The Risks We Miss

- Recent history (e.g. AIG and Fortis)
  - "The Formula That Killed Wall Street"
- What is there to stop us being in a similar position in 5 years?
- Essay by Andrew Smith "Difficult Risks and Capital Models"
  - Described experiences of a very unfortunate company
- Categorised the issues and considered what techniques can be applied to manage the risk
- Today we will talk about 2 techniques but there is discussion on other techniques on Conference website. Comments welcomed
- There will be an Actuarial Profession paper next April

The list of techniques

Prediction Intervals rather than fitted quantiles	Audit sampling techniques to test model point homogeneity
Monte Carlo sampling error	Bayesian techniques around expert judgement
Sensitivity testing for alternative assumptions and data sources	Spanning error in proxy models, including curve fitting
Gross up techniques following dimension reduction	

- Today going to talk about:
  - Judgement
  - Spanning Error

2

Agenda

- Background (Sandy)
- Modelling Judgement (Parit)
  - Modelling
  - Types of uncertainty
  - Known vs. unknown models / parameters
  - Dealing with Parameter and Model uncertainty
- Spanning Error (Laura)
  - Introduction and toy problem
  - Conclusions and future work

3

2. Modelling uncertainty and Judgement

- Modelling
- Types of uncertainty
- Known parameter and models vs…
- Unknown models and / or parameters
- Dealing with Parameter and Model uncertainty

4

Modelling

- Modelling is an inescapable part of Actuarial Life
- A model is necessarily a simplified representation of the real world!
- In this (Actuarial) context, we think of models as tools / processes that:
  - Use information from the past (history)
  - Together with knowledge about a particular problem (judgement)
  - To model future (uncertain) outcomes
  - (and hence help make decisions about the future)

5

Types of Uncertainties

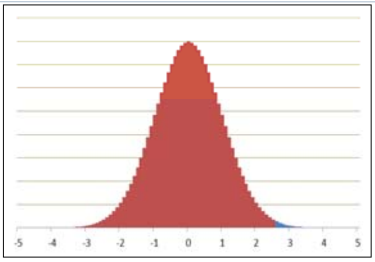
- A model is necessarily a simplified representation of the real world!
- The process of stripping down to the bare useful components and “calibrating” the resultant model has necessarily got a large amount of judgement / decisions associated with it
- These judgements / decisions manifest themselves in various different ways. Broadly speaking, some of the ways we encounter decisions over the process of actuarial modelling are:
  - Choice of overall framework for the model
  - Choosing individual parts of the model (e.g. distribution)
  - Choice of calibration methodology
  - Choice of parameters, overriding certain parameters if necessary

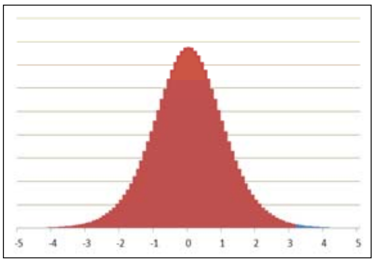
6

Model certainty…

If we know true distribution, can just read off the relevant quantile e.g.

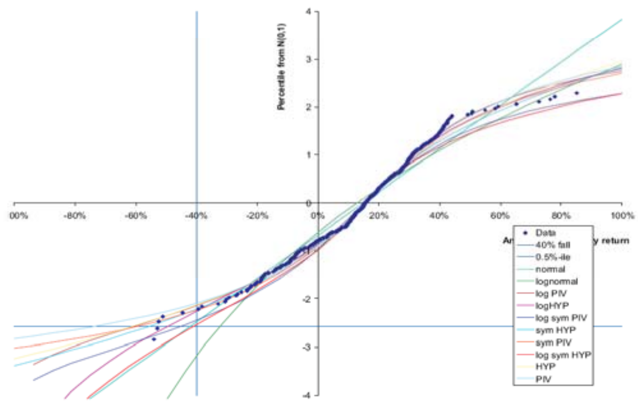
- Normal distribution with known μ, σ
  - 99.5th %tile is $\mu + \Phi^{-1}(0.995)\sigma = \mu + 2.58\sigma$
- t distribution, 10 degrees of freedom
  - 99.5th percentile is 3.17
- etc…





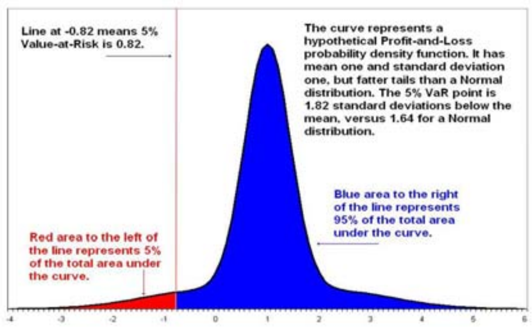
7

vs. Model Uncertainty!



8

What does '1 in 200' even mean?



- If underlying model for observations X is known, then
  - 99.5% quantile q is calculable as $F^{-1}(0.995)$, so **P** (X ≤ q)=99.5%
- However, if the model is unknown then q is unknown.
- Consider two estimators VaR(1) and VaR(2) defined by:
  - E(VaR(1)) = q i.e. VaR(1) is an unbiased estimate of q, so **P**(X ≤ VaR(2))=0.995
- Often this will mean VaR(2) > VaR(1)!

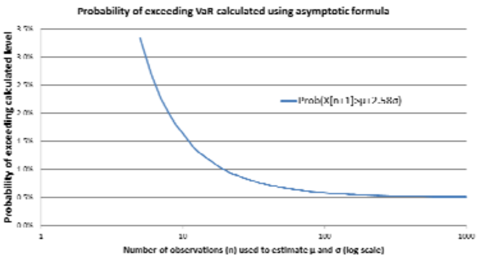
9

Example: normal distribution
Unknown parameters, n observations

- Standard unbiased estimates of mean and standard deviation
- i.e. $\hat{\mu} = n^{-1}\sum_1^n X_i$, $s^2 = (n-1)^{-1}\sum_1^n (X_i - \hat{\mu})^2$, $\hat{\sigma} = s.\sqrt{\frac{n-1}{2}}.\Gamma(\frac{n-1}{2})/\Gamma(\frac{n}{2})$
- Then $\hat{\mu} + 2.58\hat{\sigma}$ is an unbiased estimate of $\mu + 2.58\sigma$, the quantile we are after

BUT

- It is not true that the next observation has a 0.5% chance of exceeding this estimate (or that 0.5% of the next *m* observations will exceed this level)
- $\mathbb{P}(X_{n+1} > \hat{\mu} + 2.58\hat{\sigma}) > 0.05\%$



What can we do?

1. Understand the various sources of judgement / decisions
2. Ensure we allow for systematic biases where we can
3. Raise awareness that VaR is subject to model error and parameter error
4. Bayesian techniques to allow for prior views?
5. Robust techniques

Confidence intervals v prediction intervals

- Confidence interval
- $\mathbb{P}(\theta \in [\hat{a}, \hat{b}]) = p_1$
- $\mathbb{P}(\theta \leq \hat{b}) = p_2$
- Statement about parameters

- Can be Monte Carlo tested
- Parameter unobservable so untestable in practice

- Prediction interval
- $\mathbb{P}(X \in [\hat{a}, \hat{b}]) = p_1$
- $\mathbb{P}(X \leq \hat{b}) = p_2$
- Statement about observations

- Readily tested in practice
- This is exactly the VaR backtest: count 'exceptions' in a time series

Approaches to dealing with uncertainty

Situation	Responses	
Known parameters	Exact formula available	
Unknown parameters	Create estimators of known exact formula Typically hit desired probability level exactly	Calculate prediction interval for next observation Typically hit desired probability level exactly
Unknown model	**Bayesian approach** Bayesian prior over family of models Calculate desired quantile of posterior distribution Probability of observation exceeding VaR may be >0.5% for some models	**Robust approach** Family of models form an ambiguity set Determine VaR so that quantile exceeded for all models Probability of exceeding VaR may be <0.5% for some models

Practical issues

- Although this appears to work well for certain univariate cases, it may be very difficult to scale - there are 100’s to 1000’s of choices being made in a typical life-office model
- There is still judgement required on the class of models
- We would need other methods for more generalised model choices e.g. time variation in returns, number of factors to model, etc…

Summary and Conclusions

- There is a large amount of judgement / decision making involved in specifying a model
- This means that there are many aspects where the model specification could be different from reality
- In particular we need to acknowledge model error and parameter error
- We have looked at a few different tools to address this
- However, the tools do not generalise to the entire model, so one still needs to be careful when / where to use them

2. Spanning Error

- Introduction
- The toy problem
- Error analysis
- Key findings
- Direction of future research

16

Introduction: Spanning error in proxy models

- Insurers' assets and liabilities are complicated functions of millions of inputs whose future values are uncertain
- Insurers use "heavy models" to compute assets and liabilities as functions of the long list of inputs. In theory, we need a full stochastic projection to calculate the probability distribution of assets and liabilities
- Despite further (foreseeable) advances in computer calculation, a full stochastic projection of stochastic liabilities remains beyond the reach of most insurers. Instead, there is a widespread use of proxy models
- Proxy models have two main standard features:
  - A reduction in the number (or "dimension") of inputs, from millions to tens or hundreds
  - A set of selected "basis functions" from which a linear combination is selected to describe the assets or liabilities

17

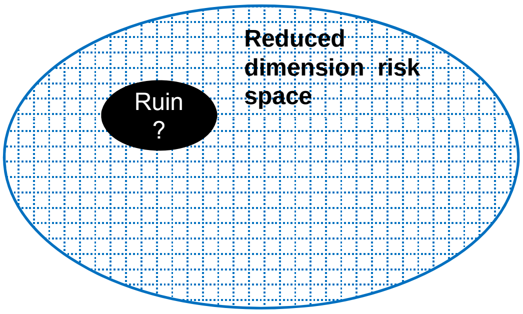
Introduction: Spanning error in proxy models

- A reliance on proxy techniques assumes that accurate functional form approximations can be made to the heavy modelled “true” values. The following are typical examples where spanning failure can occur :
  - Missing material risks in the dimension reduction
  - The true function has discontinuities e.g. due to modelled stakeholder actions but all the basis functions are continuous
  - Regions of parameter values where assets collapse or liabilities explode, but none of the basis functions exhibit this behaviour
- These are all examples of spanning error, which is any mis-statement in the required capital that arises from spanning failure

18

Options for error analysis

- Out of sample checking runs are computationally intensive. To save computational time the runs may be:
  - focused on a particular region e.g. the suspected region of risk drivers having the largest impact on the SCR
  - a sparse covering of the entire risk space to more generally determine whether the reduced dimension model gives a correct “ruin region”



- It is possible to gain some insight into the task of fitting proxy models using models with closed form analytic solutions
- By flexing the reduced proxy model we can gain some insights into different product types

19

The toy model - Analytic formula models

The following assets and liabilities are modelled with a simplified set of term independent risk drivers. In addition we look at the error in the aggregate balance sheet

- **Liability models**
- Term assurance
- Annuity
- Guaranteed equity option

- **Asset models**
- Coupon paying Government and corporate bonds
- Equity
- Cash

We have given ourselves 3 liabilities, 3 risky assets and 9 risk drivers. All risk drivers have a shifted scaled log-normal distribution and the Toy Model technical provisions can be found analytically

The Toy model

Process:

1. Decide risk drivers
2. Decide modelling parameters of risk drivers (choice of distribution, data to fit to, fitting process etc)
3. Proxy functions
   - Form of proxy functions
   - Fitting process (Number and choice of fitting points, OLS?, fitting criteria)
4. Running of model (number of sims, criteria for acceptance)

The toy model allows us to experiment with 3 & 4. We are assuming perfect knowledge of our risk drivers)

Development platforms: Excel, Mathematica (testing) and R

Toy model risk drivers

Risk driver	Term Assurance	Annuity	Guaranteed Equity	Government bond	Corp Bond	Equity
Risk free discount rate	X	X	X	X	X	
Equity price			X			X
Equity volatility			X			
Corp bond portfolio spread					X	
Liquidity premium *		X				
Mortality(Term)	X					
Mortality (annuity)		X				
Lapses (term)	X					
Lapses (Guaranteed equity)			X			

The most complex liability depends on 4 risk drivers and there is variation in the number of risk drivers

Fuller details of Toy Model in a document provided on Conference website

Parameters and fitting

If we fit to the median and upper and lower quantile stressed values for each risk then we have three fitting points per risk. Assets and liabilities dependent on one risk driver are therefore limited to a quadratic curve fit.

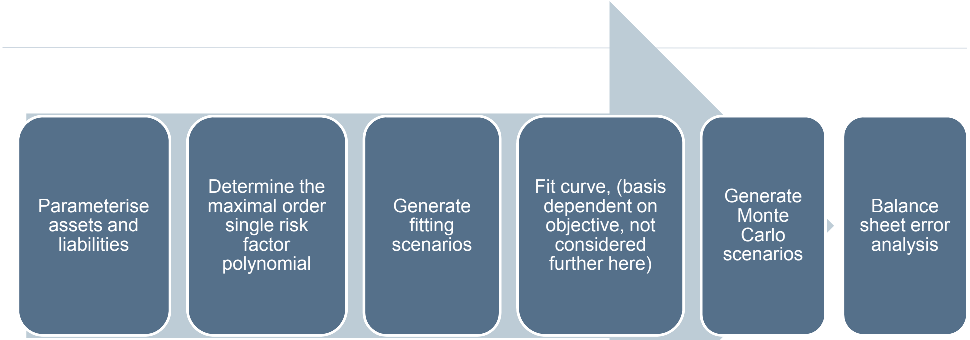
Risk Drivers	Fitting Points	Parameters in proxy function		
		Linear	Quad	Cross
1	3	2	3	-
2	9	3	5	6
3	27	4	7	10
4	81	5	9	15
5	243	6	11	21

Features of the model:

- Annuity with a discontinuous first derivative
- Guaranteed equity bond has optionality. (We could use LSMC but currently apply Black-Scholes formula)

We can investigate dimensionality (to a certain extent)

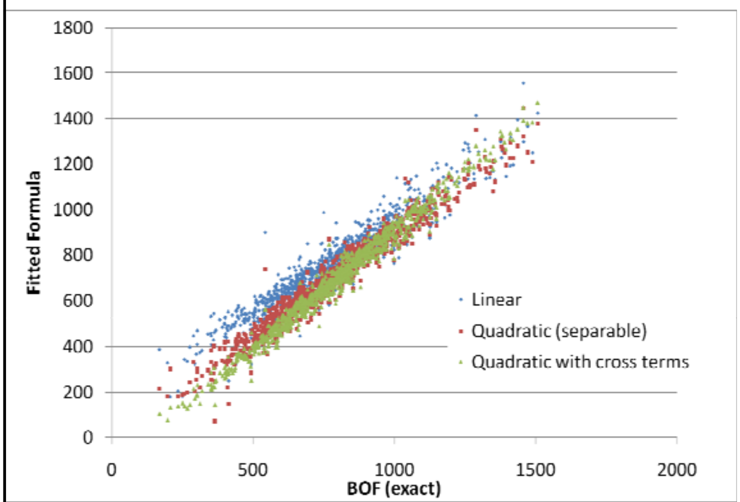
Toy modelling



Simulation methodology

- Specify the risk driver (exponential distribution used)
- Simulate outer scenarios
- Value portfolio

Experimental results



Correlation FF vs BOF (all sims)

Linear	95.0%
Quadratic	97.7%
Quadratic plus cross terms	98.9%

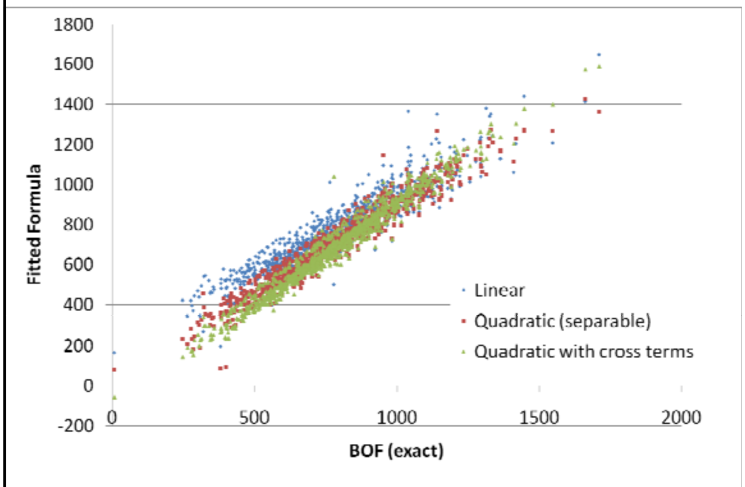
Discontinuities and management actions

- Fitting without management actions

Equity example:

- If the purchase of an equity put option is triggered to prevent a fall below 25% but this isn't included in the modelling then the quality of the fit is reduced. (Particularly in the tail)
- The same could be true in real life balance sheets for any un-modelled management actions e.g. executive bonuses
- This highlights the importance of a rule which is modelled in the stress tests but not in the original tech provision calculation
  - it might be argued that the discontinuity is an unintended consequence of inconsistencies in different parts of the model
  - if you capture the action in the TP calculation then the liability is continuous and the fit is better

Experimental results including modelled management action



Correlation FF vs BOF (all sims)

Linear	93.9%
Quadratic	96.9%
Quadratic plus cross terms	98.4%

Key preliminary findings

- Importance of cross terms
  - Without the flexibility to use cross terms the accuracy of fits is severely reduced
- In our example bringing the fitting points in from the tails increases the R squared but actually reduces the accuracy of the SCR (c.f. the “true” – no proxy- value)”
- Can we seek a point where the accuracy is best? corresponding in some sense to a 99.5% confidence
- There is a trade-off between the measures we can easily use to fit (R-squared, equivalent to OLS) and the measures we would theoretically like to use (minimax)
- Discontinuties and areas of blow up will need to be modelled as an additional layer and cannot be well modelled using polynomial functions

Areas for further research

- The relationship between R squared and the maximum error
- Location of fitting points and the error for the actual distribution
- Compare the quality of proxy fits based on R-squared versus minimax
- We would like to derive minimax statements for the toy example such as :
  - A: “The true function and the proxy function differ by at most £x over a given joint risk driver range. This range has a (real world probability) of at least 99.9%”

  therefore

  - B:” The 0.5%-ile of the true function lies between

  0.4%-ile of the proxy function- £x and 0.6%-ile of the proxy function + £x

  which leads to statements such as “the use of proxy models introduces an error of no more than 20% in the SCR”

The realities of curve fitting

- We can peek at the underlying models
- Knowledge or false knowledge of the “ruin region”
- Polynomials are an un-intuitive representation of the balance sheet. They are an unnatural fit asymptotically
- The curve fits required for an accurate fit are far higher order than the examples given. For example a “ real life” annuity product may require a 5^{th} or 6^{th} order polynomial

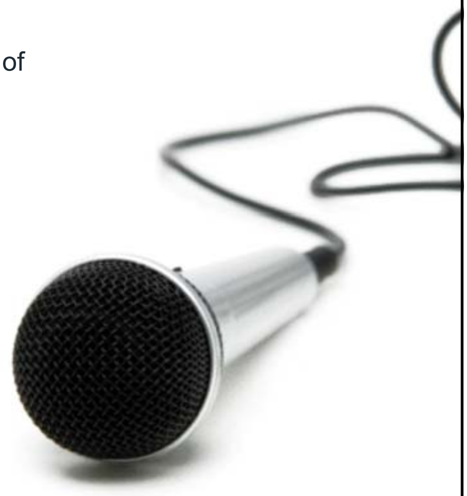
Key points

- We have created a toy model to investigate individual proxy errors and overall balance sheet errors
- Importance of cross terms
- Poor fit of curves asymptotically. Further research on error bounds planned
- Further development - tool and findings to be presented in April paper

Questions or comments?

Expressions of individual views by members of The Actuarial Profession and its staff are encouraged.

The views expressed in this presentation are those of the presenter.



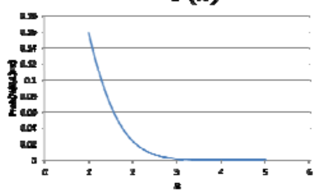
32

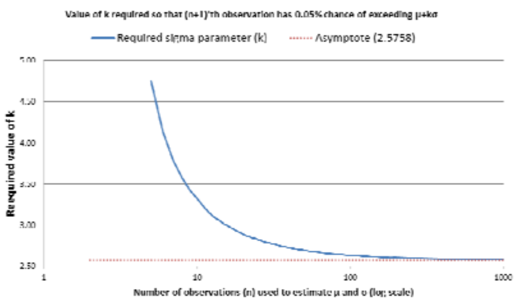
Appendix - Adjusting the estimate

- Even though $\mathbb{E}(\hat{\mu} + 2.58\hat{\sigma}) = \mu + 2.58\sigma$, we have

  $\mathbb{P}(X_{n+1} > \hat{\mu} + 2.58\hat{\sigma}) = \mathbb{E}(\bar{F}(\hat{\mu} + 2.58\hat{\sigma})) >$
  $\bar{F}(\mu + 2.58\sigma) = 0.05\%$
- This is to be expected: the tail cdf function $\bar{F}$ is convex in the region of interest and the estimator $\hat{\mu} + 2.58\hat{\sigma}$ is uncertain, so this behaviour is familiar from Jensen's inequality.

$\bar{F}(x)$





- An adjustment can be determined analytically

  VarEstimator $= \hat{\mu} + k(n)s$
- $k(n) = \sqrt{\frac{n+1}{n}}\, t_{n-1}{}^{-1}(p)$

  Such that
- $\mathbb{P}(X_{n+1} > \hat{\mu} + k(n)s) = 0.05\%$

33