Institute and Faculty of Actuaries

What does the Bootstrap Trap?

Matt Locke & Andrew D Smith

B02
13:30, Wednesday 21 October 2015

Matt.Locke@UMACS.co.uk
AndrewDSmith8@Deloitte.co.uk

Agenda

- Aims
- What is the bootstrap?
- Back-test results
- Does the bootstrap underestimate extreme percentiles?
- Conclusions

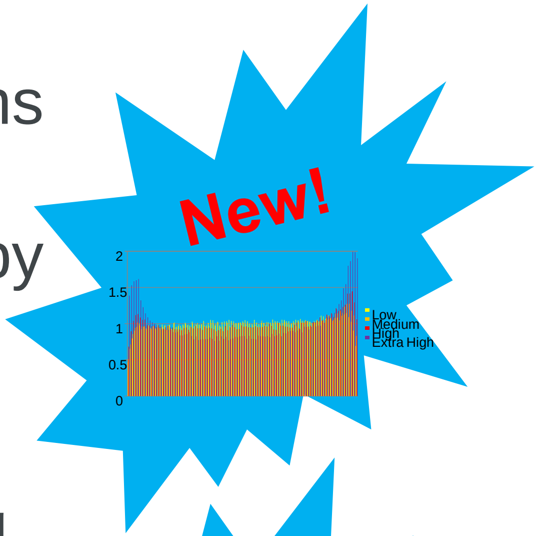
Agenda

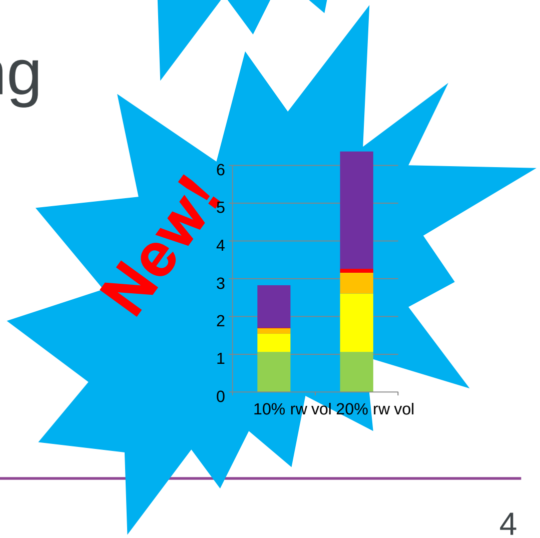
- Aims
- What is the bootstrap?
- Back-test results
- Does the bootstrap underestimate extreme percentiles?
- Conclusions

Aims of our investigation

We ***are*** aiming to:

- Get a better understanding of the strengths and limitations of the over-dispersed Poisson bootstrap (ODPB) as described by England & Verrall (2002)
- Compare the predictive distribution from ODPB against the actual outcomes, using generated data
- Investigate the robustness of the ODP bootstrap's predictions when the model assumptions are violated





Aims of our investigation

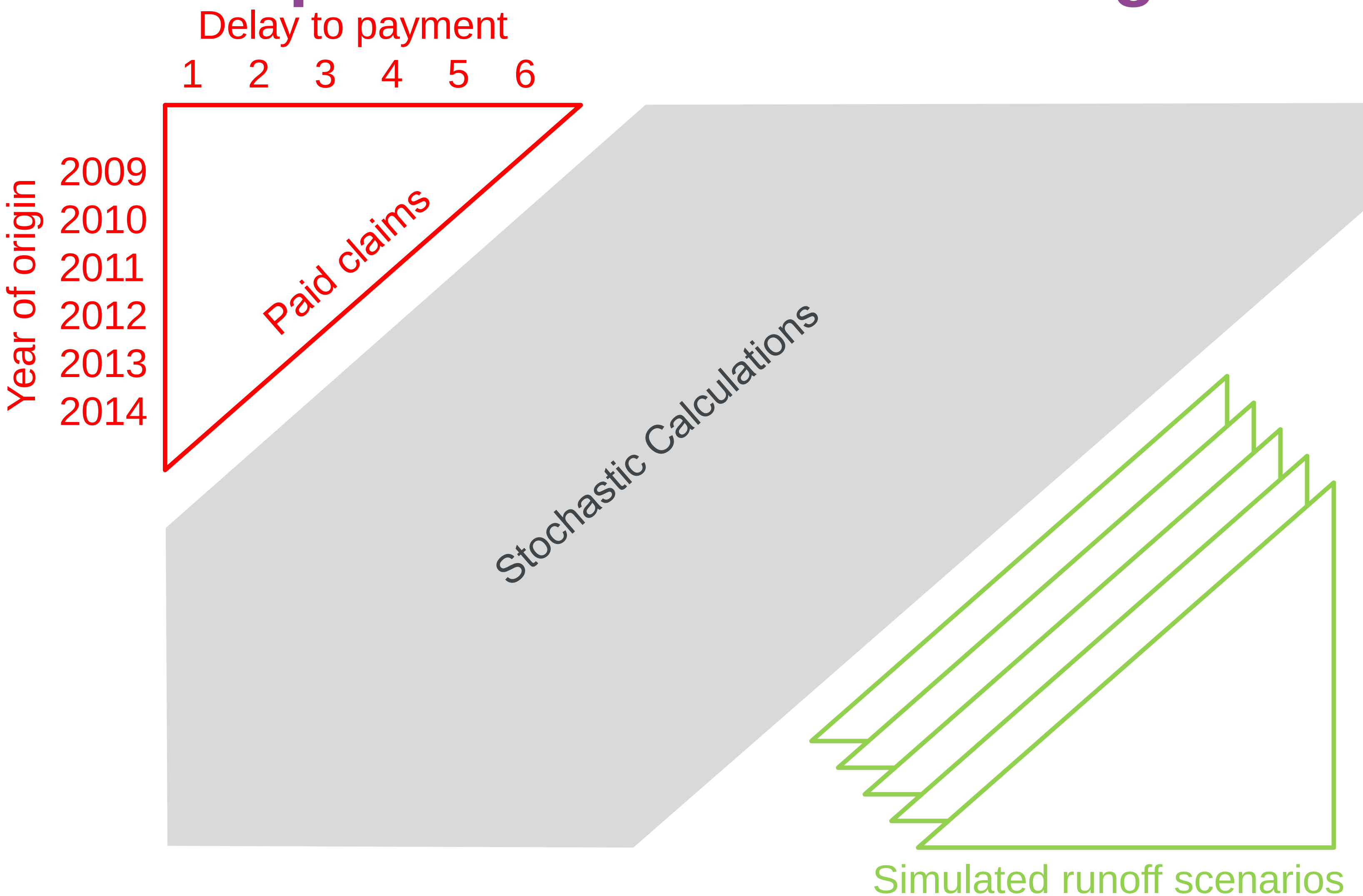
We ***are not*** aiming to:

- Compare the performance of the ODP bootstrap with that of other mechanical or judgement based methods
  - Bootstrap methods applied to *paid* claims; we do not consider incurred triangles or frequency/severity models here
- Promote or discourage the use of:
  - The ODP bootstrap
  - Other mechanical methods
  - Judgement based methods

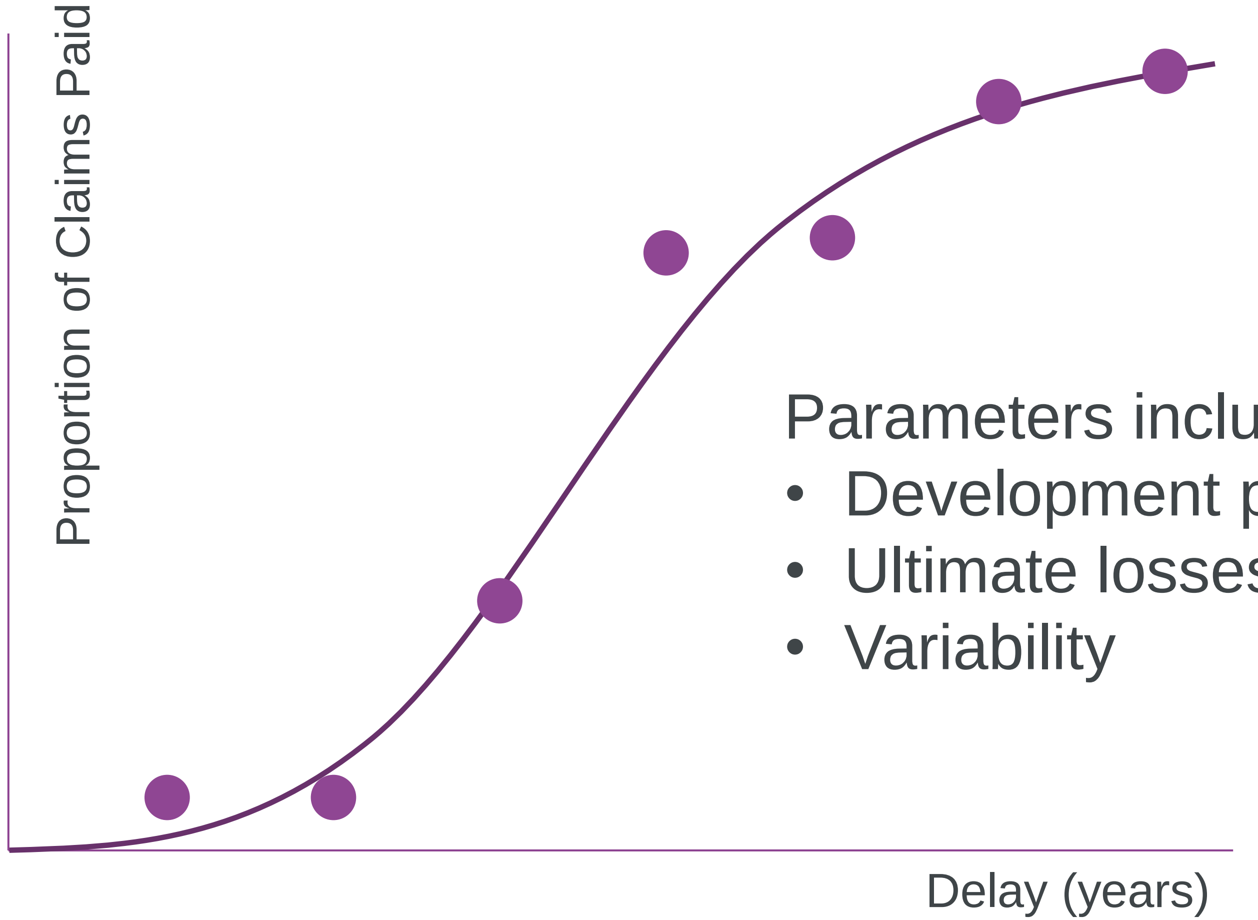
Agenda

- Aims
- What is the bootstrap?
- Back-test results
- Does the bootstrap underestimate extreme percentiles?
- Conclusions

Bootstrap in Stochastic Reserving



Bootstrap Steps: Parameter Estimates

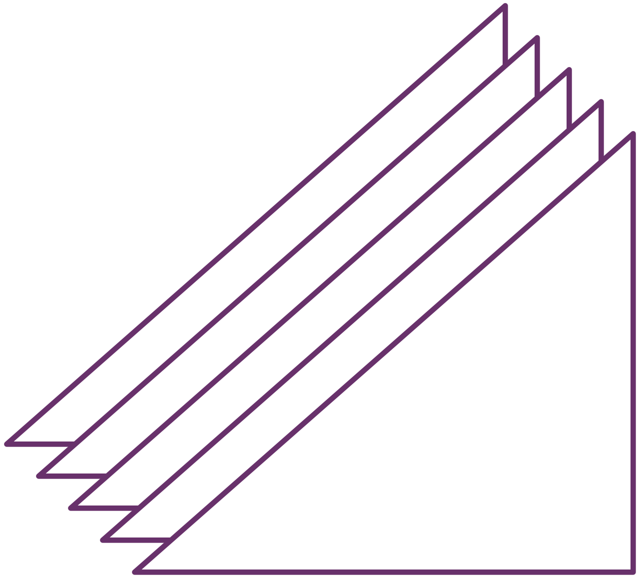


Parameters include:

- Development patterns
- Ultimate losses
- Variability

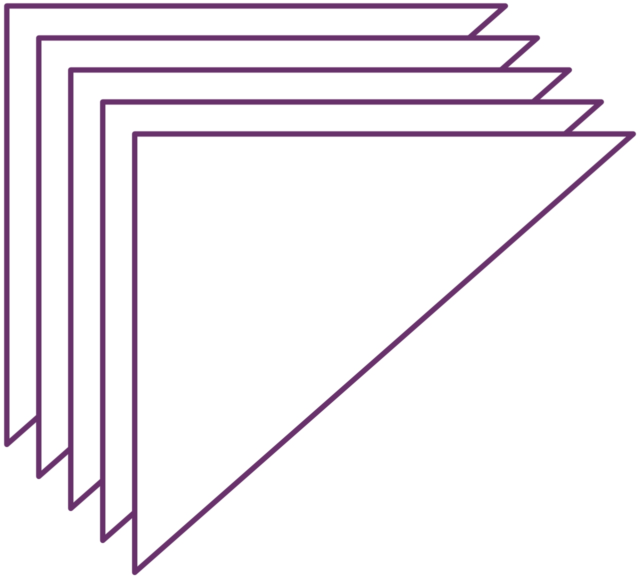
Bootstrap Steps: Forecasting

- Also called “noise” or “process error”
- Simulating one or more future claim scenarios based on estimated parameters (we use gamma distributions here)
- This is a familiar approach for many other risks besides reserving uncertainty

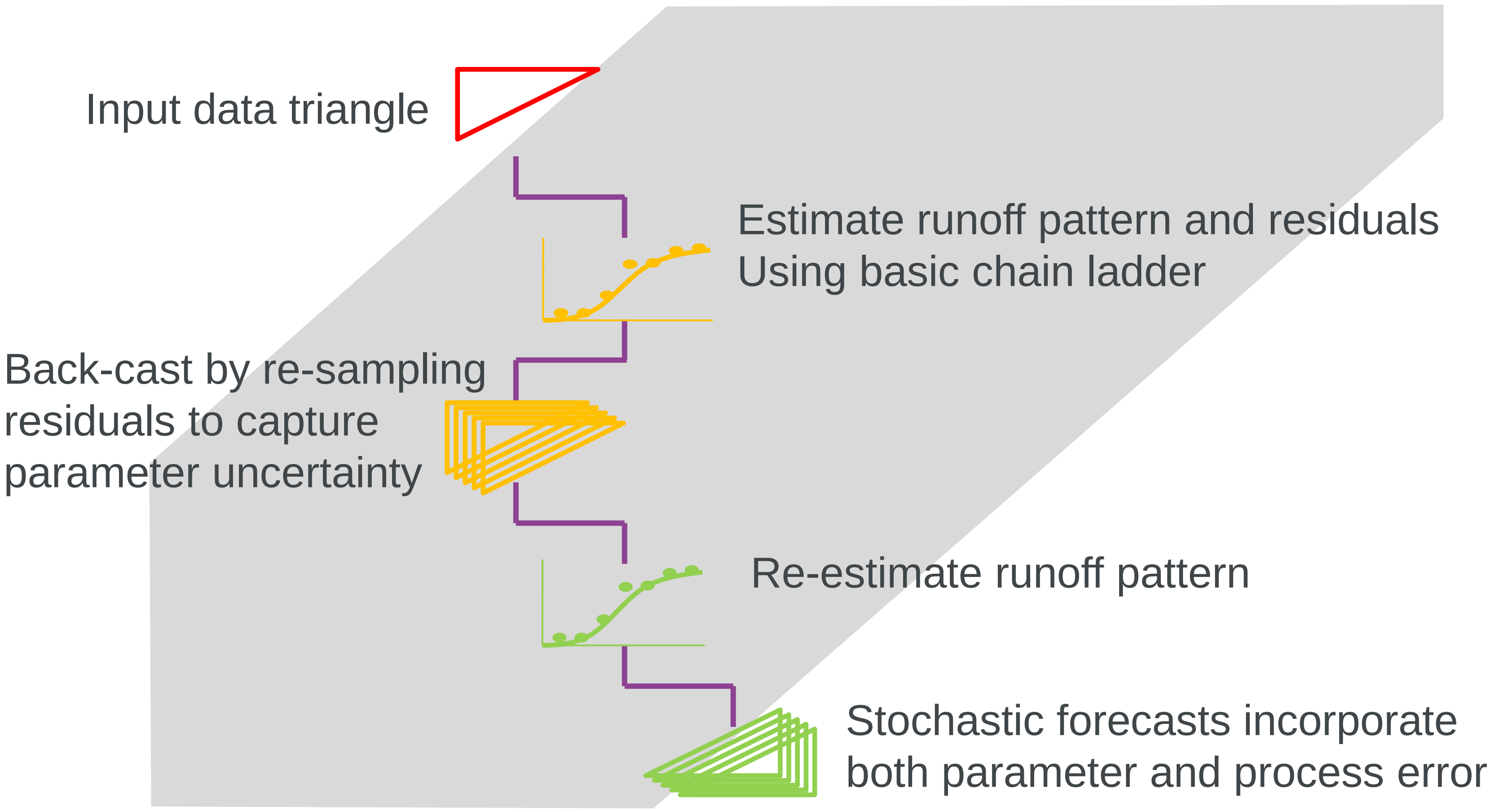


Bootstrap Steps: Back-Casting

- Re-creating hypothetical historical claim scenarios based on estimated parameters
- May use re-sampled residuals (non-parametric) or analytical distributions (parametric)



How the Steps Fit Together

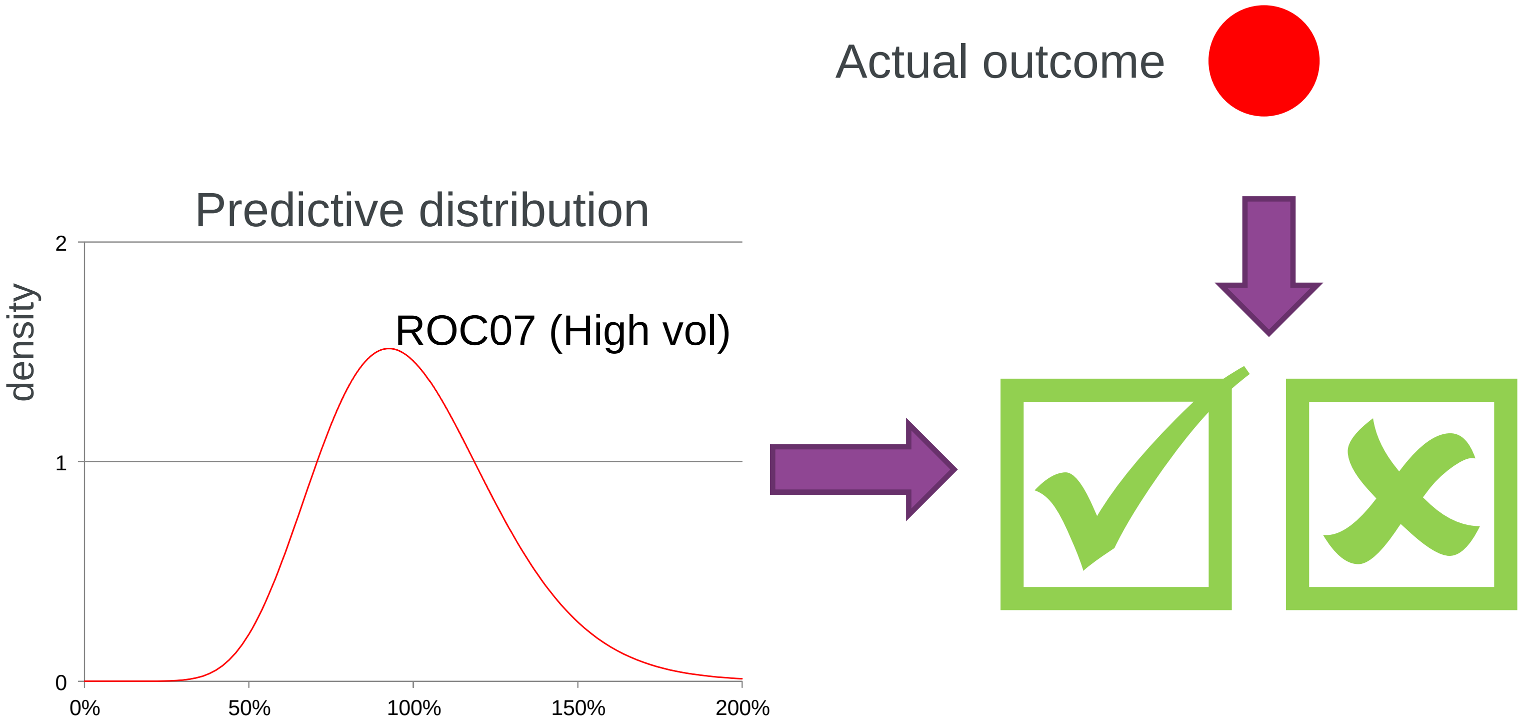


Agenda

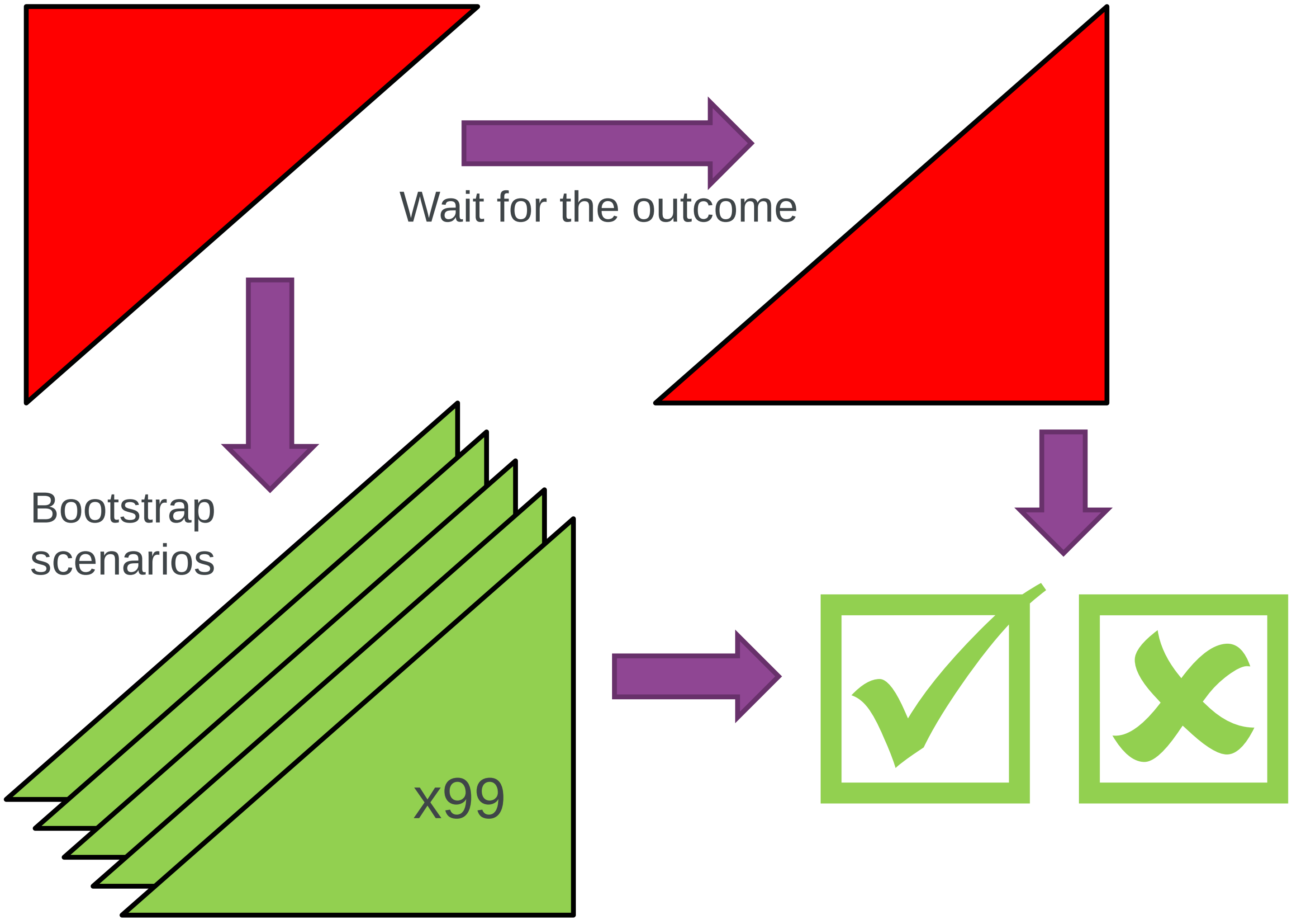
- Aims
- What is the bootstrap?
- Back-test results
- Does the bootstrap underestimate extreme percentiles?
- Conclusions

Back-Testing

- Compare actual outcomes to a predictive distribution

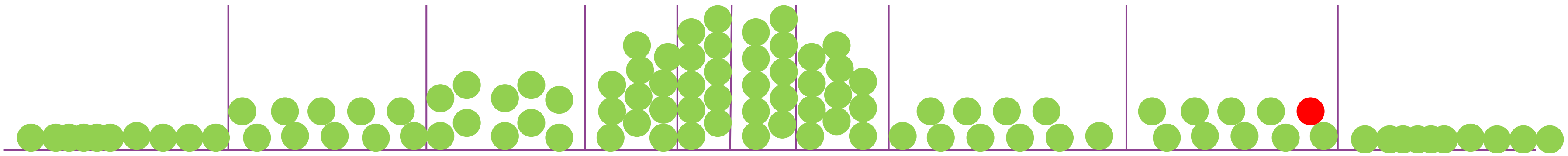


Back-Testing the ODP Bootstrap



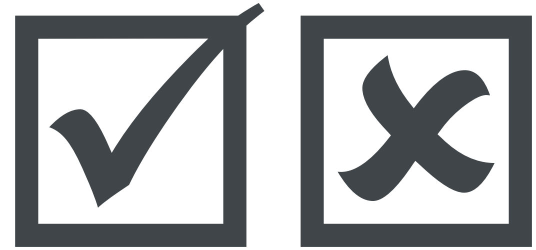
Ranking the Outcomes

- Take 100 future claim scenarios
  - 1 actual outcome and 99 from bootstraps
- Sort into increasing order of outstanding claims
- Divide into 10 buckets, each containing 10 observations



- Suppose the actual outcome and the bootstrap are independent samples from the same distribution
- Then there is 1-in-10 chance the red lies in each bucket

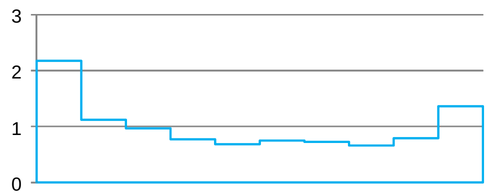
The Back-Test

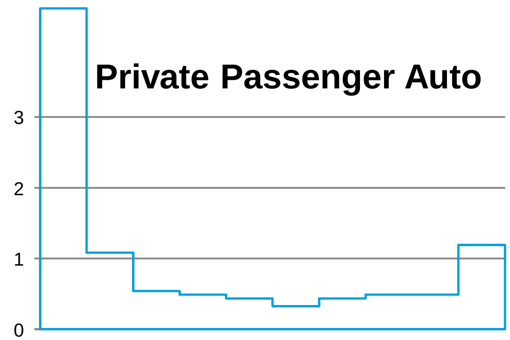


- Bootstrap multiple historical claim triangles
  - Multiple insurers
  - Multiple projection years
  - Or multiple random “realistic” triangles
- Aggregate the bucket counts across bootstraps
- Back-test passes if 10% of actual outcomes lie in each bucket
  - Within random sampling tolerance

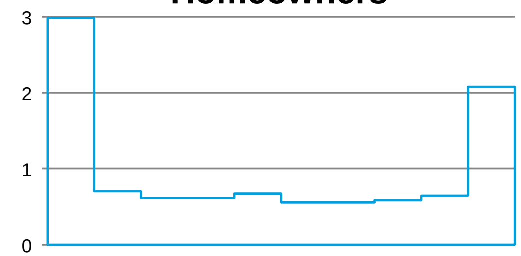
Back-test on Real Data: Leong et al (2012)

Commercial Auto Liability

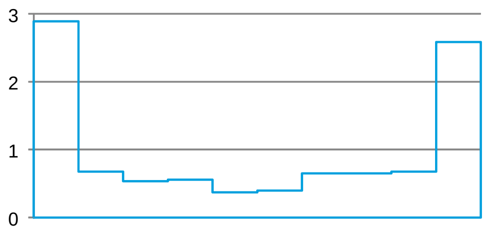




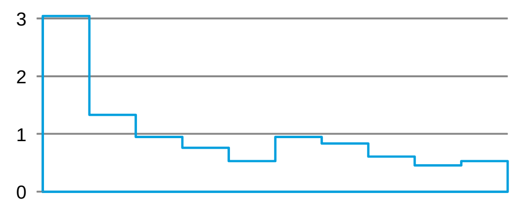
Homeowners



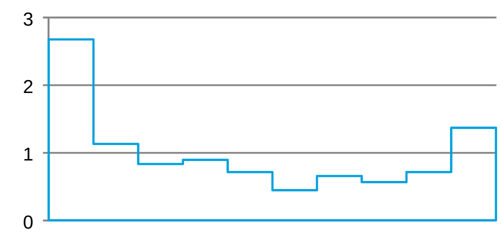
Workers Compensation



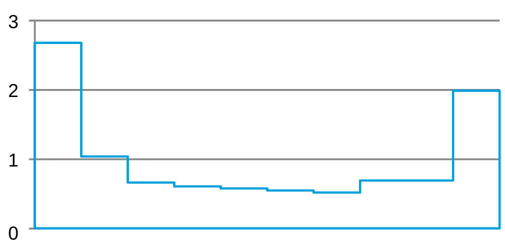
Medical Professional Liability



Other Liability



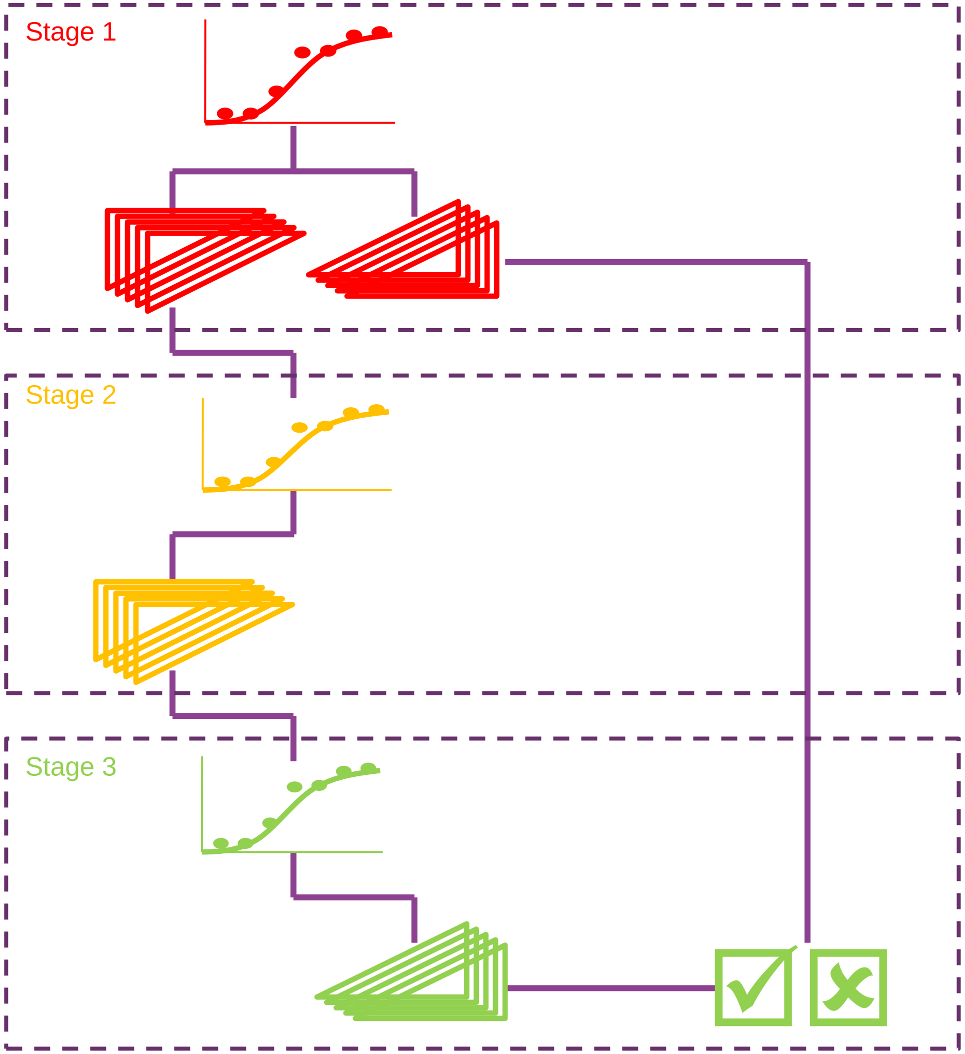
Commerical Multi Peril



Too many actual outcomes lie in the top and bottom bootstrap deciles, so bootstrap distribution too narrow

Different companies / years not independent so we do not know how significant this effect is

The Monte Carlo Back-Test (MCBT)



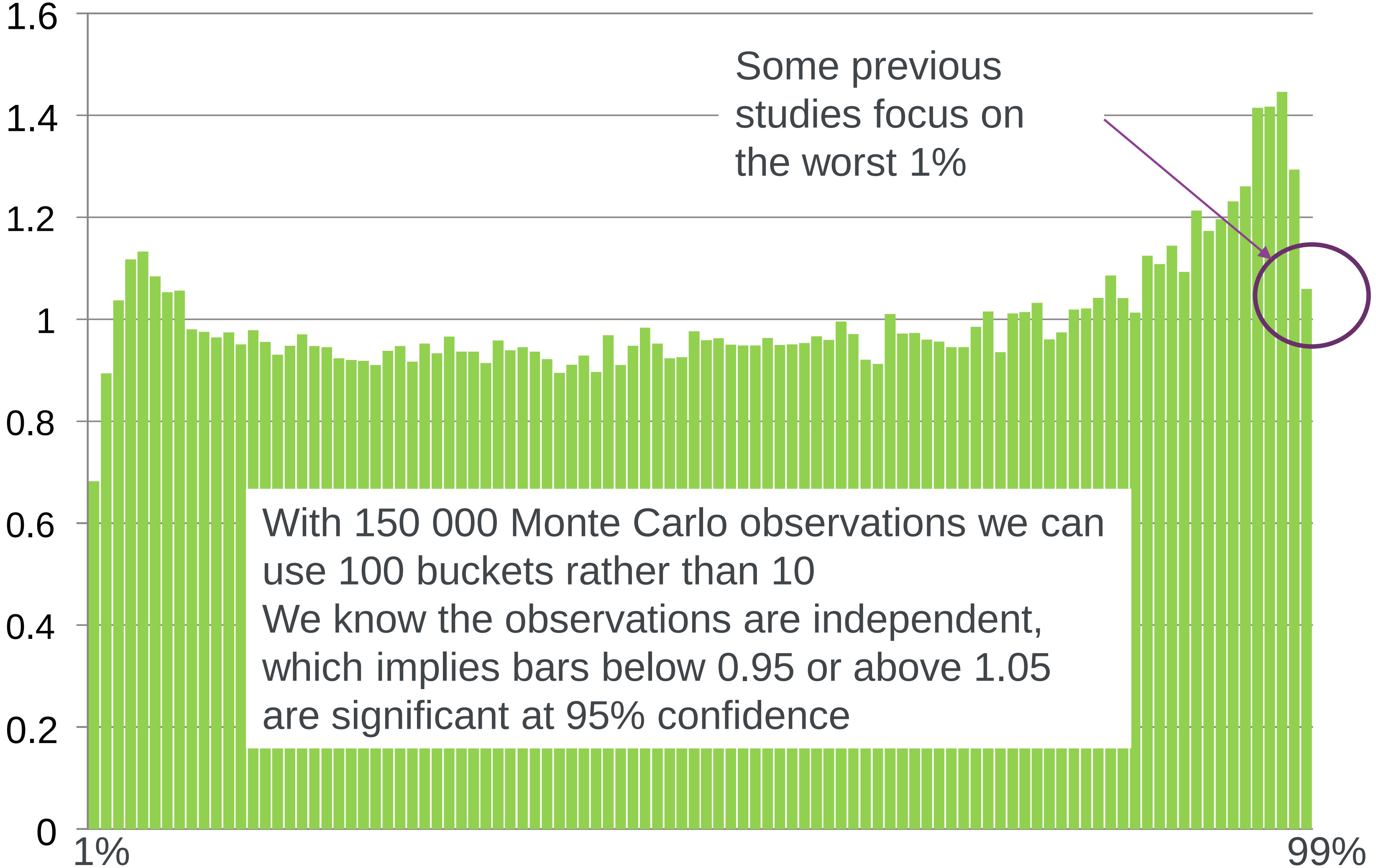
Parametric back-cast

One set of stage 2 parameters for each stage 1 back-cast triangle

Non-parametric back-cast

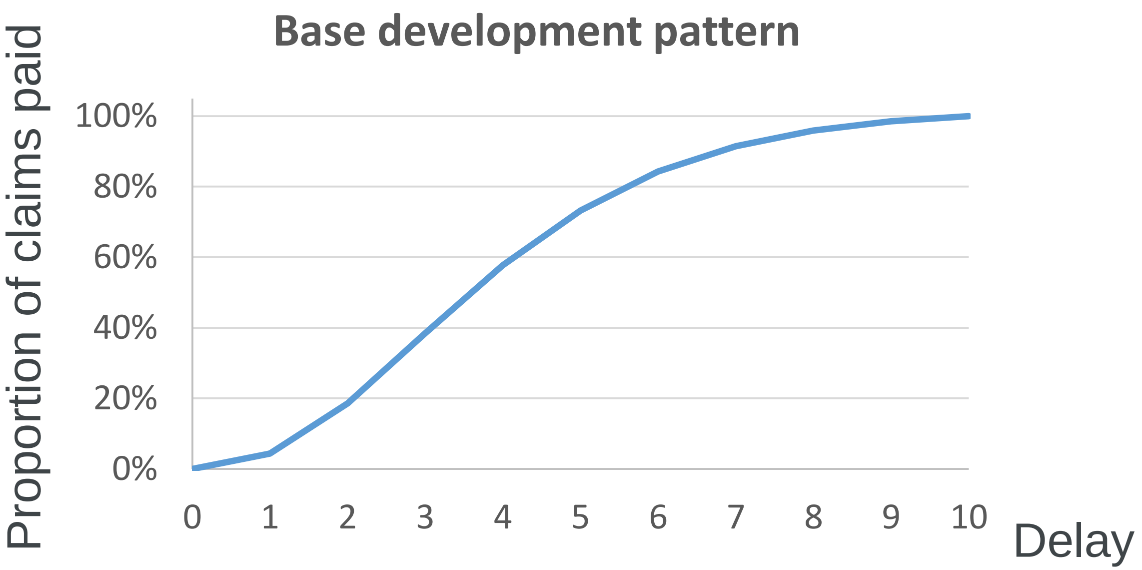
One set of stage 3 parameters for each stage 2 back-cast triangle

Example Output (Long dev, High vol)



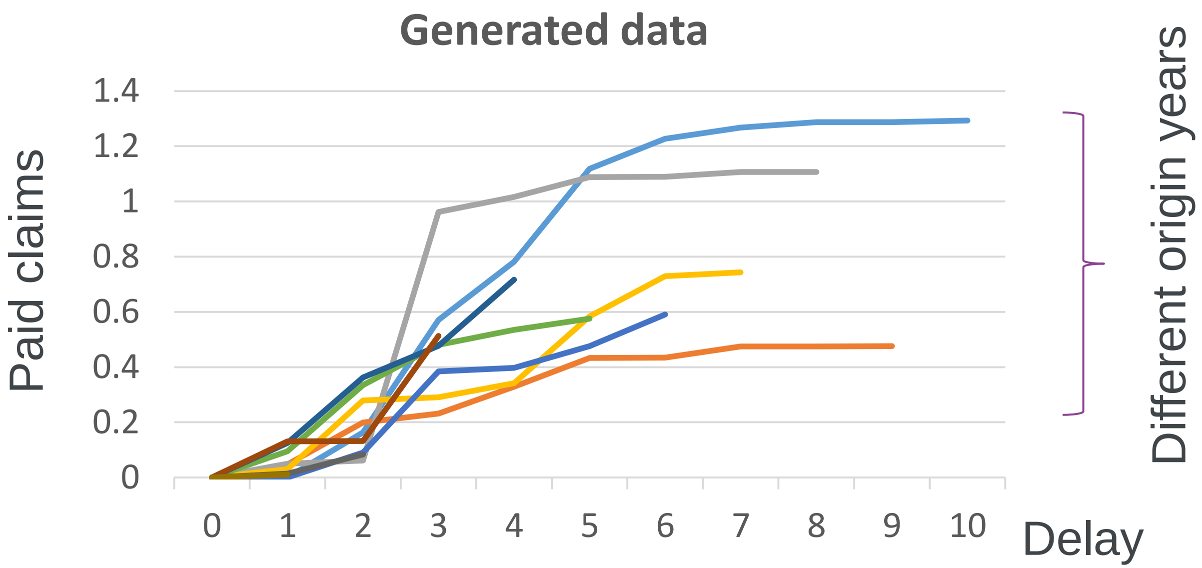
Generating triangles – simple case

- SIGNAL: Assume a base development pattern
  - Use the same pattern for all origin years
- NOISE: Incremental claims in each cell generated from a gamma distribution with mean from pattern (with specified gamma vol.)

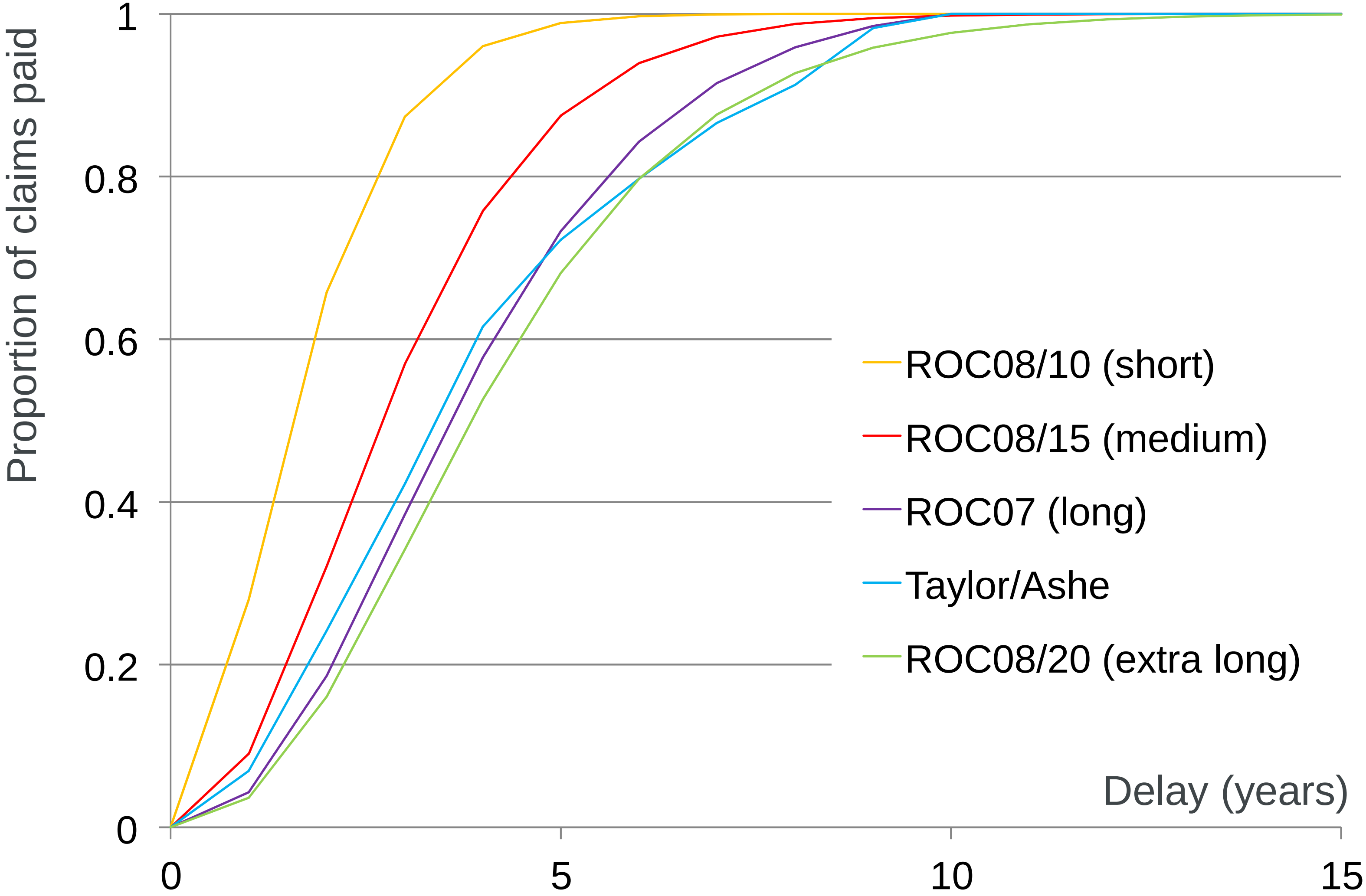


Generating triangles – simple case

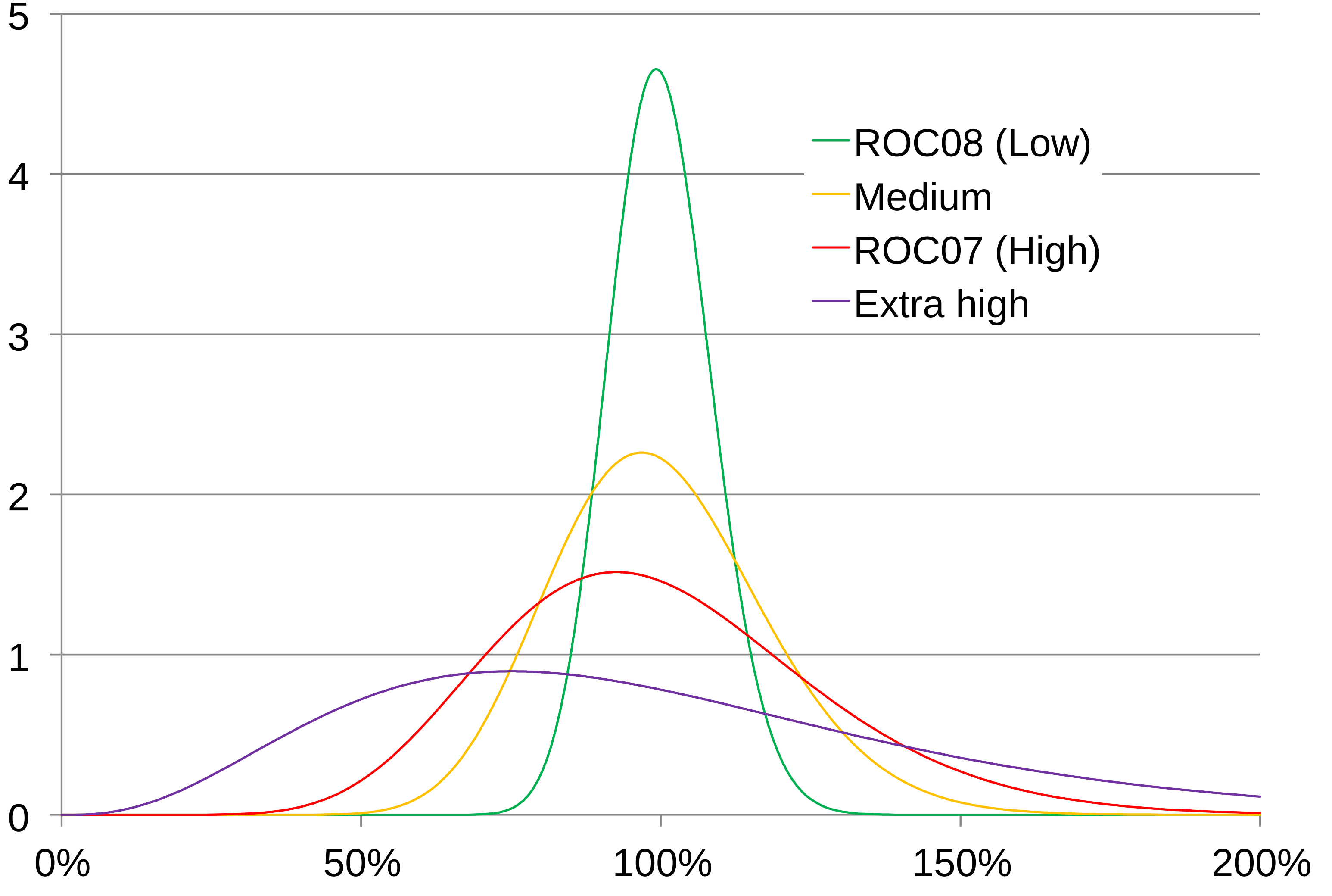
- SIGNAL: Assume a base development pattern
  - Use the same pattern for all origin years
- NOISE: Incremental claims in each cell generated from a gamma distribution with mean from pattern (with specified gamma vol.)



Claims Development Patterns



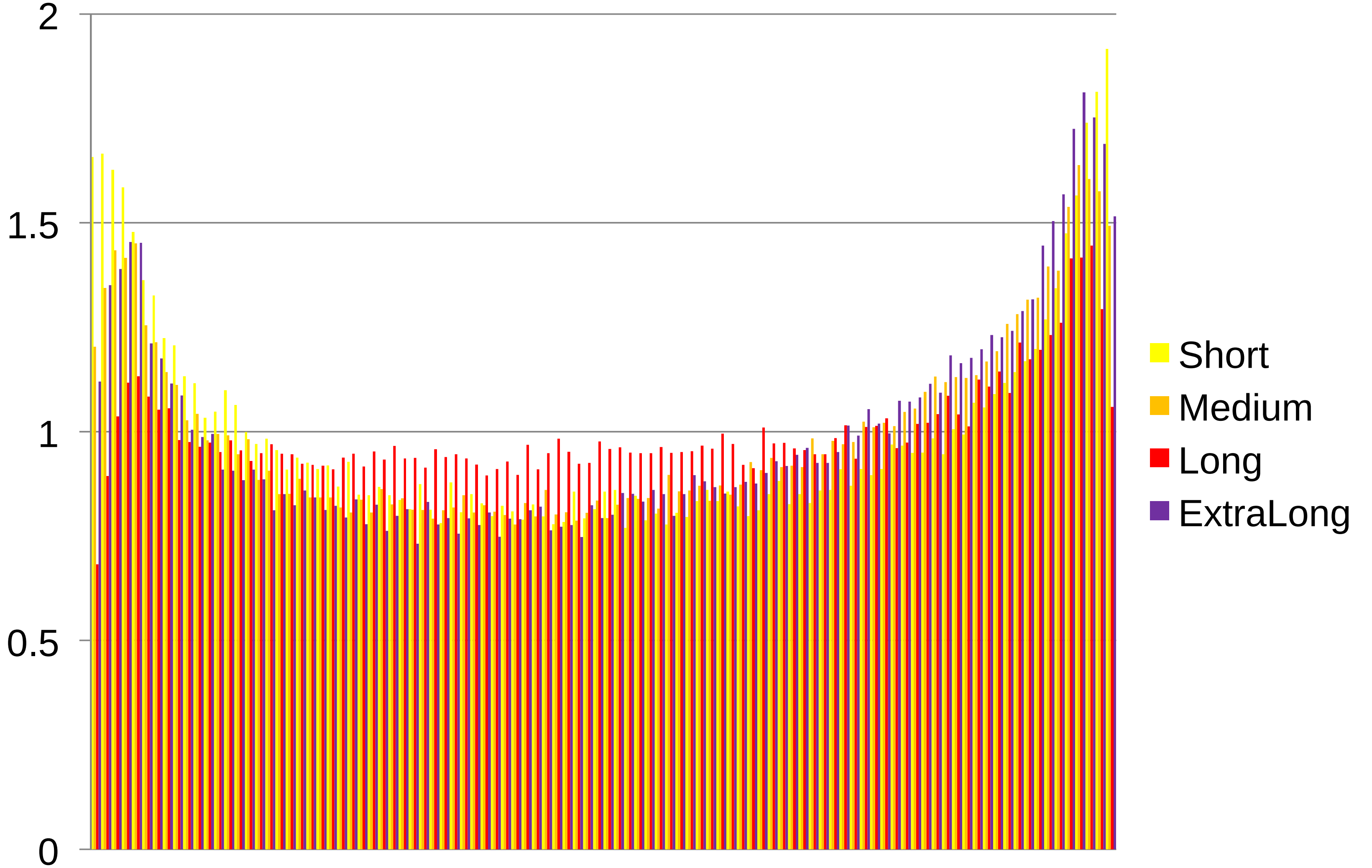
One Year's Ultimate Loss Distribution



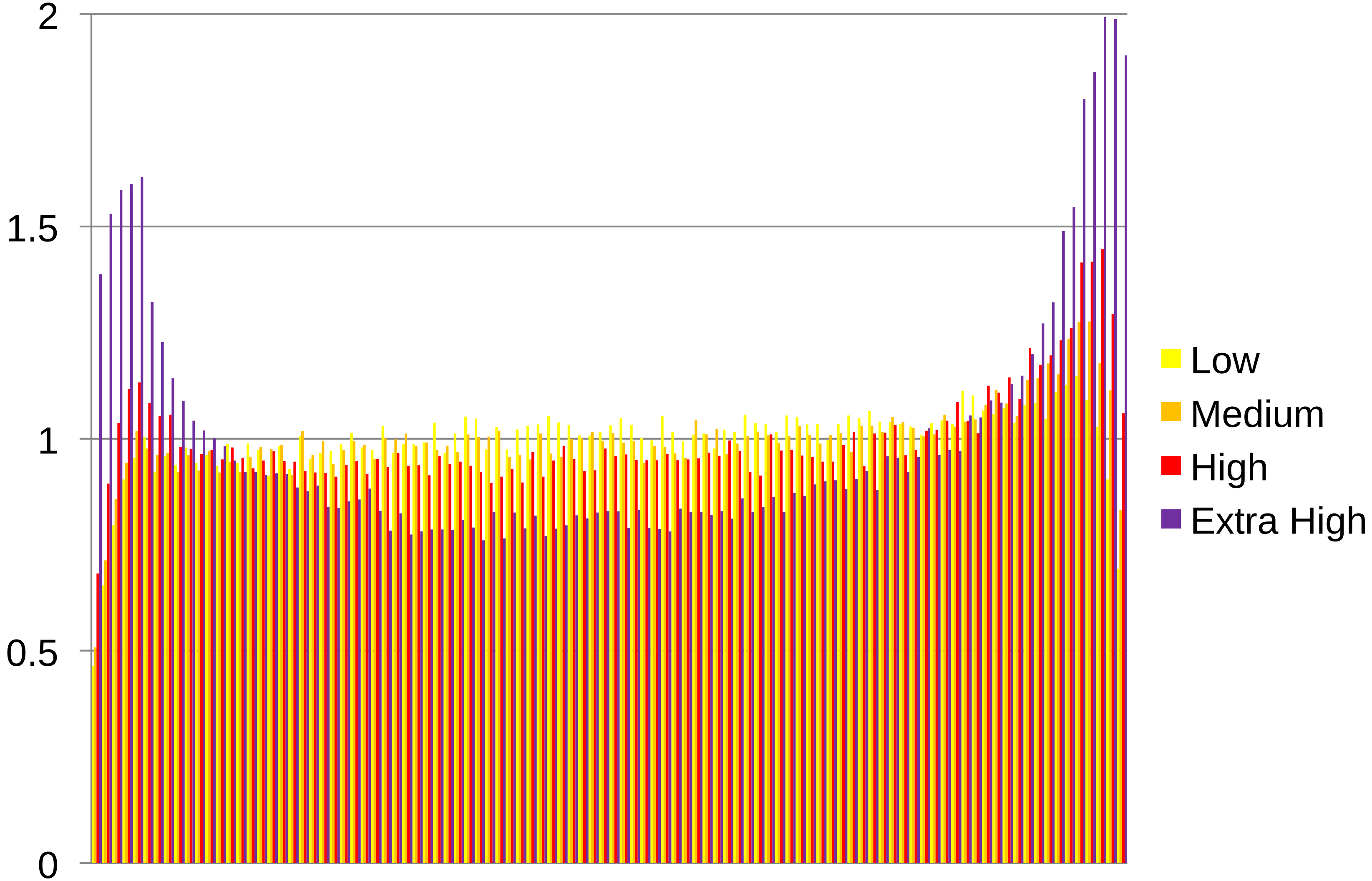
MCBT Results: Proportion > 99%-ile

	Development pattern length			
Gamma Volatility	Short	Medium	Long	Extra Long
Low	1.1%	0.7%	0.7%	0.6%
Medium	1.5%	1.1%	0.8%	1.1%
High	1.9%	1.5%	1.1%	1.5%
Extra High	3.0%	2.7%	1.9%	2.7%

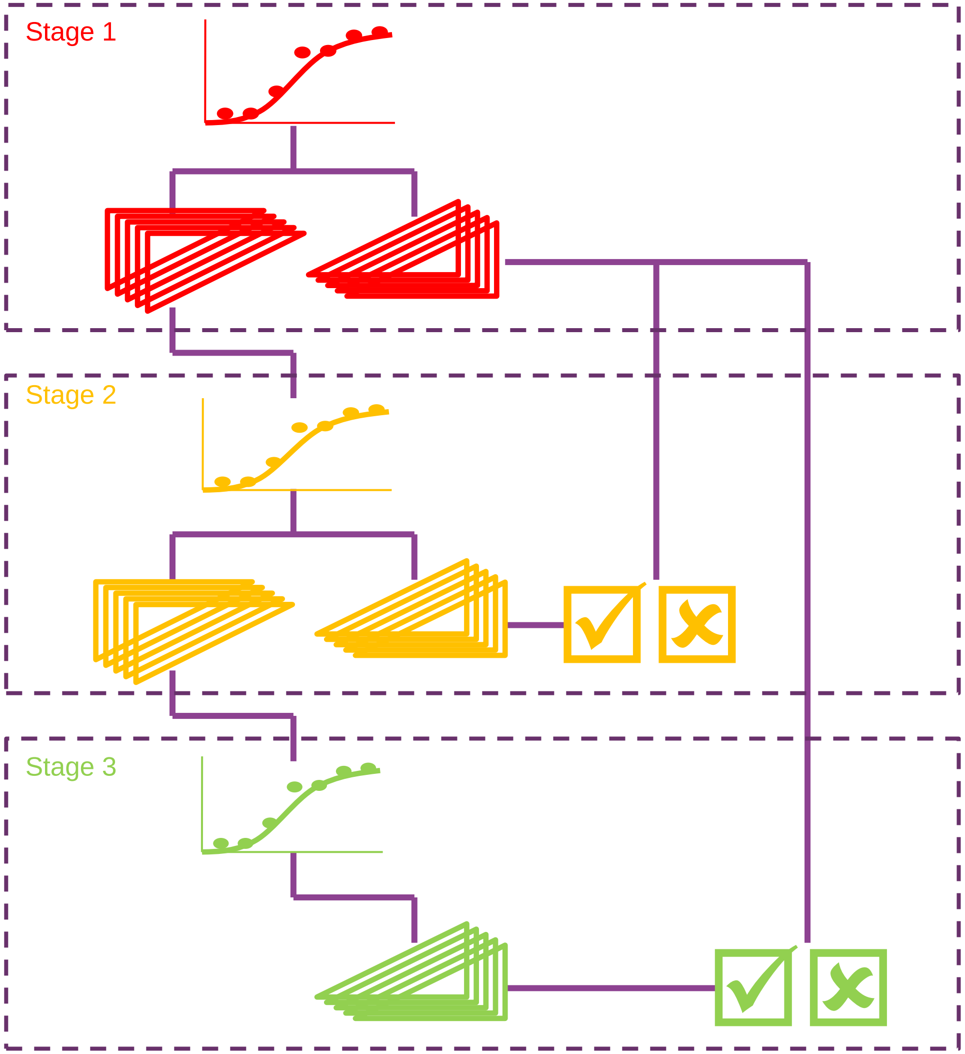
Impact of Tail Length (High Volatility)



Impact of (Gamma) Volatility (Long Tail)



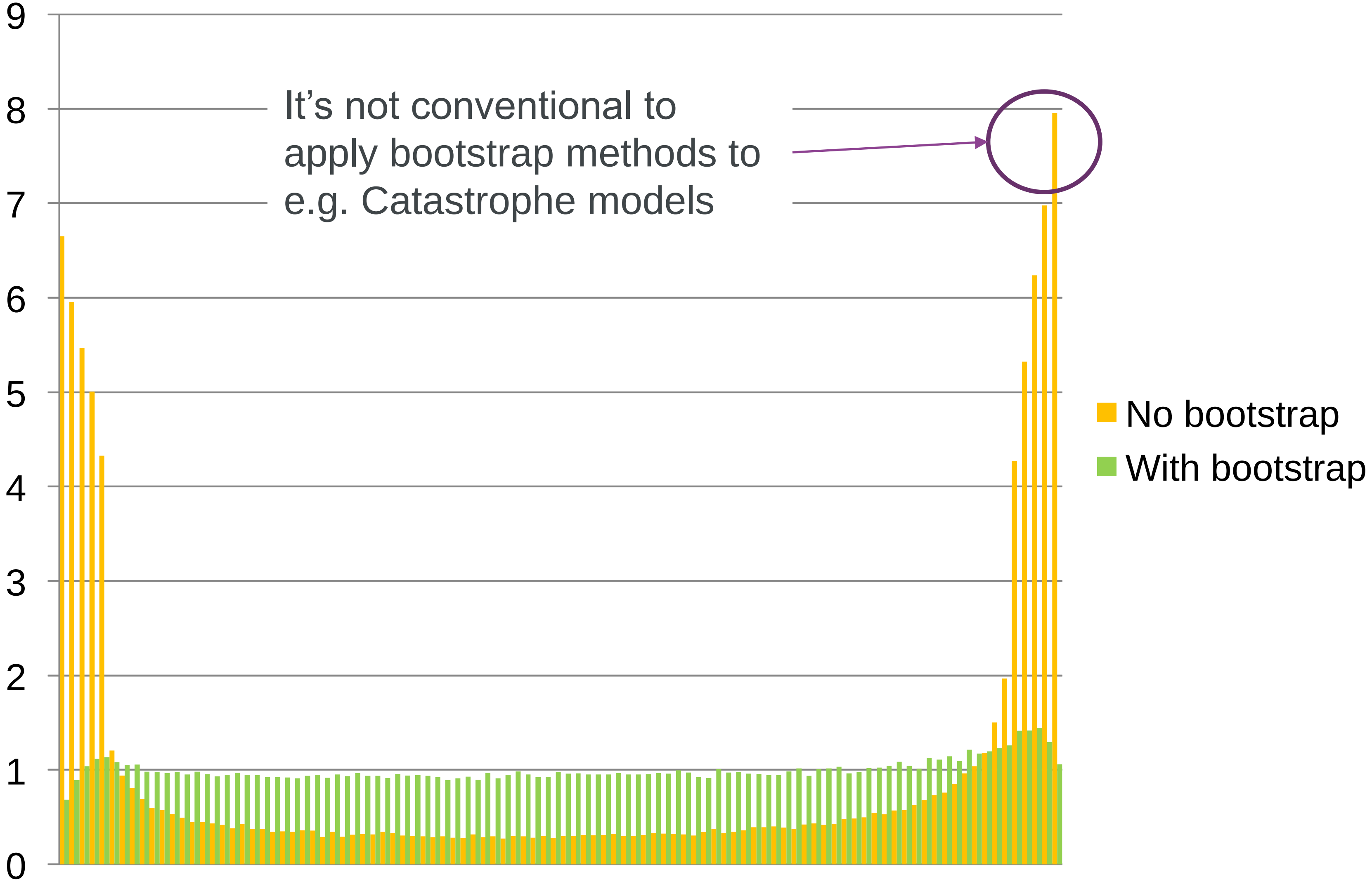
Experiment: Omitting the Back-Cast



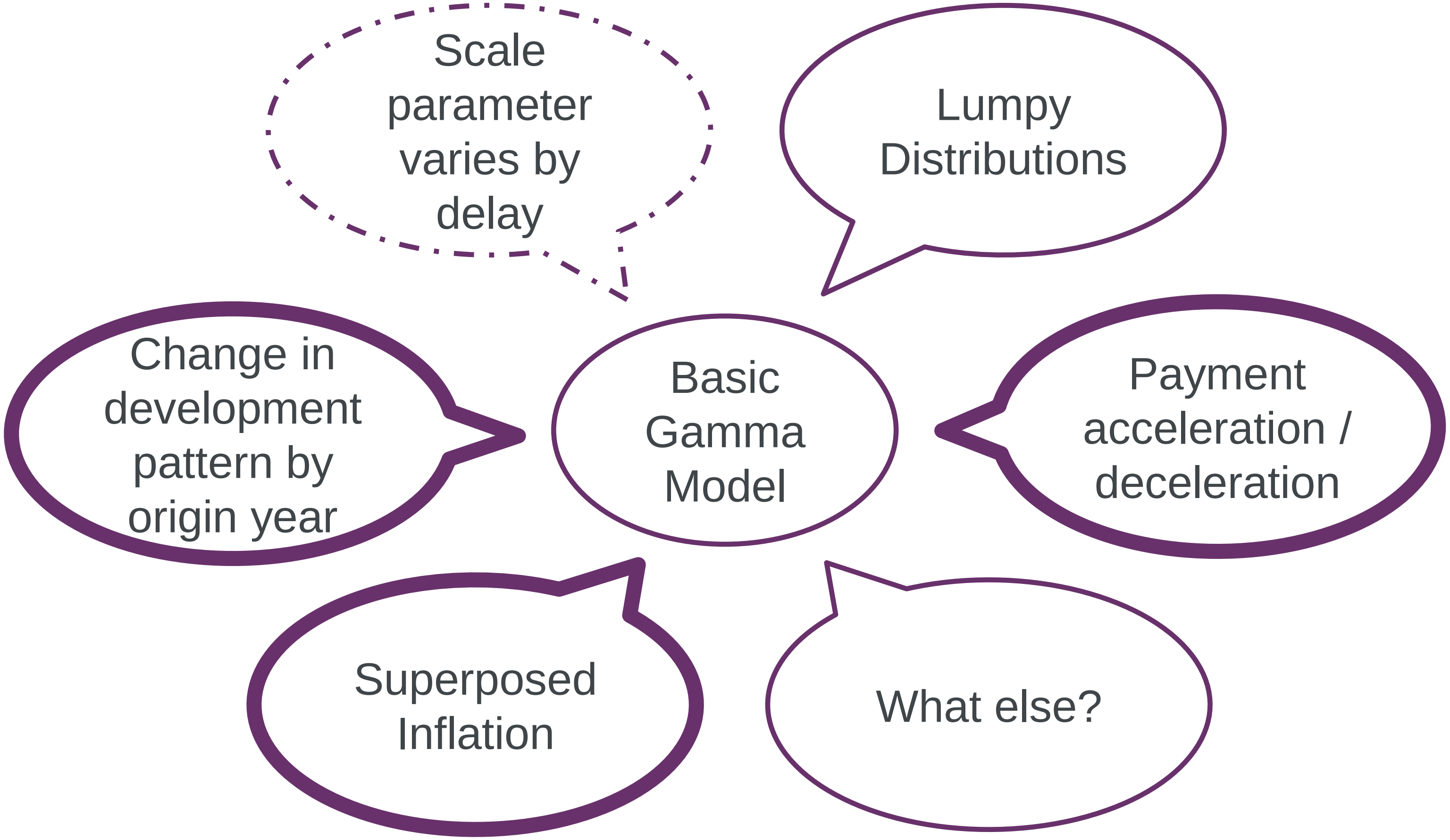
How much is the bootstrap adding?

- Perform the Monte Carlo Back Test using stochastic projection of the step 2 fitted parameters
- Therefore have no allowance for parameter uncertainty

Much Better to Bootstrap than Not

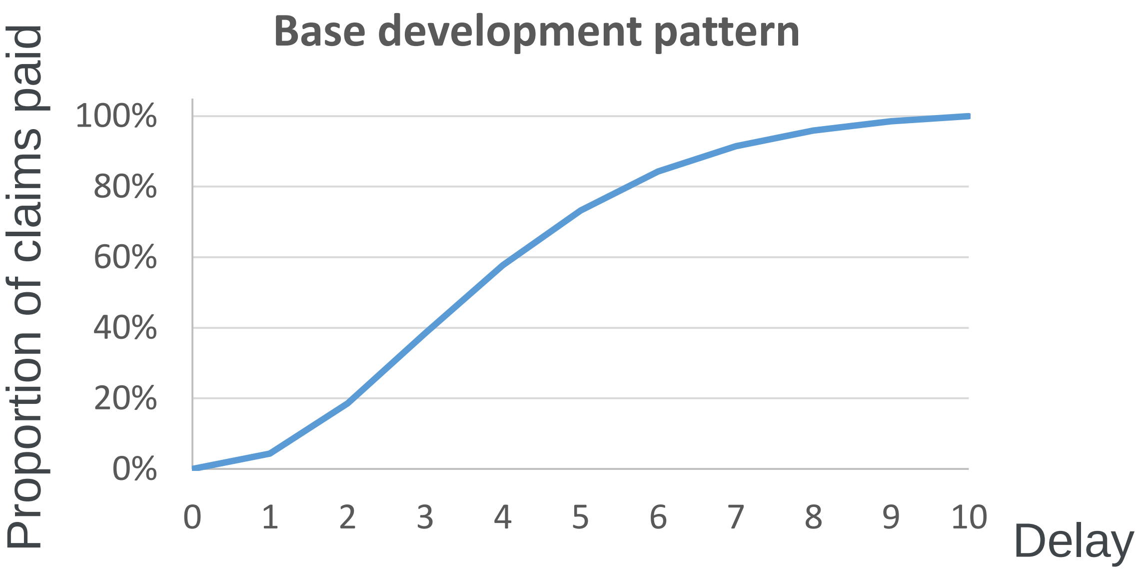


Making Triangles More Realistic



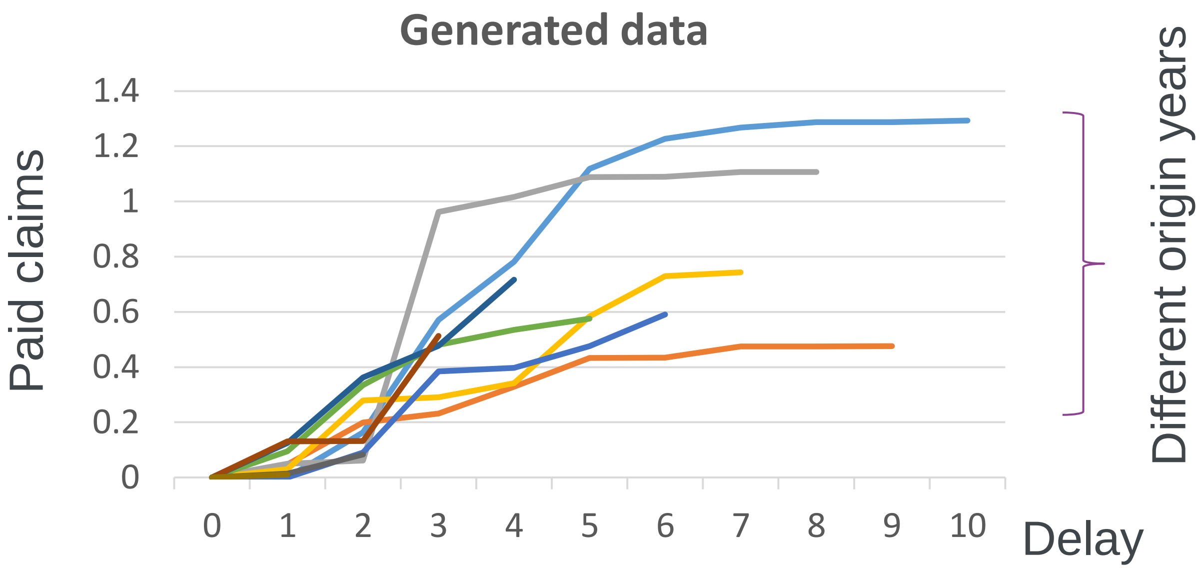
Generating triangles – simple case

- SIGNAL: Assume a base development pattern
  - Use the same pattern for all origin years
- NOISE: Incremental claims in each cell generated from a gamma distribution with mean from pattern (with specified gamma vol.)



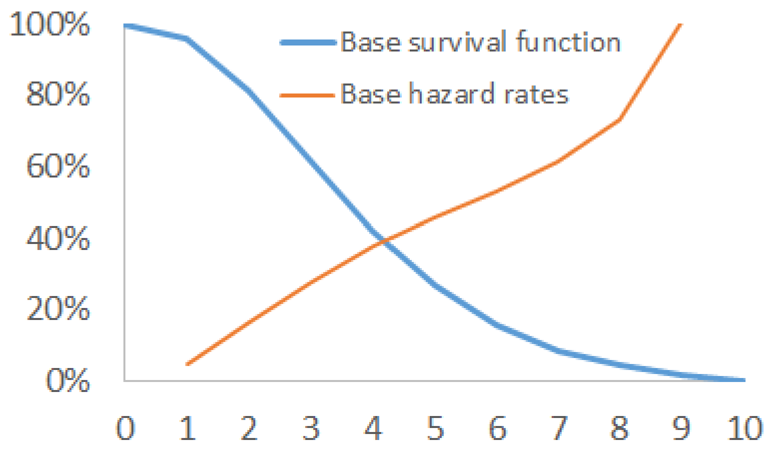
Generating triangles – simple case

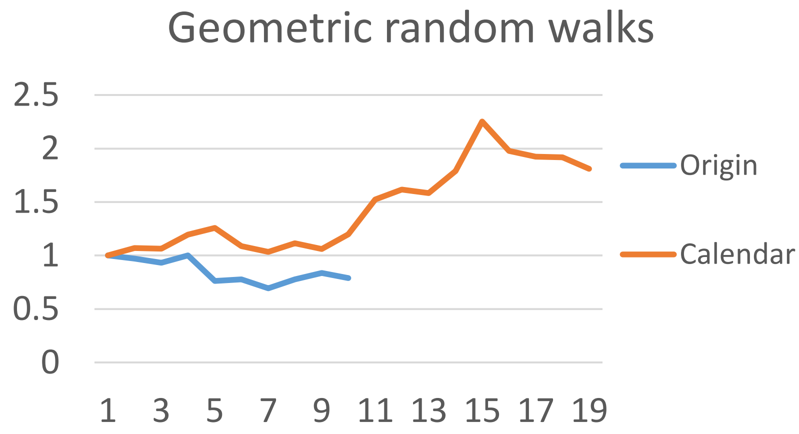
- SIGNAL: Assume a base development pattern
  - Use the same pattern for all origin years
- NOISE: Incremental claims in each cell generated from a gamma distribution with mean from pattern (with specified gamma vol.)



Making Triangles More Realistic

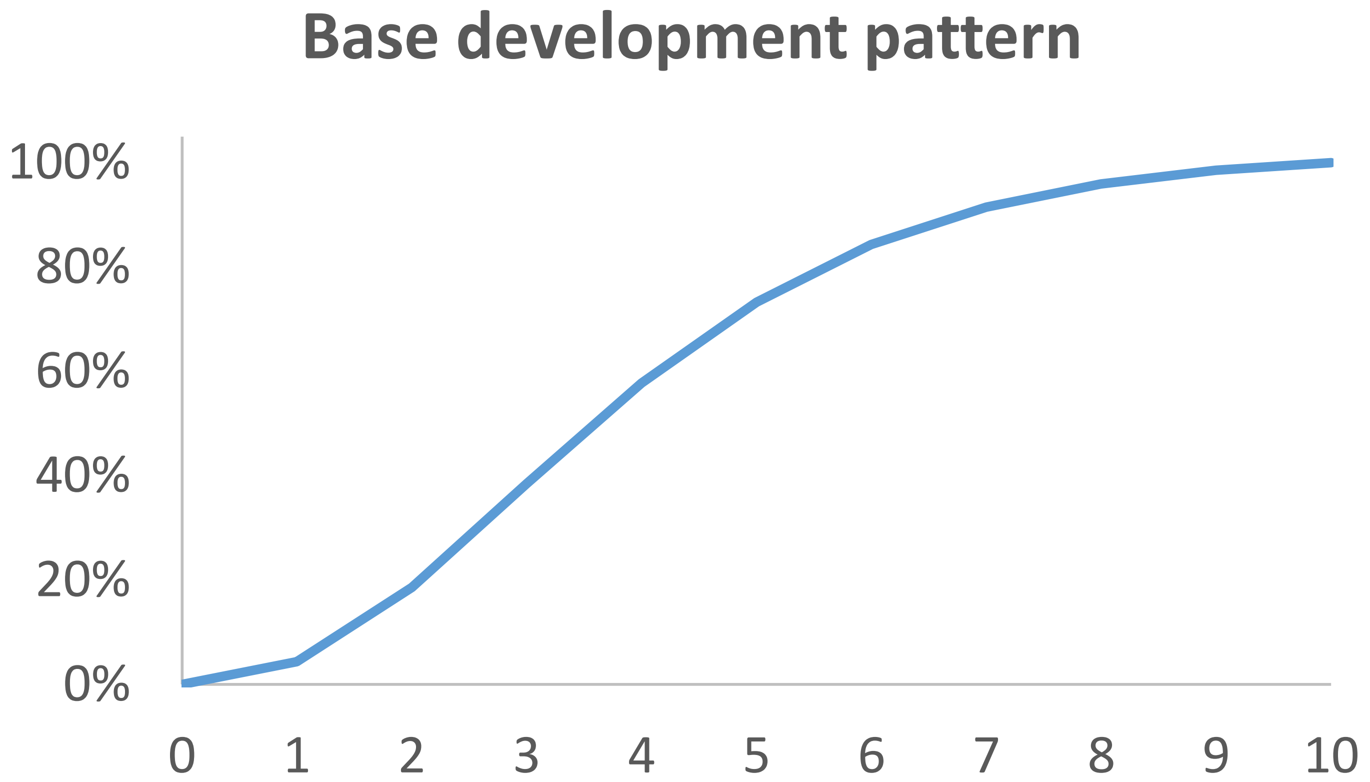
- SIGNAL
  - Express the development pattern as hazard rates (compare force of mortality μ_x = minus log of survival rate p_x)
  - Transform these hazard rates for each origin year
    - Multiply by geometric random walks for origin period and calendar (or equivalently, raise survival rates to a power)





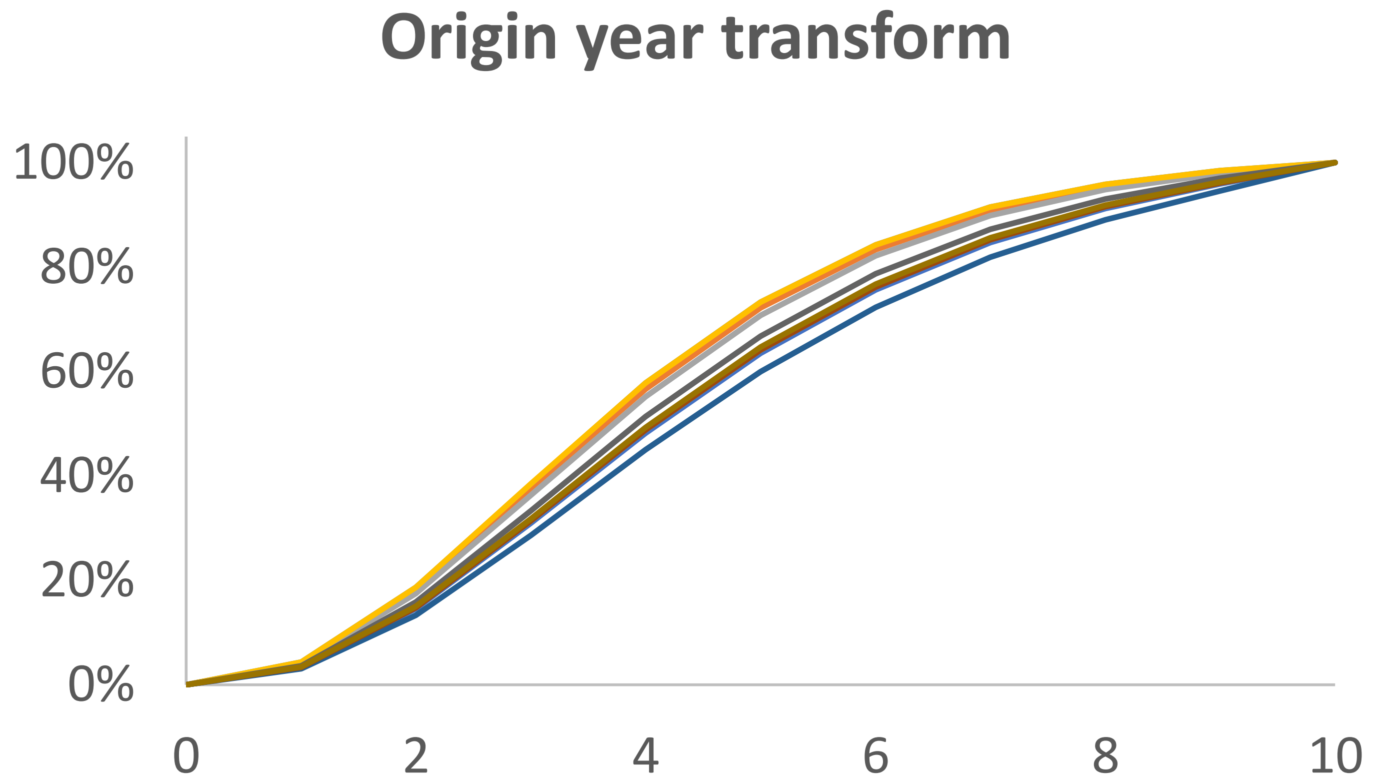
Making Triangles More Realistic

- SIGNAL: example



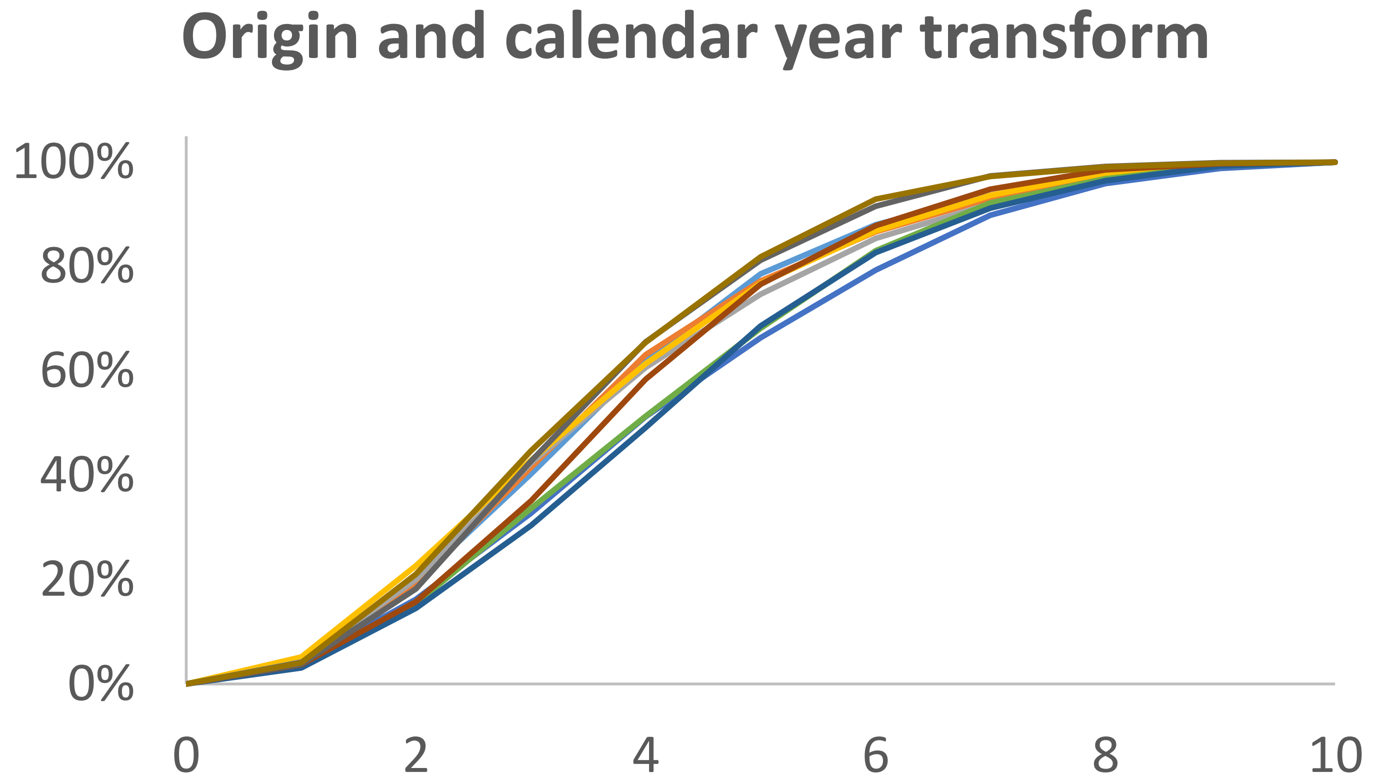
Making Triangles More Realistic

- SIGNAL: example



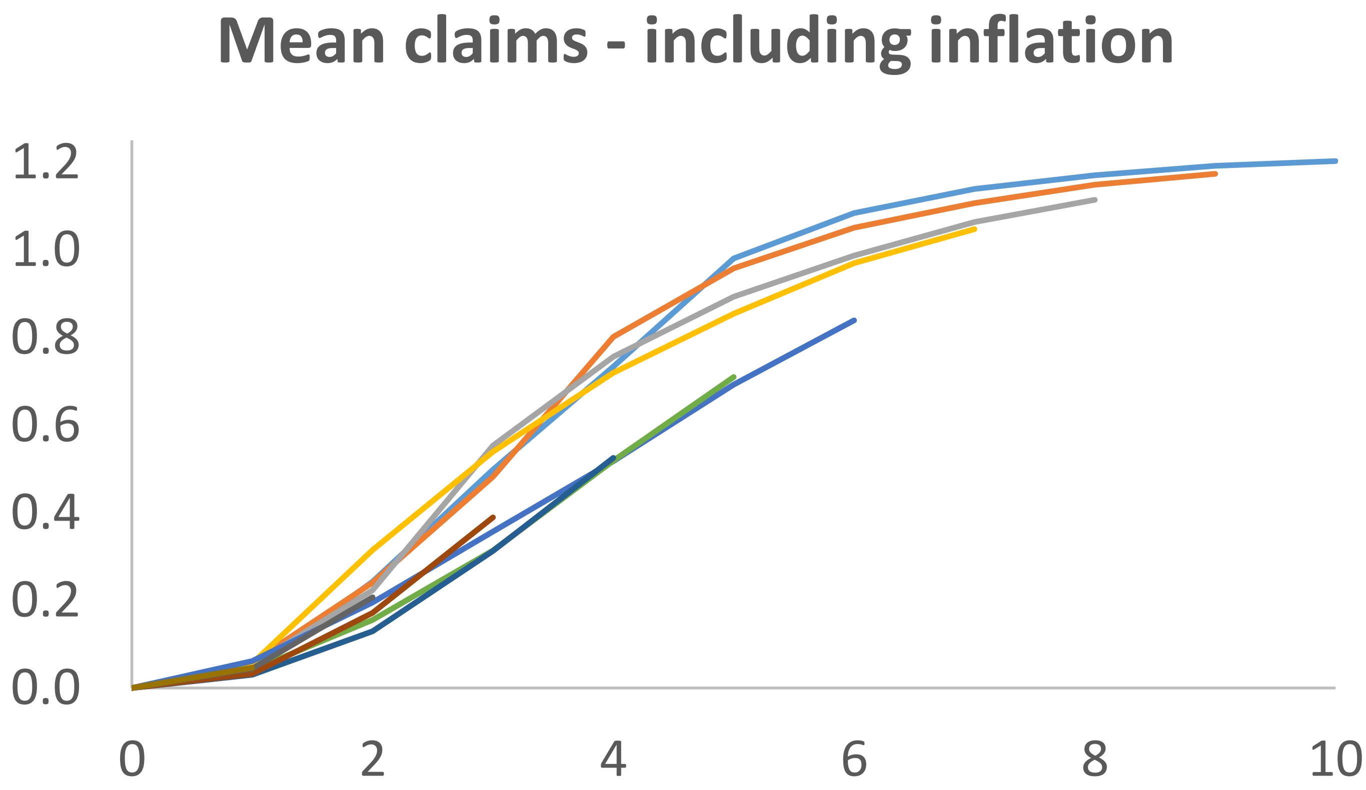
Making Triangles More Realistic

- SIGNAL: example

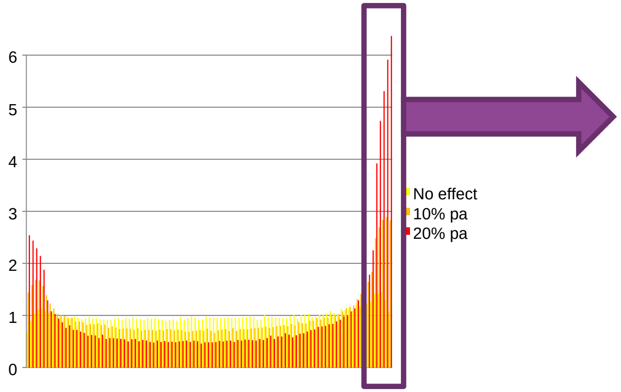


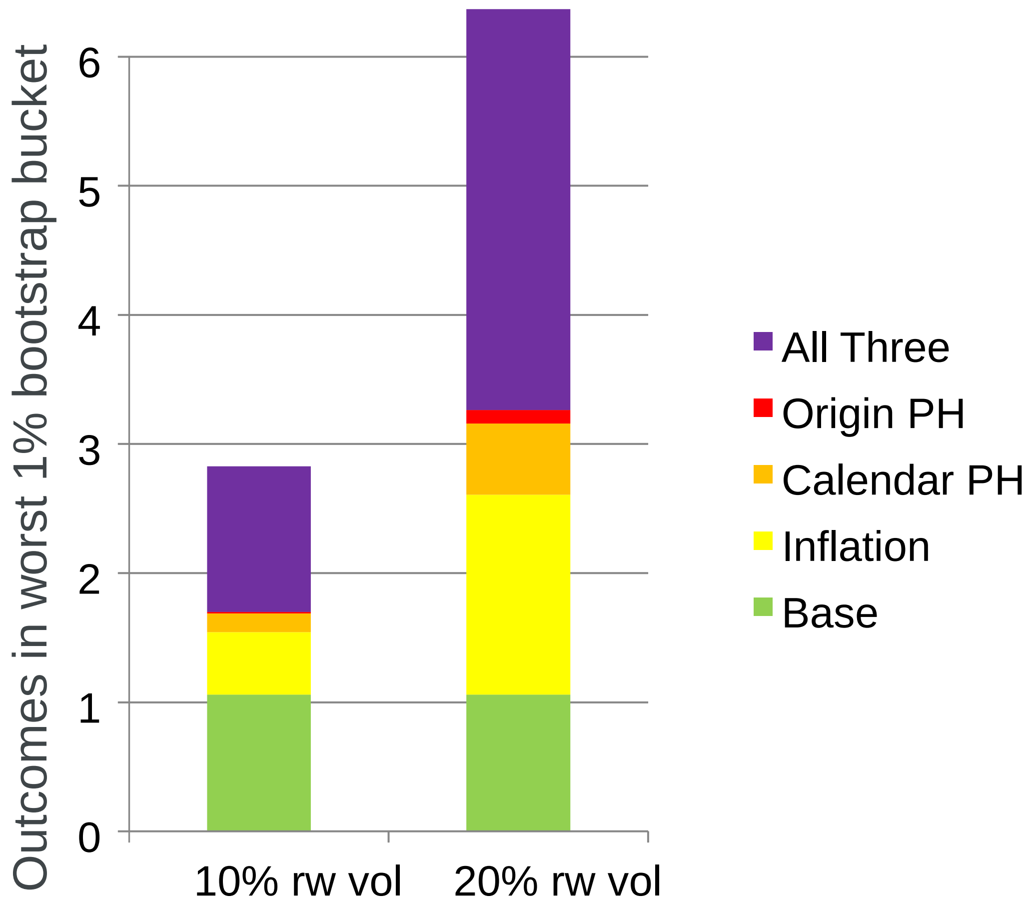
Making Triangles More Realistic

- SIGNAL: example



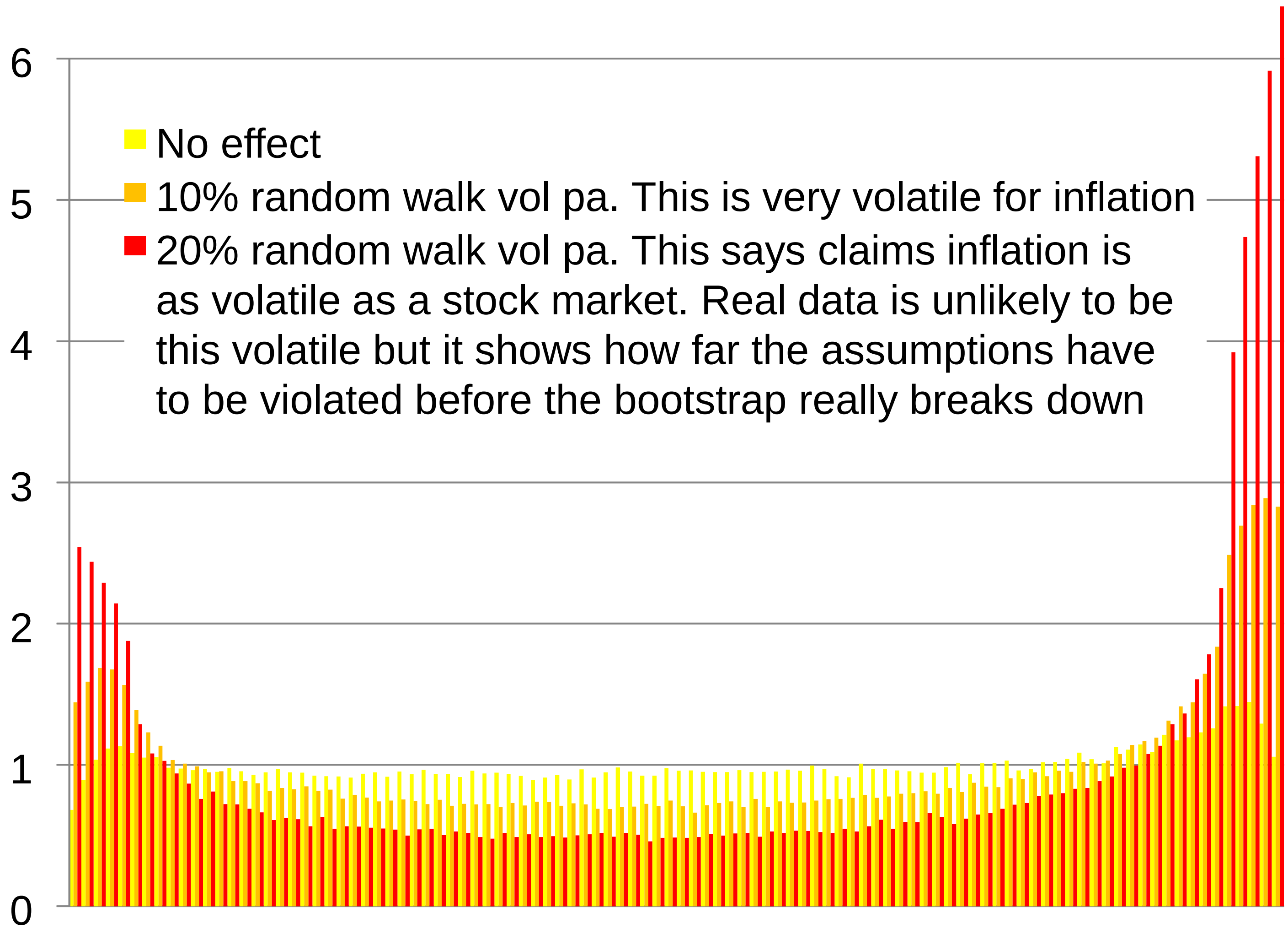
Origin Year and Calendar Year Effects



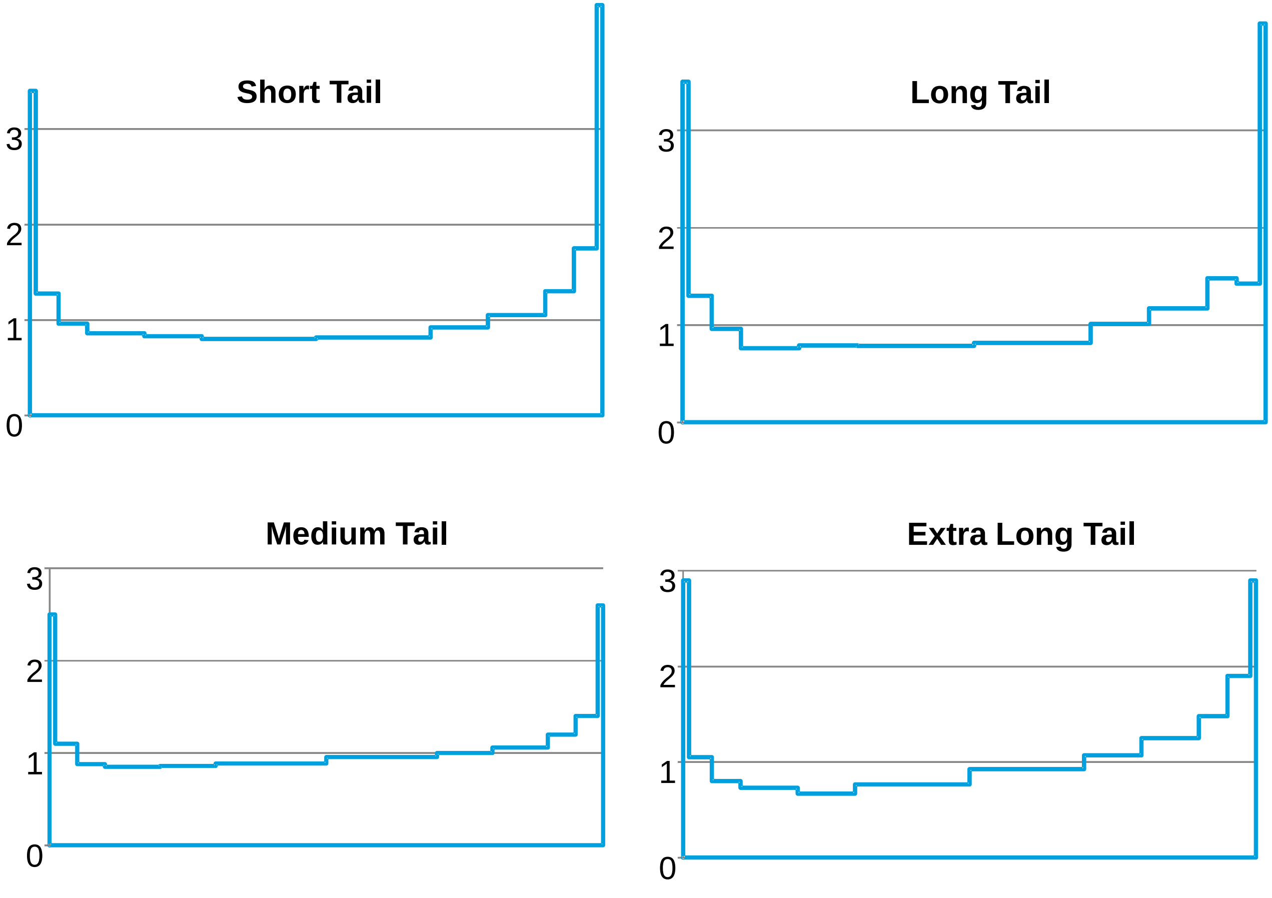


These figures relate to the long development pattern, and high gamma volatility

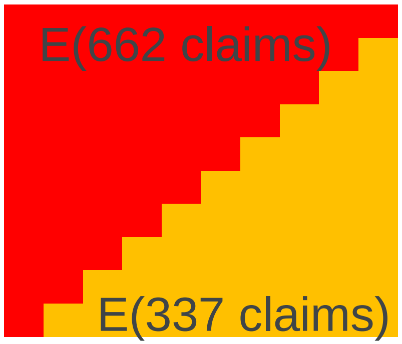
Origin Year and Calendar Year Effects

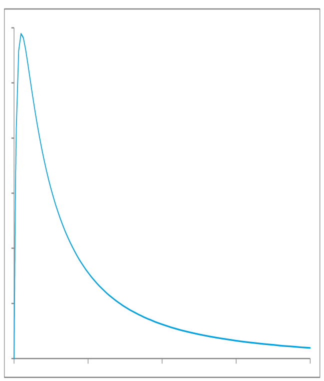


Lumpy Claims (ROC 2008, Method B)

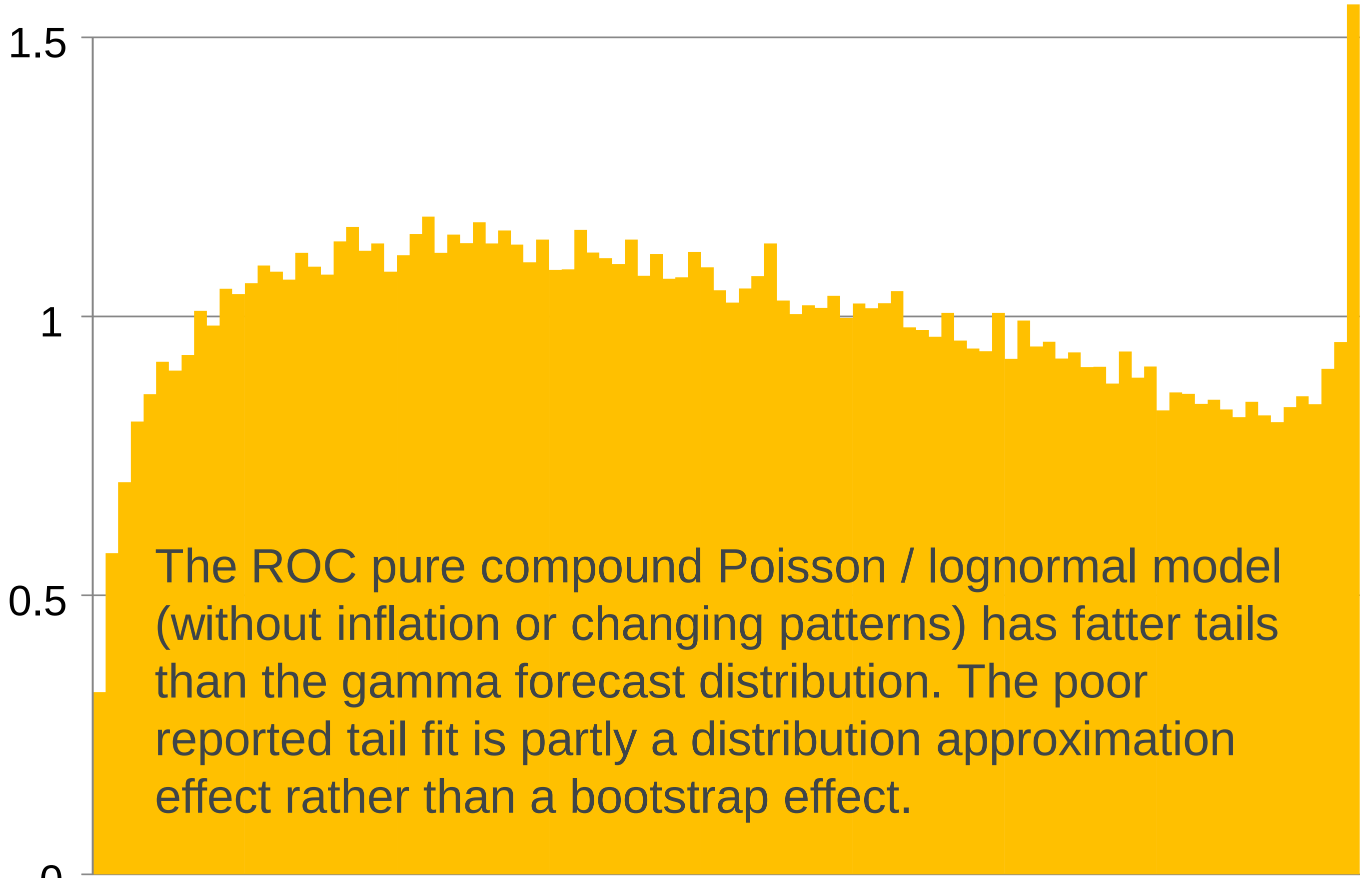


ROC Data Process: Compared to Gamma





Lognormal $\sigma^2=2$



The ROC pure compound Poisson / lognormal model (without inflation or changing patterns) has fatter tails than the gamma forecast distribution. The poor reported tail fit is partly a distribution approximation effect rather than a bootstrap effect.

Agenda

- Aims
- What is the bootstrap?
- Back-test results
- Does the bootstrap underestimate extreme percentiles?
- Conclusions

Extreme Percentile Underestimation

- Given how simple the concept is, the bootstrap does well for most of the distribution
- We replicate results of ROC and others that bootstrap does not perfectly capture extreme tails
  - In some instances the bootstrap distribution is less extreme than reality, and in others it is more extreme
  - We would not expect perfection
- Bootstrap is remarkably robust to moderate assumption violations

Some alternatives to the Bootstrap

- Bootstrap is not the only way to address parameter uncertainty
- Classical methods estimate parameters by maximum likelihood and derive standard errors from the Fisher information matrix
- Bayesian methods (England & Cairns, 2009)
- Single (non-bootstrap) forecast, with the standard deviation multiplied by an “adjustment factor”
  – Use the Monte Carlo Back Test to solve for the adjustment factor

Agenda

- Aims
- What is the bootstrap?
- Back-test results
- Does the bootstrap underestimate extreme percentiles?
- Conclusions

The Importance of Quantitative Testing

	Rhetorical	Quantitative
Mechanical Methods	Bootstrap takes account of parameter error Assumptions do not hold in practice Frequentist peg in a Bayesian hole	**Prob**{outcome > 99%-ile} Robustness to model mis-specification Unbiasedness / efficiency
Subjective methods	Use underwriting knowledge, common sense, relevant for the board, practical decisions Telling management what they want to hear, profit smoothing.	? Wanted: outcome-based tests for subjective methods

Points for Discussion (1)

- What to do about known changes (premium cycles, legal developments, settlement processes, inflation etc)?
  - Should we strip them out of the data and put them back into the forecast? Or is that part of the noise we're trying to measure and extrapolate?
  - The bootstrap allows for these mechanically, but only to the extent that these fluctuations affect the past and the future

Points for Discussion (2)

- Parameter and model errors pervade many risk models: premium risk, cat risk, reserving risk, credit risk, market risk etc.
  - For reserving risk we have the bootstrap. It’s not perfect but we’d give it 8/10 for capturing model and parameter uncertainty
  - For the other risks, we probably ignore model and parameter error, scoring 0/10
  - We can try to perfect the bootstrap, but should we prioritise other risks?

Acknowledgements

- We are grateful for the support from the Managing Uncertainty Qualitatively working party, the Managing Uncertainty with Professionalism working party and from our employers
- Special thanks to Sarah MacDonnell, Tom Wright and Peter England for detailed comments on earlier drafts
- All views expressed and any remaining errors are ours alone

Further Reading

Cairns M and England P D (2009) Are the upper tails of predictive distributions of outstanding liabilities underestimated when using bootstrapping? General Insurance Convention (presentation only) http://www.actuaries.org.uk/sites/all/files/documents/pdf/a09england.pdf

England, P.D. and Verrall, R.J. (2002) Stochastic Claims Reserving in General Insurance (with discussion). British Actuarial Journal, 8, pp 443-544 http://www.actuaries.org.uk/sites/all/files/documents/pdf/sm0201.pdf. Note that this link is to the originally distributed sessional meeting paper. The is a crucial typographical error for the residual adjustment in Appendix 3 which is corrected in the paper finally published in the BAJ, and also in this paper.

Efron B & Tibshirani, R J (1993). An introduction to the bootstrap. Chapman and Hall.

General Insurance Reserving Oversight Committee Working Party (Chair: Lis Gibson, 2007) Best Estimates and Reserving Uncertainty.

General Insurance Reserving Oversight Committee Working Party (Chair: Neil Bruce, 2008) Reserving Uncertainty.

Previous two papers available here: http://www.actuaries.org.uk/practice-areas/pages/general-insurance-reserving-oversight-committee-gi-roc-0

Leong J, Wang S and Chen H. (2012) Back-Testing the ODP Bootstrap of the Paid Chain-Ladder Model with Actual Historical Claims Data. Casualty Actuarial Society E-forum. http://www.casact.org/pubs/forum/12sumforum/leong_wang_chen.pdf

Pinheiro, P J. R; João Manuel Andrade e Silva and Maria de Lourdes Centeno. (2003). Bootstrap Methodology in Claim Reserving. The Journal of Risk and Insurance, 70: 701-714.

Shapland M R and Leong, J. (2010) Bootstrap Modeling: Beyond the Basics. Casualty Actuarial Society E-forum.

Questions

Comments