Predictive Modeling: A Day at the Zoo?
A Day at the Zoo: Speakers

Douglas Wing
AVP, Analytic Products

Hernan L. Medina, CPCU
Manager, Personal Property Actuarial

Hema Chaurasia, CPCU
Product Manager
A Day at the Zoo

- Welcome to the Zoo
- Modeling Field Station
- Visitor Center
- Wildlife Trivia
- Modeling Habitat
- Taste of Africa
- Down Under
- Sea Lion Show
- Aquarium
- Bug Parc
- Aviary
- Display Birds
- Wildlife of Britain
- Petting Zoo
- Down on the Farm
A Day at the Zoo: Agenda

• Section I: Welcome to the Zoo
  ➢ Introduction to predictive modeling

• Section II: Modeling Habitat
  ➢ Regulations
  ➢ Impact on stakeholders

• Section III: Model Evaluation Field Station
  ➢ External validation
  ➢ Model validation charts

• Section IV: Wildlife Trivia

• Section V: Visitor Center (Questions & Answers)
A Day at the Zoo: Welcome to the Zoo
Predictive Modeling in Insurance

• What is predictive modeling? Truths and myths.

• Where are predictive analytics used in insurance?

• What kinds of benefits do carriers see?
What is Predictive Modeling?

- Finding relationships through a collection of mathematical techniques
- Using historic data to better predict future events
- Finding patterns in data to make better business decisions
Predictive Modeling: Truths

• Truths:
  ➢ Better predict future events
  ➢ Support business decisions
  ➢ Uncover new insights

Applications of Analytics
✓ Box Office Success
✓ Portfolio Management
✓ Sports
✓ Weather Forecasting
✓ …………
✓ …………
✓ …………
✓ Insurance
Predictive Modeling: Myths

• Myths:
  – Patterns are easy to detect in data
  – Models will tell me the answer
Predictive Modeling in Insurance

- Predictive modeling is used in:
  - Pricing
  - Underwriting
  - Claims
  - Marketing

[Bar chart showing frequency of predictive modeling use by function]
Predictive Modeling Benefits

• Why the emphasis on predictive analytics?

• By-Peril Rating Leading the Way

  ➢ 25 carriers using by-peril plans have increased market share from **28 percent** to **34 percent** in 6 years.

  ➢ Companies rating by-peril have loss ratios **6.6 points** lower than their competition, **69.8 percent** versus **76.4 percent**.
Predictive Modeling in Insurance

- Enhancements outside of pricing/underwriting
  - Marketing—spending dollars on the right targets
  - Claims handling—create efficiencies for better processing
  - Agency scores—determine level of control each agent gets
A Day at the Zoo: Modeling Habitat
Origins of Predictive Modeling

- Code of Hammurabi, 1772 B.C.

- Inhabitants of Rhodes, first millennium B.C.
  - General average

- Great Fire of London, 1667
  - Thirteen hundred homes destroyed
**Insurers’ Real Concern – What is Really Important to Various Stakeholders**

**SALES**

- **Sales Growth**
  
  \[
  \text{Sales Growth} = \left( \frac{\text{Sales in Yr2} - \text{Sales in Yr1}}{\text{Sales in Yr1}} \right)
  \]

**SERVICE**

**DEAL goods**

**Soft Market**

**Demand**

---

An affiliate of The Institutes
Finance Division’s main concerns

Financial Reporting
fair reporting
FRAUD
Management
PROFIT
Audit
Securities Act
Financial Condition

Cash Flow
Leverage Ratio
Chief Operation Officer

- Performance
- Profits
- Revenue!!!
- Expenses
- Routine Reports
- Labor
- Operations
- Presence
- Cost
Laws, Regulation, Bulletins....

- Are you **compliant**?
Definitions

• **Predictive Modeling**: Process by which a model is chosen to *try* to best *predict* the probability of an outcome (defining the target variable).

• **Programmer**: An organic machine that converts caffeine into source.

• **Actuary**: Versicherungsmathematiker? A wizard of odds?

• **Claims Professional**: Claims to be a professional.
Coming together is a beginning; keeping together is progress; working together is success. – Henry Ford
A Day at the Zoo: Model Evaluation Field Station
Model Evaluation Field Station

• Model evaluation techniques
  ➢ Internal validation
  ➢ External validation
  ➢ Validation charts
    • Decile charts (including loss ratio charts)
    • Gini index charts
• Objectives

➢ Understand the difference between internal and external validation

➢ Learn a bit about external validation techniques

  • Decile charts and loss ratio charts

  • Gini index and Gini index charts

➢ Comprehend the meaning of some predictive modeling key words
Model Evaluation Field Station: Internal Validation

- Internal validation

- An evaluation of a model’s classification effectiveness performed using data that was used to build the model
Model Evaluation Field Station:
Internal Validation

• Bootstrap Technique
  ➢ Create a large number of data sets by taking random samples with replacement from the original data
  ➢ Use these data sets to estimate the accuracy of model parameters or goodness of fit measures
• External validation

➤ An evaluation of a model’s classification effectiveness performed using data that was not used to build the model

Key Words
Model Evaluation Field Station:
External Validation

- External validation
  - Training sample
  - Validation sample
  - Test sample
• Decile charts

Charts built by sorting validation or test data by model prediction, subdividing the data into ten groups of equal size, and comparing average actual versus actual predicted results

Key Words
Model Evaluation Field Station:
Decile Charts

• Sort by model prediction
• Subdivide into groups of equal size
• Compare actual and predicted
Model Evaluation Field Station: Decile Charts

- What to look for:
  - Increasing values
  - Close fit
  - High versus low
Model Evaluation Field Station:
Decile Charts

Sorted by Loss Cost
Underlying Current Rates

Sorted by New Model’s Predicted Loss Cost

Actual
Current
Model
• Loss ratio chart

➤ A chart built by sorting validation or test data by model prediction, subdividing the data into ten groups of equal size, and calculating the loss ratio for each group
Model Evaluation Field Station: Loss Ratio Charts

- Bar chart: Predicted Loss Cost Decile vs Loss Ratio
- Line chart: Actual vs Model Loss Ratio

CPCU Society
An affiliate of The Institutes
Model Evaluation Field Station:
Gini Index

- Named after Corrado Gini
- Developed for economics
- Adapted for model validation
Gini index = \frac{A}{A + B}
• Gini index

➢ The ratio of the area between the line of equality and the Lorenz Curve to the total area below the line of equality
Model Evaluation Field Station: Gini Index

- “Perfect” car insurance model
- Claim frequency = 5 percent
- 95 percent have $0 loss cost
- Gini index close to 100 percent

![Graph showing a Lorenz curve with a Gini index close to 100 percent]
Model Evaluation Field Station: Gini Index

- Realistic model
- Prediction = expected loss cost
- High Gini index:
  - There is loss cost inequality
  - Model reflects it well
Model Evaluation Field Station:
Resources

• Predictive Analytics:
  - The Power to Predict Who Will Click, Buy, Lie, or Die
  - Eric Siegel, Ph.D.

• The Institutes white paper on predictive modeling
  - Charles Nyce, Ph.D., CPCU, API
  - [http://www.theinstitutes.org/doc/predictivemodelingwhitepaper.PDF](http://www.theinstitutes.org/doc/predictivemodelingwhitepaper.PDF)

• And the Winner is…? How to Pick a Better Model
  - Hernan L. Medina, CPCU, AIM, API, ARC, AU
  - Dan Tevet, FCAS
  - [https://cas.confex.com/cas/rpms14/webprogram/Session7061.html](https://cas.confex.com/cas/rpms14/webprogram/Session7061.html)
• Two methods of model validation
  ➢ Internal validation
  ➢ External validation

• External validation techniques
  ➢ Decile charts (including loss ratio charts)
  ➢ Gini index and Gini index charts

• Predictive modeling key words
A Day at the Zoo: Wildlife Trivia
A Day at the Zoo: Visitor Center
Questions?