



Article from

Actuary of the Future

May 2018

Issue 42

What Does Data Science Mean for the Future of Actuarial Science?

By Dustin Verzal

Big data, machine learning and artificial intelligence have been making headlines across industries everywhere, so it shouldn't be surprising to see the impact reach the world of actuarial science as well. It is of the utmost importance for us to understand how data science can influence the industry so we may take an active role in shaping that future.

ACTUARIAL HISTORY

To understand the future, let's first revisit the past to see how the actuarial profession developed into its current state. Unsurprisingly, people have been trying to quantify risk for quite some time. The first use of the word *actuary* dates back to 1762 when Equitable Life used the term to describe its CEO.¹ All modeling at the time was deterministic. Commutation functions were used to aid in calculations since calculators had not yet been invented.² The Actuarial Society of America (ASA) formed in 1889. The magnitude of the financial and ethical implications for actuarial work meant rigorous processes were necessary to qualify professionals and uphold esteem for the profession. Actuarial exam programs began in 1897.³ The ASA later played a part in the merger with the American Institute of Actuaries (AIA) to form the Society of Actuaries (SOA) in 1949.

In the 1930s and 1940s, the foundation for the stochastic theory was developed, which gave way to probabilistic loss models.⁴ The industry saw computation methods evolve from pencil and paper to punch cards to the personal computer systems we know today. Each step forward has improved the accuracy and reliability of the analyses actuaries perform. Regardless of the techniques used at a given time, they all follow a similar pattern, using historical data, mathematical models and judgment to make projections about the future.

WHAT IS DATA SCIENCE?

Enter data science—an emerging field touting exciting new ways to analyze and derive value from data. Determining what constitutes data science has engendered a bit of disagreement, as there is no consensus on the definition of its curriculum.

Broadly speaking, it is the unification of statistics and data analysis, which grew out of the need to grapple with tremendous volumes of data that are generated on a daily basis. What has emerged from this struggle is a group of people now referred to as “data scientists” (that is, a group of curious and analytical computer gurus). Data scientists use historical data, mathematical models and judgment to predict and understand the future. Sound similar to your profession? I thought so.

WHAT WILL DATA SCIENCE CHANGE?

This new wave of data analysis techniques brings a significant degree of automation to the development of predictive models. New algorithms include mathematical feature selection to identify independent variables with a valuable signal. Advanced decision-tree-based models allow for nonlinear relationships and interactions to be captured implicitly. Neural networks mimic the human brain's ability to find abstract relationships in data. Models can scale to millions of input variables with relative ease. These methods are being applied not only to the insurance industry, but also to the markets they support. Here are a few examples:

- **Auto.** Insurers are using machine learning on data captured via accelerometers and gyroscopes to predict driver risk.
- **Health.** Researchers are developing neural networks to analyze images to help identify cancer or heart disease.
- **Life.** Insurers are exploring the use of predictive modeling to provide real-time quotes on a limited set of underwriting variables.
- **Employers.** Predictive models rank workers' compensation claims into buckets by expected severity, which are then used to target claimants with an expected high-cost trajectory.



Although data science and actuarial science are technically two different fields, they are not mutually exclusive.

HOW IS IT RELATED TO ACTUARIAL SCIENCE?

Although data science and actuarial science are technically two different fields, they are not mutually exclusive. Both sides can benefit from collaboration and cross-training in each other's discipline. As many of you know, the SOA has recently altered the exam syllabus to include predictive analytics. It recognizes the valuable opportunities gained by applying these techniques to actuarial problems. I urge you to embrace this vision and explore beyond the exam requirements. Actuaries already possess the mathematical fortitude required to learn and apply data science and have access to extensive datasets. Most important, though, they already have expert domain knowledge of the insurance industry, allowing them to connect problems to the proper solutions. The upshot is that actuaries are in a unique position to leverage data science to answer extremely interesting and valuable questions, such as the estimation of incurred but not reported (IBNR) claims, the prediction of likely hospitalization, the prediction of pharmaceutical market opportunity and the prediction of substance abuse.

WHERE SHOULD WE START?

Many sources on the internet identify how to begin the journey into data science and predictive modeling. One search engine query will return more study material than you will likely have time for. The key is not to focus upfront on which algorithms or programming language to use, as they are just a means to an end.

I believe the first step is finding a problem that excites you—something that will fuel your appetite for learning. Explore how data science is being used in the world already and think about what types of problems you want to solve. If you need some help, check out *Kaggle.com*. Kaggle is a data science playground filled with predefined problems and datasets where people can build predictive models to compete for recognition and even cash prizes. Some of the competitions have even been put together by insurance companies seeking innovative solutions. I recommend that you explore some of these competitions to get a feel for the types of problems data science is trying to solve. I suspect you'll find something that speaks to you. After that, if you're still interested, read "Learning Data Science From an Actuary's Perspective" in the 40th issue of *Actuary of the Future*.⁵ The journey into data science will not be an easy one, especially if you're working on actuarial exams, but I promise the results will be rewarding. ■



Dustin Verzal is a data scientist at Milliman Inc. in Milwaukee. He can be reached at Dustin.Verzal@Milliman.com.

ENDNOTES

- 1 Ogborn, M. 1956. The Professional Name of Actuary. *Journal of the Institute of Actuaries* 82, no.2:233–246.
- 2 Slud, E. V. 2006. *Actuarial Mathematics and Life-Table Statistics*. College Park, MD: University of Maryland.
- 3 Hickman, J. 2004. History of Actuarial Profession. In *Encyclopedia of Actuarial Science*, edited by J. L. Teugels and B. Sundt. Hoboken, NJ: John Wiley & Sons.
- 4 Bühlmann, H. 1997. The Actuary: The Role and Limitations of the Profession Since the Mid-19th Century. *ASTIN Bulletin* 27, no. 2:165–171.
- 5 Li, X. 2017. Learning Data Science From an Actuary's Perspective. *Actuary of the Future* 40:10–12.