

The Actuarial Profession
making financial sense of the future

34TH ANNUAL GIRO CONVENTION

CELTIC MANOR RESORT, NEWPORT, WALES

The Actuarial Profession
making financial sense of the future

Best Estimates and Estimating Uncertainty

GIRO 2007

Best Estimates and Estimating Uncertainty

A brief overview of what the paper covers:

- Defining / describing “best estimate”
- The “judgement” method
- List of “uncertainty” methods/models
- Surveying the profession on methods used
- Trying out methods/models on real data
- Testing methods when data “perfect”
- Next year’s objectives

The Actuarial Profession
making financial sense of the future

Testing methods on real data

GIRO 2007

Contents

- Objective
- Methodology
- Methods tested
- Data used in investigations
- Numerical results
- Qualitative results
- Summary /“Learning points”

Objective

- What we wanted to do:
 - Compare common methods on real data
 - Assess relative strengths and weaknesses of each on different criteria
 - Provide a starting point for actuaries to learn about the available methods
- What we didn't want to do:
 - Produce an up-to-the-minute comparison of the most sophisticated methods
 - Recommend one or more particular method
 - Cover all possible methods or data issues

Methodology

- Selection of methods
 - Concentrating on introductory level of methods
 - Wide spread of model type and data requirements
 - Includes subjective methods
 - Not all methods provide comparable output
- Provision of data
 - Anonymisation
 - Non-standard level of detail
 - Wide variety of classes
- Allocation of methods by group
- Investigation (qualitative and quantitative)
 - Compilation of questionnaire
 - Standardised tests
- Summary of results
 - By class
 - By model

Methods – long list

- Models (generic categories only)
 - Mack
 - Over Dispersed Poisson
 - Bayesian
 - Operational time
 - Curve fitting
 - Scenario testing
 - Individual claim
 - Proprietary methods
- Methods
 - Bootstrapping / Stochastic simulation
 - Regression
 - Analytic calculation
 - Judgement

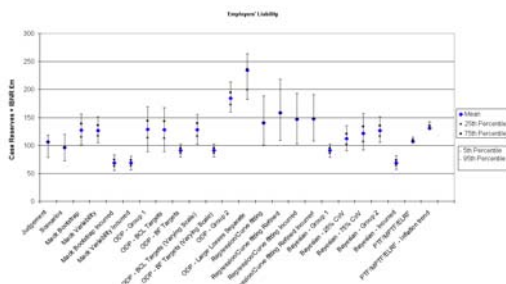
Methods – short list

- Selected methods
 - Mack (analytic)
 - ODP (Bootstrapping)
 - Bayesian ODP (Bootstrapping)
 - Judgement
 - Scenario tests
 - Curve fitting (Regression)
 - Operational time
 - Proprietary methods
 - (Transaction level – no data available)
- Software
 - Some proprietary models used for more common methods
 - Some models developed from scratch

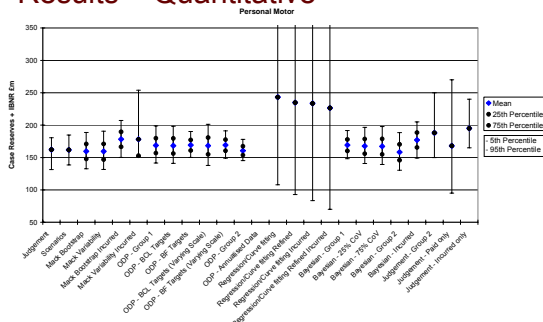
Data

- Offered from a variety of organisations
- Does not currently include pseudo-data
- Data was provided to best-estimates working party for consistency
- Data was adjusted to make source anonymous
 - Process not very complex and not consistent between datasets, hence resulted in some inconsistencies within datasets (or really, really good underwriting)
 - Process also not consistent between datasets
- Classes covered personal lines, commercial lines and Lloyd's data
- Data was provided in annual and/or quarterly periods
- Some classes had large losses excluded, others did not
- Methods applied to as many data sets as possible, but results and efforts concentrated on most promising:
 - Employers' Liability
 - Personal Motor
 - Commercial Property

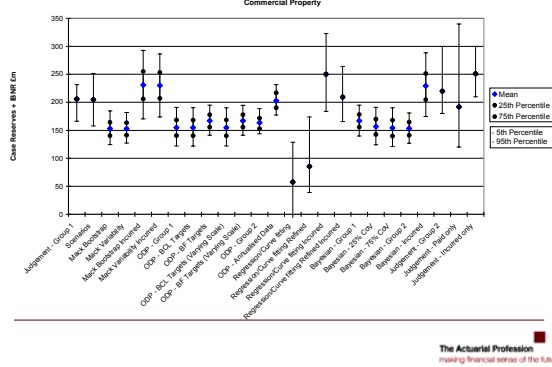
Results – Quantitative



Results – Quantitative



Results – Quantitative



Results – Qualitative

- Questions:
 - Would the method be acceptable to the Profession?
 - Ease of use and practicality of method
 - How difficult is it to apply judgement and / or amendments to the results
 - How easily would you be able to explain the method to non technicians?
 - Does the method include extreme events? (By this we mean can you allow for the sudden emergence of large individual losses, late tail kicks in incurred, surprising developments on known large losses, etc)
 - When is the method good, when is it not good, and when does the method fail?

Results – Qualitative

- Mack
 - Easy to program; some ability to apply judgement; doesn't generate a full distribution (analytic version)
- ODP
 - Well known; proprietary packages may make assumptions on how to treat imperfect data; easy to explain key principles, but more challenging technically; scaling issues to mean.
- Bayesian
 - More challenging to program and explain; can include judgement through prior ULRs
- Judgement
 - Mixed results, with it being "easy" to apply and potentially more valid for longer-tail classes, but difficult to explain sufficiently for, say, audit purposes. Peer review considered to be a key control.
- Scenarios
 - Does not provide a full distribution, but allows identification of key assumptions; is easy to explain, but suffers from same limitations as judgement methods.
- Curve fitting
 - Is easy to explain and implement, but does not give a full distribution, or even an indication of the likelihood of the range values.
- PTF
 - Requires specialist software to run

Summary/ "Learning points"

- We have:
 - Looked at some basic methods using "real" data sets
 - Investigated usability
- We intend to:
 - Look at further methods (including transactional level data and more sophisticated versions of current methods)
 - Expand the qualitative and quantitative comparisons of "core" methods
- Learning points
 - Keep targets achievable!
 - Aim for regular meetings and promote discussion on particular methods

Testing the Methods on "Perfect" Data Tom Wright

GIRO 2007

Testing Stochastic Methods by Numerical Simulation - Overview

- Key steps:
 - Generate artificial run-off triangle and 'true' ultimate.
 - Apply stochastic method to run-off triangle.
 - Compare result of stochastic method to 'true' ultimate.
 - Repeat a large number of times (30,000).
- Generation of 'data' can be done in two ways:
 - Strictly following the assumptions of the stochastic method being tested (to test performance of the method in ideal circumstances).
 - In ways that violate the assumptions of the stochastic method (to test robustness of the method in more realistic circumstances).

Merits of Simulation Approach

- Enables methods to be tested on large numbers of triangles.
- Predictions can be compared to 'true' ultimates without waiting years for development.
- Triangles can be constructed so they perfectly satisfy the assumptions of a method.
- Robustness of a method to violations in its assumptions can be tested in a controlled way.

Comparison of Stochastic Prediction with 'True' Reserve

- Method 1:
 - based on best-estimate (m) and root-mean-square predictive error (s).
 - possible for all stochastic methods.
- Method 2:
 - based on complete predictive distribution $F(r)$.
 - requires additional assumption (eg log-normal) for stochastic methods that do not give a complete predictive distribution.

Comparison of Stochastic Prediction to 'True' Reserve – Method 1

- Simulation of data to ultimate gives 'true' reserve (r) (= ultimate less latest paid).
- Stochastic method applied to simulated triangle gives best estimate (m) and root-mean-square predictive error (s).
- By definition, we should have $s^2 = E(m-r)^2$.
- Define 'standardised predictive error' $d = (m-r)/s$, then (for each fixed value of s) we should have $E(d^2) = 1$.
- Over many simulations, value of s varies. But for each sub-set with s approximately equal, we should have $E(d^2) = 1$.
- Therefore, over all simulations, we should have $E(d^2) = 1$.
- If mean value of d^2 is significantly greater than 1, then the value of s given by the stochastic method tends to be too small. In other words the method tends to understate the chances of extreme outcomes.

Comparison of Stochastic Prediction to 'True' Reserve – Method 2

- Use stochastic method to produces a complete probability distribution for the reserve.
- The simulated 'true' ultimate is a single number.
- How should these two be compared?
- Example:
 - stochastic method gives a log-normal distribution with mean = £13.1m and standard deviation = £4.2m.
 - 'true' outcome is £24.3m.
 - 'true' outcome is 99.5% point of the log-normal distribution.
 - does this mean the stochastic method performed poorly?

Comparison of Stochastic Prediction to 'True' Reserve – Method 2

- If a stochastic method works correctly, the 99.5 centile of the predictive distribution should be exceeded 1 time in every 200.
- Mathematically:
 - $F(r)$ is the predictive distribution (cdf) produced by the stochastic method.
 - r_0 is the true reserve (ultimate minus latest paid).
 - $F(r_0)$ should exceed 99.5% one time in every 200 trials (that is, with probability 1/200).
 - $F(r_0)$ should exceed $1-\alpha$ with probability α .
 - In other words: $F(r_0)$ should have a uniform distribution.
- By simulating many triangles, we obtain many values of $F(r_0)$ so can check whether these are uniformly distributed between 0 and 1.

Comparison of Stochastic Prediction to 'True' Reserve – Method 2

- To check that $F(r_0)$ is uniformly distributed between 0 and 1:
 - Look at graph of empirical distribution function.
 - Look at the proportion of simulations in which $F(r_0)$ exceeds a specified value (eg, in 1% of cases, $F(r_0)$ should exceed 99%).

Methods Tested to Date

- Mack (1993)
- Analytic ODP method (Renshaw & Verrall 1998)
- Bootstrap ODP method (England & Verrall 1999)
- Bootstrap ODP method (England 2001)
- Operational time method (Wright 1992)
- Further testing of these methods needed before definitive conclusions can be formulated.

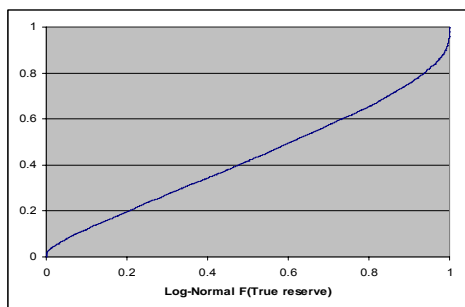
Results by Method 1: Standardised Predictive Error

Method	Mean standardised predictive error (m-r)/s	Mean square standardised predictive error [(m-r)/s] ²
Mack 1993	-0.63 to -0.50	4.3 to 6.3
Analytic ODP (Renshaw & Verrall, 1998), Pearson	-0.27	1.46
Analytic ODP (Renshaw & Verrall, 1998), deviance	-0.28	1.50
Bootstrap ODP (England & Verrall, 1999)	-0.25	1.72
Bootstrap ODP (England 2001)	-0.25	1.69
Operational time (Wright 1992), Pearson dispersion	-0.22	1.57

Method 2: Percent of Simulations with 'True' Outcome exceeding 99th centile (log-normal)

Mack 1993 (with Log-Normal)	8% to 13%
Analytic ODP (Renshaw & Verrall, 1998), Pearson dispersion	2.6%
Analytic ODP (Renshaw & Verrall, 1998), deviance dispersion	2.7%
Bootstrap ODP (England & Verrall, 1999)	3.1%
Bootstrap ODP (England 2001)	2.6%
Operational time (Wright 1992), Pearson dispersion	4.0%

Method 2: Uniformity of $F(r_0)$? Mack's Method



The Actuarial Profession
making financial sense of the future

Main Conclusions (provisional)

- All stochastic methods tested tend to understate the chance of extreme adverse outcomes.
- When reserve is under-estimated, predictive standard error also tends to be under-estimated.
- More testing needed to check whether these are generally the case or depend on the particular parameters used.

The Actuarial Profession
making financial sense of the future

Effect of Diagnostics

- In practice, diagnostics should prevent a stochastic method from being applied where its assumptions appear to be violated.
- Artificial data generated using the assumptions will produce some datasets where the assumptions appear to be violated.
- In practice the method would not be applied to these datasets (this is a "Type I error": model assumptions rejected when true).
- However, where assumptions are not satisfied, they will sometimes appear to be: method would in practice be applied to these datasets, probably leading to poor predictions (this is a "Type II error").
- So to carry out a fair test of a stochastic method: if we exclude datasets of Type I we should also include some of Type II.
- Including Type II errors is subjective (decisions on how to generate data) so we have not done this, so have not excluded Type I errors either.

The Actuarial Profession
making financial sense of the future

Best Estimates and Estimating Uncertainty

GIRO 2007
