The Actuarial Profession
making financial sense of the future

# Parameter Uncertainty and Capital Modelling

GIRO
20th October 2005
Richard Millns and Richard Weston

## Overview

- What is parameter uncertainty?
- Why is it important?
- How can it be quantified?
- What impact can it have on capital modelling?

## Sources of Uncertainty

Model

↓

Parameters

↓

Process

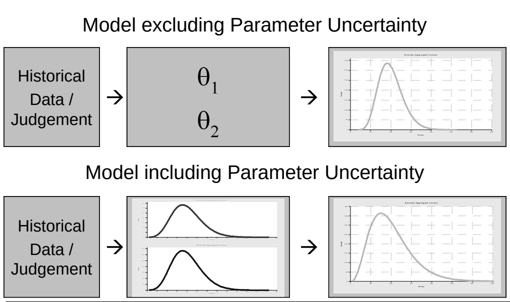
## What is Parameter Uncertainty?

- Having chosen a particular model, parameters are usually estimated from a sample of data
- There is always uncertainty associated with any estimate
- In many actuarial models, this parameter uncertainty is often ignored

## Estimating Model Parameters

- For example, maximum likelihood estimates are often used
- These are the parameters that are statistically “most likely” given the observed data
- However, we never have enough data to be sure that these are the “true” parameters. Other parameters are possible, just less likely.

## Incorporating Parameter Uncertainty

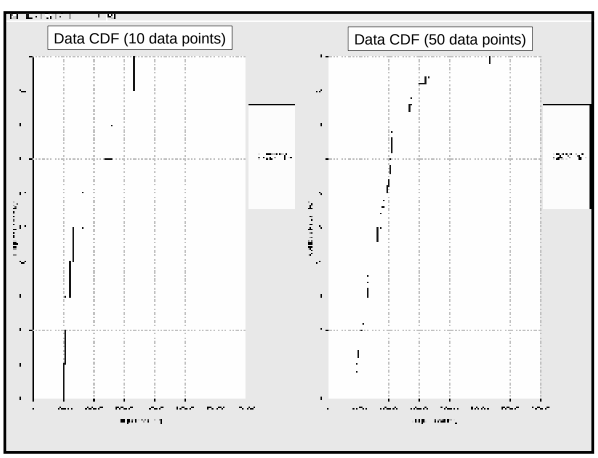


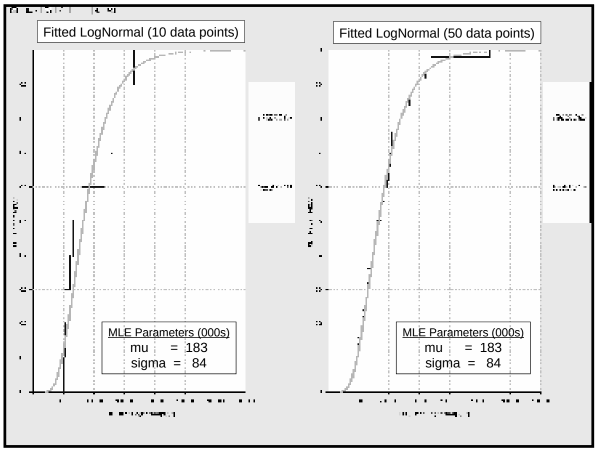
## Why is Parameter Uncertainty Important?

- Danger of underestimating variability of underlying ICA model variables (e.g. claims) where data is limited, leading to underestimation of capital, premiums etc. or other incorrect business decisions
- FSA has indicated it is keen to see allowance where there is limited data to support parameters

## Why is Parameter Uncertainty Important?

- In an ICA context, parameter uncertainty can impact models for:
  - Claims
  - Premium rates
  - Credit risk
  - Dependencies
  - Economic scenarios
  - …






## Quantifying Parameter Uncertainty

- Maximum likelihood parameters are the same for each data set
- However, intuitively we are less certain about the parameters based on the first (smaller) data set
- Just how uncertain are the parameters derived from the first data set?
- Is parameter uncertainty important for the second data set?
- How do I incorporate this uncertainty into my model?

## Quantifying Parameter Uncertainty

- "Classical Statistics"
  - Asymptotic distribution of ML estimates
- Bootstrapping
  - Resample with replacement from the original sample
  - Re-fit model to each sample
- Bayesian Statistics
  - Use Bayes' Theorem to determine posterior distribution of the parameters, given the data and prior beliefs

## Asymptotic MLE Distribution

As $n \to \infty$, $\hat{\vartheta} \to$ Multivariate Normal distribution with

Mean = $\hat{\vartheta}$

Variance-Covariance matrix = $\left( \frac{\partial^2 \ell}{\partial \vartheta^2} \right)^{-1}_{\vartheta = \hat{\vartheta}}$

The Actuarial Profession
making financial sense of the future

## Asymptotic MLE Distribution

- Pros
  - Easy to simulate
- Cons
  - Cannot always directly calculate the Var-CoVar matrix
  - Can give invalid parameter values (e.g. –ve standard deviation)
  - Need large no. of data points before Normal approximation is appropriate

The Actuarial Profession
making financial sense of the future

## Bootstrapping

- Re-sample same number of data points from original data (with replacement) many times
- Fit model to each set of “pseudo data”
- Gives (joint) distribution of the parameters
- Widely used in reserving risk models

The Actuarial Profession
making financial sense of the future

## Bootstrapping

- Pros
  - Easily understandable approach
  - No hard maths
- Cons
  - Requires parameters to be fitted for each bootstrap iteration - computationally intensive
  - Limited number of possible re-sampled data sets (can lead to problems with small data samples)

## Bayesian Method

- From Bayes' Theorem, given the data *y* and a prior distribution of the parameters $\Theta$, we produce a posterior distribution:

$$\mathrm{p}(\Theta \mid y) = \frac{1}{\mathrm{p}(y)} \mathrm{p}(y \mid \Theta)\mathrm{p}(\Theta)$$

- As $p(y \mid \Theta) = L(y \mid \Theta)$ is just the likelihood function, and $1/\mathrm{p}(y)$ is a constant, we can write the posterior distribution as:

$$\boxed{p(\Theta \mid y) \propto L(y \mid \Theta)\, p(\Theta)}$$

$$\boxed{\text{Posterior} \propto \text{Likelihood} \times \text{Prior}}$$

## Bayesian Simulation

- Select a prior distribution
  - Initial view of uncertainty in parameters
- Form posterior distribution by revising the prior in light of your data sample
  - Revised view of uncertainty in parameters taking into account the data
- Simulate parameters from the posterior distribution
  - Gives (joint) distribution of parameters
- For each simulation of the parameters, simulate process uncertainty
  - Generates predictive distribution

## Bayesian Method

- Pros
  - Uses whole of the likelihood function, not just a single point
  - Robust framework for combining data with judgement
- Cons
  - Need to specify a prior distribution for the parameters
  - Cannot always determine a standard posterior distribution for the parameters

## Defining the Prior Distribution

- Do you have a prior opinion?
- Will others see it as valid?
- Could use appropriate market data to set prior
- Can use prior to avoid “unreasonable” / undesired parameter values
- What if you have no prior opinion?
  - Could use a “non-informative” prior

## Non-Informative Prior (1)

- If we do not have any prior information on the parameter we choose a prior known as a “non-informative prior”
- The most common non-informative prior is simply a constant or uniform prior:

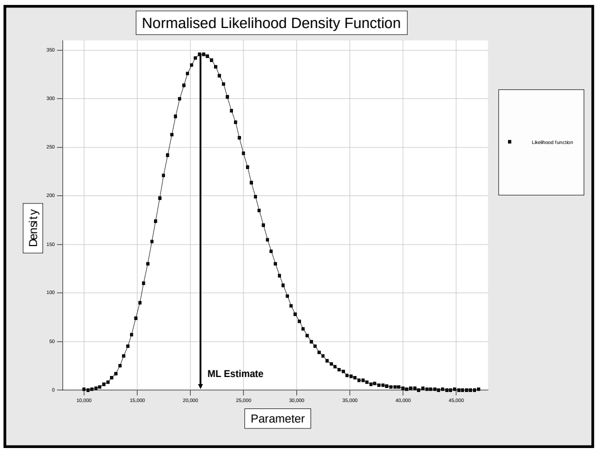
$$\mathrm{p}(\Theta) = \frac{1}{b-a} = \text{constant} \qquad a \le \Theta \le b$$

  - We then have:

$$p(\Theta \mid y) \propto L(y \mid \Theta) \qquad \boxed{p(\Theta \mid y) = \frac{L(y \mid \Theta)}{\int_{\Theta} L(y \mid \Theta)}}$$

## Non-Informative Prior (2)

- Hence, in the case of a uniform prior the posterior distribution is given simply by considering the likelihood function as a function of the parameters
- Also known as the “Normalised Likelihood” approach




## Simulating from Posterior Distributions

- Recognised Distribution
  - For certain models and priors the posterior distribution can be recognised as a standard distribution
  - Where it works the maths can be complex!
- In many cases there will be no standard distribution
  - This makes it difficult to simulate from
  - Need to use sampling algorithms
    - E.g. Rejection sampling

## Examples with Uniform Priors

- Poisson

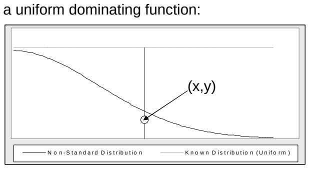
$$\lambda \sim Gamma\left(\sum_{i=1}^{n} x_i + 1,\ n^{-1}\right)$$

- Generalised Pareto

$$f(scale, shape, threshold \mid \underline{x}) \propto$$
$$scale^{-n} \prod_{i=1}^{n} \left(1 + \frac{shape}{scale}(x_i - threshold)\right)^{-(shape^{-1}+1)}$$

## Rejection Sampling - Overview

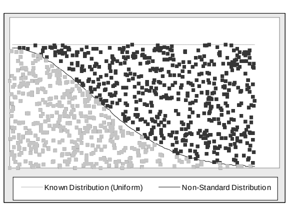
- Uses a uniform dominating function:




- Simulation Approach
  - Generate a value on the green line from a Uniform distribution, call it x
  - Independently generate a value on the red line from a Uniform distribution, call it y
  - Reject the point (x,y) if y > f(x), else Accept
  - Repeat

## Rejection Sampling - Overview

- The blue points show the ones which are accepted. The red points are rejected.

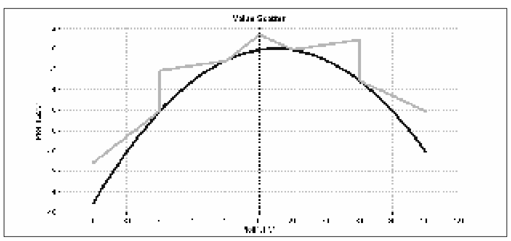


## Rejection Sampling - Overview

- Many rejections made – inefficient
- Use piece-wise exponential dominating function instead of Uniform
- Or, adapt the simulation criteria such that you reduce the number of rejections => Adaptive Rejection Sampling.

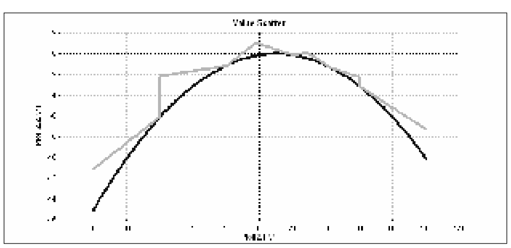
## Adaptive Rejection Sampling (ARS)

- Algorithm operates in log-space
- Density function must be log-concave



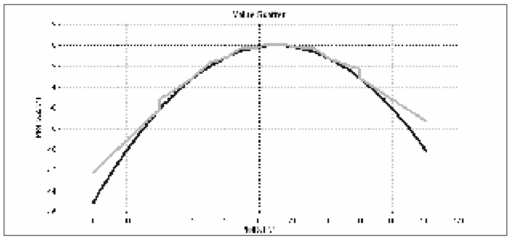

## Adaptive Rejection Sampling (ARS)

- Dominating function adapts as points are rejected



## Adaptive Rejection Sampling (ARS)

- Dominating function adapts as points are rejected



## Adaptive Rejection Metropolis Sampling (ARMS)

- ARS algorithm with an additional “Metropolis” step
- Copes with densities that are not log-concave
- The samples are no longer independent (they may have serial correlation)
- Needs more processing resources than ARS

## Multiple Parameters: Gibbs Algorithm

- Where we have multiple parameters in our posterior distribution, we need to use a method such as Gibbs sampling in conjunction with ARS/ARMS.
- Gibbs sampling is the most popular Markov Chain Monte Carlo (MCMC) method
- Consider a vector random variable $U = (U_1,...,U_K)$ with joint distribution $f(U_1,...,U_k)$
- The Gibbs algorithm allows us to simulate direct draws from this joint distribution given an arbitrary vector of starting values $u = (u_1,...,u_K)$

## Gibbs Algorithm

- An iteration of the Gibbs sampler:

$$
\begin{aligned}
\tilde{U}_1 &\sim f(U_1 \mid U_2, \ldots, U_k) \\
\tilde{U}_2 &\sim f(U_2 \mid \tilde{U}_1, U_3, \ldots, U_k) \\
&\vdots \\
\tilde{U}_j &\sim f(U_j \mid \tilde{U}_1, \ldots, \tilde{U}_{j-1}, U_{j+1}, \ldots, U_k) \\
&\vdots \\
\tilde{U}_k &\sim f(U_k \mid \tilde{U}_1, \tilde{U}_2, \ldots, \tilde{U}_{k-1})
\end{aligned}
$$

- This completes a single iteration of the algorithm and defines a transition from $U$ to $\tilde{U}$ .

## Gibbs Algorithm

- The resulting sequence of dependent draws (after sufficiently large burn-in) will almost certainly satisfy:

$$U^{(t)} \xrightarrow{d} U \sim f(U), \text{ as } t \to \infty$$

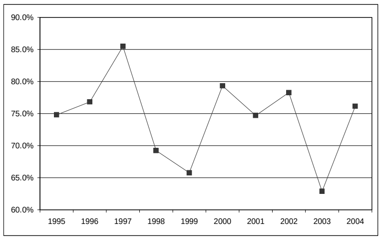
## Effect of Parameter Uncertainty - Example

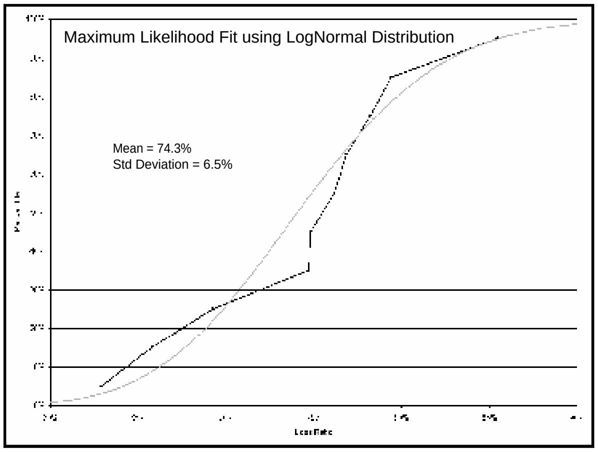
- Single line of business. Premium income £80m.
- Consider underwriting risk to ultimate for a single future u/w year
- 10 years of historical data
  - Attritional loss ratios
  - Individual large claims above £2m
  - Assume adjusted to current u/w year terms
- Assess capital required using fixed ML parameter estimates *versus* Bayesian distributions of parameters (with uniform priors)

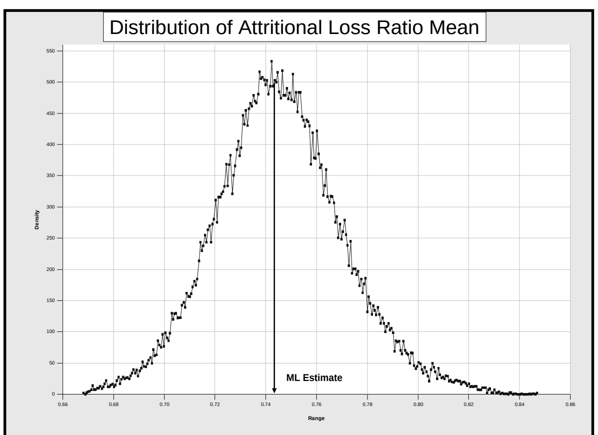
## Attritional Loss Ratio

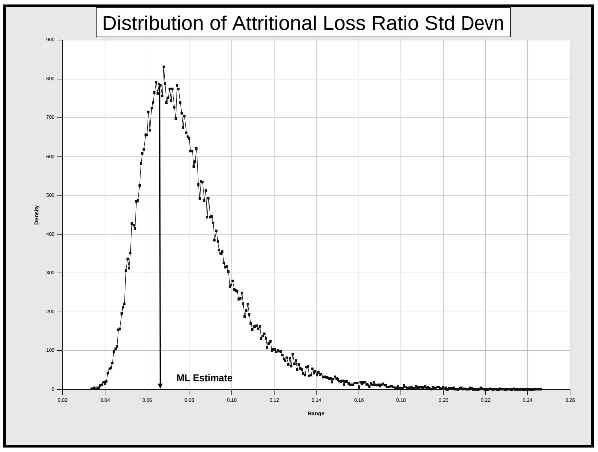
| U/w Year | ULR |
|---|---|
| 1995 | 74.8% |
| 1996 | 76.9% |
| 1997 | 85.5% |
| 1998 | 69.2% |
| 1999 | 65.8% |
| 2000 | 79.4% |
| 2001 | 74.7% |
| 2002 | 78.2% |
| 2003 | 62.9% |
| 2004 | 76.2% |

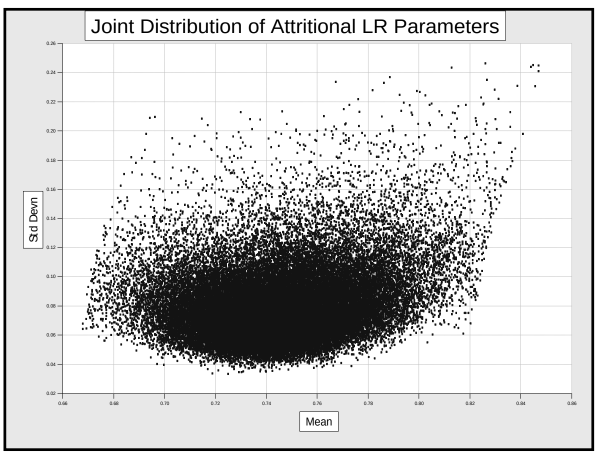
Premium = £80m

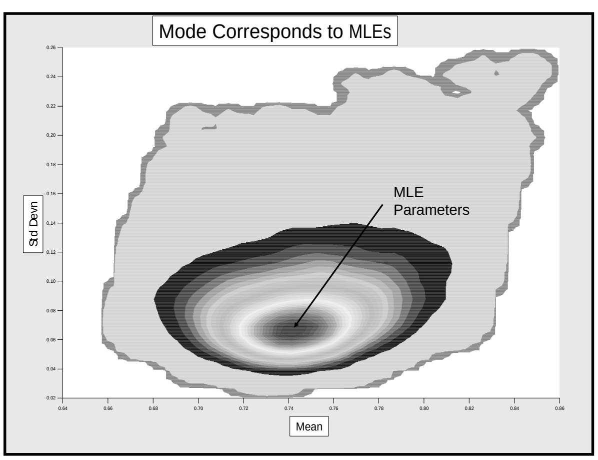


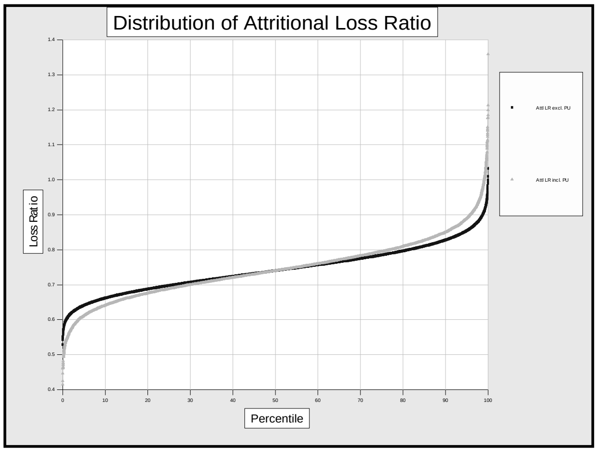



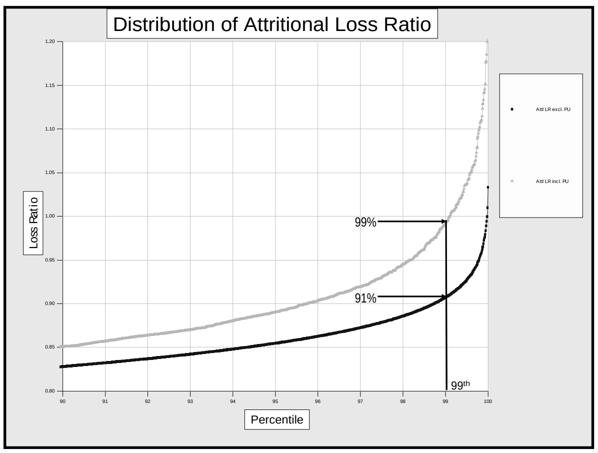




Distribution of Attritional Loss Ratio Std Devn
Density
ML Estimate
Range


Joint Distribution of Attritional LR Parameters
Std Devn
Mean


Mode Corresponds to MLEs
MLE
Parameters
Std Devn
Mean

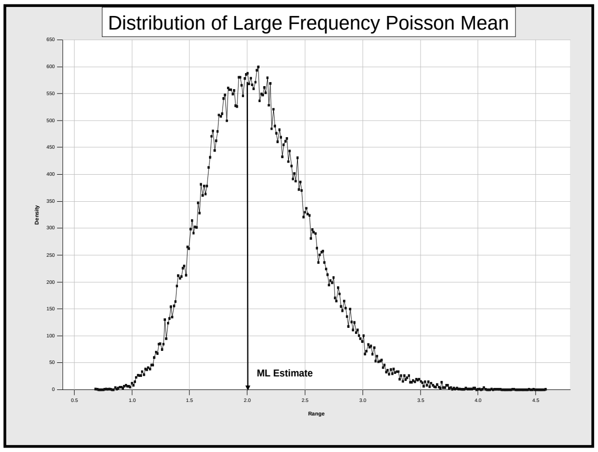


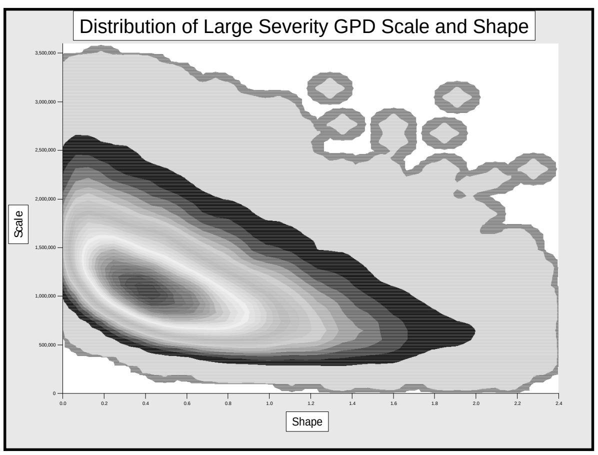



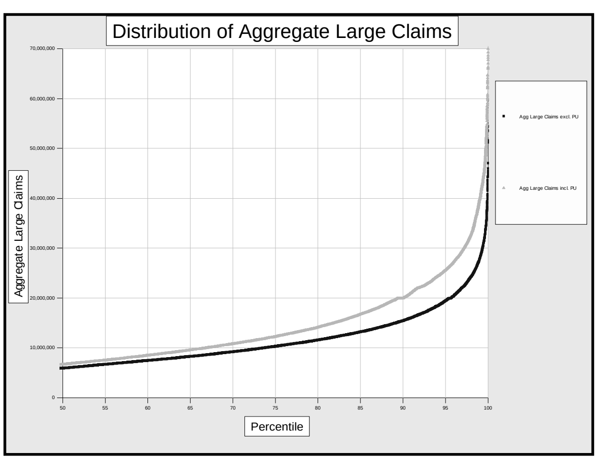

## Claims xs £2m

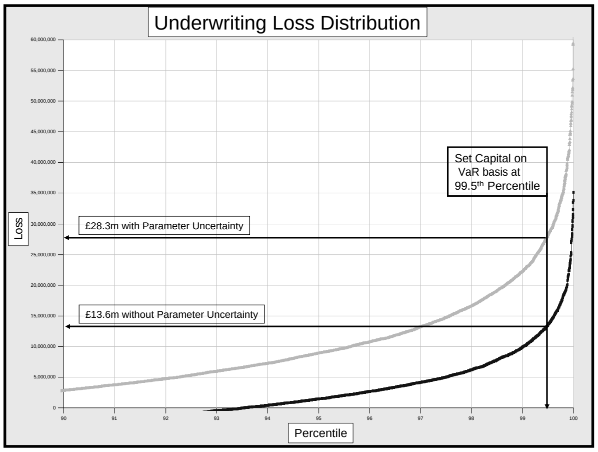
| U/w Year | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1995 | 2,910 | 2,498 | | |
| 1996 | 2,114 | 2,403 | 5,852 | |
| 1997 | 4,315 | 4,529 | 2,390 | 7,522 |
| 1998 | 2,074 | | | |
| 1999 | 11,173 | 3,243 | 2,001 | |
| 2000 | | | | |
| 2001 | 3,454 | 3,719 | 2,194 | |
| 2002 | 3,107 | | | |
| 2003 | 2,283 | | | |
| 2004 | 2,737 | 2,610 | | |

**Claim amounts in £000s**
**Assume no IBNR**


Distribution of Large Frequency Poisson Mean
Density
ML Estimate
Range


Distribution of Large Severity GPD Scale and Shape
Scale
Shape


Distribution of Aggregate Large Claims
Aggregate Large Claims
Percentile




## Summary

- Parameter uncertainty often ignored
- Statistical methods exist to quantify this uncertainty
- Bayesian techniques offer a robust solution
- Allow data to be complemented with judgement / business views
- Sometimes need to use generic sampling algorithms (ARS, ARMS, Gibbs)
- Impact on model results can be significant
  - Has largest effect with small data sets…
  - …especially in the tails of distributions

The Actuarial Profession
making financial sense of the future

The Actuarial Profession
making financial sense of the future

# Parameter Uncertainty and Capital Modelling

20th October 2005

Richard Millns and Richard Weston