Future Actuarial Tools: Understanding and Refining Simulation Models

Tuesday, March 20, 2012 10:45 am
Panel

Roger Loomis, FSA, MAAA
Senior Prophet Developer, Actuarial Resources Corp.
Peggy Hauser, FSA, MAAA
Sr. Vice President, Univita Health, Inc.
Amy Pahl, FSA, MAAA
Principal and Consulting Actuary, Milliman, Inc.
Stuart Klugman, FSA, CERA, PhD
Staff Fellow, Society of Actuaries
Agenda

• Effect of Simulation Count
• Additional Sources of Variability
• Simple Coin Example
• Simple LTC Example – Claim Terminations
  – Assess three sources of error
• More Sophisticated Example – Embedded Value
How Many Simulations are Necessary?

- Depends upon:
  - Size of block
  - Relative size of policies
  - Desired accuracy
  - Purpose of exercise
- If it’s a reasonably large block, you don’t need very many simulations
- Remember the law of large numbers and how you justify being in the insurance business
Total Paid Claims After Two Years

- **100 Sims**
- **400 Sims**
- **1600 Sims**

Future Actuarial Tools: Understanding and Refining Simulation Models
Remember Profits are Jagged
Additional Sources of variability

• Model from yesterday
  – Captured process error
  – Errors that occur even if you have accurate knowledge of all probabilities and their relationships.

• Parameter error
  – The model is correct, but incorrectly calibrated
  – May be sampling error (estimated probabilities off due to being derived from a sample, not the population)
  – May be sampling frame error (the sampled population does not match the population insured in the future)

• Model error
  – Wrong construct
  – Variables missing
Which errors need correction?

- All but process error
- The others mean premiums or reserves are not correct
- So, can we measure any of them?
- Consider grouping into three sources:
  - Process
  - Estimation
  - Sampling frame
- Model error does not fit here, but the other three are additive.
- A simple example on the next slide.
100 tosses of a possibly biased coin

- Coins are minted by a company known to have poor quality control, leading to a probability of heads not equal to 0.5. However, each year’s batch is remarkably consistent.

- We have secured a coin from the 2008 run and flipped it 200 times, getting 91 heads.

- We have been approached by someone with a 2010 coin who wants us to insure the outcome of 100 flips. We have to pay $10 for each head.

- The coin is flipped and there are 51 heads, when 45.5 were estimated. What caused the error of 5.5?
  - Process: 100 flips are random
  - Parameter: 0.455 was estimated from 200 flips
  - Frame: This is a 2010 coin
If only process error

- Then the probability of 51 heads or more comes from a binomial distribution with \( n = 100 \) and \( p = 0.455 \).
- This probability is 0.158. The distribution of outcomes is
If process and parameter error

- We can model the uncertainty about $p$ using the beta distribution. A common way to set the parameters is to let $a/(a+b)$ be the estimated value and $a+b$ be the sample size.

- In this case $a+b = 200$ and $a = 91$. The graph is

![Beta distribution graph](image-url)
Probability with sampling error

- Unconditional distribution is beta-binomial.
- Probability of 51 or more heads is 0.206 (compared to 0.158)
- Overlaid graph is
What is an extreme value?

• We could look at the 95th percentile and say heads in excess of that value are unlikely to occur by chance.

• With only process error, that value is 55 (actual probability of 55 or more is 0.036).

• With both process and estimation error, that value is 57 (actual probability of 57 or more is also, by chance, 0.036).
Simplified Application to LTC I

- Claim continuation with one stochastic variable – claim termination rate
- KISSS Insurance Company has a block of claims
  - All lifetime benefit period policies
  - 204 open claims
  - Lacking other LTC experience, predict claim continuance using Intercompany claim termination rates
  - For each claimant, projected number remaining on claim for each month after valuation date is \( I_t = I_{t-1} \times (1 - q_t) \)
  - Tables are one-dimensional, choose to use tables that vary by age at claim
# Probability of Remaining on Claim 12 Months Later

<table>
<thead>
<tr>
<th>Claimant</th>
<th>Clm Dur (mos)</th>
<th>Age at Claim</th>
<th>Probability of Remaining on Claim</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>l(t+1)</td>
</tr>
<tr>
<td>1</td>
<td>87.2</td>
<td>83.1</td>
<td>0.974</td>
</tr>
<tr>
<td>2</td>
<td>43.2</td>
<td>68.3</td>
<td>0.969</td>
</tr>
<tr>
<td>3</td>
<td>42.7</td>
<td>72.9</td>
<td>0.969</td>
</tr>
<tr>
<td>4</td>
<td>8.0</td>
<td>83.7</td>
<td>0.962</td>
</tr>
<tr>
<td>5</td>
<td>41.4</td>
<td>81.0</td>
<td>0.962</td>
</tr>
<tr>
<td>6</td>
<td>8.0</td>
<td>77.3</td>
<td>0.963</td>
</tr>
<tr>
<td>7</td>
<td>91.8</td>
<td>77.8</td>
<td>0.974</td>
</tr>
<tr>
<td>8</td>
<td>8.3</td>
<td>77.3</td>
<td>0.962</td>
</tr>
<tr>
<td>9</td>
<td>6.3</td>
<td>86.8</td>
<td>0.959</td>
</tr>
<tr>
<td>10</td>
<td>21.7</td>
<td>78.4</td>
<td>0.968</td>
</tr>
</tbody>
</table>
One Year Later….

• Of the 204 on claim last year, 160 remain on claim
• Is that okay?
• Sum of $l_{x+12} = 133$. Is 160 reasonable?
• Use a stochastic model to evaluate
  – Rather than use deterministic claim termination rates
  – Use random number each month to determine whether a claimant remains on claim
    • If random number > claim termination probability, persist (1)
    • If random number <= claim termination probability, terminate (0)
  – Run 1,000 simulations
## Stochastic Calculation

<table>
<thead>
<tr>
<th>Claimant</th>
<th>Clm Dur (mos)</th>
<th>Age at Claim</th>
<th>Stochastic Number Remaining on Claim</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>l(t+1)</td>
</tr>
<tr>
<td>1</td>
<td>87.2</td>
<td>83.1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>43.2</td>
<td>68.3</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>42.7</td>
<td>72.9</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>8.0</td>
<td>83.7</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>41.4</td>
<td>81.0</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>8.0</td>
<td>77.3</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>91.8</td>
<td>77.8</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>8.3</td>
<td>77.3</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>6.3</td>
<td>86.8</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>21.7</td>
<td>78.4</td>
<td>1</td>
</tr>
</tbody>
</table>

**Future Actuarial Tools: Understanding and Refining Simulation Models**
0 out of 1,000 simulations are $\geq 160$...
We have an error…

• Recall our three sources
  – Process – was this a random fluctuation?
  – Estimation – Intercompany data includes 250,000 claims…
  – Sampling frame – This block is different from the Intercompany data

• Model recognizing process error this is not random

• Speculate a sampling frame issue – revise the data source
  – Grab data from similar policies
  – Unlimited benefit period
  – Vary claim termination rates by age, gender, and cognitive impairment status.
10.2% of simulations are $\geq 160\ldots$
What about sampling error?

- We can apply the beta distribution in each case.
- Note that when simulating, the beta distribution is used to simulate a particular probability. If that probability affects more than one life, then use the same simulated probability each time.
- For the intercompany study, the sample sizes used were approximated and were about 1/10 the actual sample sizes.
- For the own company study, the actual sample sizes were used.
Future Actuarial Tools: Understanding and Refining Simulation Models
Effect of beta distribution

• When using the intercompany data there is no effect (except for the effect of doing a new simulation)
  – In neither case did any simulations exceed 160
  – The sample size is so large that there is essentially no sampling error
• When using the own company data, the simulations show greater spread
  – With only process error, 102 simulations exceed 160
  – With sample error added, 181 simulations exceed 160
News Flash! KISSS Exiting LTC Market!

- Sold exactly 10,000 policies
- Now exiting business
- Your company sees an opportunity
- Your owners need a 20% ROI
- How much should you bid?
Possible Deterministic Approach

• Best-estimate assumption of embedded value: $25.0M
• Add 10% Margin: $22.5M
• Bid:

$$\frac{\$22.5M}{1.2} = \$18.75M$$

• But…
  – What is the correct margin?
  – How confident are you in your best-estimate?
Simple Simulation Approach

• Use best estimate assumptions
• Run 1,600 Simulations
• Calculate embedded value for each simulation
• Bid on the 5\textsuperscript{th} percentile:
  – 95\% probability profits will be higher
5th Percentile is $24,500; Bid $20,400

Future Actuarial Tools: Understanding and Refining Simulation Models
Challenge with This Approach

• Based on assumption that the model’s rates are the actual probabilities:
  – Incidence
  – Recovery
  – Lapse
  – Mortality

• Variance reflects Process Error

• Does not capture Estimation Error: The random sample used to calculate probabilities might not reflect the true probabilities
Beta-Binomial Review

• Beta distribution has two parameters
  – Alpha
  – Beta

• When parameterizing:
  – $\alpha$ is the number of hits
  – $\beta$ is the number of misses
  – $\alpha + \beta$ is the number of trials

• As $\beta \to \infty$, the variance of the Beta distribution approaches the variance of the Bernoulli distribution
More Sophisticated Simulation Approach

• We want to modify the Embedded Value probability curve to reflect Estimation Error

• We use a beta-binomial distribution rather than a Bernoulli distribution for incidence

• We set $\beta = 4,000$

• This is because the incidence rate table was based on $\sim 4,000$ exposures for each table entry
Steps to Implement

• \( \beta = 4,000 \)
• Solve for \( \alpha \) so that:

\[
Best\ Estimate\ Incid\ Rate = \frac{\alpha}{\alpha + \beta}
\]

• To simulate
  – Simulate “true probability table” using distribution Beta(\( \alpha, \beta \))
  – Using that table and simulation, project block run out
  – Re-create simulated “true probability table” every simulation
5th Percentile is $24,200; Bid $20,200

Future Actuarial Tools: Understanding and Refining Simulation Models
Upon further contemplation…

• Variance reflects:
  – Process Error
  – Estimation Error

• What about Sampling Frame Error? The future of your block might not be homogenous with the block sampled for probabilities

• You decide $\beta = 1,000$ reflects your true level of uncertainty of estimation error and sampling error
  – Equivalent variance to each rate being based on only $\sim 1,000$ observations
  – Subjective actuarial judgment of your uncertainty
5th Percentile is $23,500; Bid $19,500
Future Actuarial Tools: Understanding and Refining Simulation Models
Tuesday, March 20, 2012 10:45 am