2019 Predictive Analytics Symposium

Session 18: AP - Using Predictive Models for Life Insurance Assumptions

RGA

Using Predictive Models for Life Insurance Assumptions

Dihui Lai, PhD, ASA

Sept. 2019

Overview

- Assumptions of life insurance
- Modeling structure: A/E model v.s. incident rate model
- Case study: B.E. model
- Case study: principle based reserve model
- Case study: pricing assumption settings

Assumptions of Life Insurance

- Pricing
- Statutory reserve
- Best estimate (B.E.)
- Lapse

A/E Model v.s. Incident Rate Model

- A/E model:
  - Use this type of models to understand the deviation of actual experience from the current assumption settings.
  - Example model structure:
    - $\overline{A} = \overline{E} * e^{X\beta}$
    - $e^{X\beta}$ is interpreted as the overall model suggested adjustment (msadj)
    - msadj $= \exp\left(X_{age}\beta_{age}\right)\exp\left(X_{gender}\beta_{gender}\right)\ldots\exp\left(X_{cmpy_{grp}}\beta_{cmpy_{grp}}\right)$
- Incident rate model
  - Use this type of models to understand mortality when no prior knowledges of assumption exist.
  - Example model structure
    - $\overline{A} = \overline{EXPOSURE} * e^{X\beta}$
    - $e^{X\beta}$ is interpreted as the adjustment needed for each factor that is included in the model

Basic Modeling Techniques

- Regression model v.s. advanced machine leaning algorithms
  - Generalized linear model
  - Random forest; neural network (are they really needed?)
- Variable selection criterion: AIC; p-value
  - Use AIC to balance the measure between model complexities and goodness-of-fit.
  - Use p-value to access the statistic significance of each individual variable.
  - Business implication, ease of implementation and compliance (e.g. degree of freedom of modeling v.s. implementation).
- Feature engineering
  - Spline; polynomial transformation;
  - Piecewise;
  - Regrouping

Case Study: B.E. Mortality Study

- Background
  - The actuarial team has an existing B.E. mortality assumption setting and would like to verify it using predictive modeling.
- Challenges
  - The current assumption setting is table-based and each table may contain adjustments on multiple variables. How can we design a modeling ‘process’ to assess the current adjustments?
  - Interpretability is critical.
  - Large data set that cannot be efficiently handled with open source R packages.
- Solution
  - A multi-stage model process to evaluate the existing adjustment tables one-by-one.
  - Hadoop based parallel computing.

Multi-Stage Model (Simplified)

Variables	Base-model	Stage-1 Model	Stage-2 Model	…
Reference Assumption	Assumption-Base	Base*Table1	Base*Table1*Table2	B.E.
Age	✓	-	-	-
Gender	-	-	-	-
Var1	✓	✓	-	-
Var2	✓	✓	✓	-
. . .	. . .	. . .	. . .	. . .
VarN	✓	-	-	-

Handling Big Data Using Parallel Computing

- Modeling process requires hundreds of iteration or even more.
- The capability of building models quickly is practically very important
- Allows real-time communication and getting feedbacks from audience.
- Speed comparison on a data of size ~ 40 GB; ~ 57 millions records

System	SQL Query	GLM Model
Terminal Server (Windows Server)	90 seconds	350 seconds
Distributed Cluster System (Hadoop Based)	5+ seconds	12 seconds

Learnings/Conclusions

- Designed a diagnostic modeling ‘process’ to assess current B.E. assumption settings.
- Discuss and accommodate changes when appropriate.
- Only a fraction of variables show statistic significance, implying the current assumption is mostly efficient.
- Certain underwriting class shows experience deviating from current assumption and the model suggests for adjustments.
- Downward trends against a few time variables, including calendar year and issue year. Does this imply mortality improvement or is it due to newer policies?
- Be patient and collect more data.

Case Study: Principle Base Reserve

Background

- Generates PBR mortality assumption for valuation team, using predictive model.
- Would like to use existing industry table as the reference line.

Challenges

- No prior works.
- A purely model-based solution may overlook the business implications and could lead to overfitting.

Solution

- An A/E model that can generate adjustment table.
- Work with the actuarial team to incorporate their insights to avoid ‘overfitting’.

Learnings/Conclusions

- Standard A/E model provides a good starting point

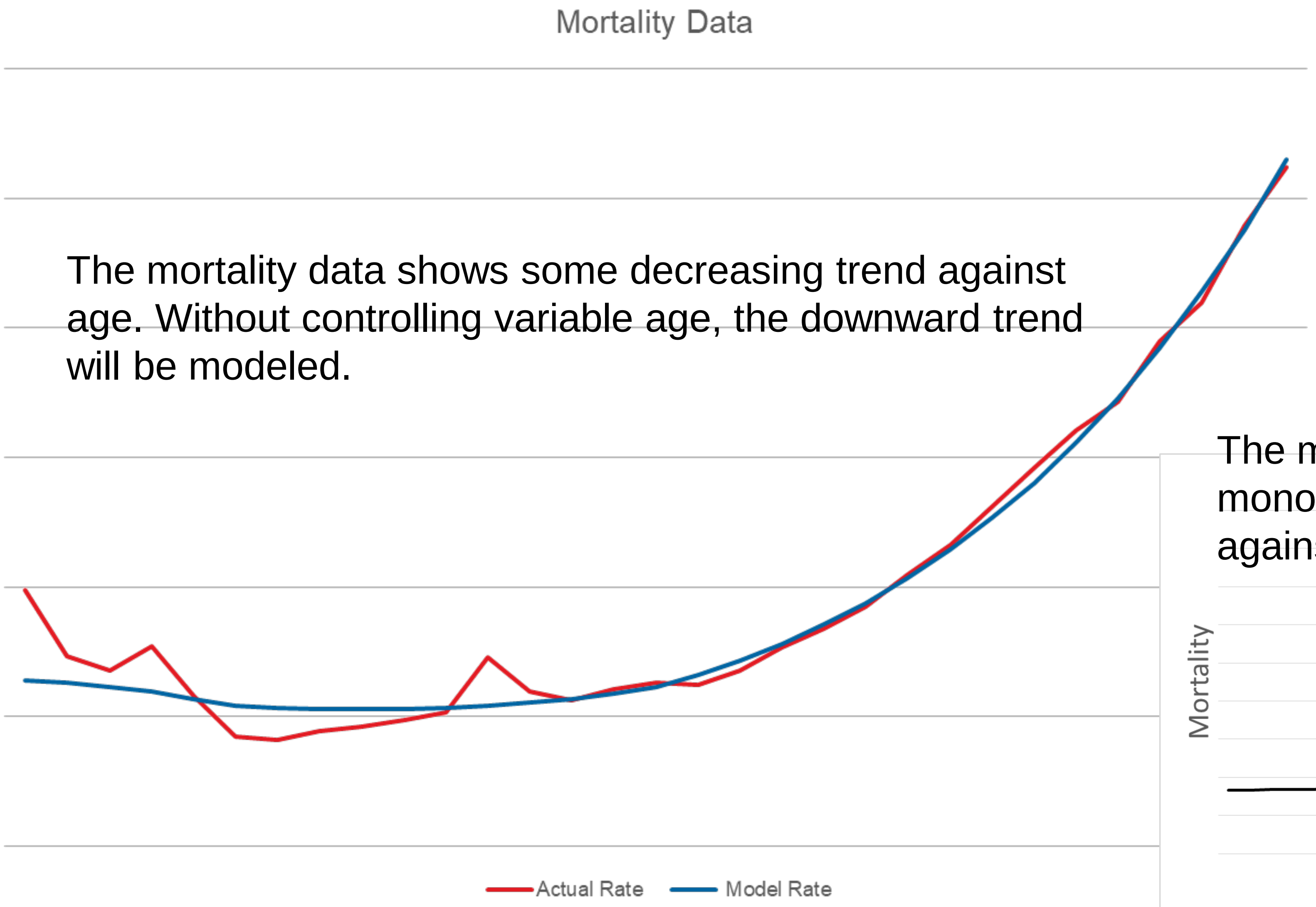
But …

- Considering model structure beyond statistics
  - Linear trend v.s. step-wise adjustment.
  - Handling data with thin exposure: theory v.s. practice.
- Seeming noise v.s. actual noise e.g. impact of anti-selection; contestable period.
- Combine modeling technique with actuarial judgement: e.g. grouping of categorical variables.

Case Study: Use Predictive Model for Pricing Assumption

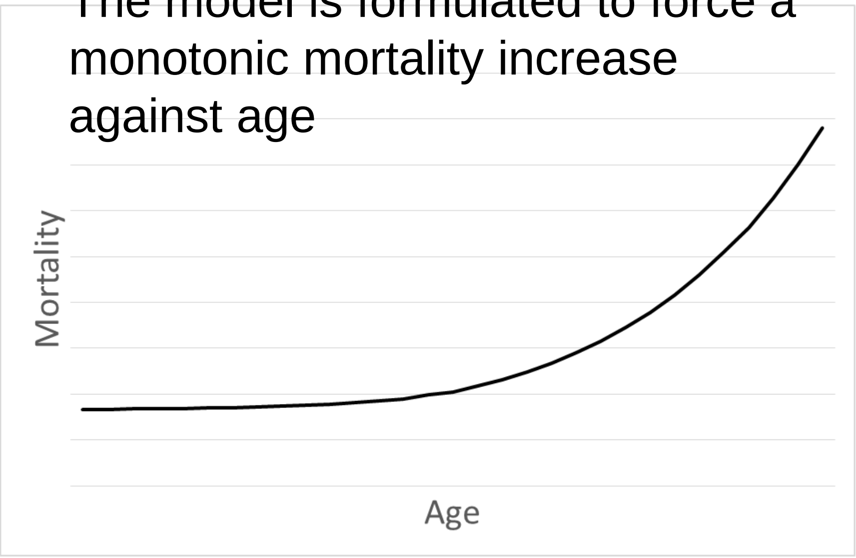
- Background
  - An actuarial group would like to use predictive model to create a new price assumption for their products.
- Challenges
  - Need to make sure the model is not crazily different from the existing assumptions.
  - Data is not clean and shows puzzling patterns that could lead to biases.
- Solution
  - Build an incident model.
  - Use statistics to smooth out small scale issues and avoid potential biases caused by data.
  - Extensive validation process to address actuarial concerns over multiple items.

Puzzling Mortality Decreasing over Age

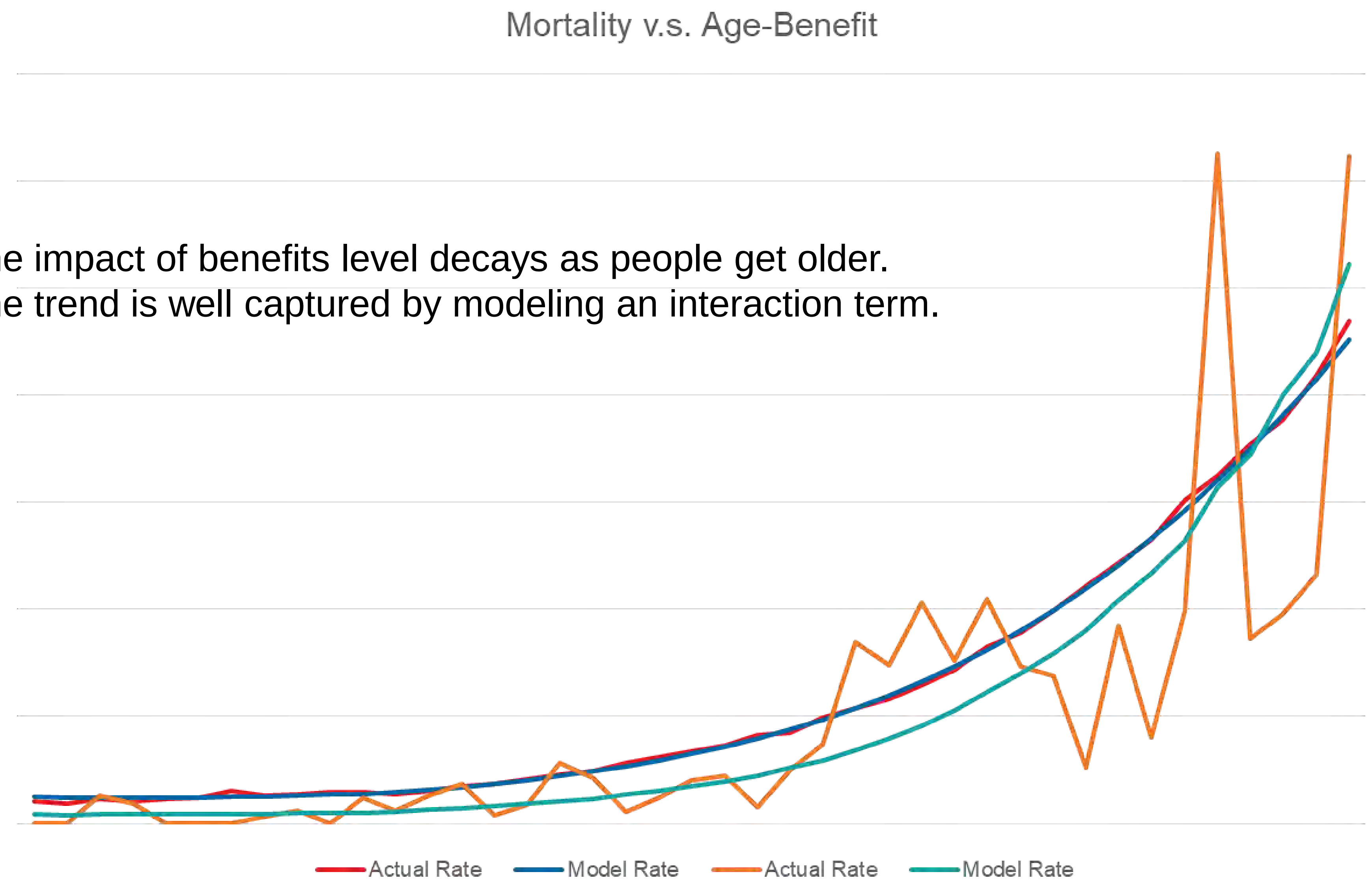


The mortality data shows some decreasing trend against age. Without controlling variable age, the downward trend will be modeled.

The model is formulated to force a monotonic mortality increase against age



The Power of Modeling Interaction Terms



The impact of benefits level decays as people get older.
The trend is well captured by modeling an interaction term.

Leanings/Conclusions

- Build two incident models.
- The mortality experience for certain age-bands shows a downward trend. The model needs to be structured to correct this absurd trend.
- There is a data cohort that contains a few thousands of valid claims but its “product” cannot be figured out. Can we simply drop the claims?
- The power of interaction terms: benefit amount : age.
- Compare existing assumption settings and be comfortable with the model: validate beyond statistic metrics.
- Assess the implication on premium/profits.

Summary

- Setting assumption is not a button-click modeling practice.
- Multiple iterations are generally needed.
- Convert ideas into mathematical language: formulate your model properly to assess the questions.
- It is critical to communicate with actuaries to build a model that can be implemented.
- Statistical significance v.s. actuarial significance.
- Make changes as needed. Be open-minded.