Loss Simulation Model: Testing and fitting

Joseph O. Marker
Marker Actuarial Services, LLC
and University of Michigan
CLRS 2011 Meeting

Testing the Model and its Benefits

- Testers set up LSM parameters for situations where the output distribution is known.
- Then tested distribution of simulated output vs. the theoretical distribution.
- If not a good fit, looked for:
  - bugs in the LSM
  - misunderstanding of how parameters work
- Important: The "R" code used for the testing is available in the various papers relating to the LSM.

Testing the Model and its Benefits

- The same "R" code used for testing can also be used to parameterize the model on an insurer's actual detailed loss data.

- To do this:

  Reformat the real data to the format of the LSM's detailed output files.

  Apply the "R" code used for testing to produce parameters for specific distributions and to measure the goodness of fit.

- Note: The detailed output files from the LSM include, for each claim, all the reserve changes and their dates.

Testing the Model and its Benefits

- What if modeler has only life-to-date claims evaluated as of one valuation date?

- Case 1: Claims are essentially at ultimate valuation.

  The LSM produces ultimate claim values before it simulates the reserve and payment transactions that lead to the ultimate values.

  The "R" code is full of tests for ultimate claim count and ultimate severity.

- Use the "R" code to fit the ultimate frequency and severity.

- Case 2: Claims are not fully developed (go to next page)

Testing the Model and its Benefits

- Case 2: Claim data is not fully developed
- The LSM uses parameters that model the development of individual claims.
- The LSM produces aggregate loss development triangles.
- The modeler may need to experiment with development parameters until the aggregate triangles look similar to the insurer's historical development.
- Kailan Shang will discuss testing of the development parameters under the topic "Case Reserve Adequacy".

An Aside: Application to Pricing

- The model produces ultimate claim counts, severities, and aggregate losses, and allows one to pick any range of accident years. The price is a function of the ultimate loss for a future accident year.
- Repeated simulation allows one to estimate uncertainty in losses.
- The LSM is a compound frequency-severity model. Once parameterized, the LSM produces simulated information needed for pricing.
- The LSM's loss development feature allows one to estimate discounted loss costs.
- Caveats:

  The exposure level needed to estimate frequency must be externally determined.

  Pricing by classification requires "covariates". The covariate feature of the model is not yet well-developed.

Model fitting (2007-08)

- Task:

  Fit distributions to real data from anonymous source.

  Use this result to design flexibility into LSM.

  Model used Richard Vaughan's previous modeling work done in APL.

- Data consisted of Auto BI and Collision claims as of 12/31/2006.
- Working party produced paper in December 2007:

  Parameterizing the Loss Simulation Model
  http://www.casact.org/research/lsmwp/BSUPaper.pdf

- Topic presented at 2008 CLRS.

Model fitting (2007-08)

- Ball State University class did work for this paper.
- Most of work done in Excel, some in "R"
- The paper documents the various areas investigated.
- "R" code used to fit models is included in paper.
- Parametric survival models were used to fit censored distributions.

- Examples of models developed:

  Univariate modeling of lags and claim size with covariates

  Correlation between Settlement Lag and Claim Size for Auto BI

  Zero modification of the Claim Size distribution.

  Effect of Deductibles on Collision Losses; Pareto Model

  Interaction of Report Lag and Settlement Lag

Model Testing and Fitting (2010)

- 2007 Paper helped design the basic model.
- Much more extensive testing was needed.
- As explained previously, code developed for testing can help parameterize the LSM on an entity's own claim data.
- The LSMWP wrote the following paper in 2010:

  Modeling Loss Emergence and Settlement Processes
  http://www.casact.org/pubs/forum/11wforum/LSMWP.pdf

- The paper was discussed at the 2010 CLRS,
  and published in the CAS Forum for Winter 2011

Model Testing and Fitting (2010)

- Test distributions of:
  - Number of claims (frequency)
  - Size of ultimate loss (severity)
  - Correlated frequencies between lines using copulas.

- Emphasis in Paper:

  Document the LSMWP's work in developing the LSM.

  Document the "R" code used in performing various tests.

  Provide references for those who want to explore the modeling further.

  Provide visual as well as formal tests: QQPlots, histograms, densities, etc.

  Document the model for users and for developers wishing to customize the model.

Model Testing and Fitting (2010)

- Example of severity testing:

  One person ran LSM to generate claims from three different LOB according to Lognormal, Pareto, and Weibull distributions.

  "Testers" fit the data for each line to severity distributions using "R", not knowing which LOB had which type of distribution.

  The entire process of exploring the data, fitting to distributions, and graphical analysis is documented in the paper.

Model Testing and Fitting (2011)

- The DRMC and Committee on Reserves recognized that much additional work was needed and issued a call for papers in 2011.

- Areas of interest for call papers:

  Applications to reserving problems.

  Enhancements to model.

  Further testing of model in specific areas.

Model Testing and Fitting (2011)

- One call paper was submitted:

  Loss Simulation Model Testing and Enhancement
  by Kailan Shang FSA, CFA, PRM, SCJP
  http://www.casact.org/pubs/forum/11sumforum/Shang.pdf

- This paper makes substantial contributions to the areas:

  Further testing of the LSM.
  Enhancement of Model.
  Reserving, through the testing of the case adequacy parameters.