/or issue age, explain away much of the variation by type of policy.

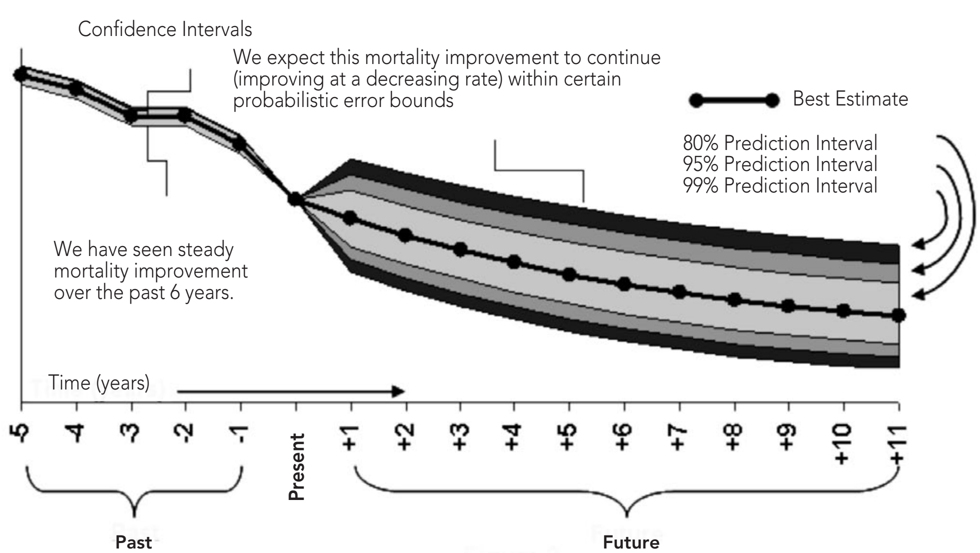
Controlling for the Environment

To predict future mortality, many mortality studies will use only the most recent years of historical experience as changes in the mix-of-business as well as changes in underwriting practices. Further, general mortality improvement over time will mitigate the extent to which past experience is predictive of future experience. However, we can carve these biases out of our data by including in our analysis direct or proxy variables, such as calendar year, which control for these changes over time. This practice then allows us to fit models using many years of data—which increases the credibility of our results—mitigating the effects of combining experience over these extended time periods. Furthermore, we are not only able to control for these historical changes, but we can extrapolate into the future general mortality change over time to gain a better estimate as to future mortality. Figure 2 shows exactly this type of analysis where we have removed the historical mortality improvement. This allows us to combine five years of experience to fit our mortality model, as well as our projection of future mortality improvement. This graph also shows intervals around our historical model fit representing our confidence in these estimates, as well as intervals around our projections giving us an indication of the possible error in our predictions. This type of analysis could easily and effectively be merged into dynamic, stochastic mortality models to develop a unified understanding of future behavior.

Optimal Variable Banding

It is a common modeling practice to combine or band together continuous, or pseudo-continuous, variables into discrete groups. In mortality analyses, examples might include banding together age (e.g., 15-to-19, 20-to-24, 25-to-29, etc.) and face amount (e.g., <25K, 25K-100K, 100K-500K, etc.). This process, loosely referred to as discretization, can be a powerful technique for deriving interpretable meaning out of an underlying dataset; however, if applied naively, it can also blur the true underlying relationship—leading to a loss in a model's predictive power. Traditional methods to banding variables include creating bands with equal interval-widths, such as those described above, or creating bands where the amount of data in each band is roughly equivalent. The former is effective at creating evenly spaced groups, while the latter is effective at ensuring that the results derived by band are equally credible. However, these methods are often not optimal because they create bands without consideration of the variable of interest (in our example, mortality) as well as interactions with other variables. Alternatively, predictive models developed using decision trees are able to optimally band together variables so as to not

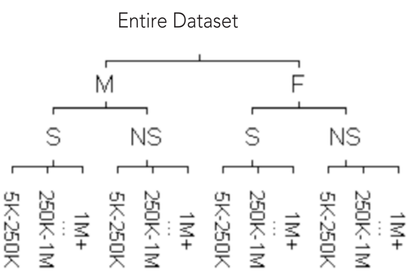
FIGURE 2



only optimize the various splits in a continuous variable (both in terms of number of bands as well as in terms of the size of various bands), but also to identify how to best group or cluster these bands with other variables. This maximizes the homogeneity of groupings (both in terms of similarities within groups as well as dissimilarities between groups), increasing the credibility of the result. Figure 3 shows just this type of analysis. And although it is not obvious from these graphs, decision tree analysis can also rank the splits, prioritizing those that are most effective and ignoring those that add little or no value.

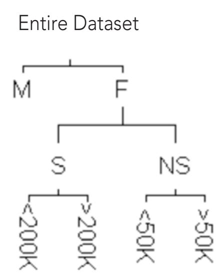
FIGURE 3

(i) Naive Analysis



(ii) Decision Tree Analysis

Note that the decision tree analysis has determined that not every split is required (i.e., mortality experience in this particular dataset does not vary by smoking status for males or face value for female/smokers)

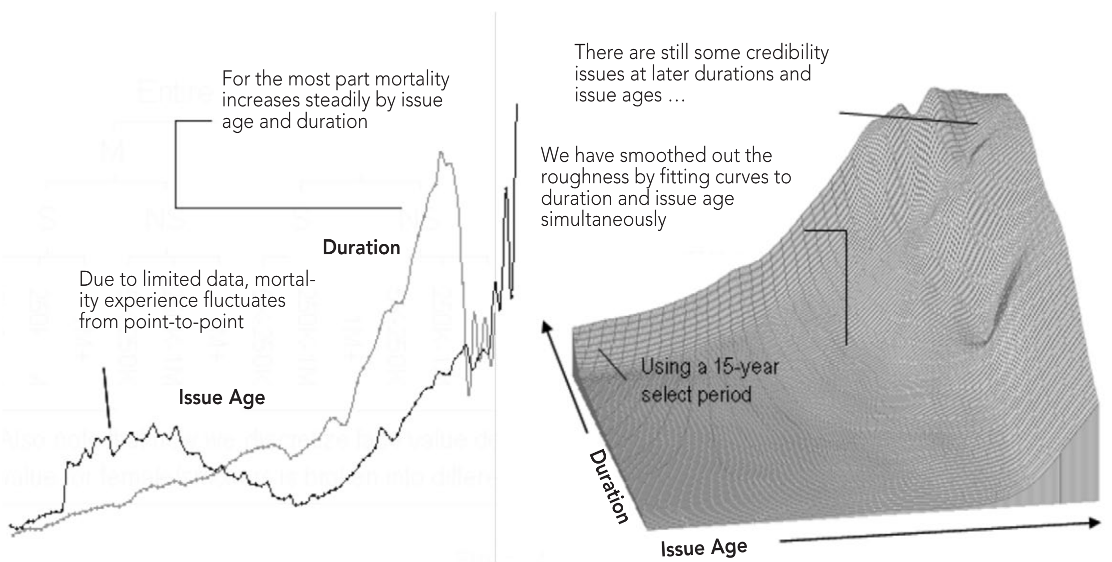


Also note that how we discretize face value depends on which cluster of data we wish to model (i.e., face value for female/smokers is broken into different groups than face value for female/non-smokers)

Removing Noise from our Estimates

To produce a workable mortality table varying by duration and issue age, it is necessary to smooth out the discontinuities in estimates of mortality by issue age and duration. These invariably result when working with imperfect data. To do this, many methods rely on first computing point estimates of mortality by issue age and duration and then adjusting these point estimates to meet some generally accepted criteria (e.g., mortality should increase by issue age, and mortality should increase by duration). The problem with this approach is that (1) it does little to suggest how much mortality should increase by issue age and how much mortality should increase by gender and (2) that this two-step procedure produces a result exogenous to the system, requiring user interaction, rather than a result that is optimized from within the system. Alternatively, rather than adjusting our point estimates of mortality at each issue age and duration point to produce a smooth mortality table, we can combine steps 1 and 2 and fit multi-dimensional functions, or curved planes, to the data. Furthermore, we can constrain these functions to represent realistic patterns, give more weight to more credible subsets of the data and be optimized over the space of all realistic functions. Figure 4 shows point

FIGURE 4



estimates at various issue ages and durations. Note that the curves produced by connecting the dots are quite jagged; combining them to produce a realistic mortality table will be no trivial task. The figure on the right, however, shows how we can fit multi-dimensional curves to smooth out the noise in our estimates and produce a workable mortality experience table.

Conclusion

The above examples conceptually demonstrate applications of predictive modeling within the life insurance industry. We have presented our examples within the context of mortality; however, the techniques described above could easily be applied to better understand and model other assumptions or lines of business such as:

- Dynamic policyholder behavior in interest-sensitive products or products with guarantees.
- Life insurers specializing in direct marketing who may greatly benefit by taking a page from the credit card companies' book. They could use customer characteristics to model propensity to buy, and subsequently propensity to persist, to selectively market to individuals who are expected to result in the highest level of profitability.
- Disability insurance claims that may best be handled using anomaly detection algorithms, which can be used to flag potentially fraudulent claims and allocate resources thusly. ◻