Predictive Modeling for Advanced Age Mortality

Lijia Guo[*]

Presented at the Living to 100 and Beyond Symposium

Orlando, Fla.

January 7-9, 2008

[*] Lijia Guo , PhD, ASA, MAAA, Algorithmics Inc., Fitch Group

Abstract

This paper introduces the predictive modeling tools to mortality research. The predictive modeling is applied to study how multiple risk drivers such as demographic characteristics and social and economic status impact the mortality improvement of the advanced age population. The paper provides both the theoretical frameworks and the application aspects of the predictive modeling process. As the result, a mortality risk score was derived in differentiating the mortality risk for the advanced age population. This process can also be used to derive morbidity risk scores upon data availability. The mortality risk scores developed in this study can also be used to enhance pricing and valuation of insurance products, marketing and insurance underwriting.

1. Introduction

In the past decade, the prospects of longer life have led to economic and social concerns over their implications for public spending on old-age support and other related topics. It is necessary to better estimate advanced age mortality for assuring the solidity of government and private pension plans; for improving life insurance and annuity pricing; for designing and pricing long term care insurance; and other actuarial practice.

Current mortality models have used a variety of mathematical techniques to generate mortality rates for advanced ages as a smooth extension of the patterns of mortality rates of septuagenarians and octogenarians. A comprehensive literature review is given by Tuljapurkar and Boe (1998). There are several approaches used to develop a basic scientific theory of mortality, including the evolutionary theory of senescence (Rose, 1991; Tuljapurkar, 1998), bio-actuarial theories (Pollard and Streathfield, 1979; Yashin, Manton and Vaupel, 1985) and hypotheses based on reliability theory (Wachter and Finch, 1997). All these approaches aim to explain the age pattern of mortality.

In addition to age, gender and smoking/non smoking, there are other risk factors that drive mortality experience, especially for seniors. The Society of Actuaries (SOA) (SOA, 2003) has linked mortality experience with insured, health status, etc. Until today, however, most mortality study hasn’t been able to capture the multiple factors and their interactions in developing mortality models.

Starting in the 1990s, many of the larger U.S. property & casualty (P&C) insurance companies began to implement predictive modeling techniques in the form of generalized linear modeling (GLM). Because of the early success realized by those companies, the vast majority of P&C companies are now starting to employ these techniques to keep up with competitors.

In insurance pricing, predictive modeling helps set base rates, quantify relationships among rating factors, enhance current models and develop special scores, such as a fire protection score. When applied to underwriting, predictive models can perform a variety of tasks, such as developing underwriting rules, performing credit analysis, creating profitability curves and determining the need for inspections. In the marketing area, predictive modeling can help determine the impact of a rate change, by incorporating customer retention in estimating the true impact on the overall book of business. In addition, predictive modeling techniques can help insurers more accurately set reserves, predict fraud and predict laws suits.

Predictive models are normally developed using rich historical data or from purposely collected data. In working with large databases, a key challenge is merging a large number of external data sources into a company's internal data. Predictive models are (normally) made up of a predictor and a number of factor variables that are likely to influence future behavior or results. Advanced statistical and data mining techniques for predictive modeling include decision trees, neural networks, generalized linear models, generalized addictive models and a combination of them.

This paper introduces the predictive modeling tools to the field of mortality research. The predictive modeling is applied to study how multiple risk drivers such as demographic characteristics, social and economic status and behavioral factors impact the mortality improvement of the advanced age population.

The paper is organized as follows. The next section introduces one of the basic predictive methods—decision trees applied it to identify the leading risk drivers in predicting mortality for the aging population. Section 3 introduces GLM and its application in senior mortality analysis. GLM is used to analyze the leading mortality risk drivers. The mortality risk score is derived in Section 4. These scores will be used for assessing the challenges and needs of the insurance products for different demographic group in different countries. Section 5 discusses some issues and techniques for projecting advanced age mortality improvement and offers some prospective thoughts for the future study.

2. Predictive Modeling: Decision Trees

The process of predict modeling starts with collecting data for the predictive variables $(x_{i1}, x_{i2}, ..., x_{ip})$. These are the drivers that affect the outcome of the target variable, Y, which we are trying to predict. Examples of the predictive variables include age, duration, gender and household income. Some example of the target variables are probability of events, profitability, loss ratio and lapse rate.

A predictive model is a process to derive the value of Y, where $Y = (y_1, y_2, ..., y_N)$ from $\{x_{i1}, x_{i2}, ..., x_{ip}\}$ based on $y_i = f\{x_{i1}, x_{i2}, ..., x_{ip}\}$. A traditional model form is

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} .$$

In the following decision tree method, one of the most popular predictive models is applied to the mortality study for the advanced age population. Decision tree is one of the basic predict modeling methods. The decision tree technique enables one to create decision trees that can classify observations based on the values of nominal, binary or ordinal targets; predict outcomes for interval targets; or predict the appropriate decision when you specify decision alternatives.

In the decision tree approach, an empirical tree represents a segmentation of the data that is created by applying a series of simple rules. Each rule assigns an observation to a segment based on the value of one input. One rule is applied after another, resulting in a hierarchy of segments within segments. The hierarchy is called a tree, and each segment is called a node. The original segment contains the entire data set and is called the root node of the tree. A node with all its successors forms a branch of the node that created it. The final nodes are called leaves. For each leaf, a decision is made and applied to all observations in the leaf. The type of decision depends on the context. In predictive modeling, the decision is simply the predicted value.

Specific decision tree methods include Classification and Regression Trees (CART) (Breiman et. al., 1984) and the count or Chi-squared Automatic Interaction Detection (CHAID) (Kass, 1980) algorithm. Both CART and CHAID are decision tree techniques used to classify a data set, and the inputs can be either nominal or ordinal. Many software packages accept interval inputs and automatically group the values into ranges before growing the tree. For nodes with many observations, the algorithm uses a sample for the split search, for computing the worth (measure of worth indicates how well a variable divides the data into each class), and for observing the limit on the minimum size of a branch. The samples in different nodes are taken independently. For binary splits on binary or interval targets, the optimal split is always found. For other situations, the data is first consolidated, and then either all possible splits are evaluated or else a heuristic search is used. The consolidation phase searches for groups of values of the input that seem likely to be assigned the same branch in the best split. The split search regards observations in the same consolidation group as having the same input value. The split search is faster because fewer candidate splits need to be evaluated.

A primary consideration when developing a tree for prediction is to decide how large to grow the tree or, what comes to the same end, what nodes to prune off the tree. The CHAID method specifies a significance level of a Chi-square test to stop tree growth. The splitting criteria are based on p-values from the F-distribution (interval targets) or Chi-square distribution (nominal targets). For these criteria, the best split is the one with the smallest p-value. By default, the p-values are adjusted to take into account multiple testing.

A missing value may be treated as a separate value. For nominal inputs, a missing value constitutes a new category. For ordinal inputs, a missing value is free of any order restrictions.

The search for a split on an input proceeds stepwise. Initially, a branch is allocated for each value of the input. Branches are alternately merged and re-split as seems warranted by the p-values. The original CHAID algorithm by Kass stops when no merge or re-splitting operation creates an adequate p-value. The final split is adopted. A common alternative, sometimes called the exhaustive method, continues merging to a binary split and then adopts the split with the most favorable p-value among all splits the algorithm considered.

After a split is adopted for an input, its p-value is adjusted, and the input with the best-adjusted p-value is selected as the splitting variable. If the adjusted p-value is smaller than a threshold you specified, then the node is split. Tree construction ends when all the adjusted p-values of the splitting variables in the unsplit nodes are above the user-specified threshold.

Tree techniques provide insights into the decision-making process, which explains how the results come about. The decision tree is efficient and is thus suitable for large data sets. Decision trees are perhaps the most successful exploratory method for uncovering deviant data structure. Trees recursively partition the input data space in order to identify segments where the records are homogeneous. Although decision trees can split the data into several homogeneous segments, and the rules produced by the tree can be used to detect interaction among variables, it is relatively unstable and it is difficult to detect linear or quadratic relationships between the response variable and the dependent variables.

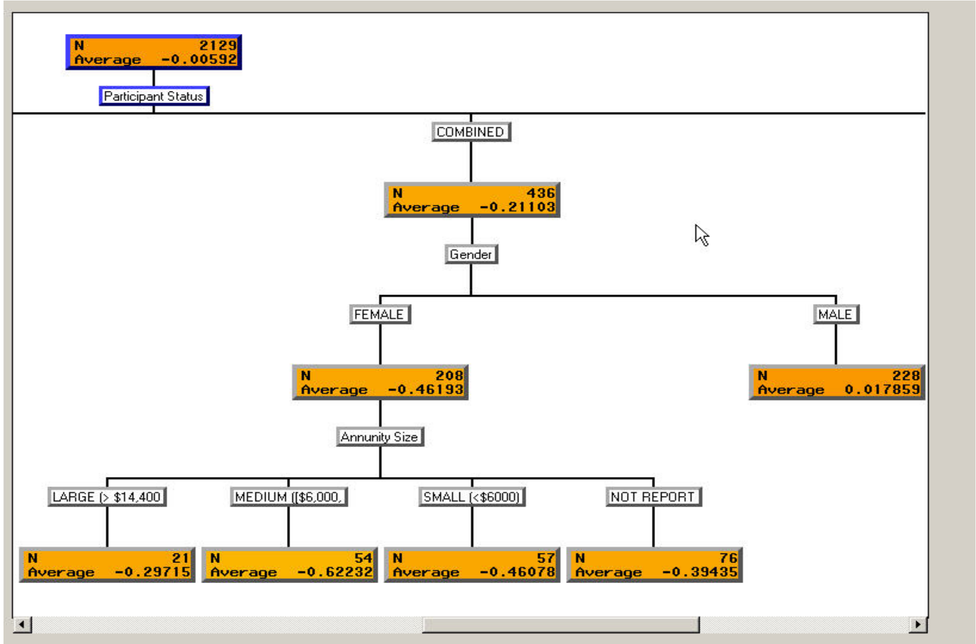
By applying the decision tree method in their old age mortality study, Guo and Wang (2001) identified some of the most important risk factors in driving the advanced age mortality. In their study with the Society of Actuaries' (SOA, 2001) data, ranking of the importance of mortality factors for older age mortality is determined as shown in Table 1. The interactions of these factors are also captured.

Within each segment by the most important factor, the decision tree can be applied to identify the relatively important risk drivers and their interactions with the segment before the construction of the predictive model. For example, within the “combined” segment, “gender” is found to be the most important variable followed by the “annuity size” variable.

TABLE 1
The Rank of Variables’ Importance

Risk Drivers	Importance
Participation Status	1.00
Gender	0.75
Annuity size	0.43
Pay Type	0.21
Union	0.18
Collar	0.00

FIGURE 1
Tree Analysis



Based on the analysis, the mortality distribution for this segment is determined by “age,” “gender,” “annuity size” and their interactions. Notice that although “annuity size” seems not important for male, the model for this segment should include both “annuity size” and “gender.” Since the “annuity size” is not important for male, this is an indication for the interaction between “gender” and “annuity size.” Both “pay type” and “union” do not appear on any part of the tree. This implies that they are not significant when studying the log odds ration between categories.

Interactions among the risk drivers are also identified with decision tree methods. Their study revealed that the male retirees’ mortality model and the female retirees’ mortality model depend on different variables. Based on the analysis, the collar and pay type are two irrelevant factors for male retiree mortality rate and should not be included in the male mortality model: Male retiree mortality depends on age, annuity size, union and their interactions. On the other hand, for the female retiree mortality model, both the annuity size and union status are insignificant risk factors based on the current database. In addition, the study showed the female demography changed in the past three decades; variables such as annuity size and union will play a more important role in determining the female mortality. The gender factors play a much-reduced role in determining beneficiaries’ mortality models when additional risk factors are considered due to the interactions.

Variable selection is an essential part of effective predictive modeling. As shown here, the decision tree method is a very effective technique for identifying and selecting the most significant predicting variables to be included in the model. The next section introduces the generalized linear model and applies it to develop a mortality risk score for predicting mortality risk for seniors.

3. Predictive Modeling: GLM

Generalized linear models (GLMs) extend linear regression models to accommodate both non-normal response distributions and transformation of linearity. GLMs include a wide range of models with linear models as a special case.

3.1 GLM Framework

A formal definition is provided as follows.

GLM Definition. A regression data set containing responses y_i and covariates x_i is said to follow a generalized linear model (GLM) if

- The responses $\{y_i\}$ are independently observed for fixed values of covariates $(x_{i1}, x_{i2}, ..., x_{ip})$, and the covariate variables may only influence the distribution of the response y_i through a single linear function

$$\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} . \tag{3.1}$$

- The mean of the response $\mu_i = E(y_i)$ is linked to the linear predictor η_i by a smooth invertible link function

$$h(\mu_i) = \eta_i , \tag{3.2}$$

while its inverse function $g(t) = h^{-1}(t)$ is called the inverse link function.

- The distribution of the response y_i is from one of the exponential family with density of form

$$f(y_i \mid \beta, \phi) = \exp\left[\frac{A_i}{\phi} \left\{ y_i \vartheta_i - \gamma(\theta_i) \right\} + \tau(y_i, \frac{\phi}{A_i}) \right] , \tag{3.3}$$

where

ϕ is a scale parameter called dispersion parameter,

A_i is a known constant, and

$\theta(\cdot) = \theta(\eta_i)$ is a function of linear predictor η_i .

The exponential family is a broader class of distributions including Normal, Poisson, inverse Gaussian, gamma, binomial and exponential distributions.

Notice that GLM is fully determined by the choices of the link function h and the form of response distribution (i.e., the form of the function γ).

One can interpret that the slope β_k as the expected amount increases (or changes) in $h(E(y))$ with a unit increase (or change) in the k^{th} covariate.

Remark: It's easy to show that $E(y_i) = \mu_i = \gamma'(\theta_i)$.

The variance is $Var(y_i) = \frac{\phi}{w_i} V(\theta_i)$, where

$V(\theta_i) = \gamma''(\theta_i)$ is called the variance function with ϕ scales the variance while A_i is a constant that assigns a weight, or credibility, to the observation i.

Distribution. A number of familiar distributions in the exponential family are: the Normal, Poisson, binomial, gamma, and inverse Gaussian:

- Normal Distribution: we can write $\phi = \sigma^2$ and $V(\theta_i) = 1$.

So, $\theta_i = \mu_i = \eta_i$ and $\gamma(\theta_i) = \frac{\theta_i^2}{2}$.

- Poisson distribution: we can take $\phi = 1$ and $V(\theta_i) = \theta_i$.

So, $\theta_i = \eta_i$ and $\gamma(\theta_i) = e^{\theta_i}$.

- Binomial: Distribution: we can take $\phi = 1$ and $V(\theta_i) = \theta_i(1-\theta_i)$.

So, $\theta_i = \eta_i$ and $\gamma(\theta_i) = \log\left(\begin{pmatrix} n_i \\ n_i y \end{pmatrix}\right)$.

- Inverse Gaussian: $V(\theta_i) = \theta_i^3$.

- Gamma: Distribution: we can take $\phi = \frac{1}{\alpha}$ and $V(\theta_i) = \theta_i^{\,2}$.

So, $\theta_i = \eta_i$ and $\gamma(\theta_i) = \log(\theta_i)$.

Details on the probability functions and their moments can be found in Appendix 5, Bowers, etc., 1997.

Link Functions. Some commonly used link functions are listed in the following:

- Logit

$$h(t) = \log(\frac{t}{1-t})$$

- Probit

$$h(t) = \Phi^{-1}(t) \text{ where } \Phi\ (t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{\frac{z^2}{2}} dz \,.$$

- Log-link

$$h(t) = \log(t) \tag{3.4}$$

- Square root

$$h(t) = \sqrt{t}$$

- Inverse

$$h(t) = \frac{1}{t}$$

Remark: Traditional linear regression requires that y_i be additive in the covariates. GLM only requires that some transformation of y_i, written as $g(y_i)$ be additive in the covariates.

The combination of the response distribution and the link function is called the family of GLM.

See Table 2 for the typical GLM models used in actuarial applications.

TABLE 2
GLM Models Used in Actuarial Science

y_i	Number of Deaths	Average Death Benefit	Probability
Link Function $h(x)$	$\ln(x)$	$\ln(x)$	$\ln(\frac{x}{1-x})$
Error	Poisson	Gamma	Binomial
Scale Parameter ϕ	1	1	1
Variance Function $V(x)$	x	x^2	$x(1-x)$
Prior Weights A_i	1	No. of claims	1

McCullagh and Nelder (1989) provide more detailed discussion on GLMs.

3.2 GLM Example

To illustrate how GLM works, consider a	Higher Income	Lower Income
Male	50	80
Female	20	40

The target variable (value to be predicted), Y, is the average number of death. The two risk drivers, income level and gender, each have two levels.
The classical linear model describe Y as a linear combination of four variables (male (X_1), female (X_2), lower income (X_3) and higher income (X_4), plus a Normal error random variable ε with mean zero and variance:

$$\eta = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \varepsilon \qquad (3.5)$$

where $\varepsilon \sim N(0, \sigma^2)$.

Equation (3.5), however, is not uniquely defined. To make it well defined for solving parameters β_i, consider instead the following model:

$$\eta = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon \tag{3.6}$$

Equation (3.6) assumes the number of death is an average for male (β_1) and an average for female (β_2), with the effect of being at lower income level having additional additive effects (β_3); notice that β_3 is the same regardless of gender. The classical linear model solution is derived by minimizing the sum of squared errors (SSE):

$$SSE = \sum \varepsilon_i^2 \;= (80 - \beta_1 - \beta_3)^2 + (50 - \beta_1)^2 + (40 - \beta_2 - \beta_3)^2 + (20 - \beta_2)^2$$

With the solution:

β_1=52.5, β_2=17.5, and β_3=25

and the number of death is predicted as:

	Higher Income	Lower Income
Male	52.5	77.5
Female	17.5	42.5

Using Poisson distribution for the error and $\ln(x)$ for the Link Function in GLM (Equation (3.4)), the predictive model yields:

$$\eta = E[Y] = g^{-1}(X\beta) = \begin{bmatrix} e^{\beta_1+\beta_3} \\ e^{\beta_1} \\ e^{\beta_2+\beta_3} \\ e^{\beta_2} \end{bmatrix} \tag{3.7}$$

For the error term, the Poisson distribution has the density function:

$$f(y_i \mid \mu) = {}^{e^{-\mu}\mu^{y}}\!\big/_{y!} \,. \tag{3.8}$$

Its log-likelihood function is:

$$\sum \ln f(y_i \mid \mu) = \sum(-\mu_i + y_i \ln \mu_i - \ln(y_i!) \,. \tag{3.9}$$

With the log-link function $\mu_i = e^{\sum_j X_{ij}\beta_j}$, GLM maximizes the following function:

$$-e^{(\beta_1+\beta_3)} + 80(\beta_1 + \beta_3) - ee^{\beta_1} + 50\beta_1 - e^{(\beta_2+\beta_3)} + 40(\beta_2 + \beta_3) - e^{\beta_2} + 20\beta_2$$

and the solution is:

β_1=3.8690, β_2=3.0958, and β_3=0.5390.

The predicted values are:

	Higher Income	Lower Income
Male	47.89	82.11
Female	22.11	37.89

In this simple exercise, GLM provided greater mortality risk differentiation for male seniors at different income levels than traditional linear regression. The GLM model also reveals that income level is a less significant risk drive for female seniors.

3.3 GLM Application

Next, we use GLM to predict the mortality risk for the advanced age population.

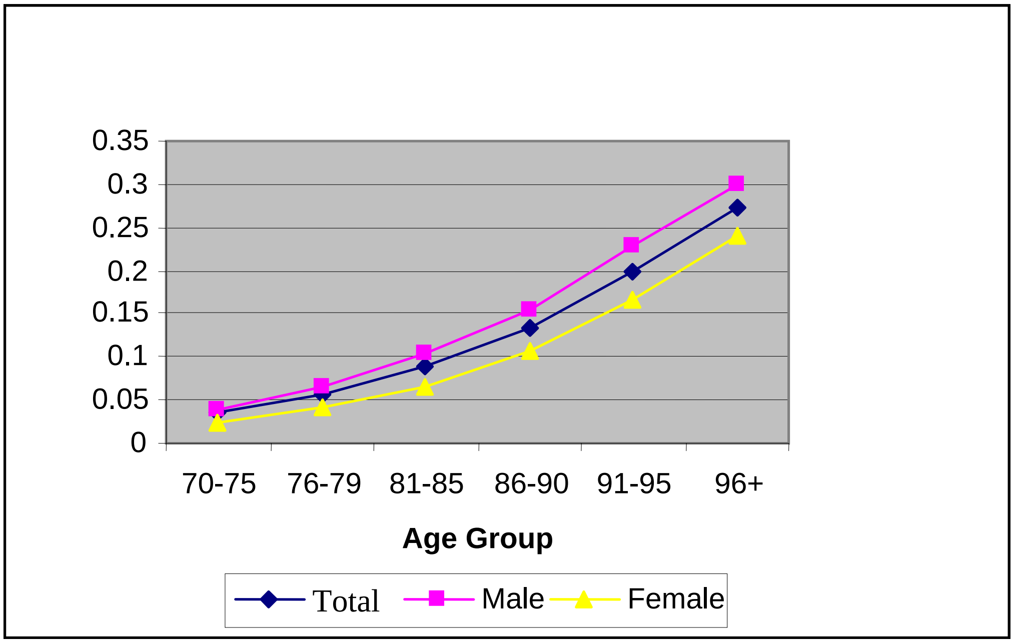
As with any data mining process, data understanding and the data preparation stages are among the most important steps. Preparing the dataset that contains the most information available for predicting variables is one of the key steps. Predictive modeling practice almost always involves merging data from different sources (policy data, underwriting data, external data, etc.). For demonstration purposes, we use SOA RP-2000 Mortality Tables for all the lives above age 70 as the demo data.

The risk factors in the dataset include age, gender, occupation information, financial well-being (income level measured by annuity size), disability status and union status, among the others listed in Table 1 in Section 2.

Age and gender have long been used to model the mortality for all populations. The impacts of these risk drivers for seniors are in the demo data as well, as shown in Figure 2.

FIGURE 2

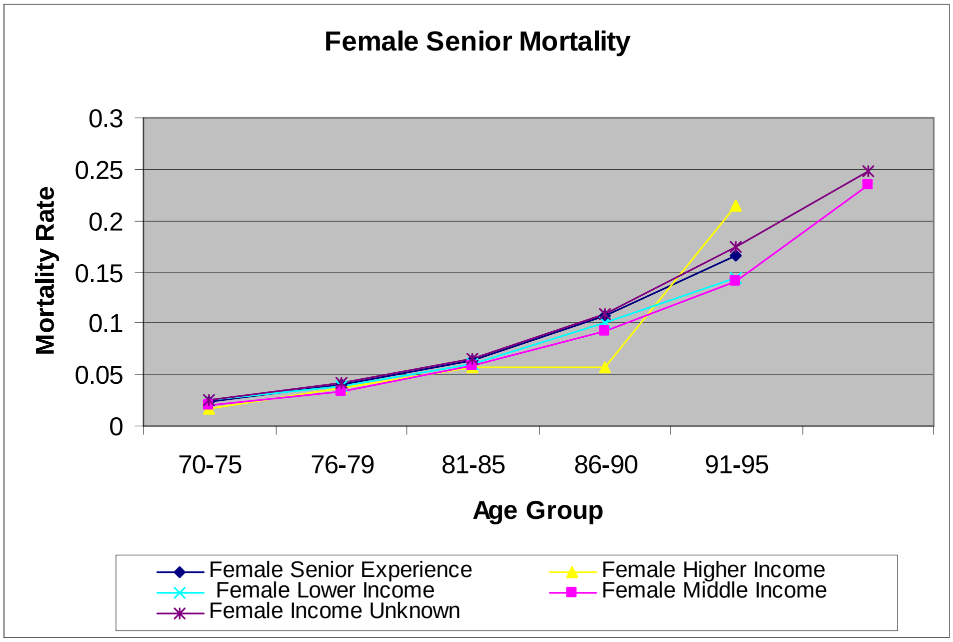
Mortality Experience for Advanced Ages



With predictive modeling, we examined all the available information in the dataset to select the key drivers for predicting advanced age mortality. Interactions and correlations of the risk drivers are also captured. For example, we consider the gender, age and income levels and their impact on the mortality distribution. As shown in Figure 3, income level is not a significant driver of mortality for the female senior population except for the very advanced age group (age 86 or older). The fact that the age 86-90 group of female seniors has a significantly lower mortality rate if financially sounded implies that financial wealth plays a very important role in older female seniors, but not much for the younger seniors.

FIGURE 3

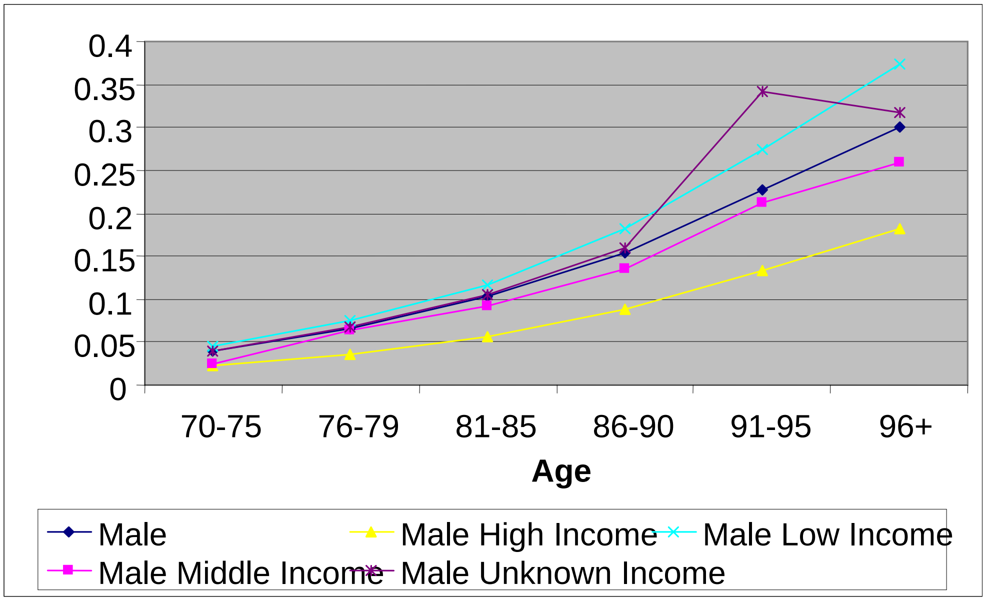
Mortality Experience for Female Seniors



On the other hand, income level is a very important driver in differentiating mortality risks in the male senior population across all age groups, as shown in Figure 4.

FIGURE 4

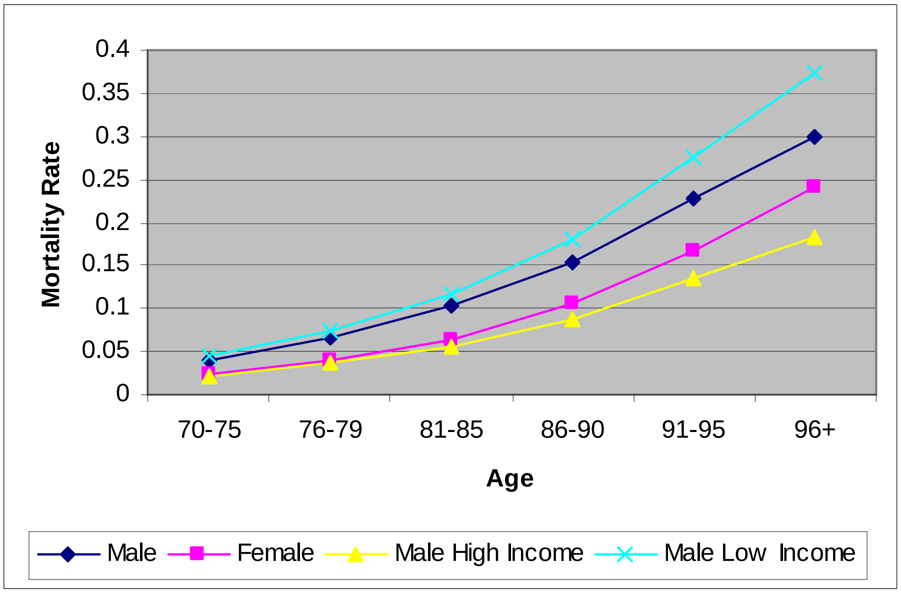
Mortality Experience for Male Seniors



To find out how much the income level affects the male senior mortality, we compare the male senior mortality experience with different income levels with the female senior mortality. Figure 5 shows that the male senior with high incomes fares better than the female senior population.

FIGURE 5

Mortality Drivers Interactions



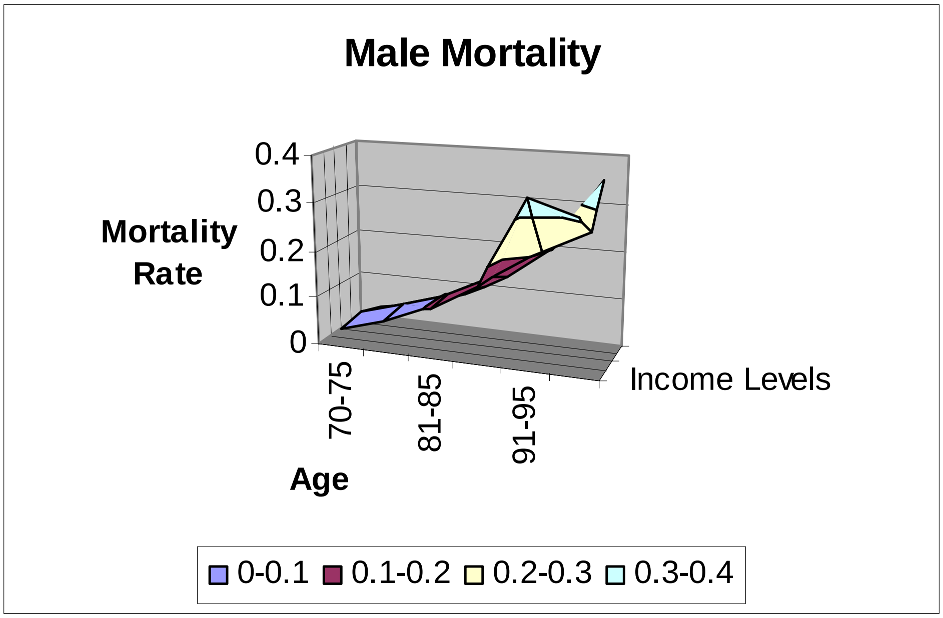
The finding suggests that the interaction of the age, the income levels and the gender should be captured in predicting senior mortality.

Using SAS, the GLM method as described in Equation (3.1)-(3.3), with Poisson distribution and Log-link function, is applied to 40091 observations.

GLM is applied to predict male mortality rate using age, income level and their interactions, shown in Figure 6.

FIGURE 6

GLM for Male Senior Mortality Prediction.



Our analysis also reflects the mortality improvement over the years. Figure 7 shows that not only the senior population mortality has improved over the years; the male senior mortality improvement has been more significant than female seniors' improvement. Figure 8 displays the mortality experience improvement over various age groups.

FIGURE 7

Senior Mortality Experience

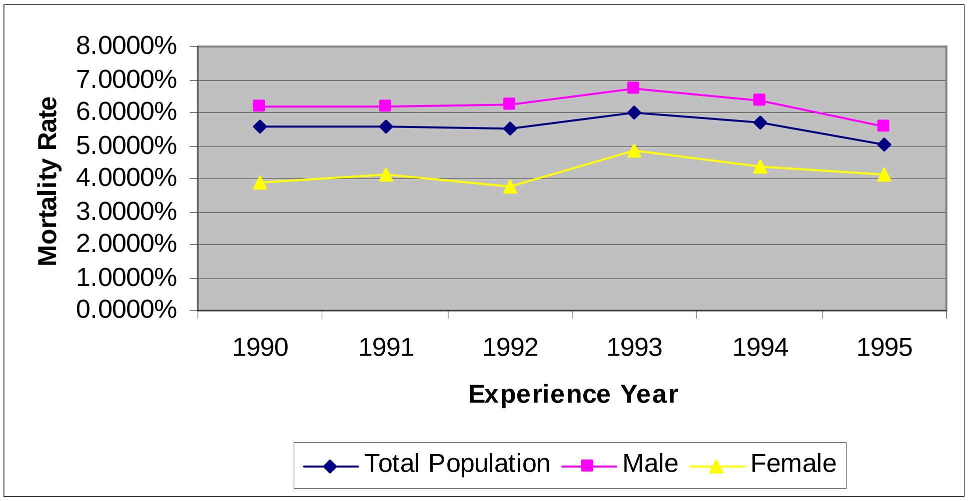
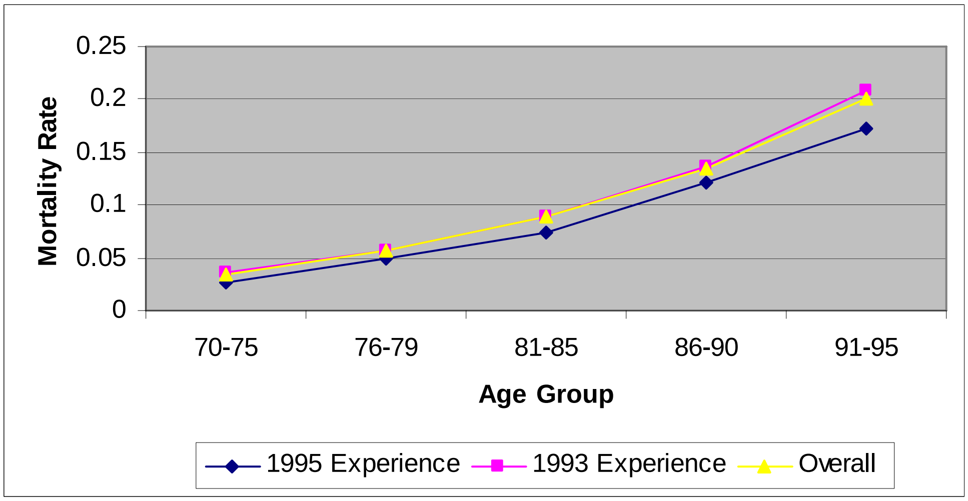


FIGURE 8

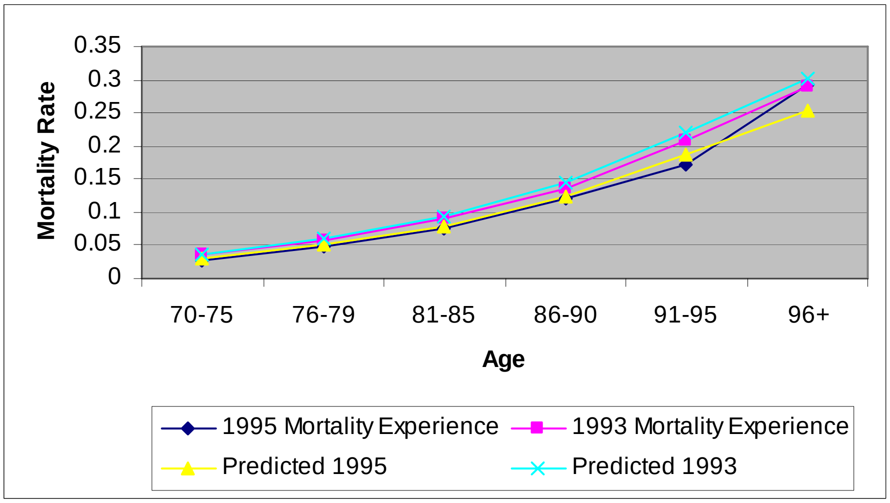
Mortality Experience for Senior Population.



GLM models are developed using experience year as a driving factor to capture the mortality improvement, shown in Figure 9.

FIGURE 9

Senior Population Mortality



Overall, GLM provides a reliable predictive model for the senior mortality. The validation population has 5.77 percent as the average mortality rate. Model validation results are shown in Figure 10, and the comparison is listed in Table 3.

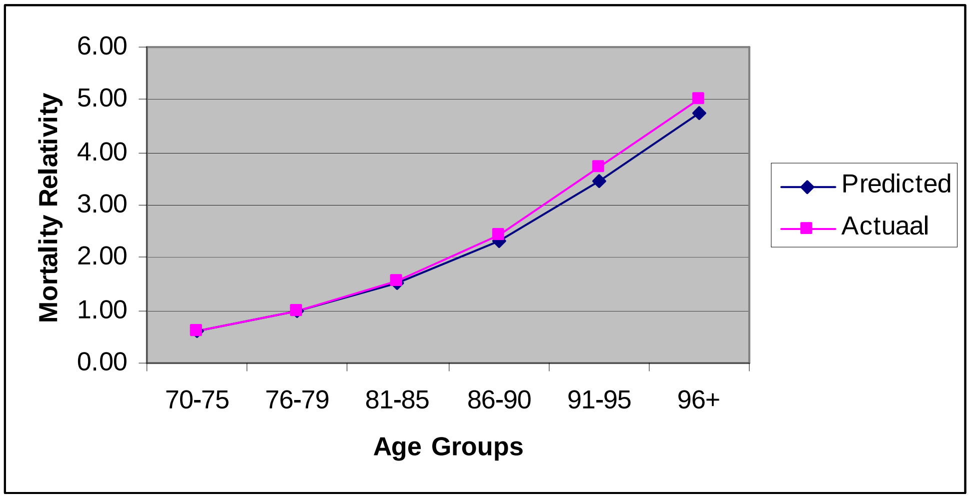
TABLE 3

Relative Mortality Risk Comparison

Age Group	Predicted Risk Relativity	Experience Risk Relativity
70-75	60.78%	60.67%
76-79	100.00%	99.01%
81-85	157.24%	152.78%
86-90	244.00%	232.42%
91-95	370.32%	346.48%
96+	502.54%	474.65%
Mean	5.77%	5.77%

FIGURE 10

GLM Predicted Mortality Relativity



4. Mortality Risk Score

Finally, a preliminary mortality risk score to predict the mortality risk for the senior population using multiple mortality drivers in the demo dataset is presented in this section.

Using the decision tree method discussed in Section 2, key risk drives are selected and the GLM method (Equations (3.4), (3.7)-(3.9)) is applied to 40094 data points.

Table 4 displays the relative weight for each individual risk drive, which reflects its impact on the mortality risk within the multiplicative model. The mortality risk score for each individual is calculated by multiplying the assigned weights for each risk factor and the average risk.

TABLE 4

Senior Mortality Score

Mortality Risk Factors	**Risk Class**	**Weight**
Age	70-75	0.121
	76-79	0.199
	81-85	0.313
	86-90	0.486
	91-95	0.737
	96+	1.000
Gender	Female	0.619
	Male	1.000
Income Level	High	0.745
	Middle	0.969
	Lower	1.024
	Unknown	1.000
Disabled	Y	1.560
	N	1.000
Occupation	High Risk	1.135
	Non Professional	1.107
	Professional	1.000
Union	C	1.013
	N	0.866
	U	1.000

For example, a 72-year-old non-disabled female professional, with high income level and non-union, belongs to the lowest mortality risk group among all the seniors. The lowest mortality risk group has the relative mortality risk score 0.0483 (=0.121*0.619*0.745*1.000*1.000*1.000*0.866). It implied that the said senior's mortality risk is about 5 percent of the average mortality rate (0.0577) for the seniors in the demo dataset. The highest mortality risk group has the relative mortality risk score

1.9799 (=1.000*1.000*1.024*1.560*1.078*1.135*1.013), which is 198 percent of the average senior's mortality risk (0.0577) in the demo dataset. The mortality risks for the different senior groups range from 0.00279 (lowest) to 0.11424 (highest).

The mortality risk score derived in this study is a multifactor predictive model that effectively separates the good risk from the bad risk. It can also be used in pricing and valuation of life insurance products as well as annuity products. Another significant application of the mortality risk score, as used in the P&C insurance industry, is in underwriting to gain significant competitive advantages for insurers.

5. Summary

This paper introduces the predictive modeling method to investigate multiple risk drivers and their impacts on the advanced age mortality. Predictive modeling has significantly increased the economic values for P&C insurers as well as the health insurance industry.

This paper presents two most useful predictive modeling methods—decision trees for identifying leading risk drivers and GLM for deriving the mortality risk score for the advanced age population. The risk score for the senior mortality not only helps us to understand how the mortality risk factors and their interactions impact the senior mortality, it also helps insurers in gaining a competitive edge in life insurance and annuity products pricing, valuation and enterprise risk management.

As with any data mining process, data understanding and the data preparation stages are among the most important steps. Preparing the dataset that contains the most information available for predicting variables is one of the key steps. In predictive modeling practice, it almost always involves merging data from different sources (policy data, underwriting data, external data, etc.).

In the follow-up study, the predictive models are developed using an expanded database. The demo data is derived by appending health-related information from the Surveillance, Epidemiology and End Results (SEER) to the demo dataset used in this study. By including more drivers (such as cancer data and geographic data) for mortality risk predicting, the mortality risk scores can be used to model the senior mortality more effectively. The techniques and the process of merging and combining the dataset used for the predictive modeling will be presented in the forthcoming paper.

Among the challenges facing actuaries adopting predictive modeling techniques is the selection of new tools, such as the right statistical package. Predictive modeling improves accuracy, but it also brings the need for training, the requirement of complicated explanations to customers and the expansion of data needs.

References

Bowers, N., Gerber, H., Hickman, J., Jones, D., and Nesbitt, C. 1997. *Actuarial Mathematics.* Schaumburg, IL: Society of Actuaries.

Breiman, L., Friedman, J., Olsen, R., and Stone, C. 1984. *Classification and Regression Trees.* New York: Chapman & Hall.

Cheeseman, P., Kelly, J., Self, M., Stutz, J., Taylor, W., and Freeman, D. 1988. "Auto Class: A Bayesian Classification System." *5th Int'l Conf. on Machine Learning*, June, Morgan Kaufman.

Gavrilov, L.A., and Gavrilova, N.S. 1991. *The Biology of Life Span: A Quantitative Approach.* Chur, Switzerland and New York: Harwood Academic Publishers.

Guo, L. and Wang, M. 2001. "Data Mining Techniques for Mortality at Advanced Age," Living to 100 and Beyond, Society of Actuaries.

Hastie, T., and Tibshirani, R. 1990. *Generalized Additive Models*. London: Chapman and Hall.

Heligman, L., and Pollard, J.H. 1980. "The Age Pattern of Mortality." Journal of the Institute of Actuaries 107: 49–75.

Kass, G.V. 1980. "An Exploratory Technique for Investigating Large Quantities of Categorical Data," Applied Statistics 29: 119-127.

McCullagh, P. and Nelder, J.A. 1989. *Generalized Linear Models*, 2nd edition, Chapman & Hall/CRC.

Pollard, J., and Streathfield, K. 1979. *Factors Affecting Mortality and the Length of Life*, Number 197. North Ryde, Australia: Macquarie University, School of Economic and Financial Studies.

Rose, M. 1991. *Evolutionary Biology of Aging.* New York: Oxford University Press.

Society of Actuaries. 2003. "THE RP-2000 MORTALITY TABLES," http://www.soa.org/research/rp2000.html

Tuljapurkar, S. 1998. "Forecasting Mortality Chance Questions and Assumptions." North American Actuarial Journal 2(4): 127-135.

Tuljapurkar, S., and Boe, C. 1998. "Mortality Change and Forecasting: How Much and How Little Do We Know?" North American Actuarial Journal 2(4): 13-48.

Wachter, K., and Finch, C. (eds.). 1997. *Biodemography of Aging.* Washington, D.C.: National Academy Press.