SOCIETY OF ACTUARIES Forecasting and Futurism Section

ISSUE 7 | JULY 2013

Forecasting & Futurism

NEWSLETTER

If I only had a brain

By Dave Snell

Actuaries
Risk is Opportunity.®

ISSUE 7 | JULY 2013

Forecasting & Futurism

NEWSLETTER

Published by the Forecasting and Futurism Section Council of the Society of Actuaries

This newsletter is free to section members. Current issues are available on the SOA website *(www.soa.org)*.

To join the section, SOA members and non-members can locate a membership form on the Forecasting and Futurism Section Web page at *www.soa.org/forecasting-futurism.*

“If I only had a brain”

By Dave Snell

This often-quoted line from the 1939 movie, The Wizard of Oz, is part of a famous scene where a Scarecrow (played by Ray Bolger) sings to Dorothy (the young Judy Garland) about what he would do if he only had a brain. It is long past copyright protection so feel free to view it on the Internet (e.g., *http://www.youtube.com/watch?v=nauLgZISozs&noredirect=1).*

This issue is packed with brainy articles to get those neurons of yours firing away with ideas for exciting new actuarial applications of new techniques and tools that are now available.

We begin this set of offerings with an introduction to a neural networking technique called Naïve Bayesian Networks. Jeff Heaton has written an easy-to-follow example, “Bayesian Networks for Predictive Modeling,” describing how to use simple conditional probability (remember our old pal Bayes, from your introductory actuarial exams?) enhanced by computer programs to infer an outcome from a collection of past events. He also included an Excel workbook as a model to enable you to substitute your own events and make your own inferences from them. After you enjoy his article, read about his books on neural networking and his free and open source engine, ENCOG, at his website *www.Heaton-Research.com* which is currently getting about 100,000 hits per month from the AI (artificial intelligence) community.

Predictive Modeling (PM) has become the rage lately, as Big Data seems to have saturated the media. Richard Xu gives a helpful overview of various PM techniques and their relative strengths and weaknesses. Quoting from Richard’s article, you will learn about techniques for “maximizing the value of data to improve business processes and customer experiences.” If you ever wondered whether to use linear regression or the generalized linear model (GLM), or decision trees or a CART (classification and regression tree), or perhaps a clustering model versus a support vector machine (SVM), see his article **Predictive Modeling** to better understand what makes sense for your applications and what indicators may suggest a different approach.

One of those approaches is Hidden Markov Models. Brian Grossmiller and Doug Norris teamed up to write “Hidden Markov Models and You,” an article about how to utilize these less obvious (hidden?) extensions to the classic Markov Chain models you saw in study notes. Brian and Doug start with the basics of what is meant by a Hidden Markov Model (HMM) and then they put an HMM to work on a claims activity example. In the HMM it is important to figure out what state you are in currently, and what state you will be in after your transition matrix. If that state stuff doesn’t make sense to you yet, then Toto, “We are not in Kansas anymore” *(http://www.youtube.com/watch?v=vQLNS3HWfCM)*. Read the article and these two wizards will explain what is behind the curtain in HMMs.

Prior to our next issue (December), we’ll lose Clark Ramsey as our Chairperson, as he rolls off his three-year term on the council, and we’ll miss his guidance a lot. His chairperson article, “Dark Side of the Moon,” is another delightful combination of science, actuarial career advice and Pink Floyd. He describes the issue of navigating through too much or

CONTINUED ON **PAGE 4**

too little data and the resultant challenges facing an actuary to maintain the integrity of pricing and reserving models; and somehow, Clark makes them all fit together. I appreciate his eclectic interests and I hope our next chairperson will be able to carry on this tune.

Scott McInturff continues the PM lesson with his cogent review of Nate Silver's best seller, *The Signal and the Noise: Why So Many Predictions Fail—But Some Don't.* Nate has had an amazing record of correctly predicting the past two presidential election outcomes on a state-by-state basis several months before the actual elections. He also gained fame and some fortune with his accurate baseball predictions. Scott has a knack for writing an engaging review; and after reading this one it seems nearly unthinkable to me that you won't want to rush out and get a copy of the book to read.

A question that naturally comes to mind after reading about these techniques is how actuaries have done in the PM arena. What is our track record? Ben Wolzenski gives us a recap of that with his summary of an actuarial set of predictions made more than a decade ago. It was quoted back then in the *Wall Street Journal*. In "Delphi Study 2000—Predictions for 2010 and 2050," Ben describes what we thought then, and how well we did (or didn't) predict the future of health care expenses, cause-specific death rates, employment and where actuaries will reside. Who could have imagined that the mean estimate of total life insurance in force in the United States a decade later would be spot on at $18.43 trillion! OK, that's after rounding to the nearest 0.01 trillion dollars; but it is still an amazing feat.

That Delphi study was reviewed and judged after a decade; but you have the opportunity to gain fame and fortune (OK, an iPad) this year! Simply enter our genetic algorithm contest and submit the entry that best solves a problem of an actuarial nature. This contribution to the profession will be rewarded (or at least announced, if the winner is not present) at the 2013 SOA Annual Meeting in San Diego. Read the contest announcement by Alberto Abalo, "Forecasting & Futurism 3rd Annual iPad Contest—Build a Genetic Algorithm" and enter to win the appreciation and adoration of your peers (plus, did I mention you get an iPad?). Brian Grossmiller and I will try to incorporate some ideas from the entries in our workshops on genetic algorithms at that same meeting.

Delphi studies are now used by many actuaries; but several other professions are using similar and supplementary techniques and there is no shortage of predictions about the future. Jon Deuchler summarizes an excellent report by the National Intelligence Council (NIC): "Global Trends 2030: Alternative Worlds." The NIC is a coalition of 17 agencies and organizations within the Executive Branch of the U.S. government and is a kind of clearinghouse of gathered intelligence. Jon distills more than 130 pages (see the full report at *www.dni.gov/nic/globaltrends*) of predictions and scenarios to an amusing and irascible summary that compelled me to at least skim the full report.

Ah, but lest we forget, there is more to the art and science of PM than just computers and numbers! Those carbon unit humans don't always follow the rules we cast in silicon. David Wheeler, a recent graduate who majored in Behavioral Economics, reminds us of that in "Behavioral Economics: Implications for Actuarial Science and Enterprise Risk Management," in which he leads us through some exercises with surprising results. As David writes, "Behavioral Economics is emerging as the leading decision science for economics, psychology, sociology, biology and neuroscience." We make a lot of best estimate assumptions in our actuarial models; and we need to do this with eyes wide open to the ways that humans differ from logical sequential cyborgs.

Some actuaries are willing to open their eyes and their minds to nonconventional thought patterns; and nowhere is that more evident than in the Actuarial Speculative Fiction contest. This is an annual event that F&F (Forecasting & Futurism) cosponsors (with the Technology and the Actuary of the Future sections) and this year's entries were even more creative than usual. I was privileged to be one of the judges and I have to admit that we had a tough time voting on just one winner. Mike Lindstrom describes the winning entry in "The Weight of Certainty—Selected Stories of Steve Mathys—The 10th Speculative Fiction Contest Forecasting and Futurism Section Prize Winner." Mike describes his interview with Steve Mathys, who has been writing stories for several years and first entered our speculative fiction contest in 2003. His winning entry, "Calibration," "describes a day in the life of Stuart, a businessman in Capitol City. Stuart subscribes to the services of a company called Precision Dynamics, which provides a list of custom probabilities in regards to a series of today's personal events submitted by the customer the night before." It's a thought-provoking read.

It is very helpful, in fact necessary at times, to step out of our immediate frame of reference and view our modeling efforts from a different perspective. One of the perspectives we sometimes don't appreciate enough until a calamity or an audit (or perhaps both) occurs is that of controls. Spreadsheets are ubiquitous in actuarial work and Darrick Fulton, a professional auditor, gives us some horror stories of what can happen as a result of errors in a spreadsheet; and he provides some expert guidance on ways to avoid these problems. His article, "Spreadsheet Controls … How to Prevent a Fire" has an amusing lead into the idea of what to do when your car … or your spreadsheet catches fire; and how simple and logical maintenance can avoid these problems.

Finally, I started out with the desire of the Scarecrow in The *Wizard of Oz* to have a brain. I had the thrill recently of holding a human brain in my hand; and I want to share my excitement about that in "I Held a Human Brain!" This article is about some alternative sources of learning available for free or at low cost if you wish to round out your F&F education about machine learning, artificial neural networks, genetic algorithms and the millennia of biological advances we draw from when we use these techniques.

I am going to close this issue's introduction with a shameless advertisement that seems in order here. Several people have been asking how to get these F&F newsletters. I tell them that after a month or so, each issue is posted to the SOA website; but if they want to see it on a timely basis, all they have to do is join our section as either a member (SOA members) or an affiliate member (anyone else). If any of your friends want to see these hot articles before the herd, they can do so for a mere $25 per year—and additionally be eligible for our iPad contests and lots of other member benefits. Quoting the television ad on drugs, "a brain is a terrible thing to waste." Utilize your opportunity for an upgrade today at F&F. ▼

Dave Snell

Dave Snell, ASA, MAAA, is technology evangelist at RGA Reinsurance Company in Chesterfield, MO. He can be reached at *dsnell@rgare.com*

Bayesian Networks for Predictive Modeling

by Jeff Heaton

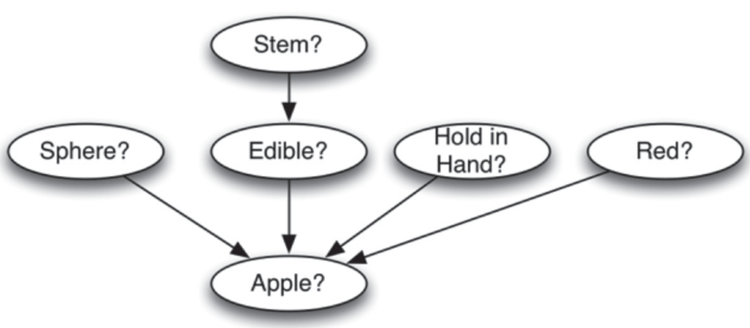
Bayesian Networks represent a convergence of Artificial Intelligence (AI) and Statistics. These networks function by creating a probabilistic model that can be used to query possible outcomes from input data. Bayesian Networks were created by Dr. Judea Pearl in 1983. Pearl is well known for championing the probabilistic movement in the field of Artificial Intelligence. Bayesian Networks can be used for predictive modeling, pattern recognition, classification and regression.

INTRODUCING BAYESIAN NETWORKS

Bayesian Networks model probabilistic relations among random variables. A random variable is a term from Statistics that means the items involved vary in some random, or unexplained manner. A Bayesian Network assigns probability factors to various results based upon an analysis of a set of input data. Like many other Machine Learning algorithms a Bayesian Network is taught using training data. Once trained, a Bayesian Network can be queried to make predictions about new data that was not represented by the training set. To see how a Bayesian Network functions consider if a trained Bayesian Network were presented with the following input data describing an object:

"I am holding an object that is red, editable, essentially spherical, fits in the palm of my hand and has a stem," you tell the network.

Figure 1: A Bayesian Network



The Bayesian network would report the probability that you are holding an apple. This probability is not 100 percent, since a cherry also fits this description. Figure 1 shows how this Bayesian network might be represented.

The above diagram indicates some relationships between the random variables. For example, the variable edible is dependent on stem. Many of the objects in our world that are edible have stems, so this makes sense. Additionally, the probability of the object being an apple is dependent on sphere, edible, handheld and red. The arrows in Figure 1 represent probabilities. The above network implies the following total probability.

```
Pr(Sphere) Pr(Edible|Stem) Pr(Hand)
Pr(Red) Pr(Apple| Sphere, Edible, Hand,
Red)
```

Because all of these variables are discrete, their probabilities can be represented by truth tables. Discrete random variables have a finite number of values. For this example each random variable has only two values—either true or false.

It is very common to represent random variables as real numbers. For numeric values there are two common methods for representation. First, real numbers can be banded into discrete ranges. Second, real numbers can be represented continuously using a probability distribution.

WHY USE A BAYESIAN NETWORK?

A Bayesian Network fills a role very similar to other Machine Learning algorithms such as an Artificial Neural Network (ANN), Decision Tree, or Support Vector Machine (SVM). However, a Bayesian Network has several unique advantages over some of the other Machine Learning algorithms.

First, a Bayesian Network handles missing values very well. Consider the previous example. You might only tell the Bayesian network that the object was red and had a stem. In this case the probability that the object is an apple would be lower. Not all Machine Learning algorithms handle missing data so well. Many Machine Learning algorithms require that incomplete data be eliminated or extrapolated.

Second, a Bayesian Network can be queried. An SVM or ANN is typically trained to predict a specific outcome given specific input. Following on the same example, you might tell the Bayesian network that you are holding an apple and that it has a stem. But this time, you query to see if the object is red. The random variables do not have fixed roles of input or output.

PREDICTIVE MODELING EXAMPLE

Now I will show you an example of how Bayesian networks can be used for predictive modeling. For this example, we will attempt to predict the direction that a particular exchange rate will move. To keep this example simple we will keep the amount of data relatively small. I will use the daily highs of the big six currency pairs. These pairs are given here.

- EUR/USD
- USD/JPY
- GBP/USD
- AUS/USD
- USD/CHF
- USD/CAD

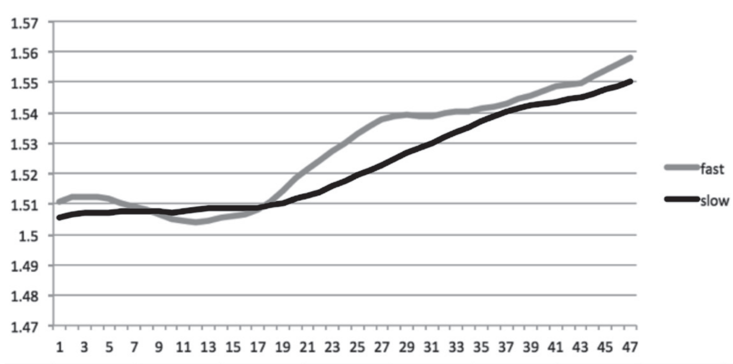
The data that I had available for these six currencies were minute bars since 2001. This is a fairly large dataset. Quite a few minutes have passed since 2001! Using SQL I decreased the granularity of the data down to just the daily highs for each of the six currencies. It is possible to deal with the individual minute bars. It is even possible to deal with the underlying ticks that made up the minute bars. Using such large amounts of data has the potential to yield better results. However, dealing with individual ticks over such a time period is a Big Data problem. Big Data often poses problems dealing with the immense size of the data set. Big Data is beyond the scope of this article, so I will stick with daily highs.

We are not going to present the data in its raw format to the Bayesian network. Like many Machine Learning problems the first step is to determine how to represent the data. Excel is a great program for preprocessing the data for the Bayesian network. There are only about 4,000 rows in my data set. Excel can easily handle this.

The goal of this Bayesian network is to use the current exchange rates of the big six exchanges and provide the probability that the EUR/USD pair will go up over the next 10 days. I would like the Bayesian network to use the rise and fall of the daily heights of all six. This considers that the various currencies may influence each other. I do not want the Bayesian network to consider the actual numbers, just the trends. Ideally the indicator of an upward trend will signal this regardless whether the exchange rate is 1.0 or 0.5.

To do this we will preprocess the data into two Simple Moving Averages (SMA). SMA is simply the average of a fixed number of days. We will make use of a fast (10-day) and slow (20-day) SMA. We are particularly interested in when these two SMAs cross. This is a well-known, and relatively

Figure 2: Fast and Slow SMA for EUR/USD



CONTINUED ON PAGE 9

simple, means of predicting trends. However, we are adding the extra complexity of using this data from six exchange rates. Figure 2 (pg. 7) shows these two SMAs.

As you can see from the chart above, the fast SMA was above the slow SMA for most of the chart. There were two crossovers.

The training data will consist of a total of 13 columns. The first column indicates the direction of the EUR/USD pair over the next day, either up or down. Each of the six currents also provides two columns. For each currency we track if the fast or slow SMA is greater. We also track if a crossover occurred on that day. You can see some of the training data in Figure 3.

Figure 3: Select Training Data for the Bayesian Network (select columns)

EURUSD Direction	EURUSD SMA	EURUSD Cross	USDCAD SMA	USDCAD Cross
Down	Slow	Cont	Slow	Cont
Down	Slow	Cont	Slow	Cont
Down	Slow	Cont	Slow	Cont
Up	Slow	Cont	Slow	Cont
Up	Fast	Cross	Slow	Cont
Up	Fast	Cont	Slow	Cont

For brevity only, the EURUSD and USDCAD exchange rates are shown in Figure 3. Likewise, only six days are shown. The training data shown above corresponds to just before the first crossover shown in Figure 2. This is all of the data provided to the Bayesian Network. No dates are provided, nor is the order of the rows. All short-range temporal position is encoded by the fact that we are encoding two averages and the crossover point.

TRAINING THE BAYESIAN NETWORK

Training is the process where training data is used to construct a Bayesian network. The Bayesian network is constructed so that the probabilities produced by the final network closely match the training data. It is very rare that you will get a Bayesian network that can exactly reproduce the training data. The success of the Bayesian network at producing results consistent with the training data is called the error rate of the Bayesian network. Training seeks to minimize this error.

Training a Bayesian network is typically divided into two distinct stages. The first stage creates the structure of the Bayesian network. The second stage creates the probabilities between the Bayesian network's nodes. The structure of the Bayesian network and the probabilities between nodes make up the entire memory of what the Bayesian network "knows."

There are several different techniques for creating the structure of the Bayesian network. Some of the most popular methods are listed here.

- Force Naïve Bayes
- Genetic Algorithm
- Hill Climbing Algorithm
- K2 Algorithm
- Simulated Annealing
- Tabu Search

The K2 algorithm is the most common, and is what I used to generate the Bayesian network for this article. Another common method is simply to force a Naïve Bayesian structure. The Naïve Bayesian structure is a very effective structure for Bayesian networks. Naive Bayes was the structure that K2 chose for this article's Bayesian network. You can see the structure of the Bayesian network in Figure 4.

In figure 4 you can see our 13 columns represented as random variables in a probability structure. A Naïve Bayes network contains only connections from leaf nodes to a central node. For Figure 4, the EURUSD next-day direction is dependent on the other six exchange rates slow and fast SMA position and crossover. There are no other dependencies. In reality there may well be such dependencies, however, the Bayesian network remains Naïve to them.

After the structure is in place, the probabilities between the variables must be determined. There are several different methods for performing this estimation.

- Bayes Model Averaging
- Multi-Nominal Bayes Model Averaging
- Simple Probability Estimation

For this article I used simple probability estimation. This allowed the Bayesian network to train to a theoretical accuracy of 62 percent. Of course this is only an example, and does not constitute investment advice. This is merely an example of how you might apply a Bayesian network for predictive modeling. This Bayesian network was modeled on the Encog Machine Learning Framework (http://www.encog.org).

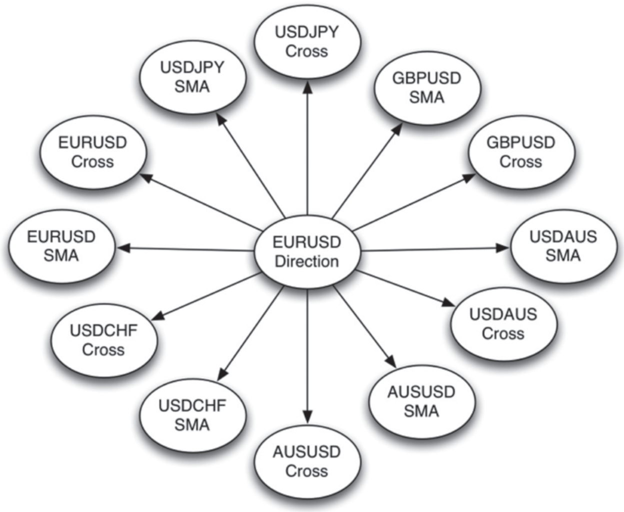
QUERYING THE BAYESIAN NETWORK

Once training is complete, the Bayesian network is ready to use. Using the Bayesian network typically involves querying the network. To query the network, you provide values for the known variables and the desired values for any output variables. You will be provided with the probability that the given input will produce the desired output.

Because we are dealing with a Naïve Bayesian network we can use the following relatively simple formula to perform the query.

$$\Pr(C|A_1, A_2 \dots A_n) = \frac{\left(\prod_{i=1}^{n} \Pr(A_i|C)\right) P(C)}{P(A_1, A_2 \dots A_n)}$$

Figure 4: A Naïve Bayes Network for Currency Predictive Modeling



The above formula is a generalized form of Bayes Theorem. It shows how to calculate a case C based on several conditions A. For the currency example the case is the direction of EURUSD, and the conditions are the other 12 variables from Figure 4.

You can see a sample of this worked out in an Excel file (www.soa.org/BayesianNetworks/). Let's assume you enter the following values.

EURUSDSMA=slow, EURUSD_CR=cross,
USDCADSMA=fast, USDCAD_CR=cont,
USDCHFSMA=fast, USDCHF_CR=cont,
USDJPYSMA=fast, USDJPY_CR=cont,
AUDUSDSMA=slow, AUDUSD_CR=cont,
GPBUSDSMA=slow, GPBUSD_CR=cross

The Bayesian Network in this case predicts the EURUSD exchange rate direction as up with a probability of 63 per-

CONTINUED ON **PAGE 10**

cent. Once again, I am not making any guarantees here. Anyone who has spent hours shoveling a driveway full of snow after a "partly cloudy" weather forecast understands that probabilities are not certainties. Your model accuracy will depend upon the validity of your basic assumptions, the type and quantity of data you use for training, and external influences.

For a non-naïve Bayesian Network there are many different ways to implement querying of a Bayesian network. One of the most simple is called enumeration. Enumeration is a brute force iteration through all values of the truth tables. This process can be streamlined by various techniques.

CONCLUSIONS

For the purposes of illustration this article used a simple dataset for predictive modeling. For real-world use you would most likely do considerably more preprocessing of the data. Additionally, we only tested how well the Bayesian network performed on its training data. It is also important to use techniques such as cross validation to ensure the Bayesian network is not simply "memorizing" the training data.

The convergence of artificial intelligence and Statistics is behind some amazing advances in technology. Bayesian networks are just one of probability inspired Machine Learning algorithms. Hidden Markov Models, Particle Filters and Bayesian networks show that probability has a very important place in the field of AI.

REFERENCES

Artificial Intelligence: A Modern Approach (3rd Edition) by Peter Norvig & Stuart Russell, Prentice Hall, December 11, 2009

Data Mining: Practical Machine Learning Tools and Techniques, Third Edition by Ian H. Witten, et al, Morgan Kaufmann, January 20, 2011

The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, and Emerged Triumphant from Two Centuries of Controversy by Sharon Bertsch McGrayne, August 6, 2012

Illustration of the K2 Algorithm for Learning Bayes Net Structures by Carolina Ruiz, *http://web.cs.wpi.edu/~cs539/s05/Projects/k2_algorithm.pdf* ▼

Jeff Heaton

Jeff Heaton is EHR informatics scientist at RGA Reinsurance Company in Chesterfield, Mo. He can be reached at *jheaton@rgare.com.*

Predictive Modeling

By Richard Xu

As predictive modeling (PM) draws more and more attention from the actuary community in life and health sections, actuaries may already read many high-level introduction articles and presentations about PM. These materials usually discuss PM background, potential benefits, general requirements, etc., so readers can understand what PM is and what applications are in insurance industry. Seldom did we see an article and discussion about technical aspects of PM, which may be pertinent to actuaries' educational background and their work. However, too many technical details and complicated statistical equations may also intimidate actuaries and turn them away. This article will focus on the mathematical side of PM so that actuaries can have a better understanding, but with a more balanced approach such that readers will get deeper understanding in mathematics terms, but not too many arduous mathematic equations.

The advancement of statistics in the past decades has provided us with abundant choices of techniques that could find their applications in actuarial science. Yet, due to the uniqueness of the insurance business and data structure, we will only find that certain tools are applicable in business, while others are hard to apply in insurance. In this article, I will focus on these techniques that have been proven to be useful in practice.

MODEL BASICS

As was pointed out in the introduction section, there are countless statistical techniques that can be utilized as PM tools in insurance applications. Generally speaking, any statistical model that relies on variables to explain variance of a target variable can potentially be used for the purpose of predicting future outcome. In the language of mathematic terms, we like to build a model as

$$\boldsymbol{y_i = f(x_{ij}, \beta_j) + \varepsilon_i} \qquad \text{(i)}$$

where y_i is called the response variable, dependent variable, or target variable. This is the variable that has been observed in experience and is to be predicted by the model. x_{ij} are called the explanatory variables covariates, input variables, or independent variables. These are variables that have been observed in historical data, and will be observable in the future for the purpose of the forecast. β_j are coefficients to be estimated in the model-building process. ε_i is the error term, which is very important for modeling, but usually not so for prediction, because in most cases we are interested in expected mean values.

TYPES OF MODELS

Linear regression and Generalized linear model

The most common and simplest model is a linear regression model. This is the bread-and-butter model that is taught in almost all colleges, and anyone with an undergraduate degree has probably had at least some exposure to it. The model essentially says the target variable is a linear combination of independent variable(s)

$$\boldsymbol{y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots}$$
$$\boldsymbol{+ \beta_n x_{in} + \varepsilon = \Sigma_j \beta_j x_{ij} + \varepsilon_i} \quad \text{(ii)}$$

To make a valid linear regression in this basic form, several assumptions are needed. A linear relationship between response and explanatory variables is obviously one, but usually this is not a problem. Either the relationship is inherently linear, or it can be well-approximated by a linear equation over short ranges. In addition, the error term ε_i must follow a normal distribution with mean value at zero and a constant variance, i.e. $\varepsilon_i \sim N(0,\sigma^2)$. Other requirements include: y_i is representative of population, observations are independent from each other, and x_{ij} is error-free, etc.

A common method of estimating β_j is the least squares method, in which β_j is chosen such that $RSS = \sum_i (\hat{y}_i - y_i)^2 = \sum_i (\Sigma_j \beta_j x_{ij} - y_i)^2$ is at its minimum, where RSS stands for Residual Sum Square, and $\hat{y}_i$ is the fitted value. There are close form solutions for β_j in matrix form. The other estimation method to find β_j is maximum likelihood in which the product of probability at all data points is at its maximum. Under the normal distribution, it can be proven in mathematics that both estimations will give the same result.

... THE LINEAR REGRESSION MODEL IS A NATURAL PART OF GLM.

Unless given a very small data set, it is not feasible to build a real model just with pen and paper. You have to rely on computing software to find β_j. The choice of statistical software is quite abundant; options include R, SAS, SPSS, MatLab, MiniTab, etc. In fact, for a very small simple application, one can even use Microsoft Excel's built-in function by selection "Data" -> "Data Analysis," although it has the limit of only 16 explanatory variables. For a large or complicated model, computing software is the only viable choice. Among the actuarial community, the two most commonly used are R and SAS. The R is free software under GNU license, while the latter one is a commercial product. R is unique, not only because it is free, but also because there is a large online community and a core statistics team to support it. You have a wide choice of educational and academic materials about R, and there will never be a shortage of statistic tools in R to build any particular model. As of now (April 2013), there are close to 4,500 packages available on top of the already abundant basic tools that come with the R system, and the number is still growing.

A linear regression is very basic, yet very powerful and efficient. You can easily find a wide range of applications in almost all industry fields. However, you can hardly find any real application in the insurance industry. The main reason is not because of the ignorance of actuaries, but the unique business model and data structure of the insurance industry in which the assumptions of linear regression model are no longer valid. For example, we know the number of claims in a certain group over a period of time is a Poisson distribution where the variance is not a constant, but equal to the mean value. In this case, a linear model can not be used to describe the process why a certain number of claims are observed. Other examples may include claim amount, which follow a Gamma distribution, or mortality rate on binomial distribution.

Luckily, the advance of statistics in the past few decades have prepared us with another model called generalized linear model (GLM). As the name indicates, this model is a natural extension of linear model. We can write the model as

$$g\big(E(y_i)\big) = g(\mu) = \eta = \Sigma_j \beta_j x_{ij} \quad or$$

$$E(y_i) = \mu = g^{-1}(\eta) = g^{-1}\big(\Sigma_j \beta_j x_{ij}\big) \quad \text{(iii)}$$

where g(…) is called the link function which links the expected mean value of target variable and the linear combination of independent variable(s).

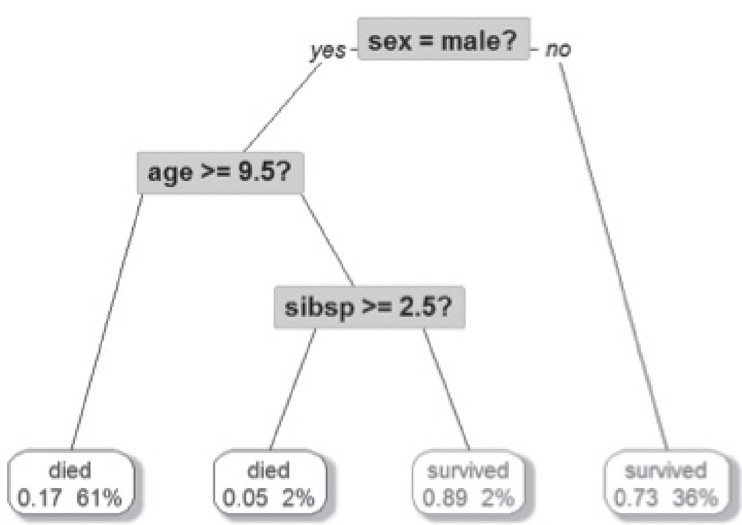
Compared to the linear model, the assumption of normal distribution is no longer needed. Instead, y_i is required to belong to the exponential family of distributions, which is broader and includes most distributions we find in insurance application, such as Poisson, binomial, Gamma, etc. The expansion of distributions also accommodates the variance structure that comes naturally with the distribution. For example, in the Gamma distribution, the variance is proportional to the square of mean. The introduction of the link function makes it possible to drop the strict linear relation between y_i and x_{ij}, resulting with a very flexible model. It is worthy to point out that logarithm could be used as a link function for various distributions. The unique feature of the logarithmic function is that the inverse function is an exponential function such that the additive linear combination in its original form now becomes multiplicative factors. This makes GLM a very powerful tool in the insurance industry as many applications traditionally have multiplicative factors to account for various parameters, such as risk class, gender, industry, locations, etc. Of course, normal distribution is also a member of exponential family, and the linear regression model is a natural part of GLM.

CONTINUED ON **PAGE 14**

Table 1 GLM: link function, variance and application

Distribution	Link Function	Variance V(μ)	Sample Application
Normal	Identity	1	General Application
Poisson	Log	μ	Claim Frequency / Counts
Binomial	Logistic	μ(1-μ)	Retention, cross-sell, UW
Gamma	Log	μ^2	Claim severity
Poisson/Gamma Compound	Log	$\mu^p, p \in (1,2)$	Pure Claim Cost & Premium
Inverse-Gaussian	Log	μ^3	Claim cost

As GLM covers most distributions that are found in insurance and includes various link functions, it is powerful and versatile, and currently is the main focus of PM in insurance. Its applications cover almost all aspects of the insurance business, such as underwriting, actuarial applications (pricing, reserves, experience study, etc.), claims administration, policy management, sales and marketing, etc. Please refer to Table 1 GLM: link function, variance and application.



Decision tree/CART

Besides GLM, another type of model that you may often hear of is an algorithm that is based on a decision tree. In its simplest form, data are split into smaller sections, called leaves, such that data in each leaf will be homogeneous to a certain degree and the variance in data can be explained by a chain of splits on a series of variables. Certain criteria are used to determine which variable to split and at which value so that the split will be optimal.

The most popular decision-tree-based model is the Classification And Regression Tree, also referred to as CART. As the name indicates, you can use this model for both regression and classification. For regression, the target variable is a continuous amount and the model is used to calculate the expected mean value. In this case, the sum of the squares error is used as a criterion to select the split point. In classification, the goal of the model is to separate data into two or more groups. There are several options to accomplish this, such as Gini measure, entropy, etc.

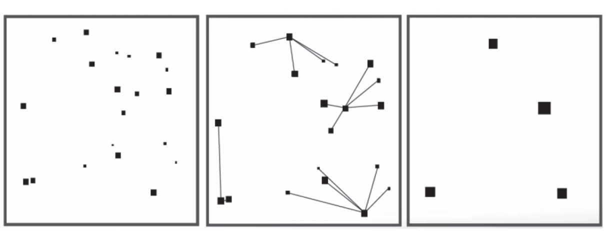
The main advantage of CART model is its intuitiveness and simplicity. When you lay out the tree diagram and present it to your audience, it is very easy to understand and discuss. For example, the Figure 1 A CART model shows a CART model to exaplain the difference of survival rates for Titanic passengers. The decimals at the bottom of each leaf are the probabilities of survival, and the percentages are fractions of the observations. Considering how split variables are chosen and at what value to split, the model itself is quite sophisticated, yet the result is intuitively simple for your audience to grasp the essence of the model without complicated math involved. Other advantages include the non-parametric nature in which you do not have to specify a distribution as assumption, and the automatic handling of possible missing variables. As no model is perfect, the main issue with using the CART model is its low efficiency in dealing with linear relation and its sensitivity to random noise.

Actually, we have already seen this type of model in the insurance business. Think of the process in underwriting, where the information about an applicant will go through an array of decision-making points and finally reach to its final underwriting results. This is exactly the same idea of CART model, although the underwriting processes are built based on experience study and business expertise, not on statistical algorithms. We believe the current underwriting model can be further improved with the help of a decision tree algorithm.

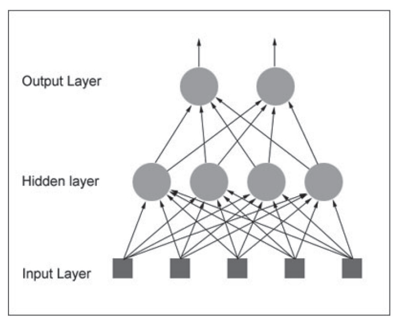
Besides the CART model, there are some other algorithms that are based on the decision tree, but instead of only one decision tree, a group of decision trees are built to extract more information from data. These algorithms are usually much more advanced and sophisticated, but also harder to interpret and gain business insights from. Examples include random forest, and Ada-boosting.

Other models

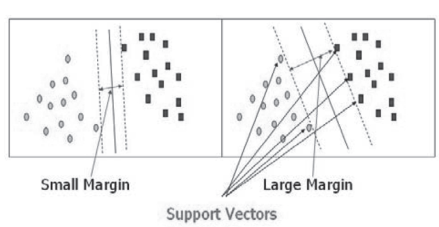
The advance of statistics has brought us more sophisticated models than are discussed above that will potentially find their ways to the insurance applications. Many of them have been utilized in other industries under such names as "business analytics," "big data," or "data mining." Some of them may well be suitable for application in insurance, and a few examples are presented here for illustration.



Clustering. This algorithm is to organize data points into groups whose characteristics in each group share similar distributions. It is an ideal candidate model for applications in classification, especially when the target variable is unknown or not certain. There are many different algorithms to form clusters, but the most popular and simplest is based on Euclidean distance in multidimensional space. You may apply the clustering for market segmentation to find a group of customers that will buy similar merchandises, for identification of effective advertisement for different consumer groups, for recommender systems, etc. In actuarial science, clustering is a very useful tool for in-force cells compression or scenario reduction, especially when a detailed seriatim study is needed or a large number of scenarios have to be simulated.



Neural Network. Also called artificial neural network, the neural network model has its deep root in biological neural networks. The algorithm mimics the interconnected biological neural cells and uses weights for each connection to model patterns in data or relation between inputs and outputs. This model is very powerful in mathematics such that it can replicate any distribution in theory. Its applications date back to the 1990s and today you can find its usage in almost every industry. The neural network is essentially a black-box approach, and it is very hard to interpret the model once it is built. Although its effectiveness and predictive power have been proven in practice, the model cannot help to better understand the business and provide insightful clues to improve it, which limits its application in real business.



SVM. This refers to support vector machine. The basic idea is to split data into two groups in such a way that the separation margin between them is at a maximum. The real algorithm is much more complicated than the

CONTINUED ON PAGE 16

simple idea, with multidimensional nonlinear feature space mapping and inclusion of regression as well as classification. This model is generally more accurate than most other models, and is very robust to noise and less likely to have over-fitting problems. Although it is not totally a black-box algorithm, it is still hard to interpret the model and may take a long computing time for a complicated model. Nevertheless, it has great potential in applications in insurance.

The choice of a model that is the best fit to a specific business purpose does not have to be limited to the models that have been briefly discussed here. There are certain rules to follow when selecting a model, but there is also a combination between science and art when you have the freedom to choose between varieties of options. The most advanced and sophisticated model is not necessarily the best choice for a particular business situation. More often than not, some simple models such as GLM may well meet the accuracy requirement and produce desirable results. As long as a model can meet the demand of real business, it will be much more effective to choose a simple model than a complicated one.

CONCLUSION

It should be clear by now that predictive modeling provides a wide range of potential applications for insurance companies. Whether it is a logistic regression model for an underwriting process, a Cox proportional hazard model for a mortality study, or a CART model for pricing, the same core objectives are sought—maximizing the value of data to improve business processes and customer experiences.

To build a successful predictive model that has business value in practice, statistical skill is certainly a very important part of the equation. Actuaries need more education and training in statistical modeling skills. Far too often the statistical nature of the models creates uneasiness for a vast majority of actuaries. On the other hand, mastering of statistical techniques alone does not guarantee a fruitful PM project. Statistical experts often lack the topic specific experience of the businesses for which models will be applied. The merge of these two sets of skills will need a high degree of collaboration between the statistical modeling teams and the business unit experts in order to maximize the potential of PM in business. Actuaries could play an irreplaceable role in the process. ▼

Richard Xu

Richard Xu, FSA, PhD, is senior data scientist and actuary at RGA Reinsurance Company in Chesterfield, MO. He can be reached at *rxu@rgare.com*

Hidden Markov Models and You

By Doug Norris & Brian Grossmiller

This is the first of a series of articles exploring uses of Hidden Markov Models in actuarial applications. In this introduction, we will go over the basics of Hidden Markov Models along with some brief illustrative examples.

WHAT IS A HIDDEN MARKOV MODEL?

A Hidden Markov Model (HMM) is a method for evaluating, and finding patterns within, time series. This can be a very useful way to model data under the right circumstances, such as when individual data points could be swayed by different influences. For instance, the data in Figure 1 shows some clear low and high periods, indicating that some manner of mixture model could be a good fit.

Some interesting questions arise from patterns like these, particularly when we see them in claims or sales volume data. How do we know if we are in a high or low state? What are the odds of staying high or low? What do we think the next data point will look like?

HMMs are similar to the Markov Processes covered in the current actuarial syllabus. The critical difference is that the matrix of state transitions is hidden, and we have to infer it from our data. Fortunately, the transition matrix typically is quickly estimated using a computer; we'll get into a basic example after we review some assumptions and definitions.

STARTING WITH THE BASICS

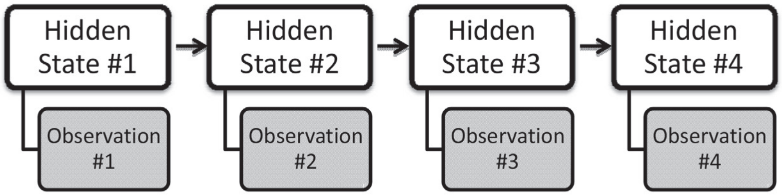
When our data contains categorical information, we can construct a sequence of states easily. In the absence of defined categories, the real power of HMMs comes into play, because we have to use the data values themselves to determine these states or categories. A set of observations associated with the sequence of hidden states can infer the (hidden) transition matrix. A basic illustration of the hidden states and resulting observations is shown in Figure 2.

In order to apply an HMM, we need to assume that each observation is drawn from a given distribution. Discrete distributions, such as the Poisson, Binomial, or Negative Binomial, are attractive due to their mathematical tractability. Using a computer, several distributions can be constructed rapidly, along with statistics measuring how well they fit. In practice it may prove convenient to model them all and select a distribution on the back end.

In addition to the matrix of transition probabilities between states, the data also can determine the parameters for the assumed probability distributions in each state. We'll illustrate this with an example later in the article, but for now we'll go

Figure 1: Example of data showing high and low patterns.

Figure 2: Hidden States and Associated Observations



CONTINUED ON **PAGE 18**

over some of the benefits of using an HMM in the context of our three earlier questions.

What state are we in?

Decoding is the process of using our data to determine the most likely sequence of hidden states we have observed. When we are looking at claims or sales volume, we will be interested in the state we are observing at each point (known as local decoding).

Conversely, in applications where the hidden state is an in-or-out type category (such as the presence or absence of a chronic disease), we are more interested in what state all of the observations collectively imply. This is known as global decoding and is the method used in many applications of speech and facial recognition technology.

What state will we be in?

Our decoded data can give us a useful estimate of what the current state might be, and combined with the transition matrix we can estimate what state we are likely to be in over the near term. If we estimate a large probability of transitioning into a different state at the next time point then the most recent data may be a poor estimator.

In the forthcoming example, we will assume that the transition at any point in time only depends upon the state immediately preceding it. In practice, you might see higher order dependencies. For instance, visual inspection may suggest that your data shows a "high" state lasting for about three time periods before transitioning. In this case, modeling the transitions to depend upon the past three states would make intuitive sense.

THE EM ALGORITHM IS AN ITERATIVE METHOD, WHICH PRODUCES MAXIMUM LIKLIHOOD ESTIMATES FOR MISSING PARAMETERS IN AN HMM MODEL SOLUTION.

What do we think the next data point will look like?

In addition to the transition matrix, constructing an HMM also requires estimating the probability distribution in each state. If we have assumed a convenient distribution, then calculating the expected values of each state is straightforward. Our estimate of the probabilities of future states can serve as weights to determine the value we expect to see at that point in time.

Now, let's take a look at an actual example to get an idea of how HMMs work in action.

ESTIMATING HIDDEN MARKOV MODEL SOLUTIONS

Because Hidden Markov model solutions are not easily solved in a closed form, many numerical analysis algorithms have been built to estimate HMM parameters. The most popular of these algorithms is known as the EM algorithm (or the Baum-Welch algorithm, after its inventors).

The EM algorithm is an iterative method, which produces maximum likelihood estimates for missing parameters in a HMM model solution. This algorithm involves repeatedly invoking an "E step" (estimating the conditional expectation of the functions generated from the missing data) and an "M step" (maximizing the likelihood, where the functions are replaced by those conditional expectations), until the algorithm converges upon a solution. The solution space is typically non-linear, and will depend upon our initial estimate of the HMM parameters.

These algorithms are programmatically straightforward; however, it is always nice to have a head start. The results in the following example were built using the "R" programming language, with a publicly-available HMM package whose algorithms are based upon the text "Hidden Markov Models for Time Series," by Walter Zucchini and Iain MacDonald.

PUTTING HMM TO WORK

One of the great challenges in actuarial work is teasing out the effects of seasonality on claim patterns. In this example, health actuary Albert Franken is having a bit of trouble estimating the number of claims that will be seen in the next month for a small rural clinic. He knows that seasonality effects are significant, as influenza and other maladies spring up in bunches and spike utilization. Fortunately, he has seven years' worth of claim history and a graph of the history readily shows a seasonal pattern in Figure 3:

Upon inspection of the data, Franken theorizes that the clinic's claim activity can be estimated by a Hidden Markov Model with three states:

- A low level of claim activity, where claim volume can be modeled as a Poisson distribution with parameter λ_1.
- A medium level of claim activity, where claim volume can be modeled as a Poisson distribution with parameter λ_2.
- A high level of claim activity, where claim volume can be modeled as a Poisson distribution with parameter λ_3.

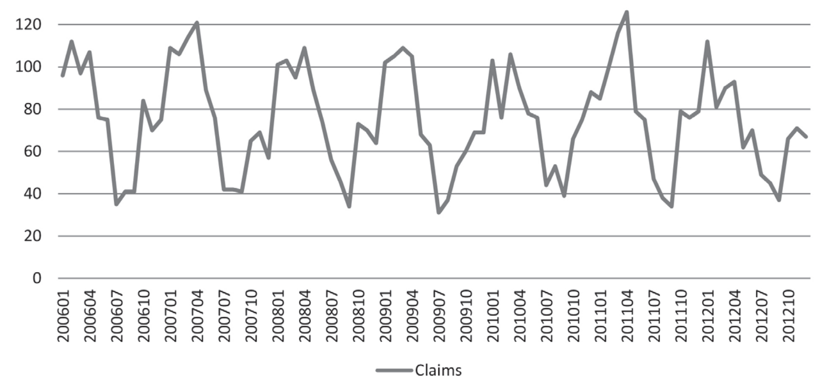
Estimating the three levels of claim activity at 53 claims (in a low month), 75 claims (in a medium month), and 93 claims (in a high month), Franken uses an EM algorithm to produce a maximum likelihood estimate Hidden Markov Model to describe his clinic's data.

The resulting HMM gives us a transition matrix of:

65.3%	34.7%	0.0%
20.3%	58.1%	21.5%
0.0%	30.2%	69.8%

Figure 3: 7 years of claims data showing a seasonal pattern.

Claims by Month (General Clinic)



And Poisson parameter values of:

Low	41.74
Med	71.91
High	102.89

Therefore, the HMM suggests that there are three states of activity—a low level of claims (with about 42 claims per month), a medium level of claims (with about 72 claims per month), and a high level of claims (with about 103 claims per month). Moreover, if a given month is a month with low activity then the following month will either be low as well (with 65 percent probability) or will be medium (with 35 percent probability).

In the last month of clinic data there were 67 medical claims. How can the HMM help our actuary to estimate what could happen next month? He first needs to estimate which activity state we are currently in, and with 67 claims we have a 99.7 percent probability of being in the "medium" level of activity (along with a 0.2 percent probability of being in the "low" level of activity and a 0.1 percent probability of being in the "high" level of activity).

CONTINUED ON **PAGE 20**

Feeding this information into the HMM (after remembering how to do matrix multiplication), Franken finds that next month will be in a "low" level of activity (with 20 percent probability), a "medium" level of activity (with 58 percent probability), or a "high" level of activity (with 22 percent probability). Overall he should expect about 73 claims from the clinic next month:

$$41.74 * 20\% + 71.91 * 58\% + 102.89 * 22\% = 72.69$$

OBSERVING THE UNOBSERVABLE

We hope that you enjoyed our introductory article to Hidden Markov Models. In the December newsletter, we plan on putting together some more in-depth examples in Excel to help inspire your own applications. Stay tuned for more on the most fun models you never saw!

REFERENCES

Zucchini, Walter and MacDonald, Iain, Hidden Markov Models for Time Series. Upper Chapman & Hill/CRC, 2009. Print. ▼

Doug Norris

Doug Norris, FSA, MAAA, PhD, is a consulting actuary at Milliman Inc. in Denver, Colo. You can reach him at *doug.norris@milliman.com*

Brian Grossmiller

Brian Grossmiller, FSA, FCA, MAAA, is provider contract evaluation manager at Moda Health in Portland, Ore. He can be reached at *Brian.Grossmiller@yahoo.com*

Dark Side of the Moon

By Clark Ramsey

Welcome to the latest issue of the newsletter of the Forecasting & Futurism Section. But before I begin to talk about your section, please join me for a brief detour …

It was not until 1959 that Luna 3, a robotic Soviet spacecraft, rounded the moon and provided mankind its first glimpse of the far side of the moon, which is often less accurately referred to as the "dark" side of the moon. For all the preceding millennia that humans had walked the Earth and gazed up at our nearest and brightest celestial neighbor, it was always the same side of the moon that they saw. However the moon came to be the Earth's outsized satellite, it likely only required a few million years for Earth's gravity to tidally lock the moon to the Earth, so that the moon's orbit around the Earth takes almost exactly the same amount of time as the moon takes to rotate once on its axis. The first human eyes to see the far side of the moon live were those of the astronauts on the American spacecraft Apollo 18 in 1968. Five years later, Pink Floyd released their epochal album *Dark Side of the Moon*. Perhaps the fact that I will almost certainly be hearing selections from *Dark Side of the Moon* at my upcoming 40th high school reunion is what sent me on this digression.

Prior to Luna 3's poor, grainy 1959 photos, humans collectively had essentially zero data points about the far side of the moon. Those pictures may have been low-resolution, but they provided our first real data, and were sufficient to reveal that there were striking differences between the near and far sides of the moon. Later spacecraft, both robotic and manned, have added immense volumes of data, but in many important ways we remain largely ignorant about the far side of the moon and the reasons that its mountainous, relatively mare-free terrain stands in such stark contrast to the side we have known for so long.

Clark Ramsey

Our knowledge of the far side thus stood at essentially zero until a little more than half a century ago, and now that knowledge is extensive in some ways and yet still quite deficient in others. This is not unlike what we may see in an actuarial career. In our actuarial lives, we sometimes face situations that require forecasting from little or no data, perhaps in pricing or reserving for a new product or for an existing product in a new market. The brave actuaries who first priced long-term care insurance lacked data on voluntary lapse rates for this new coverage, so they often made the assumption that experience would be similar to life insurance products. Before Luna 3, astronomers had in a similar manner assumed that the far side of the moon would be the same as the visible side. In both of these cases, subsequent data proved otherwise—the far side was very different from the near side, and long-term care lapse rates did not look much like those on life insurance.

Other times we may have an abundance of data, but the route to meaningfully interpreting that data in a manner that facilitates preparing whatever forecasts are desired may not be clear. How do we navigate through these varied circumstances? Having a wide variety of forecasting and futurism tools at our disposal would serve us well, and that of course is the reason for the Forecasting & Futurism Section's existence.

Your section council has been very active; here are some of the ways that we are striving to help you expand your forecasting and futurism toolkit:

Complexity science is one of the primary focus areas for the section, and **genetic algorithm** techniques in turn are at the center of several section efforts. Many of you were able to attend our webcast on genetic algorithms last December. We will sponsor two genetic algorithm sessions at the Annual Meeting; it is not unlikely that a new webcast will follow from these sessions. We are also once again sponsoring an iPad contest, with this year's topic being "Build a Genetic Algorithm"; Albert Abalo writes about the contest in this issue. Give the contest a try; you might advance actuarial science and walk away with a new iPad to boot!

Together with the Joint Risk Management section, we sponsored a webcast on Emerging Risks. I hope you enjoyed it as much as I did.

Clark Ramsey, FSA, MAAA, is chief actuary at Employers Reassurance Corporation in Overland Park, Kan. He can be reached at *clark.ramsey@ge.com.*

We are very active with the Delphi technique, one of the judgmental forecasting methods that your section champions. Impending regulatory changes for long-term care created an immediate need to address what factors might contribute to an ideal solution to anticipated funding shortfalls. We are partnering on a Delphi study with the Long Term Care Think Tank, which is striving to use the Delphi study results to influence the future of long-term care in the United States. We are also participating in a Delphi study on future roles for actuaries, something which should be of interest to all of us.

Another branch of complexity science, **agent-based modeling,** will be the topic at sessions we will sponsor at both the Health Meeting and the Annual Meeting. In the coming months you will see Forecasting & Futurism involved in even more of the newer techniques useful for actuaries. In this issue we will introduce **hidden Markov models, Bayesian networks,** and **behavioral economics** and provide additional insights into **Delphi studies,** and **predictive modeling** choices.

I am also excited about the increased size of our team of volunteers this year. What used to be a small band of zealots doing just about everything, has grown to a lot of new contributors who bring their own special skills along with their enthusiasm. Forecasting & Futurism is one of the few sections that seems to be thriving during a general slide in section membership. We thank you for your continued support.

As always, the council welcomes your suggestions on how we can continue to make section membership more valuable to you. Please let me or any of the councilmembers know your thoughts and suggestions, and if you see me at a Society meeting, please feel free to introduce yourself.

My term as chair of the Forecasting and Futurism Section Council will end before our next newsletter, so I would like to take this opportunity to thank my fellow councilmembers and friends of the council. It has been an honor to be associated with this fine group of volunteers, and it is my sincere hope that I am able to work with each of them again on future volunteer efforts. My time spent dealing with them and with Meg Weber and Christy Cook from the SOA staff has contributed to my professional and personal growth and has been a delightful experience.

IN OUR ACTUARIAL LIVES WE SOMETIMES FACE SITUATIONS THAT REQUIRE FORECASTING FROM LITTLE OR NO DATA.

Also rolling off the council this year are Brian Grossmiller, Jon Deuchler and Donald Krouse. Brian has served as secretary/treasurer and Health Meeting coordinator, as well as handling the section's member outreach and new member recruitment. He has also frequently served as a speaker at Society meetings and has enlivened our council calls with his unique sense of humor. Jon's term originally expired last year, but he graciously volunteered to serve another year when we had an unexpected opening on the council and has offered different perspectives and wise counsel in our deliberations. Donald preceded me as chair of the council and leaves a legacy of achievement for the section. I have a great deal of confidence in the remaining councilmembers, who will continue their terms into the next year. They provide a solid nucleus for the next council and will continue to move the section forward.

Enjoy your newsletter. As for me, I think I will read it while playing a little music in the background, perhaps a cut or two from ***Dark Side of the Moon.*** ▼

The Signal and the Noise: Why So Many Predictions Fail—but Some Don't

By Nate Silver
Book Review by Scott McInturff

The Signal and the Noise: Why So Many Predictions Fail — but Some Don't by Nate Silver is a book every actuary who is interested in forecasting should read. If Silver were writing about fly casting, rather than forecasting, upon finishing his book the reader would want to immediately hand-tie a fly, wade into a wide river and cast a line repeatedly in hopes of using newly learned methods to hook a whopper lurking at the bottom of a deep pool. After reading this book you will most certainly be using some new techniques in your next foray into forecasting.

Not only is Silver's book entertaining, it is informative. This is an archetypal introductory text on forecasting that is worth the time spent reading it. Filling his narrative with anecdotes and examples, novice and expert forecaster alike will enjoy reading Silver's proclamations concerning the approaches and pitfalls of political, financial, geological and social forecasting. Your forecasting toolbox will be expanded by reading Silver's book.

Silver has been a forecaster in many disparate arenas. He won more than $400,000 as a late night clandestine on-line gambler, using Bayesian logic to know when to hold 'em and know when to fold 'em. He used his love of baseball and knowledge of statistics to develop a superior commercially viable forecasting system to evaluate baseball players by de-emphasizing performance measures that rely on luck and focusing on those that gauge skill. Silver gained his greatest celebrity by accurately predicting the outcomes of the 2008 elections at national and state levels through his blog FiveThirtyEight. He then repeated this accomplishment after his book was published in the 2012 election cycle.

Though there's an undercurrent of hubris in his writing, this is not a chest-thumping autobiography. While he'd be forgiven if he boasted, Silver downplays his personal forecasting successes, omitting key details of the accuracy with which he predicted the outcomes of the 2008 political elections. He knows it is not his accomplishments that make taking time to read his book worthwhile. This book is a gem because Silver is able to describe in detail, citing example after example, how forecasting should be approached and completed. This is a classic textbook on forecasting that is masquerading as an engaging best seller.

Each chapter in Silver's book is chock full of forecasting tips. While it is easy to get caught up in his storytelling, the lessons on how to forecast are impossible to miss. Subheadings, lists, italics, and bolded words are all used to make sure that the reader doesn't mistake this informative textbook for a fast moving novel. Any actuary with an aptitude for forecasting will expand his or her toolbox by reading the entire text.

Silver's secret to forecasting success is a very Bayesian one. He always couches his forecasts based on his assessment of the probability of their being accurate. Not only does he clearly describe Bayes' theorem and give examples of how it is used, he implores all forecasters to think probabilistically every time they forecast. Because forecasts are dynamic predictions, Silver advises that every prediction be revised as new information becomes available. He states that the biggest trap in forecasting is holding a position when information has changed and gives examples of political pundits who make a living by refusing to budge from a position even when it has eroded.

The book's chapters are easily managed as each stands on its own and contains many useful forecasting nuggets that are not to be missed. The book as a whole is heavily researched. The Notes section alone spans 56 pages. Silver uses footnotes not only to reference all his source material, but also to make interesting and insightful comments on points he had made in the various chapters. One certainly would become bogged down if one attempted to read every one of Silver's footnotes as it was encountered in the text. However, the footnote section should not be skipped when the reader finishes the final chapter. I have never found a book with a more interesting and informative footnotes section than in Silver's offering.

Scott McInturff

Scott McInturff, FSA, MAAA is vice president, Retrocession Pricing at Pacific Life Re in Boston, MA. You can reach him at *scott.mcinturff@pacificlifere.com*

Though it is a poor cousin to the Notes section, the Index, at only 20 pages, is highly useful. It is complete enough to allow the reader to readily use Silver's book as an invaluable forecasting reference guide for years to come.

If you are interested in forecasting, I encourage you to read The Signal and The Noise by Nate Silver. If it doesn't make you a better forecaster, it may at a minimum make you a better Texas Hold 'em player. It certainly will give you a better appreciation for the uses of forecasting in everyday life and the knowledge of how to forecast using some of the best basic techniques available. If you take the time to read it, I can almost guarantee that like the novice fisher exposed to a well written text on fly casting, you'll be searching for a quiet place to practice the art of forecasting.

Delphi Study 2000—Predictions for 2010 and 2050

By Ben Wolzenski

ABSTRACT

Futurism section members participated in this Delphi study, which had multiple choice questions on a range of life, health, economic and professional issues. This article provides background on the study and comments on the predictions of the participants for 2010 (and 2050) in light of experience to date.

BACKGROUND

In the spring of 1999, members of the Futurism section (renamed to Forecasting & Futurism) were invited to participate in a Delphi study[1] on "varied topics of interest to actuaries." The study was part of the SOA's 50th anniversary celebration. In the first Round, 115 responses to the 25 multiple choice questions were mailed back to Albert E. Easton, FSA, MAAA. At that time, Al was chair of the Futurism Section Council and conducted the study. Those responses were tabulated and shared with the participants for their consideration of a second response to the same questions: "Round 2" in Delphi parlance. The results were not just shared with the participants; the SOA staff worked with Easton and the council to develop a media kit to let journalists see the survey results quickly. The *Wall Street Journal* put a "Business Bulletin" item about the study on its front page, which was picked up by the Associated Press. Subsequently, dozens of articles in lay and industry press picked up on the story.

HOW ABOUT THOSE 2010 PREDICTIONS?

First, an explanatory note is needed. Participants chose answers within ranges, and mean responses were computed, generally by using range mid-points. Because range widths varied widely, sometimes the percentage picking the correct range was more informative, and other times the mean was a better indicator. So this article refers to some of both. There were two rounds in this study, but since the second round results (when participants could see aggregate first round responses) varied little from those of the first round, only the final (second round) responses are described.

Ben Wolzenski

Ben Wolzenski, FSA, MAAA, is managing member at Actuarial Innovations, LLC in St. Louis, Mo. He can be reached at *bwolzenski@rgare.com*

Generally (predictably?), the actuaries who participated did best when predicting modest change or the continuation of a trend, and least well when there was substantial change or a change in direction. Here are a few of the highlights.[2]

HEALTH CARE EXPENSES AND INSURANCE

Most participants' predictions were directionally correct, if not in the precise range, for each of the four questions. Total health care expenditures as a percentage of GDP were 17.9 percent in 2010 compared to a mean prediction of 16.6 percent, an increase from the 13.6 percent in 1993-1996 that participants had as a reference. The average annual increase in cost of prescription drug benefits was 10.4 percent over 1999-2008 (the latest I could find online) against a mean prediction of 10.5 percent for 1999-2010.

CAUSE-SPECIFIC DEATH RATES

Predictions were a mixed bag. Only 26 percent of participants correctly predicted that the death rate from heart disease would fall below 200 per 100,000 in 2010. The rate per 100,000 in 1995 had been 271, and it fell to 194 in 2010. On the other hand, 68 percent correctly predicted a drop in the death rate from cancer below 200 per 100,000; it was 201 in 1995 and 186 in 2010.

EMPLOYMENT AND RESIDENCE OF SOA MEMBERS

In 1997, 39.9 percent of employed SOA members worked for consulting firms. Only 6 percent of participants correctly predicted that that percentage would not be at least 40 percent in 2010, when it was 37.9 percent. In 1997, 8.6 percent of SOA members lived outside the United States and Canada; 66 percent correctly predicted that that percentage would be at least 10 percent by 2010, when it was 12 percent.

BEST PREDICTIONS

Fourth best: In May 1998 there were 437 electricity producing nuclear reactors in operation worldwide. The actual number on 12/31/2010 was 441. This compares to a mean prediction of 438. But wait, that was arrived at by weighting the percentages of answers in two very wide bands, 100-500 (predicted by 68 percent) and 500-1,000. So I downgraded this to fourth best.

Third best: The percentage of the labor force in 1990 who were women was 45.3 percent. Eighty-five percent predicted the correct range for 2010, 45-50 percent, and the mean prediction of 47.6 percent was close to the actual, 47.2 percent.

Second best: The expectation of life at birth in 1997 was 76.5 years. Eighty-two percent predicted the correct rather tight range of 77-79 years for 2010, when the actual was 78.7. The mean prediction was 79 years.

Best: In 1997, the total life insurance in force in the United States was $13 trillion. Only 68 percent predicted the correct range ($15-20 trillion) for 2010. But the "best" moniker goes here because the actual amount was $18.43 trillion and the mean prediction was identical at $18.43 trillion!

WORST PREDICTIONS

Fourth worst: Participants were informed that the greatest ever one-day percentage loss in the DJIA was 22.6 percent on 10/18/87. Only 6 percent of participants predicted that the correct range for the worst one-day percentage loss between 2000 and 2010 would be "less than 10 percent." The actual greatest one-day loss in that period was only 7.9 percent, on 10/15/08.

Third worst: The average net investment income rate on the general account assets of U.S. life insurance companies in 1996 was 7.75 percent. Only 4 percent of participants predicted the correct range (less than 6 percent) for the actual 2010 value of 5.37 percent.

Second worst: The unemployment rate in the United States in 1997 was 4.9 percent. Only 3 percent of participants predicted the correct range (greater than 7 percent) for the value in 2010, which was 9.6 percent.

Worst: The death rate from AIDS in the United States in 1995 was 13 per 100,000 population. Only 2 percent of participants predicted the correct range (less than 5 per 100,000) for the value in 2010. The actual value was only three per 100,000.

AND WHAT ABOUT THE PREDICTIONS FOR 2050?

Of course, we do not know how accurate the predictions will be … but here are three of them on which there was broad consensus:

- 83 percent of participants predicted that the average annual increase in the U.S. Consumer Price Index from 1999-2050 will be between 2 percent and 5 percent. (The average increase from 1999-2010 was 2.5 percent.)
- 96 percent of participants predicted that the expectation of life at birth would be 80 or greater in 2050, and a majority of those predicted 82 or greater.
- 83 percent of participants predicted that in 2050, 15 percent or more of employed SOA members would work outside insurance companies or consulting firms. In 2010, 10.75 percent did. ▼

ENDNOTES

1 For background on the Delphi technique, readers can access the article "The Delphi Method" by Scott McInturff in the September 2009 issue of the Forecasting and Futurism newsletter by entering "The Delphi Method" in the Quick Search field of the SOA website home page. The first result of the search is a link to a pdf copy of the article.

2 For a copy of the full questions, answers and actual 2010 results, please contact the author of this article.

FORECASTING & FUTURISM THIRD ANNUAL IPAD CONTEST

Build a Genetic Algorithm

By Alberto Abalo

The Forecasting & Futurism Section has been at the forefront of educating actuaries in the area of complexity. Over the course of the past few years, our section has presented its members with insights into applications as diverse as predictive modeling, agent-based modeling, artificial intelligence, and genetic algorithms. We now look to you to advance our research efforts.

Our third annual iPad contest is looking for new actuarial applications of genetic algorithms.

In addition to bragging rights, the winner will receive a new iPad (or cash equivalent). More importantly, your work will serve to advance the actuarial state of the art. The winning submission (and solid runner-ups) will be profiled in a future edition of our newsletter and potentially be presented at future SOA events.

"WHAT IS A GENETIC ALGORITHM?"

To quote our former chairperson, complexity and the study of complexity have been around a LONG time. That said, complexity science as we think of it today is fairly new. While many of the subtopics of complexity science emerged at different times, one of the landmarks was the invention of the genetic algorithm by John Holland in the 1960s at the University of Michigan.

Genetic algorithms sound very complicated, but they don't have to be. A genetic algorithm is simply another technique to find solutions to problems that we do not know how to solve deterministically. They are not a panacea for all problem solutions; but they can be very useful for problems that exhibit the following characteristics:

1. There is no direct deterministic solution known; or, if there is, it is prohibitively difficult to implement in real time.
2. The number of possible solution sets is very large—too large to try them all in the desired time frame.
3. We can devise a sufficient scoring system or ranking scheme to quickly compare the relative value of any two possible solutions.


Alberto Abalo

Alberto Abalo, FSA, CERA, MAAA, is a principal at Oliver Wyman in Atlanta, Ga. He can be reached at *alberto.abalo@oliverwyman.com*

"WHERE DO I BEGIN?"

Several resources, including examples are available.

- ***Forecasting & Futurism Newsletter***
  - "Genetic Algorithms—Useful, Fun and Easy!" by Dave Snell on page 7 of the Dec. 2012 issue. A workbook is also available.
  - "Are Genetic Algorithms Even Applicable to Actuaries?" by Ben Wadsley on page 6 of the July 2011 issue.
- ***The Actuary Magazine***
  - "Complexity or Simplicity" by Dave Snell on page 16 of the December 2012/January 2013 issue.
- 2012 SOA Annual Meeting Presentation
  - Complexity Science: Genetic Algorithm Applications to Actuarial Problems by Dave Snell & Brian Grossmiller

"HOW WILL MY ENTRY BE JUDGED?"

Entries will be judged winner on the basis of several criteria:

- Usefulness to the actuarial profession—solving some meaningful problem. This will be the most heavily weighted criterion.
- Sophistication of the algorithm—what techniques are covered (simple in/out genes, range of discrete or continuous function values, case tables or some other recognition of ambient conditions).
- Clarity of implementation—an actuary familiar with (Excel, VB, R, NetLogo or whatever language is used) should be able to follow the logic.

- Extensibility—to other similar or non-similar types of problems.
- Speed to a solution—in generations, or calendar time.

"HOW DO I ENTER?"

Please submit your final entries by Aug. 31, 2013, to Christy Cook at ccook@soa.org. The only requirement is that you are a member of the Forecasting & Futurism Section. Entrants should verify section membership status on the online Actuarial Directory.

The winner will be announced at the 2013 SOA Annual Meeting Forecasting & Futurism Section breakfast (attendance at annual meeting is not required to win).

RULES

- Entrants must be current members of the SOA Forecasting & Futurism Section.
- The Forecasting & Futurism Section Council reserves the right to not award any prize in the event all entries are wide of the mark.
- The Forecasting & Futurism Section and the Society of Actuaries may choose to use information about any of the entries submitted in publications or other venues of the SOA without further involvement of the entrant.
- The Society of Actuaries reserves the right to substitute the cash equivalent value of the contest prize.
- The competition winner is responsible for taxation issues as they are appropriate to his/her region.

If you have any questions on the contest rules, feel free to contact Alberto Abalo at *alberto.abalo@oliverwyman.com* or Doug Norris at *doug.norris@milliman.com.* ▼

Global Trends 2030: Alternative Worlds

Report by the National Intelligence Council summary by Jon Deuchler

There is a quote by John Maynard Keynes that reads: "… the idea of the future being different from the present is so repugnant to our conventional modes of thought and behavior that we, most of us, offer a great resistance to acting on it in practice." Many actuaries, this one included, take that to heart in our modeling and pricing. We often lock into one set of best estimates to project expected results and test only minor perturbations in the drivers to come up with a range of results. We forecast mortality improvement, investment returns and interest rates and sometimes do stochastic projections and attach high credibility to our ability to understand the variances.

Global Trends 2030 is a report produced by The National Intelligence Council (NIC) and released in December 2012. This is the fifth such report, the first being produced in 1997 and roughly every four years thereafter. This is really a fascinating report discussing global trends and speculation on demographic, economic, natural resource and power shifts to come in the next 20 years.

The NIC is a coalition of 17 agencies and organizations within the Executive Branch of the U.S. government and is a kind of clearinghouse of gathered intelligence. This report adds the perspective of experts outside the government (and outside the intelligence community) and was presented to policymakers in the Executive and Legislative Branches to aid in long-term planning. This is not meant to be a prediction of the future. It's more like a list of possible situations that deserve attention, thought and discussion. It proposes some approaches for solutions or responses, but strives not to make value judgments as to the "best" solution. I found that I didn't always agree that the issue was as major or minor as proposed in this report, but I also found it very interesting that my view of the world was quite different than presented here.

Jon T. Deuchler

Jon T. Deuchler, ASA, MAAA, is associate actuary at RGA Reinsurance Company in Chesterfield, Mo. You can reach him at *jdeuchler@rgare.com*

The report is a deeply written and detailed document of more than 130 pages. It is both a fascinating read on global politics, societal changes, economic upheavals, technological advances and the future role of current super powers and an almost dry, fact (or accepted as fact) driven discussion of situations deemed certain to happen. I found it extremely interesting and sometimes disappointing or somewhat disturbing that the changes and situations forecast could be accepted almost matter-of-factly.

There are two pieces to this publication. The first is an abstract of four pages and the second is the larger, more complex document (over 130 pages). The abstract summarizes the larger document, separating the major points into a menu like format. The starting courses are the megatrends, items that are considered to be relatively certain of happening. These include individual empowerment, demographic swings, and diffusion of power and natural resource issues (food, energy, climate change and fresh water). Chapter 1 of the larger report is devoted to discussing these items in great detail.

Individual empowerment is the idea that the middle class will grow dramatically over the next 15-20 years due to poverty reduction, education opportunities and better health care (recall the perspective, this is a global analysis report). While the lives of many are expected to improve, the opportunities for individuals and small groups to be very disruptive will also be enhanced, due to cyber weapons, bio-terror or nuclear weapons or other forms of large scale violence.

Demographic swings are four trends that will affect the internal economic and political climates of many countries and how all countries interact with each other. The four trends are: aging, migration, growing urbanization and youthful (or relatively so) states and societies. It's no secret that many countries are faced with an aging population. This presents challenges for many resources, including health care, productivity and housing. At the same time, relatively youthful states will lose population by migration to states with better job opportunities (or simply higher wages). This will put a strain on both states, on the former to replace or strive to keep their population and on the latter to house and feed the influx. Growing urbanization refers to both of these issues.

Diffusion of power refers to the shift of the major population centers to Asia, China in particular, in addition to the growth of the GDP in Asia countries. Further complicating this is the consideration of what actually constitutes power in an era where the individual becomes more highly prized. Authoritarian governments may not be able to leverage their economic power if their control of their population is lowered or limited.

Natural resources issues seem almost too easy to include. It's obvious to most that food, fresh water and energy needs will continue to grow. If the middle class does dramatically increase, then so will their consumption patterns. Can the existing world economic meet those needs? Are we headed for shortages in critical items? What will be the effect of climate change? All of these questions seem to me to suggest that the growth in the middle class will not be as great as suggested in the first item.

The main courses are called critical game-changers. This is a very broadly drawn list of six items that interact with one another and that will drive the megatrends. The six items are: the global economy, current governments (and current forms of government and their response to the megatrend changes), the potential for increased conflict, regional instability expanding into global instability, the impact of new technologies and the future role of the United States. All of these elements are discussed in Chapter 2 in just about half (66 pages) of the report.

As this is the largest and most detailed section, trying to summarize it would not do it justice. There are many interesting and insightful sections that describe the rationale behind the conclusions. To say that I consider this required reading for all actuaries would not be sufficient: everybody everywhere should read this section. It's that good and that important. And, if it doesn't make you think long and hard about the future, you're in the wrong business.

The dessert section describes four scenarios (mostly bleak) and how they can affect the game changers. To really destroy the metaphor here, the idea is that the dessert can make or break the dinner. The four scenarios are, "Stalled Engines, Fusion, Gini Out of the Bottle and Nonstate World." Stalled Engines is a scenario where the United States, Europe and China turn inward and retreat from a position of leadership in the world. Thus, slower economic growth leads to more conflicts in lesser-tier countries and there is greater instability in certain regions. This scenario also postulates a pandemic that ravages the poorer countries. Because the richer countries are walled off, they are better able to withstand the pandemic and emerge even further ahead economically. Since this flies in the face of every zombie apocalypse movie or television show that I've seen (even those soon to come), I consider this to be a real Black Swan scenario.

Fusion refers to a scenario where cooperation between the United States and China leads to a more cooperative and prosperous world. The United States and China essentially share the role of policeman to the world and intervene in conflicts so as to limit their impact and cost. This scenario results in a rise in the fortunes of virtually all countries. Since this seemed like the most common sense scenario, I immediately dismissed it as being a fairy tale. It's just too good to be true. However, it's just possible that enough people will read this and realize what potential there is in the world. I wish.

Gini out of the Bottle seems to be a scenario that is pretty much status quo. The inequalities between rich and poor nations continue to grow, leading to increased resentment and conflict. While the global economy grows (better than stalled engine but less than fusion), it's country centric and not even across all countries. I'm not sure I agree that the economy would perform as described in the write-up; it seems to me that more conflict would be more inhibiting of economic growth. But the report concludes that this is an intermediate growth scenario.

CONTINUED ON **PAGE 32**

Nonstate World describes a scenario where nongovernmental organizations take the lead in managing the world and addressing global issues. These organizations include supra-country organizations, multinational business, academic institutions, even wealthy individuals. In some cases, the list might include mega-cities within a single country (crossing state lines) or across country borders. The role of centralized national governments is lessened in this scenario, a surrendering of power and influence that seems far-fetched to me. Economic growth is expected to be ahead of the Gini scenario but behind Fusion. Criminal and terrorist networks are projected to thrive (given the shift in power and authority for centralized governments), but that would seem to impact economic growth more than reflected in the discussion.

This is a remarkable report and it deserves a very large audience. Regardless of your politics, this report will give you much to think about.

www.dni.gov/nic/globaltrends ▼

Behavioral Economics: Implications for Actuarial Science and Enterprise Risk Management

By David Wheeler

Behavioral Economics is emerging as the leading decision science for economics, psychology, sociology, biology and neuroscience. Currently though, it is not widely employed by actuaries. Perhaps that is unfortunate. Actuaries are great at making models that incorporate best estimate assumptions, covariances, and lots of sophisticated mathematical techniques to quantify risk and to predict future scenarios in probabilistic terms. You base these on facts, not impressions. Yet, impressions are often important. Sometimes they turn out to be as important as facts.

Behavioral economics addresses the decision process: heuristics, biases, decision frames, environmental cues, and culture. It is not solely focused on the actual decision or outcome. Some actuaries may consider these items too subjective to model, and therefore not of interest to them. They may be more comfortable with statistics and statistical tests, such as the 'F-Test' for hypothesis testing.

Let's consider another type of 'F-Test.' This one involves a lot less complicated mathematics. In the phrase below, just count the number of occurrences of the letter F:

FINISHED FILES ARE THE

RESULT OF YEARS OF SCIENTIFIC

STUDY COMBINED WITH

THE EXPERIENCE OF YEARS

If you counted three, don't feel badly; that's the most common answer when this is shown to a group of people. The correct answer is six; but since the word of is often processed as one piece of information, those three F's disappear in the mind.

This is not to imply that humans are not able to count. It just shows that we sometimes tend to be

JUMPING TO

[redacted]

without even realizing it. I'll get back to that point shortly.

Heuristic evidence offers insight into the brain's biological superstructure, but we are left wondering how this configuration produces, influences, and—in some cases—limits behavior. The public good scenario serves as an example of how collective interest and private pursuits are often in conflict with one another. Rational-choice theories predict that outcomes of such dilemmas tend to be sub-optimal from the social perspective. Experimental evidence challenges this assumption and supports a central conviction of behavioral economics: individual variance in decision-making is evident and can lead to non-rational, socially-optimal behaviors.

Public good scenario: Imagine your town plans to build a new public park, fully funded by the local populace. The expanse and layout of the park is dependent on the total amount invested by all township members. You have the choice of investing anywhere between $0 and $20. Returns on investment and saving are as follows:

- Each $1.00 kept for private savings (i.e., not invested) earns you $1.00.
- Each $1.00 invested earns each player $0.50. This can be thought of as an increase in an individual's utility through social returns.

As always, there is a catch: Investments are kept private.

Mainstream economics, in line with rational-actor and self-interest theory, suggests that an individual is unwilling to invest in the park and attempts to reap the benefits of others' investments. In this way, an individual keeps his $20 and collects the return on social utility ($0.50 for every dollar invested by others). This is a multi-player game, however, and with all other actors behaving alike, standard game theory brings us to an equilibrium at $0 invested for every player. Under these assumptions, no park is constructed, no social return is gained, and no one ends up better off.

Alternatively, we can imagine a social model, where behavior is driven by socially optimal outcomes. With each player willing to invest the full amount toward the public good,

CONTINUED ON **PAGE 34**

we reach equilibrium at $20 invested for each player. Using this model, the best possible park is constructed, the highest social return is achieved, and everyone ends up better off.

The above scenarios represent the extremes of human behavior. This experiment has been carried out countless times, and as one should expect, the results do not neatly align with either perspective. The apparent fact that emerges is that decision-making, along with attitudes toward selfish and social outcomes, vary from person to person.

The traditional interpretations of human behavior for financial models have been assuming that we are completely rational and that we tend to make our decisions based on economic gain. In a sense, we are viewed as Homo economicus—with actions governed by economic gain or loss. The new behavioral perspective shows humans are actually more human than originally thought. We make our decisions and exhibit behavior along a spectrum from what might be called human economicus all the way to a state we might call human reciprocans—one who is more inclined to reciprocate in kind to others than to try to maximize personal gain. In effect, we tend to be what Eric Beinhocker, in his excellent book, The Origin of Wealth, calls conditional cooperators and altruistic punishers. We will go out of our way to deal in good faith when we feel the other person is treating us fairly; and to dig in our heals and take a sometimes irrational stand if we feel we have been wronged.

In contrast to top-down Neoclassical assumptions, the behavioral movement offers a bottom-up approach to understanding decision-making. It provides alternative methods for firms to obtain, interpret, and utilize data. As a science supported by mounting experimental evidence, behavioral economics recognizes variation at the level of the individual and draws conclusions about how the aggregate of these variations plays out at the macro level.

In summary, behavioral economics can be a useful extension to our more mathematically oriented ways of building models. It can help us avoid

JUMPING TO

GQNSLHSIQNS

without realizing we are doing so. ▼

David Wheeler

David Wheeler is a management associate at Synodex in Hackensack, N.J. He can be reached at *dwheeler@synodex.com.*

THE 10TH SPECULATIVE FICTION CONTEST FORECASTING AND FUTURISM SECTION PRIZE WINNER

The Weight Of Certainty—Selected Stories Of Steve Mathys

By Mike Lindstrom

Recently the Forecasting and Futurism Section has enjoyed being among the sponsors of the Speculative Fiction Contest. The Contest invites those actuaries who are so inclined to submit a short, original work of speculative fiction. These works are distributed to the general SOA membership to read, and are also considered for awards by different panels. For the past several years the section has awarded a prize for the best use of forecasting in a story.

On behalf of the section I would like to congratulate the 10th Speculative Fiction Contest Forecasting and Futurism Section Prize winner, Steve Mathys for his story, "Calibration."

"Calibration" describes a day in the life of Stuart, a businessman in Capitol City. Stuart subscribes to the services of a company called Precision Dynamics, which provides a list of custom probabilities in regard to a series of today's personal events submitted by the customer the night before. If the probability provided by Precision Dynamics is less than the a priori probability for a given event, then it can be assumed that the given event will take place that day with a high degree of certainty. On the morning in question, Stuart receives a very troubling number—for the probability that he will die that day. The story unfolds as Stuart responds to what he believes to be his imminent demise.

I recently had the opportunity to speak with Mathys, who has been writing stories since before college. He first entered the Speculative Fiction Contest back in 2003, acting on a suggestion from a coworker who knew he enjoyed writing. He ended up winning a prize that year for the story "Antiquity in their Midst." More stories for subsequent contests followed, such as "Condemnation of Fate," the detective story "Eight Seconds, or Maybe Nine," which became his favorite from the contests, and the Ninth Speculative Fiction Contest Forecasting and Futurism Section Prize winning story, "The What Ifs."

As in "Calibration," the lead character of "The What ifs" has her life greatly altered by a forecast of the future. After projections of her athletic career show a large probability of early injury, she is not picked for the draft. In a chance encounter she meets the modeler who ran the projection and they develop a relationship. Eventually he gives her the opportunity to have access to the model and she must decide whether or not to use it. That is, she has the choice between knowing with a high probability which life decisions will be optimal, or she can continue to live her life without guidance. Mathys believes it is not the certainty that is scary, but it's the loss of the opportunity to change things that might prevent someone from wanting a perfect forecast.

Whether writing a story or doing actuarial work, Mathys enjoys the process of "making something out of nothing," and feels that creativity has a place in his day job as well. He says actuaries should always be searching for a better solution, rather than just relying on what was done previously. His reluctance to be overly influenced by one source is also evident in his reading, as he limits himself to only one book by the same author in any given year. He admits that even in a fiction contest for actuaries, the story has to come first. Let's hope he continues to put it there. ▼

Michael Lindstrom

Michael Lindstrom, ASA, MAAA, is associate actuary with Federal Life Insurance Company (Mutual) in Riverwoods, Ill. He can be reached at *mlindstrom@federallife.com*.

Spreadsheet Controls … How to Prevent a Fire

By Darrick Fulton

The first car I drove after getting my license was my parents' 1982 Pontiac Grand Prix. This car still had a carburetor to mix air and gas in the engine (instead of the fuel injectors which all cars have today). The carburetor needed repair, so when I came to a stop, I had to carefully put my left foot on the break to stop and my right foot on the gas, to keep the car from dying. Instead of investing in a repair for an aging car, my Dad offered me the following good advice: "If the car ever catches on fire, get away from it and watch it burn." Fortunately, my car never caught on fire; however, this story from my youth illustrates the approach that many companies take to spreadsheet controls. Action is often delayed with the hope of avoiding a fire in the future.

The following instances demonstrate the danger of ignoring spreadsheet controls and some "fires" that have resulted:

- The J.P. Morgan trading losses related to the "London Whale," were attributed in part to an incorrect calculation in a spreadsheet model which resulted in more risk than intended.(1)

- A spreadsheet cut and paste error cost TransAlta $24 million on contract bids.(2)
- Hidden fields in an Excel spreadsheet left Barclay's with more Lehman assets than expected.(2)

Spreadsheets can also be manipulated to conceal fraud:

- Falsely linked spreadsheets resulted in a $691 million currency trading fraud at AllFirst Bank.(2)
- A rouge trader at Societe Generale manipulated spreadsheet reports to hide fraudulent trading activity.(2)
- The Ex-CFO of Voyager Learning Company used hidden rows in spreadsheets, which did not appear when printed, to conceal false accounting entries.(2)

Spreadsheets play an active role in the operation and financial management of most businesses. The benefits of using spreadsheets lie in their functionality, ease of use, and wide availability.(5) While the benefits of spreadsheets are obvious, they are powerful tools which should be used with caution. Economics Blogger James Kwak has correctly stated that spreadsheets "don't tell you when they break, they just give you the wrong number."(1) The following are risks associated with spreadsheet end-use:

- Access by Unskilled Users—Individuals who are not properly trained or not related to the process may have the ability to review or change data.
- Errors in the Spreadsheet—Cells either are not locked or output is not validated. Formulas or calculations are inaccurate.
- Lost Data—Critical data can be lost if periodic backups are not performed.
- Change Management—Changes to calculations are not properly reviewed and approved.(2)

In light of these risks, what steps can be taken to reduce spreadsheet risks at your organization? The following steps are considered to be key aspects of an effectively controlled spreadsheet environment:

1. DEVELOP A SPREADSHEET INVENTORY

All spreadsheets within the business that support important financial amounts and processes should be inventoried. This step will help ensure the population of spreadsheets in the organization is defined and subject to review.(6)

2. EVALUATE THE PURPOSE AND COMPLEXITY OF SPREADSHEETS

After the inventory is completed, the next step is to evaluate the use and complexity of each spreadsheet. This involves determining how a spreadsheet is used and assigning and documenting a level of complexity to the spreadsheet (low, moderate or high).(4)

Spreadsheet Complexity Increases Risk: Working in the insurance industry, I have had the privilege of working with many actuaries. As a non-actuary, one thing you notice about actuaries is their love for complex spreadsheets with a multitude of formulas, links, copies, tabs, vlookups, and pivot tables. With increased complexity, comes increased risk of error, and the need for increased controls.

3. DETERMINE CONTROLS NEEDED FOR THE SPREADSHEET

The appropriate combination of controls will help mitigate the risks inherent in a spreadsheet environment.(4) Examples of spreadsheet controls include:

- Applying version control techniques.
- Limiting user access to files (consider using a secured drive).
- Locking cells to prevent unauthorized changes.
- Using passwords to limit access to spreadsheets.(5)

A listing of required controls should be developed as appropriate for your organization.

4. EVALUATE EXISTING CONTROLS

The listing of required controls should then be compared to existing controls to determine if any gaps exist.(4)

5. PERFORM A SPREADSHEET BASELINE REVIEW

As part of evaluating existing controls, a spreadsheet baseline review should be conducted on high risk spreadsheets. A baseline review consists of testing the spreadsheet to ensure the data inputs and the overall design of the spreadsheet formulas are working as intended.(6)

6. CORRECT CONTROL GAPS

As these steps are completed, missing or ineffective controls may be identified. An action plan should be developed to correct the control gaps. The action plans should ensure the necessary controls are put in place and operating as intended, with the highest risk spreadsheets receiving top priority.(4)

Implementing these steps can help ensure that your organization has an effective spreadsheet control environment … and help reduce the risk of a spreadsheet fire.

(1) Kwak, James, "The Importance of Excel," www.baselinescenario.com, Feb. 9, 2013.

(2) "The Problem with Spreadsheets, A Business Case for Spreadsheet Management and Controls," www.sarbox-solutions.com, CIMCON Software, Inc., pages 5-7.

(3) Metz, Larry "Five Common Spreadsheet Risks and Ways to Control Them," www.theiia.org, Internal Auditor Magazine, October 2007.

(4) "The Use of Spreadsheets: Considerations for Section 404 of the Sarbanes-Oxley Act," www.auditsoftware.net, Price Waterhouse Coopers, July 2004, pages 4-6.

(5) "Implementing Effective Spreadsheet Controls," www.McGladrey.com, April/May 2011.

(6) Burdick, Tim, "Improving Spreadsheet Audits in Six Steps," www.theiia.org, Internal Auditor Magazine, March 2008. ▼

Darrick Fulton

Darrick Fulton, CPA, is manager, Insurance Advisory Services at the accounting firm Brown Smith Wallace. He can be reached at *DFulton@bswllc.com.*

I Held a Human Brain!

By Dave Snell

There I was, wearing surgical gloves, and holding a real human brain! This was a healthy one, as described by the professor of neurology teaching this part of the lab, and soon afterwards I got to hold and inspect others with various stages of dementia. Perhaps some of you readers might not find this a thrilling experience; but I was elated. After many months of studying what I could about biology and genetics to better understand artificial neural networks and genetic algorithms, I was finally seeing first hand, literally, the marvelous organ I had been trying to replicate.

During that same evening, over the course of three hours, I moved on to see and hold and feel a human heart, lungs, liver, and several other organs. Seeing a healthy lung and feeling its elasticity, and then feeling and seeing the ulcerated lung of a heavy smoker was enough to scare almost anyone thinking about smoking. In the life insurance world, we assume essentially standard mortality after a certain number of years of not smoking; but it was visibly clear here (and confirmed by the professor of pulmonology) that these holes and ugly scars would never heal—ever.

Why is this article in the F&F newsletter? Because I want to let you know about the many low cost, and often free, classes and other educational opportunities available to actuaries and anyone else interested in learning more about machine learning, and the biological systems we try to mimic with our various F&F modeling techniques.

Dave Snell

I have just completed all 24 weeks of the Mini Medical School classes at Washington University's Medical School—one of the finest in the country. The total cost was $150 for each of my three eight-week semesters and for that really low amount, I had personal interaction with world experts in various medical fields. Plus, I had hands-on labs on suturing (both laparoscopic and microscopic—with a 150 nanometer needle), a tour of their world-famous genome center, and many more experiences beyond my expectations.

Dave Snell, ASA, MAAA, is technology evangelist at RGA Reinsurance Company in Chesterfield, MO. He can be reached at *dsnell@rgare.com*

Lots of medical schools across the country offer this type of public service to their respective communities. Search for MiniMed and your school or city name to find what may be available in your area.

Note that this is not something I recommend as an alternative to the wonderful Med School for Actuaries that the SOA sponsors. I have attended one of those too; and the two types of schools complement each other rather than replace each other. Med School for Actuaries is much more condensed, and focuses on more-direct applications to health insurance claims and cost management.

The various medical training available now to the layperson far exceeds what used to be available only to those dedicated to many years of medical school, internship, residency, etc. to the exclusion of almost all other aspects of their lives during those years.

Likewise, the caliber of free courseware from Stanford, MIT, SFI, and other prestigious schools is amazing! I enrolled in a Machine Learning course from Stanford, a matrix algebra course from MIT, a complexity science course from the Santa Fe Institute taught by Melanie Mitchell (the author who rekindled my interest in complexity sciences years ago in her book Complexity: A Guided Tour), a neural networking course from the University of Toronto taught by Geoffrey Hinton (one of the pioneers in the field), and they all were free! Well, free of tuition cost anyway. The hours spent studying were hard and I sometimes start a course and decide afterwards that I just can't justify the time to do the assigned problems and keep up with the other items I want to enjoy in my life; but that's OK! Learning a little about a particular subject is better than learning nothing about it. I made some progress and kept that grey matter active in the learning process. Current thoughts about the plasticity of the brain suggest that continual learning helps defer the impact of dementia. I want my skull to be holding the healthy brain.

I recommend Coursera (www.coursera.com) and MIT's Open Courseware (www.ocw.mit.edu) as good starting points for seeing the thousands of courses available now for free. These tend to be at an undergraduate or graduate level and sometimes they assume some heavy mathematical or programming background. Sometimes, you can get a head start on that background with the free classes from the Khan Academy (www.khanacademy.org) which offers thousands of easier courses. Salman Khan earned three degrees from MIT and an MBA from Harvard, and then left his job at a hedge fund firm to provide free tutoring for his cousin, her family and friends, and ultimately, the world. Sal charges no fees, accepts no advertising revenue, and is funded by philanthropists such as Bill Gates, who takes Khan Academy courses along with his son, Rory, and who calls Khan "his favorite teacher." Obviously, Bill Gates can afford any teacher he wants. Tuition is no obstacle for him. You can have the same teacher—for free!

In my presentations on complexity science, I often stress the value of inductive reasoning to supplement our typical deductive reasoning. In those talks, I say that if you dissect a brain (the top-down deductive approach), you end up with mush; but if you build a brain (inductively, from simple rules and the interaction of adaptive agents), you gain insights and create knowledge.

I held a brain in my hand! You can build a brain—your brain! I hope you study complexity science and the associated machine learning and artificial intelligence topics I find so fascinating. But the opportunity is here now for you to study almost anything you wish—beyond just SOA study notes. Please do it! Hold a brain in your hand … or at least in your head. ▼

SOCIETY OF ACTUARIES Forecasting and Futurism Section

Forecasting & Futurism

NEWSLETTER

475 N. Martingale Road, Suite 600
Schaumburg, Illinois 60173
p: 847.706.3500 f: 847.706.3599
w: www.soa.org