

# The Process of Predictive Modeling

Dorothy L. Andrews  
Merlinos & Associates



Conceive



Develop



Test



Implement



Analyze

# The Process of Predictive Modeling



## My Background

- Actuary, ASA, MAAA
- MA – Mathematical Statistics
- MA – Mathematics & Education
- BA – Mathematics
- Everett Curtis Huntington Prize for Best Actuarial Research Paper
- Outstanding Teaching Award 2002 – BU

## Predictive Modeling Practices

- United States Department of Agriculture
- Transamerica Reinsurance Company
- Harleysville Insurance Company
- Wachovia Corporation
- Deloitte & Touché LLP
- John Hancock Life Insurance

- PA Surveillance Model for the USDA
- Pathogen & Residue Sampling Techniques for Food Safety
- Big Dataset Handling: Banking & Reinsurance
- Scorecard Development to Evaluate Borrower Credit Risk
- Text Mining of Adjuster Notes for WC Predictive Modeling
- Analysis of Credit Data: Experian and Dun & Bradstreet
- Development of Enterprise Modeling Database
- Big Dataset Handling: Banking & Reinsurance
- Develop & Implement Commercial Auto Predictive Model
- Annuity Relational Database Development for Actuaries

# The Process of Predictive Modeling



## Today's Agenda

- Importance of Stakeholders
- Model Building Framework
- Model Validation & Monitoring
- Current Technologies for Modeling
- The Regulatory Environment

# The Process of Predictive Modeling



Vs.



# The Process of Predictive Modeling



## Important Company Stakeholders & Partners

- C-Suite & Other Senior Management
- Actuarial Pricing Department
- Underwriting Department
- Information Technology Group
- Agents and Brokers

# Underwriter Survey Results



## Underwriter's Perspective on Predictive Modeling:

- Welcome the opportunity to participate in review of the variables
- A small focus group to examine the intuitiveness of the new model
- The models should free underwriters to solve hard problems
- Agents need to be convinced that the models are valid for pricing
- The models need to identify previously undiscovered opportunities
- The underwriter needs to believe the models to get more rate

# Impact of Excluding Underwriters



- Unclear on the utility of the model and most ignored model results
- Their field experience was not considered in the model design
- Variable results ran counter to how policyholders behave
- Little was communicated on the practicality of the model
- Some believed there is value but didn't understand what it was
- Implementation of models was fractured across underwriters

# Effective Partnering with Underwriters



- Underwriters do more than process information
- Recognize their insights as adding value to the process
- Partner to develop underwriting guidelines and in model building
- Use their intuition to interpret and “reality check” model results





# Effective Partnering with Underwriters



- Discuss the strengths and weaknesses of the model with them
- Get the simple cases right so underwriters can focus on emerging risks
- Use their frontline observations of risks in variable development
- Don't overlook the small stuff! – Doing that can lead to big problems!





CASUALLY ACTUARIAL SOCIETY



CPCU  
SOCIETY™

An affiliate of The Institutes

## Seminar: Expanding The Toolset – Underwriting Collaboration

October 22 – 23, 2015  
Boston, MA



### Sample Topics:

1. Underwriting & Actuarial Collaboration
2. Underwriting Decision-Making
3. Tools for Risk Assessment & Evaluation
4. Model Risk Management
5. Predictive Analytics & Pricing
6. Catastrophe Underwriting
7. Telematics & Usage-Based Insurance
8. Emerging Risk Assessment & Planning

# Pre-Model Construction Considerations



1. Understanding the market environment for model
2. Understand the way participants think about risk factors
3. Identify variables to treat stochastically vs. deterministically
4. Develop qualitative interrelationships among variables
5. Identify industry data on independent variables
6. Identify company obstacles to successful model execution

\*Source: Goldman Sachs, Quantitative Strategies Research Notes, “Model Risk,” April 1996

# Pre-Model Construction Considerations



7. Identify external systematic risks impactful to results
8. Develop actions to mitigate adverse risks
9. Demonstrate the financial benefits of predictive models
10. Identify hurdles to organizational knowledge transfer

# The Process of Predictive Modeling



## Predictive Modeling Phases:

Phase 1: Define Problem & Financial Impact of Solution

Phase 2: Understanding the Product to be Modeled

- Coverages
- Distribution Channels
- Underwriting System
- System Data Captured

# The Process of Predictive Modeling



## Predictive Modeling Phases:

### Phase 3: Identify Internal & External Data

- Major Issue: Legacy Systems

### Phase 4: Formulate Pre-Modeling Parameters

- Policy Term vs. Policy
- Time Period of Data
- Credit Data Proxy Development
- Variable Development & Analytics

# The Process of Predictive Modeling



## Predictive Modeling Phases:

### Phase 5: Iterative Data Scrubbing & Cleansing

- Missing & Mis-specified Data Fields
- Combine Internal & External Data
- Control Total Reconciliation
- Data Accuracy Criteria
- Visual & Descriptive Analytics
- Legal & Underwriting Review of Variables

# The Process of Predictive Modeling



## Predictive Modeling Phases:

### Phase 6: Model Construction

- Frequency x Severity
- Tweedie Distribution
- Cluster/Principal Components
- Machine Learning Using C4.5
- Generalized Linear Models
- Stepwise Variable Elimination



# The Process of Predictive Modeling



## Predictive Modeling Phases:

### Phase 7: Interpretive Model Support Analytics

- Correlation Matrices
- ROC & Decile Lift Curves
- Disruption Analysis
- Probability Plots
- Model Fit Metrics
- Validation Review

# The Process of Predictive Modeling



## Predictive Modeling Phases:

### Phase 8: Systems Integration, Testing & Implementation

- Dashboard Analytics
- Reason Code Development
- Convert Scores to Business Rules
- Independent Validation of IT Implementation
- Development of Monitoring Metrics
- Develop Documentation & Training Material

# The Process of Predictive Modeling



How do we know the model is “right”?



Recognizing the “wrong” model is easier than qualifying the “right” model.

Model validation can help assess whether a model is a reasonable representation of the phenomena under study. But remember the model is only a model of the phenomena and not the real thing. This means we can’t remove the human element from the modeling process.

# Types of Model Risk



Model Risk -“The risk of loss by using a model to make financial decisions.”

– Goldman Sachs

1. Inapplicable Model
2. **Incorrect Model**
3. Correct Model, Incorrect Solution
4. Correct Model, Inappropriate Use
5. Badly Approximated Model
6. Software and Hardware Bugs
7. Unstable Data



# Signs Model May Be Incorrect



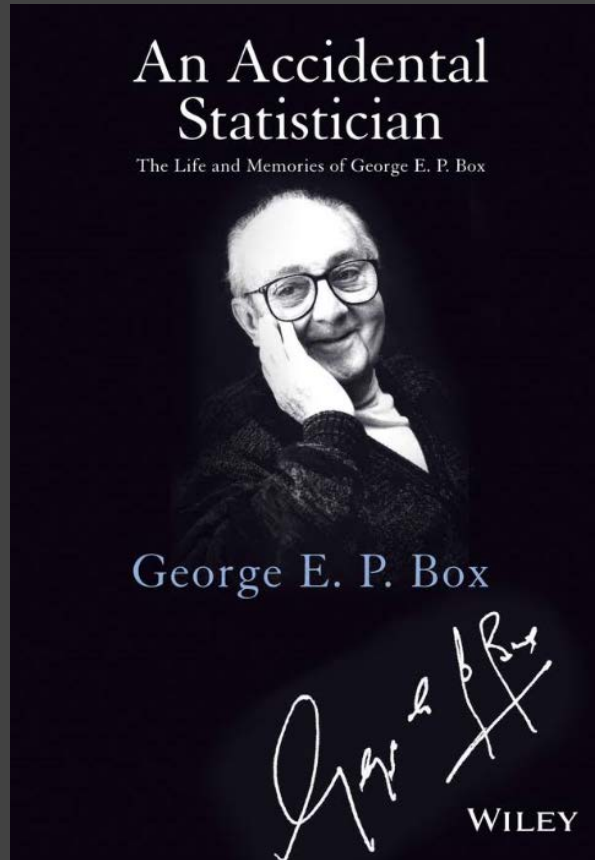
1. Important factors not considered
2. Stochastic variables estimated with deterministic variables
3. Incorrect dynamics assumed for a factor
4. Model inappropriate for current conditions
5. Correct under ideal circumstances

# Signs Model May Be Incorrect



6. Only approximate under real-life situations
7. Correct in principle but limited in the short term
8. Underlying data not robust enough to fully model phenomena
9. Model reasonable, but world unstable

# The Process of Predictive Modeling



“Essentially, all models are wrong,  
but some are useful.”

- George E.P. Box

# The Process of Predictive Modeling



## Mitigating Model Risk:

1. Regard models as interdisciplinary endeavors
2. Know what constitutes a warning sign
3. Test complex models using the simple cases first
4. Test the model's boundaries



# The Process of Predictive Modeling



## Mitigating Model Risk:

5. Don't ignore small discrepancies
6. Provide a good user interface
7. Diffuse the model slowly to an expanding group
8. Pride of Ownership – Model builders should like to build models

# Current Technologies for Modeling



Technology	Programming	GLM Method	Machine Learning	Documentation
R	Yes	Yes	Yes	Yes
Python	Yes	Yes	Yes	Yes
SAS	Yes	Yes	Yes	Yes
DataRobot	No	Yes	Yes	Only on Screen
Skytree	No	No	Yes	Only on Screen
Talon	No	Yes	Yes	Yes – Excel Files
Emblem	No	Yes	No	Yes

# The Mission of the Regulatory Process



- Prevent unfair discrimination in risk selection & pricing
- Regulatory are keenly interested in
  - Model variables
  - Reliance on other models
- Any model relied upon must be included in the filing
- Regulators disallow unfairly discriminatory variables & models
- Consult with legal counsel on variable selection

# The Process of Predictive Modeling



## Data Sources to Avoid in P&C Modeling?:

1. Consumer Transactions – B2C Data
2. Pharmaceutical Transactions
3. Political Affiliations Data
4. Aggregate Household Information
5. Individual Household Information
6. Customer – Focused Marketing Data

Acxiom's data and technology have transformed marketing – giving our clients the power to successfully manage audiences, personalize customer experiences and create profitable customer relationships.

[www.acxiom.com](http://www.acxiom.com)

## Marketing Data Sources:

- [www.infousa.com](http://www.infousa.com)
- [www.nielsen.com](http://www.nielsen.com)
- [www.infogroupmediasolutions.com](http://www.infogroupmediasolutions.com)
- [www.wpcurve.com](http://www.wpcurve.com)

# Questions?

