UNSW Business School/ Risk & Actuarial Studies

Data Analytics in Actuarial Education

Andrés M. Villegas
School of Risk and Actuarial Studies
UNSW Business School, UNSW Sydney

SOA Actuarial Teaching Conference
27 Jun 2017, Pittsburgh
Big Data, Data Science, Predictive Analytics, Data Analytics ...
Agenda

• Overview of data science and data analytics techniques

• The skills of data scientists

• Data scientist vs. Actuaries

• What the actuarial profession is doing

• UNSW’s experience incorporating data analytics in the curriculum

• Questions for discussion
Data Analytics

• Recent massive growth across many industries
• Driven by increased data availability, particularly “Big Data”
• Computer scientists developing software for data analytics (machine learning)
• Statisticians developing methods and packages for data analysis (statistical learning)
• Marketing and other business disciplines developing business applications
• Software availability with open-source packages for data analytics (Python, R etc)
• Increased number of books on techniques, particularly by practitioners
• Increased number of Masters degrees in Business Analytics, Data Science or Data Analytics
• Professional associations specializing in the area (Certified Analytics Professional CAP – INFORMS)
The Data Science Universe

Source: “An introduction to the data science universe” by Mark Lee
The Data Science Universe

Source: “An introduction to the data science universe“ by Mark Lee
Machine/Statistical Learning

Supervised Learning
Data with a dependent variable/classification

- Training Data
- Learning Algorithm
- New Data
- Model
- Prediction

Unsupervised Learning
Data exploration

- Data
- Algorithm
- New Representation

Source: “An introduction to the data science universe“ by Mark Lee
What's different about Data Analytics

Rigorous approach to model selection and validation
• Data visualization and exploration
• Allocating data to model selection (training set) as well as evaluating model performance (includes resampling techniques)
• Over fitting versus predictive power

Using data to make decisions, relying less on model assumptions
• Unsupervised and supervised machine learning techniques

Emphasis on predictive models rather than explanatory models

Potential to incorporate “Big data”
• Individual and less structured data
• More “features”, dimension reduction and PCA
• Model robustness, cross validation
Data science skills

Data science skills

- Computer science
- Math
- Statistics
- Machine learning
- Domain expertise
- Communication and presentation skills
- Data visualization
Do actuaries need to train?

Google trends

- Actuary
- Data scientist
Back in 2013 – Best jobs in America, Forbes

Source: https://www.forbes.com/pictures/efkk45eifhm/no-1-best-job-actuary/#485a3db7e159
In 2016 – Best jobs in America, Forbes

1. Data Scientist
Annual Median Salary: $128,240
Growth Outlook: 16%

Source: https://www.forbes.com/pictures/570fbcc9e4b01d5c5870282f/1-data-scientist/#2e2f79e3bd99
In 2017 – Best jobs in America, Forbes

1. Data Scientist
   - Job Score: 4.8
   - Job Satisfaction Rating: 4.4

Source: https://www.forbes.com/pictures/5886923c4bbe6f1f20eb79d9/1-data-scientist/#13009a9e5f82
Are actuaries the data scientist of insurance?

+ Core technical training
+ Analytical thinking and quantitative skills
+ Ongoing training
+ Communicating insight to the business
+ Domain knowledge

- Limited skills to deal with unstructured data
- Unlikely to be better programmers
- Lacking the visual representation ability

Are actuaries the data scientist of insurance?

Actuary + Data & Tech = Data Scientist

What is the profession doing?
IAA Syllabus

Data and Systems

**Aim:** To enable students to *apply methods from statistics and computer science* to real-world data sets in order to *answer business* and other questions, in particular with application to questions in long and short term insurance, social security, retirement benefits, healthcare and investment.

**Data as a resource for problem solving**
- Aims, stages, tools, sources (including extremely large data sets)
- Common data structures and data storage systems
- Tools for cleaning, restructuring and transforming data

**Data analysis**
- exploratory data analysis, summary statistics and undertake exploratory data visualizations
- PCA, distributions, linear regression, survival models, GLMs (Including implementation in a computer)

**Statistical Learning**
- Supervised and unsupervised, types of problems,
- Commonly used techniques, neural networks and decision trees

**Professional and Risk Management issues**
- Ethical and Regulatory issues, data governance, risks

**Visualising data and reporting**
- Visualisation to communicate key conclusions
- Reproducible research and reproducible data analyses
Predictive analytics has spread to most areas of actuarial practice.

Actuaries need to know more than the basic regression and time series methods.

Incorporate predictive analytics topics to better prepare SOA trained actuaries for today’s opportunities.
What is the profession doing?
ASA and CERA Curriculum changes

EXAM SRM – Statistics for Risk Modeling
• Basics of Statistical Learning (7.5-12.5%)
• Linear Models (40-50%)
• Time Series Models (12.5-17.5%)
• Principal Components Analysis (2.5-7.5%)
• Decision Trees (10-15%)
• Cluster Analysis (10-15%)

Exam PA – Predictive Analytics
• Model Building Process (10-20%)
• Problem Definition, Exploratory Data Analysis, and Initial Model Selection (15-25%)
• Model Selection (20-30%)
• Model Validation (15-25%)
• Communication of Results and Uncertainties (15-25%)
UNSW’s School of Risk and Actuarial Studies
UNSW’s School of Risk and Actuarial Studies
UNSW’s School of Risk and Actuarial Studies
UNSW’s School of Risk and Actuarial Studies
Our faculty

- A/Prof. Anthony Asher
- A/Prof. Benjamin Avanzi
- Prof. Hazel Bateman
- A/Prof. Ram Bhar
- Dr. Katja Hanewald
- Dr. Katja Ignatieva
- Dr. Kevin Liu
- Prof. Michael Sherris
- Prof. Greg Taylor
- Dr. Andrés Villegas
- A/Prof. Bernard Wong
- Dr. Jinxia Zhu
- Dr. Jonathan Ziveyi

- 4 Postdoctoral researchers
- Eric Cheung, JK Woo, Qihe Tang joining in July 2017
UNSW’s School of Risk and Actuarial Studies

Our Programs

Undergraduate
- Bachelor of Actuarial Studies
- Bachelor of Actuarial Studies (Co-op)
- Bachelor of Actuarial Studies / Bachelor of Commerce
- Bachelor of Actuarial Studies / Bachelor of Economics
- Bachelor of Actuarial Studies / Bachelor of Laws
- Bachelor of Actuarial Studies / Bachelor of Science
- Bachelor of Actuarial Studies / Bachelor of Science (Advanced Maths)

Postgraduate
- Master of Actuarial Studies
- Master of Commerce (Risk Management Specialisation)

Research
- MPhil in Risk and Actuarial Studies
- PhD in Risk and Actuarial Studies

Professional recognition
- Exemptions for Part I and II of the Actuaries Institute Professional Examinations (Australia)
- Exemptions for Core Technical courses of the Institute and Faculty of Actuaries (UK)
- Validation by Educational Experience credit for the Society of Actuaries (USA)
- SOA Center of Actuarial Excellence since 2016
UNSW’s School of Risk and Actuarial Studies

Our Students (Undergraduates)

Total Students

1096

Distinct Headcount by Citizenship Country

Students in 2017 Cohort

329

Distinct Headcount by International or Local

International 220 (20.07%)

Local 876 (79.93%)

Distinct Headcount by Attendance Type

Semester 1

Enrolled Full-Time 96%

Enrolled Part-Time 0%

Distinct Headcount by Age Group

a. <21 83.05%
b. 21-25 16.5%

Distinct Headcount by Gender

Female 30.93%

Male 69.07%
Integration of Data Analytics in the Curriculum

- Led by recent advances in the profession
- Follow recommendation from our Industry Advisory Board
- We are implementing the following initiatives
  1. Redesign of course ACTL1101 Introduction to Actuarial Studies
  2. Systematic integration of R and Excel throughout the curriculum
  3. New course ACTL3142 Actuarial Data and Analysis
ACTL1101 Introduction to Actuarial Studies
Course Summary

• 300+ first year students
• 2017 is first cohort where R will be fully integrated (and assessed) through out their studies
• 2 hour lecture with the whole group
• 1 hour Tutorial (25 Students)
• 1 hour Lab (25 Student)

“This course is designed to provide an introduction to actuarial studies. It covers the fundamental modelling tools used by actuaries (probability, statistics, financial mathematics), as well as some of the basic actuarial models in areas such as insurance, superannuation or financial risk management, and which will be studied in great depth during the remainder of the degree. The main areas of actuarial practice and research are also introduced and discussed. Finally, labs will provide a foundation in programming, as well as data manipulation and visualisation, with a particular focus on R.”

A/Prof. Benjamin Avanzi
Associate Head of School
1. Evaluate and apply basic principles of probability, statistics and financial mathematics;
2. Evaluate the fundamental principles underlying risk management and insurance
3. Evaluate and apply fundamental actuarial mathematics techniques;
4. Describe how the actuarial profession is organised, its code of conduct, its main practice areas, as well as its current challenges and opportunities;
5. Interpret and create basic algorithms and control loops in pseudocode;
6. Communicate data insights effectively;
7. Perform efficient computation, as well as manipulate data, in R;
8. Work effectively in teams;
ACTL1101 Introduction to Actuarial Studies

Learning activities and content

Lectures:
• Fundamental tools and actuarial techniques
• Discuss practice areas, professionalism, and recent issues in the actuarial world

Labs
• Foundation in programming
• Data manipulation and visualisation
• Focus on R

Lab Content (Selected)
• Algorithms and pseudo-code
• Introduction to R and R Studio
• Data structures, import, export, and control flow in R
• Vectorisation, operations and distributions in R
• Visualisation: theory
• Basic visualisation tools and management of plot in R
• Advanced visualisation with ggplot2 and Shiny
• R functions and matrix operations
• Numerical and symbolic methods, and efficient calculations in R
ACTL1101 Introduction to Actuarial Studies

Additional points

• Introduce practice areas with interviews of past UNSW graduates with senior industry positions
  - **Predictive analytics:** Jon Shen, FIAA, Senior Manager, Data Science at Suncorp Chief Data and Transformation Office

• Individual lab mini-assignments to be solved using R

• Major group assignment
  - Develop skills in communication, R and teamwork
  - Tasks involves exploratory data analysis and visual communication
ACTL3142 Actuarial Data and Analysis
Course Summary

- Course developed by A/Prof. Bernard Wong
  - Accreditation Actuary
  - Member of the Australian Actuaries Institute Working Group
  - Head of School from July 2017
- First offered in 2016 as a level 3 elective for undergraduate and postgraduate students (50+ Students)
- Part of BActSt major in Actuarial Risk Management and Analytics and new UNSW degree in Data Science and Decisions

“This course covers the techniques in data analysis including techniques for mortality, health, and insurance data used in actuarial analysis and decision-making. Particular focus will be on techniques often referred to by ‘predictive analytics’ and ‘statistical learning’, and their applications to actuarial work.”
ACTL3142 Actuarial Data and Analysis
Education and Strategic aims

• Provide knowledge of **modern data analytic techniques** to future actuaries

• Raise interest and develop students’ capabilities for **lifelong learning** in this developing area - which is moreover viewed as the **future of actuarial science**

• Allow the school and UNSW to position itself as the **leading provider of actuarial education in data analytics**
ACTL3142 Actuarial Data and Analysis
Student Learning Outcomes

1. Understand aspects of the theory and practice of predictive analytics/data analytics for insurance and financial applications
2. Assess models used for predictive analytics/data analytics in practice and their advantages and shortcomings
3. Estimate and apply various statistical learning models for practical applications
4. Understand and explain ethical and regulatory issues associated with the use of data and analytic techniques
5. Use effective presentation, discussion and report writing skills for explaining risk-modelling concepts used in quantitative risk management
ACTL3142 Actuarial Data and Analysis

Course content

1. Linear Regression techniques: Simple/Multiple; K-Nearest Neighbours
2. Classification Methods: Logistic regression, Discriminant Analysis
3. Resampling Methods: Cross Validation and Bootstrap
4. Linear Model Selection and Regularization: Subsets, Shrinkage, and Dimension Reduction
5. Non-Linear methods: Polynomial Regression, Splines, and Generalized Additive Models
6. Decision Trees: Bagging and Boosting
7. Support Vector Machines and Classification
8. Unsupervised Learning Methods: PCA and Clustering
9. Professional Ethics in Data Modelling
ACTL3142 Actuarial Data and Analysis
Course textbook
ACTL3142 Actuarial Data and Analysis
Course textbook

- Web site  
  http://www-bcf.usc.edu/~gareth/ISL/
- Excellent text, great for undergraduate course for actuarial students
- R labs with detailed explanations
- Several MOOCS based on the book:
ACTL3142 Actuarial Data and Analysis
Tailoring of material to actuarial students

- Links to different topics in actuarial science
  - Survival models and life tables
  - GLMs for insurance pricing
  - Curve fitting and yield curves in financial mathematics
- Provision of direct actuarial applications
  - e.g. fraud detection and pricing
- Mathematical sophistication was made in line with the actuarial prerequisites
- Professional industry presentations by actuaries working in data analytics.
  - Analytics in General Insurance, Jon Shen, Suncorp group
  - Data Science at Quantium
- Ethical aspects of data analysis
ACTL3142 Actuarial Data and Analysis
Learning and teaching approach

Focus on deep understanding of key concepts
• Focus of this course lies not in the mathematical proofs of the results per se

• Focus in the understanding of the concepts and key ideas underlying modern data analytic techniques and their relation to actuarial methods and applications

• Mathematics are used as a tool to aid understanding

• Enable deep learning and facilitate lifelong learning and continuous professional development of the students.
ACTL3142 Actuarial Data and Analysis
Learning and teaching approach

Delivery via Active Learning

• The course is structured using a combination of lectures and labs

• The lectures are structured such that significant amount of time is spent on exploration of intuition and active learning

• Labs involve students applying techniques.

• A major assignment: students learning about a new technique from the academic literature (elastic net), and then applying it (along with other techniques discussed in class) and providing a report on the findings of their analysis
ACTL3142 Actuarial Data and Analysis
Learning and teaching approach

Practical Application/Implementation using R

• Application of course concepts in software

• Actuarial judgement and communication of results

• Tutorials are held in labs, and code and video of programming sections are also provided
ACTL3142 Actuarial Data and Analysis
Student feedback

Student feedback for the course has been exceptional

“Practical content presented in concise manner”

“Easy to understand the concepts as we are slowly eased into them. The focus on qualitative work is much appreciated as it helps us to understand actuarial work better”

“One of the more practical courses that is becoming more important in recent times. Went into detail of things that could actually be used in application in the real world, especially in combination with the tutorial classes using a statistical program R”
Discussion

What role do data analytics concepts play as part of the core actuarial skill set?
What, where and how much should we be included in the actuarial syllabus?
What opportunities are there for application within actuarial areas of practice and broader areas?
How important is the availability of more individual data versus aggregate data for data analytics in the actuarial syllabus?
Is data analytics a core part of the actuarial skill set? What differentiates the profession from other data analytics professionals?
What are the interesting research questions for actuarial researchers in data analytics? What are the best methods to use for actuarial data?