Institute
and Faculty
of Actuaries

# Quantitative Approaches to Model Uncertainty - Update on Recent Research

Andreas Tsanakas
Andrew Smith

29 October 2014

# Presentation Overview


- What we can know and what we cannot know 4
- Biases in Statistical Estimation 8
- Ambiguity and Robustness 23
- Validation test Materiality 29
- Social Model Validity 34
- Conclusions 37


Institute and Faculty of Actuaries

*A blind man in a dark room seeking for a black cat — which is not there.*

Definition of metaphysics attributed to Lord Bowen (among others)

*We consider it a good principle to explain the phenomena by the simplest hypothesis possible.*

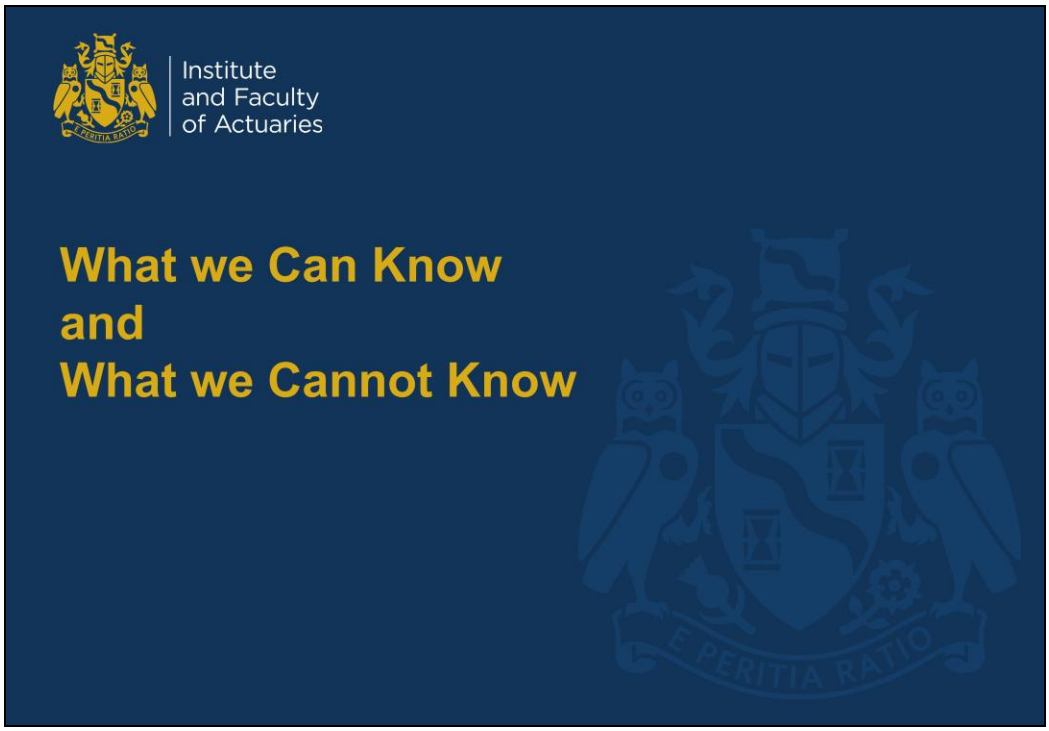
Attributed to Ptolemy, 2nd century CE

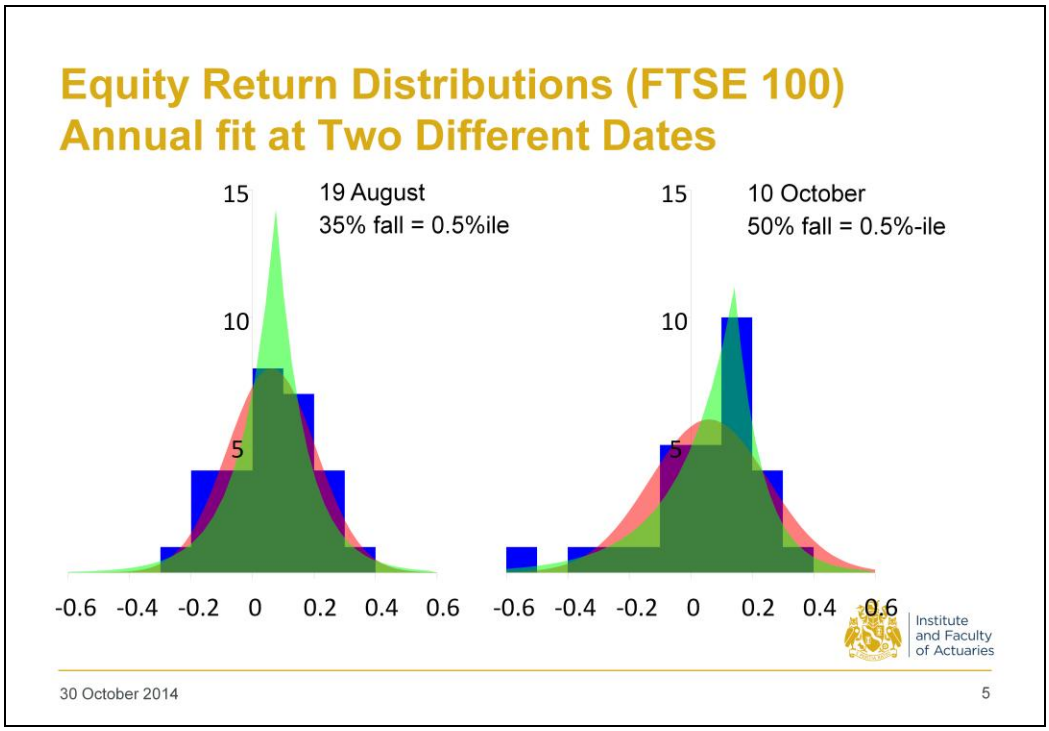
Institute and Faculty of Actuaries

30 October 2014

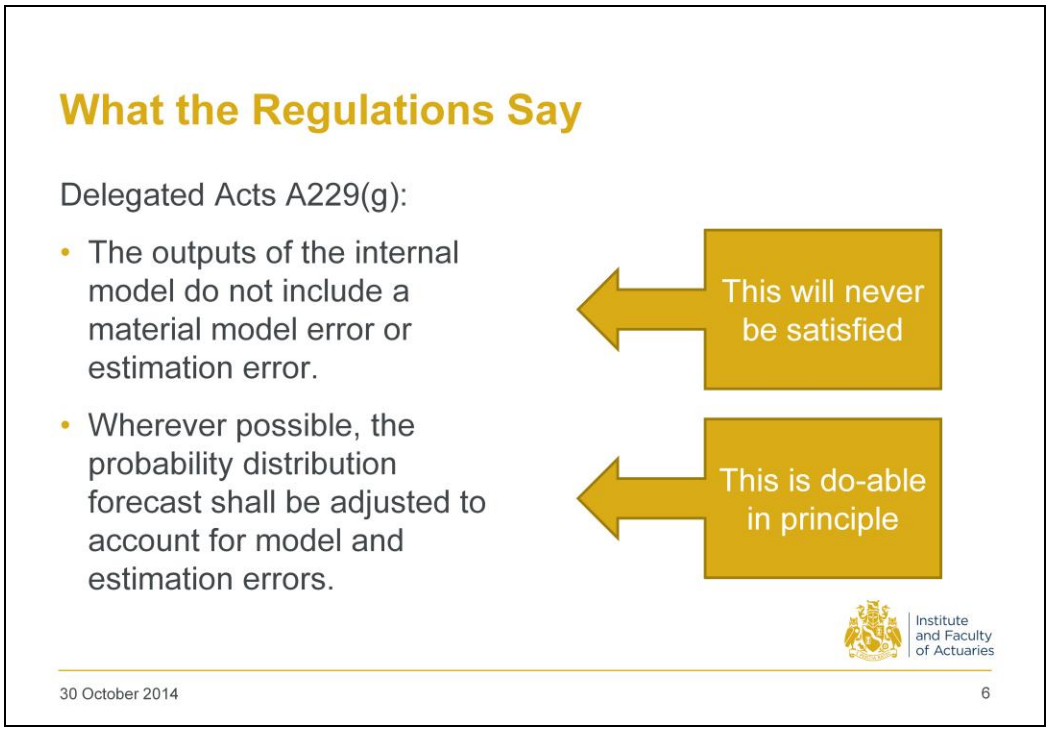
3

When looking for the “correct model” we do indeed resemble the proverbial blind man in a dark room searching for something that is not there. But that is a definition of metaphysics – not something that we are either trained to do or is in our job descriptions. To do statistics, we need to simplify and maybe give up on the idea of a “correct model”, even one unknown to us. Simplification is necessary to turn our struggle into a scientific study rather than an exercise in metaphysics. But simplification creates its own problems.


Institute
and Faculty
of Actuaries
What we Can Know
and
What we Cannot Know
E PERITIA RATIO




This is the distribution of annual equity returns FTSE100, based on 30 observations. The red density is a fitted normal, while the green density is a fitted tilted Laplace. An arbitrary change in the starting date of the calculation changes the estimated 1:200 drop in asset values from 35% to 50%.




Regulation places demands upon us on to control model and estimation error. We cannot banish the possibility of model error (an ill-defined idea anyway in the context of deep uncertainties). But what we can do is adjust the probability distribution – more generally the estimation procedure used – to control such errors.

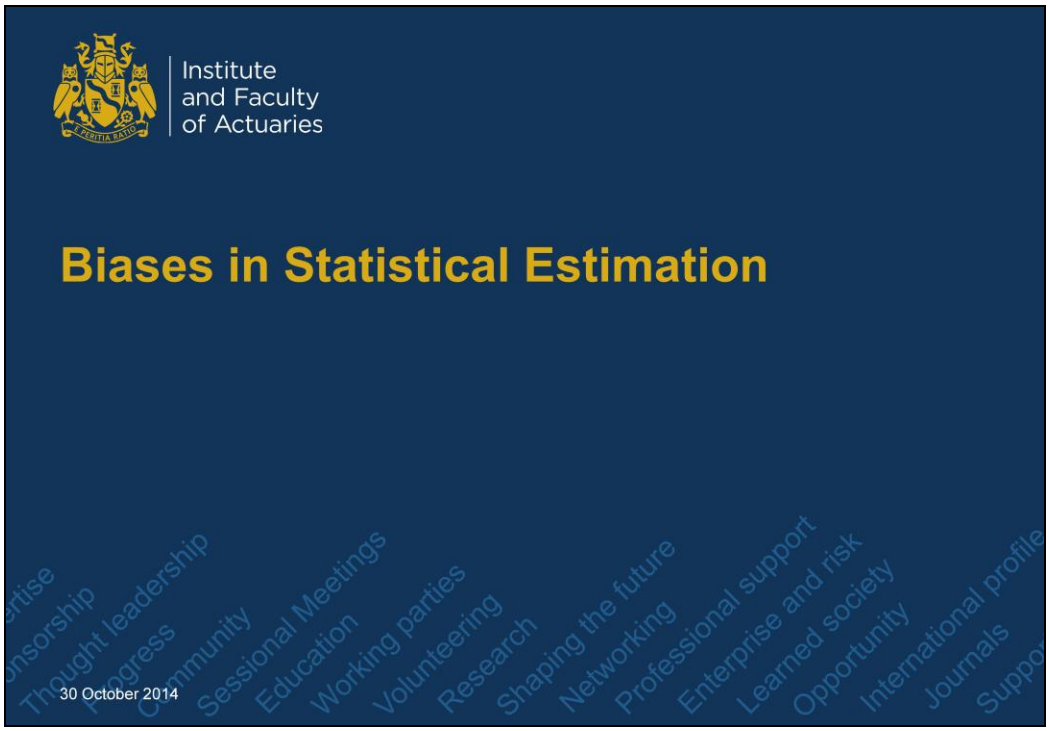
## Model validation?

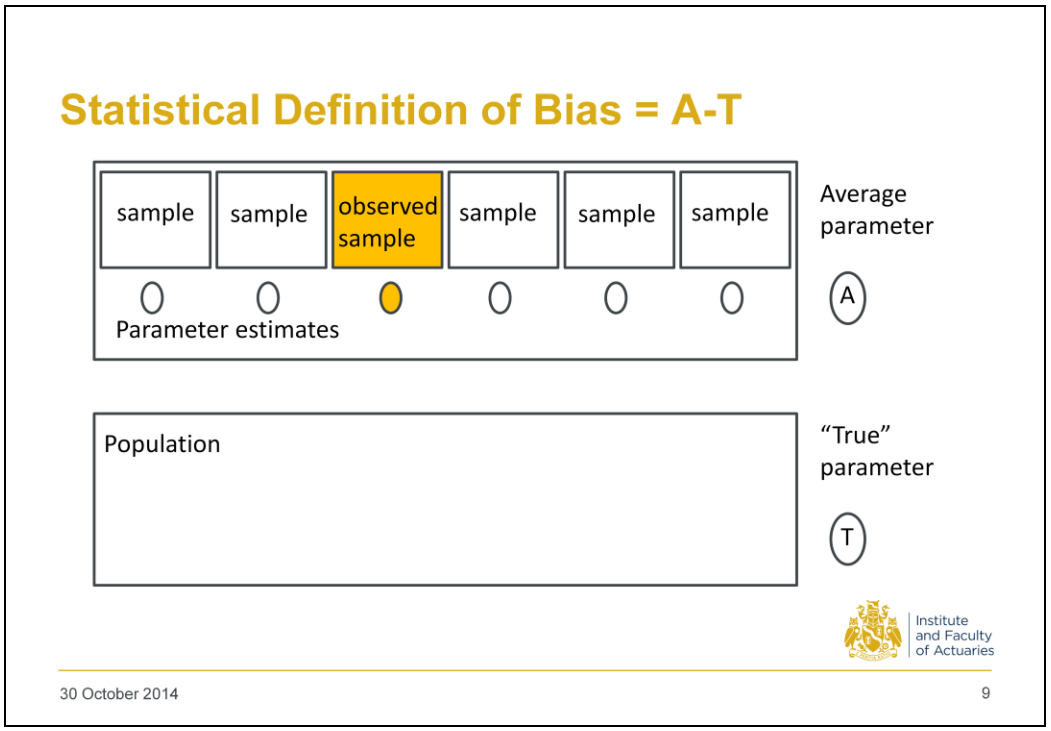
- We are commonly asked to work out extreme losses
  - 1:100, 1:200…
- We **cannot** know that our estimates are “correct”
- We **can** say that the way our estimates were worked was good
  - Uncertainty reflected in estimates
  - How?

Institute and Faculty of Actuaries



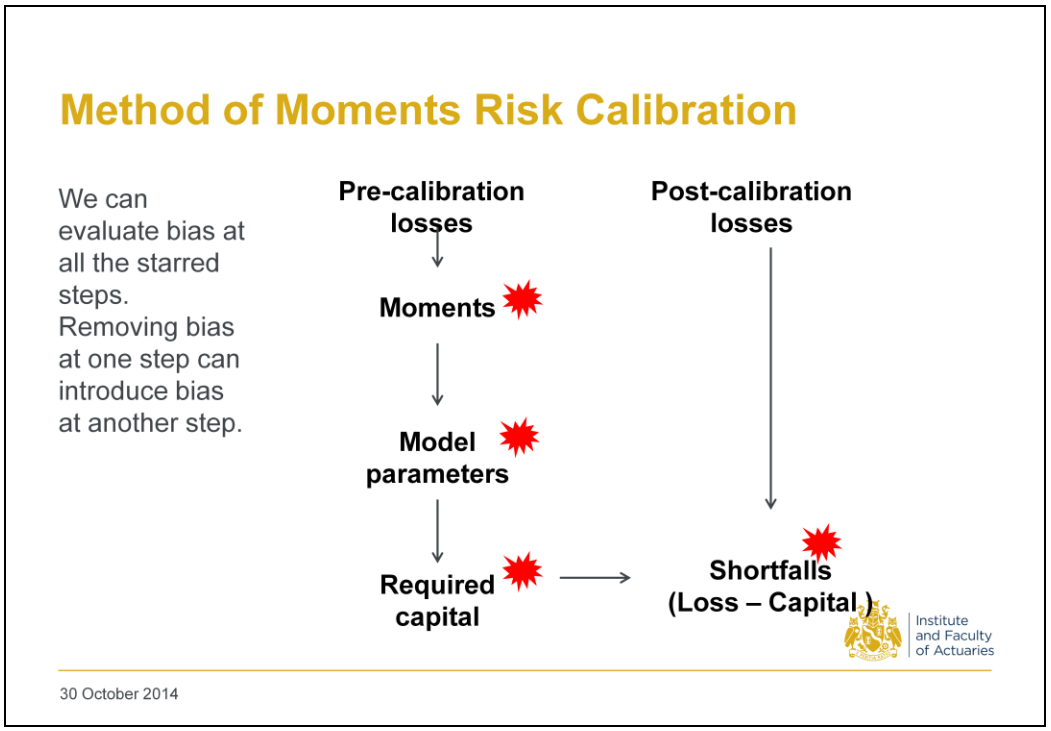
In statistics we can NEVER say that our estimates are accurate. What we can say is whether our estimation METHOD and our adjustments for uncertainty were good. “Good” here means they usually work - not always. To explain what we mean by “they work” and what we mean by “usually” we need to think statistically.


Institute
and Faculty
of Actuaries
Biases in Statistical Estimation
30 October 2014
Thought leadership
Progress
Community
Sessional Meetings
Education
Working parties
Volunteering
Research
Shaping the future
Networking
Professional support
Enterprise and risk
Learned society
Opportunity
International profile
Journals



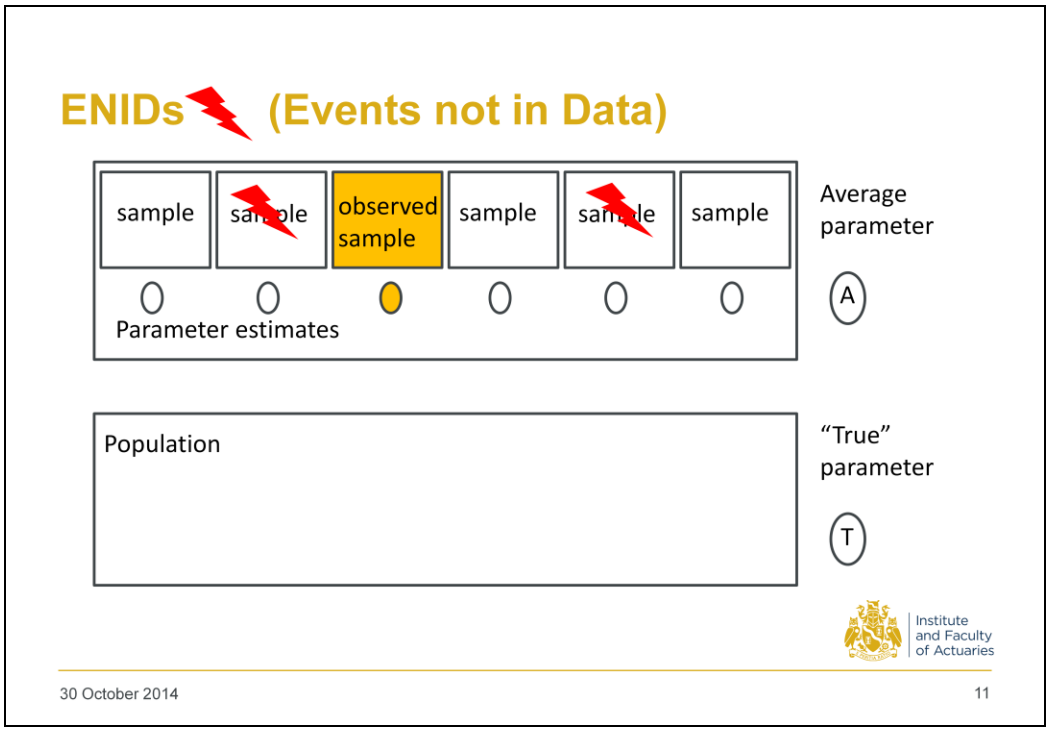

Bias is a fundamental idea in statistics. Let's say that we want to estimate a "true parameter". We observe one particular sample of the data and we use that to estimate the parameter. But now we have to stop and think. Our sample was itself generated by a random process. There could have been different samples generated by the same process. Each sample would have given a different parameter estimate. This idea is in some way counterintuitive. In probability modelling we are used to thinking about alternative futures. In statistical inference we need to think about alternative histories.

An estimator is unbiased when the average value of the parameter estimate across alternative samples equals the true parameter. This appears a desirable property - but is this the sort of unbiasedness we are really interested in?




Once we collect data, we can usually find estimators of moments that are unbiased or of model parameters or of the required capital (VaR). However they can never all be unbiased at the same time – for example, requiring that model parameters are unbiased, as usual statistical practice would dictate, implies that the capital estimate is actually biased.

We argue that the key quantity to consider is the difference between future (post-calibration) losses, a random quantity, and estimated capital, a quantity that can be viewed as random due to its dependence on random sample. This difference reflects the exposure to model uncertainty from a solvency perspective and this is the quantity we propose to study.




Events not in the Data refer to unusual and severe events that may not appear in risk calibration data. Parameter uncertainty can be one reason for the existence of such events. If they are absent from the observed sample, we may under-estimate capital. If they are present, then they are overrepresented and may overstate capital. Thus ENIDs can be to an extent characterised by thinking of the alternative data we could have seen – a standard thought experiment of statistical estimation. So how are we going to consider ENIDs in a VaR estimation context?

## Setup for Analysing Shortfall Bias

- We are interested in calculating the 99%VaR of some random variable Y
- We have collected a random sample **X**, from the same distribution as Y
- Estimation procedure: some function g, such that

$$\text{VaR}^{\text{est}} = g(\mathbf{X})$$

- The function g stands for any way – simple or complex – for transforming the data **X** to a VaR estimate

Institute and Faculty of Actuaries

## A technical quality criterion relevant to VaR

- For the true but unknown $VaR^{true}$, we have

$$\Pr\left(Y \le VaR^{true}\right) = 99\%$$

**Exceptions should only happen once every 100 periods**

- Shortfall unbiasedness means:

$$\Pr\left(Y \le VaR^{est}\right) = 99\%$$

  - Now both sides of the inequality are random variables
  - $VaR^{est}$ is a *predictive limit* (Barndorff-Nielsen & Cox, 1996)
  - Any discrepancy of the LHS from 99% is a shortfall bias.

Institute and Faculty of Actuaries

30 October 2014 13

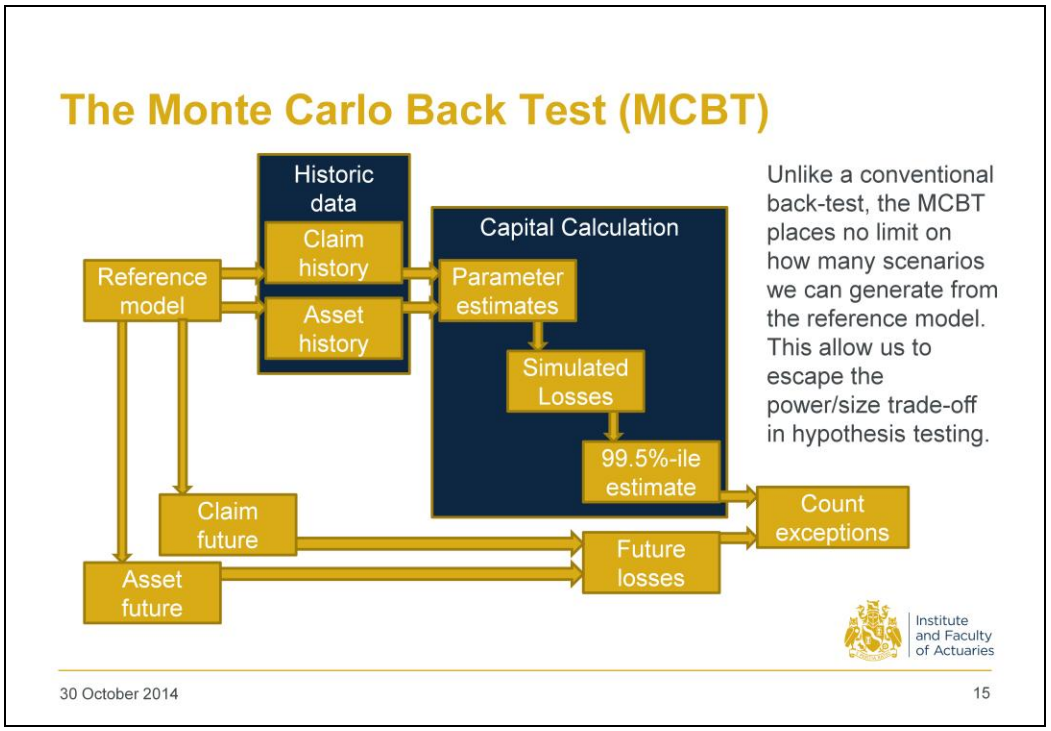
As a quality criterion, we require that the probability that Y, the future loss, is lower than $VaR^{est}$, the estimated VaR, is equal to the nominal confidence level of the VaR measure, in this case 99%. In the second displayed equation, the inequality is thus between two random variables – the randomness of $VaR^{est}$ reflects parameter uncertainty.

Such a requirement for VaR estimates has been formulated independently by Gerrard and Tsanakas (2011) and by Frankland et al (2013).

# Interpretation

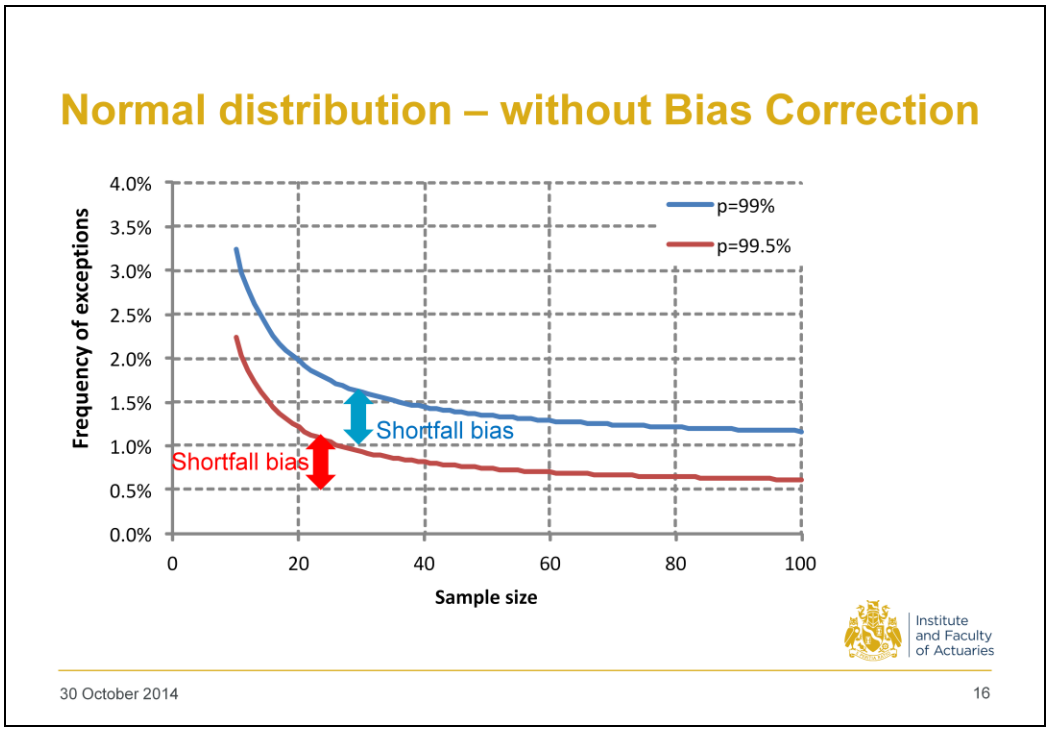
- Back-testing
  - "Fewer than 1 in 100 VaR estimates lower than the loss that was subsequently observed" ie exception frequency ≤ 1%
- "Monte-Carlo back-testing" – a thought experiment
  - Simulate large number of histories from **X** and Y from given distribution
  - Calculate $VaR^{est}$ for each history
  - Estimate $\Pr\left(Y \le VaR^{est}\right)$ by frequency of (non-)exceptions
- A different thought experiment
  - Independent agents in a market, using same estimation procedure
  - Probability $\Pr\left(Y \le VaR^{est}\right)$ represents (non-)default frequency

Institute and Faculty of Actuaries



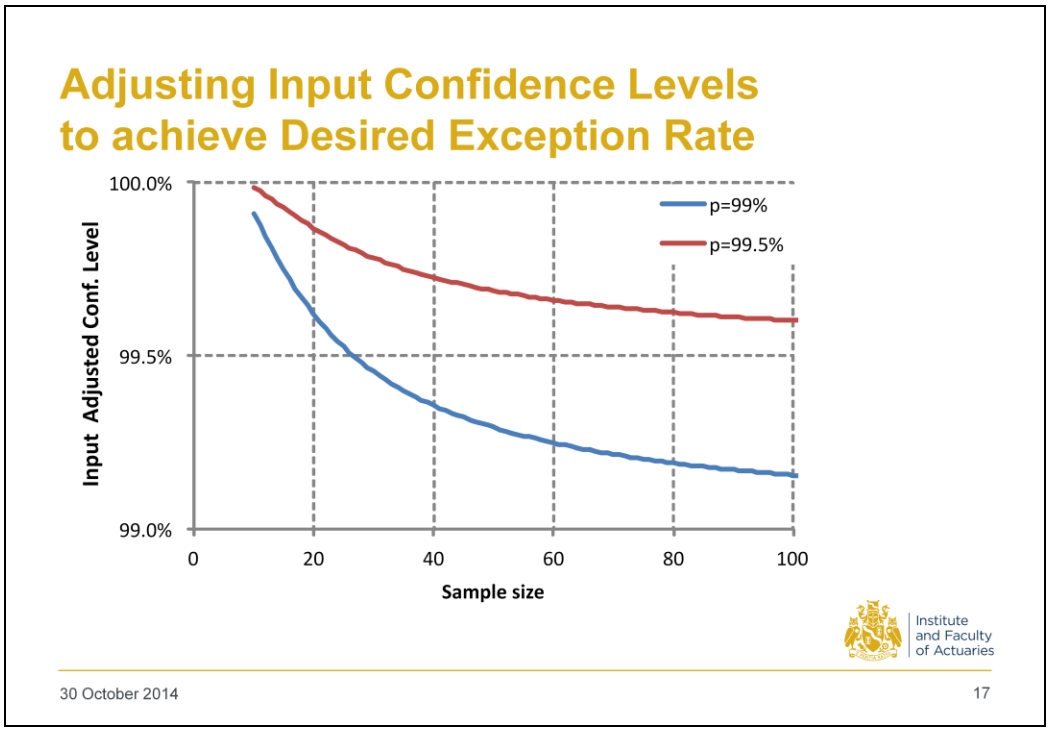

In standard backtesting, historic VaR estimates are compared to realised losses, to calculate the historical frequency of violations/exceptions. However, a very large volume of data is required to establish with confidence that the VaR estimation approach followed actually under- or over-states VaR.

For the MCBT, a reference model is specified from the outset. The model is used to simultaneously simulate past data histories and realisations of sources of uncertainty affecting the future loss positions. The two parallel simulation histories allow us to model the variability of the VaR estimator and of the future loss: therefore it allows us to model the distribution of the Shortfall = Loss - Capital.




In a simple illustration, consider a normally distributed loss with parameters estimated by MLE. The vertical axis corresponds to the expected frequency of exceptions under the reference model – in other words the probability of Loss > Capital. The horizontal axis shows the sample size. For small samples, the frequency of exceptions is much higher than the nominal level of the VaR confidence level. Focusing on p=99%, the difference between the blue curve and the horizontal line at 1% reflects exactly Shortfall Bias we are interested in. The curve converges to the nominal level p with an increasing sample size, but the bias does not quite go away.

This example is from Gerrard and Tsanakas (2011).




Staying within the normal model, we can actually change the capital estimation method in order to yield the correct level of exception frequency, that is, to make shortfall bias equal to zero. One possibility is to increase the confidence level of the VaR estimator, such that VaR is estimated at a higher than nominal level. In that case, small data sizes are penalised, as they require a higher adjusted confidence level and thus, more capital. Hence an allowance for parameter uncertainty produces a very explicit capital add-on.

Note that this adjustment does not guarantee that the estimated capital will actually be “right”. When estimating extreme percentiles from small data set we are of course very likely to get it “wrong”. What the adjustment introduces is unbiasedness: we will get the exception frequency right on average, across alternative data histories and futures.

## Can we satisfy the quality criterion for VaR?

- Yes, but
  - It helps if the distribution is **nice**:
    - (log-)normal, (log-)t, (log-)logistic, exponential / Pareto
  - The distribution **must be known**
    - Errrr...
- How?
  - Increasing confidence level, depending on data-size
  - Using Bayesian predictive distribution
  - Stressed parameter estimates
  - Gerrard & Tsanakas (2011), Frankland et al. (2013)

Institute and Faculty of Actuaries



The previous example was based on a very simple model. In fact there is a large number of distributions for which such a capital estimation adjustment can be carried out; crucially the adjustment then does not depend on the "true" parameters of the reference model. Even for distributions where this is not possible, well performing adjustments can be carried out using bootstrapping or Bayesian procedures.

However, when we don't actually know what the family of distributions is (e.g. Normal, t etc), that is, we move from parameter to model uncertainty, then the above approaches don't work. Model uncertainty is much harder to address.

## What if we include Goodness of Fit tests?

1. Collect data
2. Specify a family of distributions
3. Estimate parameters, to give *fitted distribution*
4. Test fitted distribution for *goodness-of-fit*

   [If it fails, return to 2. and specify a different family]
5. Calculate VaR as *percentile* of fitted distribution
6. Apply any *add-ons and adjustments*

In reality the distribution family used will itself be selected using the observed data. A possible way to do this - and one often used though not necessarily endorsed by statistical textbooks - is to proceed sequentially. First a distribution family is chosen, parameters are estimated and Goodness-of-Fit testing performed. If the distribution is not rejected once sticks with it. If it is rejected, one moves to the next distribution.

## From parameter to model uncertainty

### Frequency of exceptions, sample size=30

| Estimation method | p=99% | p=99% |
|---|---|---|
| | *reference: normal* | *reference: t(4)* |
| Normal MLE (adj. confidence level) | **1.0%** | 1.7% |
| Normal MLE | **1.6%** | **2.2%** |
| Sequential estimation (KS – Lilliefors) | 1.5% | **2.2%** |

Adjustment is important

Sequential estimation doesn't help

Sequential estimation does not add much to just fitting a normal. This is because the sequential test seldom rejects the first proposed model.

This table summarises the results of a simulation experiment. There are two reference models (normal and t(4)) corresponding the two columns, which are used to generate data histories and future losses. Three different methods are used for estimating 99%-VaR, each corresponding to a row in the table:

(1) Estimate normal parameters by MLE and use the percentile of a normal distribution with those parameters, at adjusted confidence level such that the expected frequency of exceptions would be 1%, if the data were normal.

(2) Estimate normal parameters by MLE and use the 99$^{th}$ percentile of the estimated distribution (no adjustment).

(3) Select a model sequentially as in the previous slide, first fitting a normal, then a logistic, then a t(4). As a goodness of fit test, Kolmogorov-Smirnov is used, with Lilliefors adjustment (to reflect the impact of parameter estimation on the test's error probabilities)

Cells (1,1) and (2,1) in the table: This shows the effect on shortfall bias (exception frequency) of not performing the adjustment for parameter uncertainty.

Cells (2,1) and (2,2): This shows the further increase in shortfall bias (doubles from 0.6% to 1.2%) when we have model error.

Cells (2,2) and (3,2): This shows that using a sequential estimation approach as described above does not really help with moderating the effect of model error. Because the dataset has few observations, it is difficult to reject any model (the KS test has low power). As the first distribution tested is a normal, in nearly all scenarios the normal distribution is used to estimate capital: the correct t(4) model is almost never selected.

## From VaR to TVaR / Expected Shortfall

- For an estimation procedure satisfying our criterion:

$$\mathrm{VaR}\left(Y - \mathrm{VaR}^{\mathrm{est}}\right) = 0$$

  - Cost of model uncertainty in £

- Can use this to extend idea to general risk measures, using simulated shortfalls:

$$\mathrm{TVaR}\left(Y - \mathrm{TVaR}^{\mathrm{est}}\right) = 0$$

  - Calculated iteratively by the parametric bootstrap technique
  - Left hand side is a form of shortfall bias
  - Bignozzi & Tsanakas (2014)

Institute and Faculty of Actuaries

30 October 2014 21

In fact the reasoning applied so far can be extended in order to specify quality criteria for estimators of other risk measures. To see how this works, consider the first displayed equation. This states that the VaR of the shortfall (future loss minus random estimated capital) should be zero. This is mathematically equivalent to the requirement the expected frequency on non-exceptions is 99%. Now to move to different risk measure, we need only to change the VaR in the first equation into any other risk measure, e.g. TVaR/Expected Shortfall.

More detail on this idea can be found in Bignozzi and Tsanakas (2014), where methods for adjusting capital estimates to satisfy this criterion are discussed.

## Example Calculations of Unbiased Shortfalls

| Scale factor k applied to sample standard deviation | Value at Risk (99.5%) | Expected Shortfall (99%) |
|---|---|---|
| # years data | Annual | Annual |
| 10 | 3.5718 | 3.7568 |
| 20 | 2.9582 | 3.1077 |
| 30 | 2.8436 | 2.9478 |
| 50 | 2.7239 | 2.8250 |
| 100 | 2.6461 | 2.7347 |
| Infinite | 2.5758 | 2.6652 |

Figures based on Monte Carlo backtest, assuming a normal random walk.

Institute and Faculty of Actuaries

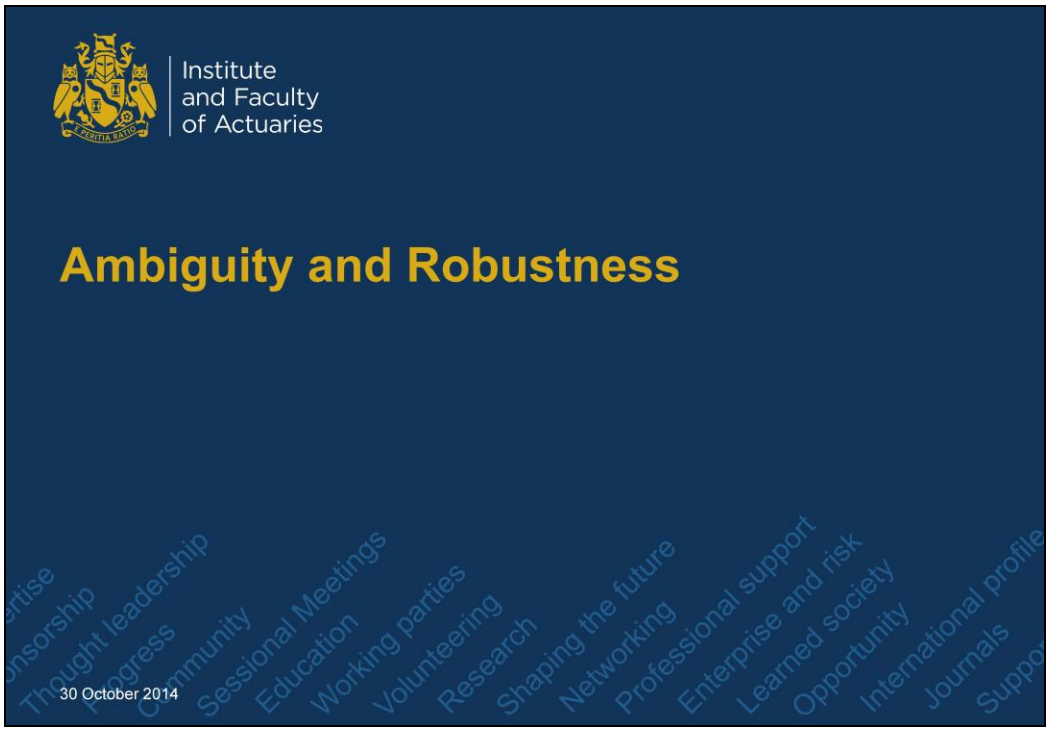


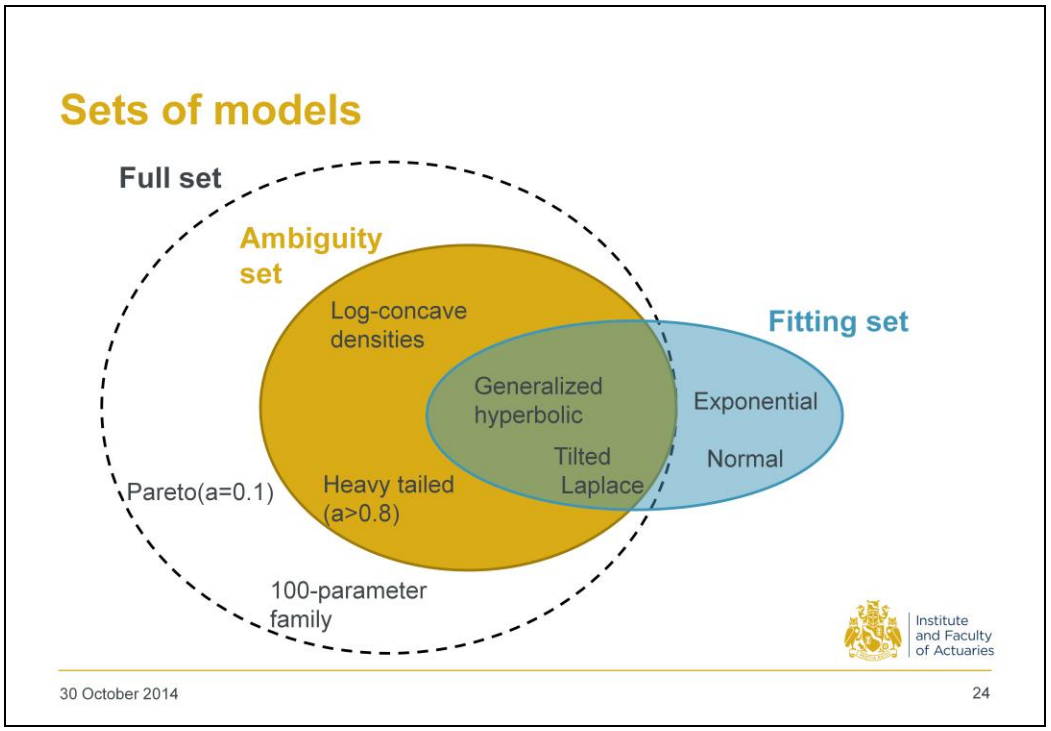
Here, in with a standard normal reference model, we calculate what multiple of sample standard deviation the estimated capital must be, in order to satisfy the shortfall unbiasedness condition for either 99.5% VaR (predictive limit interpretation) or 99%-Expected Shortfall / TVaR (construction from Bignozzi & Tsanakas (2014) as in the previous slide). Comments:

(a) If you haven't got much data, you need to allow for this by using a larger multiple of standard deviation than with a larger data set

(b) Might need to increase capital by 40% if you've only 10 years data

(c) That's true for VaR as well as TVaR


Institute
and Faculty
of Actuaries
Ambiguity and Robustness
30 October 2014
Thought leadership
Progress
Community
Sessional Meetings
Education
Working parties
Volunteering
Research
Shaping the future
Networking
Professional support
Enterprise and risk
Learned society
Opportunity
International profile
Journals




One difficult problem in statistics is to know how many models to consider. If you restrict attention to a narrow set of models, you can have easy, stable fitting but little chance that your set contains the true model. If you cast your net more widely, you have a better chance of containing the true model but estimation is unstable; answers change wildly with each new data point.

The discipline of robust statistics offers us a way out of this dilemma. We separate the notion of a *fitting set* - the class of models we fit, from the *ambiguity set* - a class of models that might generate the data. We then need to investigate what might go wrong when historic data comes from a model in the ambiguity set, but we fit a model from the fitting set. Even though we're fitting a wrong model, the answer might be good enough for some purposes.

## Sets of models

- Current approaches usually report results as if:
  - a) "Fitting set = Single distribution"
    - Ignore other models not inconsistent with the data
  - b) "Fitting set = Ambiguity set"
    - Ignore scale of uncertainty, but produces "reasonable" parameter error
    - Lets practical constraints dictate which models we concern ourselves with
    - How does fitted distribution work on the whole of the ambiguity set?
- We must be explicit about the scope of the estimation exercise
  - Reflect uncertainties that matter to us
  - Admitting to model graveyards

Much of what we're used to, uses small ambiguity sets. We fit a model, go through some form of validation process, and the model, to us, becomes the truth. We lose sight of all the other plausible models.

Even when we follow textbook practice, and estimate parameter errors using Fisher information matrices, we're still kidding ourselves that the data could only have come from a model in the fitting set.

Nassim Taleb – the Black Swan author – sometimes talks about *model graveyards*, that is, all the models we considered but didn't use. We should be sceptical of a model that has been cherry picked to give a particular conclusion from a large graveyard of less favourable models. Good statistics means being honest about the process we have followed. Andreas' example with normal and T4 distributions in slide 20 shows how we need to test the whole inference trail, not just the final fitted model. We might need to do a lot of digging in the model graveyard.

## Ambiguity and robustness

- Confidence level adjustment requires that the distribution family is known
  - Ambiguity and fitting sets reduced to the same parametric family
  - What to do when we have a larger ambiguity set?
- Design estimation procedure such that

$$\Pr\left(Y \le \mathrm{VaR}^{\mathrm{est}}\right) \ge 99\%$$

for **all distributions** in the ambiguity set

- For very small samples, model selection is illusory
  - Because tests have low power to reject a wrong model

How can we say a fitted model works, when we know the model is wrong? The answer is that a model doesn't have to be perfect; it has to be good enough for a particular purpose. If we want to calculate VaR at 99% confidence, we need to show that there is at least a 99% chance the VaR will be enough to cover next year's loss. I say *at least* because, for a large ambiguity set and a small fitting set, you won't be able to hit 99% spot on, for all circumstances. Robustness demands that we can be at least 99% confident; for some more benign models in the ambiguity set, we might be more than 99% confident. This one-sided test means that robustness entails some degree of conservatism. Maybe we'd prefer a method that's 99% confident on average, but that would require some notion of how likely each model is to be correct, to give us weights for the average. Frequentist probability cannot answer this question; we can only talk about probability of outcomes given a model.

## How to achieve robustness?

- Bayesian Approach: Using multiple models

  (Bignozzi & Tsanakas, 2013 – sponsored by the IFoA)

  - Works well when fitting set = ambiguity set
  - Not robust to fitting set ≠ ambiguity set
- Frequentist approach
  - Fit from a small set of models; test a larger ambiguity set
  - Stress parameter estimates and confidence level for VaR calculation
  - What is the *cost of robustness*?

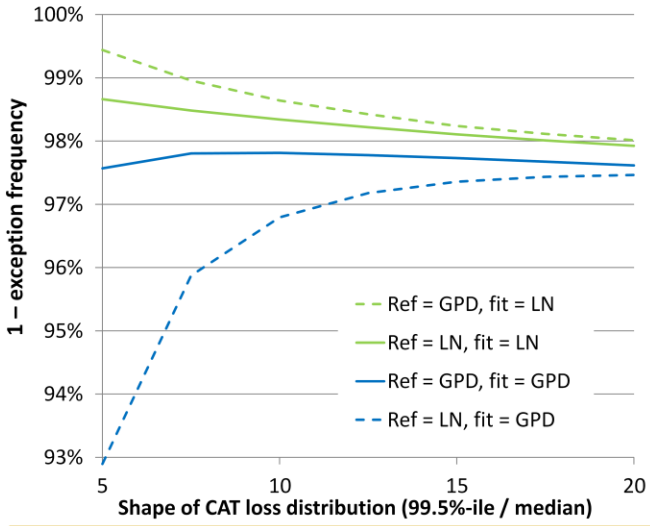
Institute and Faculty of Actuaries



In a Bayesian world, we can talk about the probability of a model being the right one. Valeria Bignozzi and Andreas Tsanakas investigated what happens if you use Bayesian method, but then tested the approach using our frequentist test of counting how often losses exceed required capital. They found good performance – but only if your Bayesian average includes the true model.

The alternative to get a robust method is to take a frequentist approach, but tweak it in some way until it passes the tests for a sufficiently broad ambiguity set.

## Robust Example (Nicholson & Smith, 2013)




The chart shows the probability that the estimated 99.5%-ile exceeds the next observations, based on 40 years' CAT data. This shows that if you want VaR at 98%, fitting the lognormal gives more robust answers than Generalised Pareto. Eshun et al (2011) found similar results for market risks

Institute and Faculty of Actuaries

30 October 2014

28

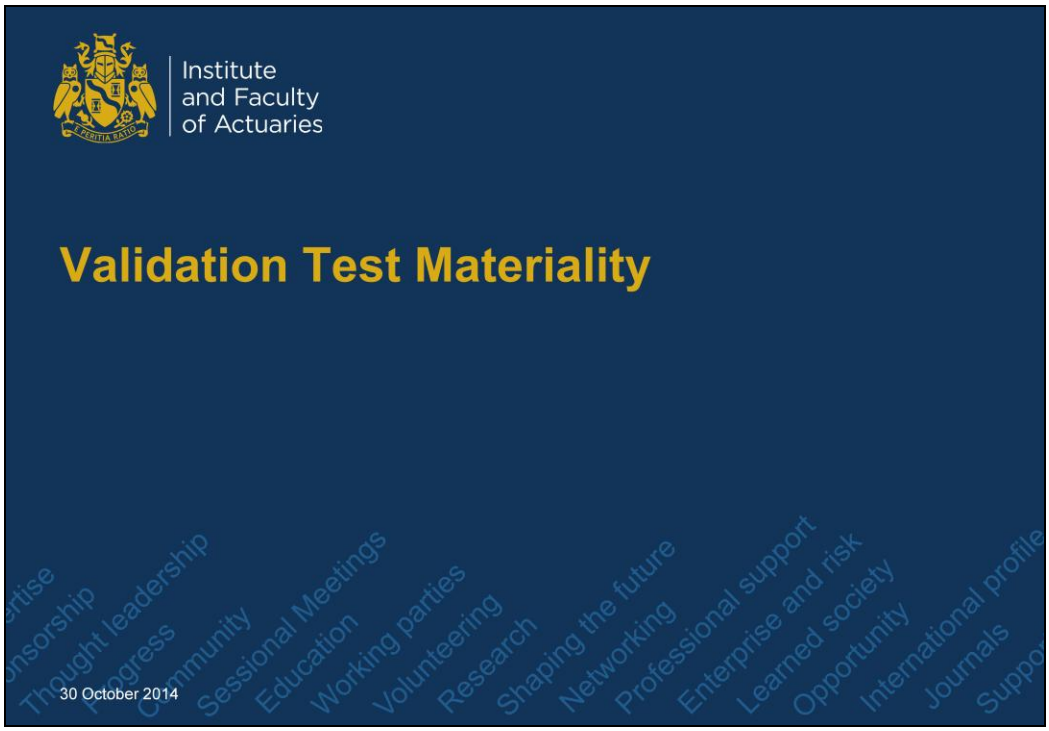
Here is a worked example I prepared earlier – it can be found in Nicholson and Smith (2013).

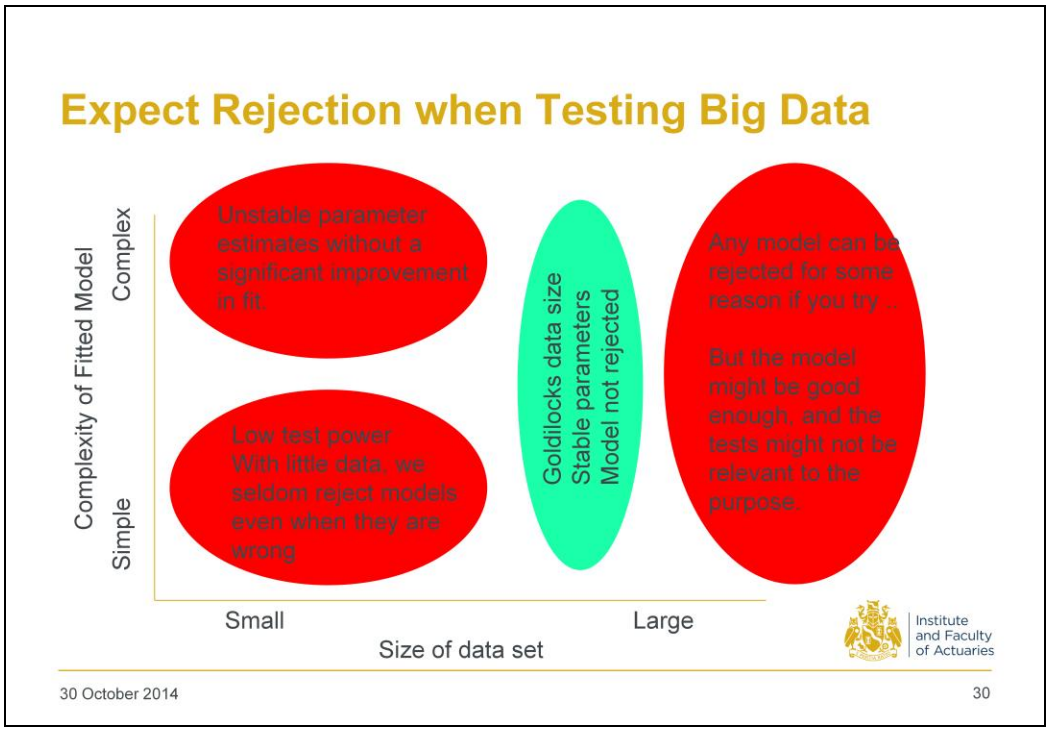
Suppose we wanted robust value-at-risk with 98% confidence.

The green lines on this chart show what happens if your fitting set is a lognormal, but the ambiguity set consists of lognormal and generalised Pareto distributions. The horizontal axis is distribution shape, with fatter tailed distributions to the right.

The chart shows firstly that fitting a distribution and reading off the 99.5%-ile does *not* give us a 99.5% chance of having enough capital. The discrepancy is because the model we have fitted does is not necessarily the same as the one that generated the data, because of sampling and mis-specification error. However, this *does* show robustness relative to a 98% VaR standard, even if we take parameter and mis-specification into account, with some conservatism for thinner tailed distributions.

The chart also shows (blue line) what happens when we fit GPD. Now, the lognormal is essentially an arbitrary choice while the GPD might be justified by appeal to Extreme Value Theory (the Pickands-Balkema-de Haan Theorem). Our results, however, do not support the EVT approach. Scored in terms of robustness, fitting GPD is much less robust than fitting lognormal, at least for the ambiguity set we have considered. That shouldn't be a surprise; in the same way someone might advocate using normal distributions on the strength of the central limit theorem but we all know that using normals lacks robustness to fat tailed models in the ambiguity set.


Institute
and Faculty
of Actuaries
Validation Test Materiality
30 October 2014
Thought leadership
Progress
Community
Sessional Meetings
Education
Working parties
Volunteering
Research
Shaping the future
Networking
Professional support
Enterprise and risk
Learned society
Opportunity
International profile
Journals



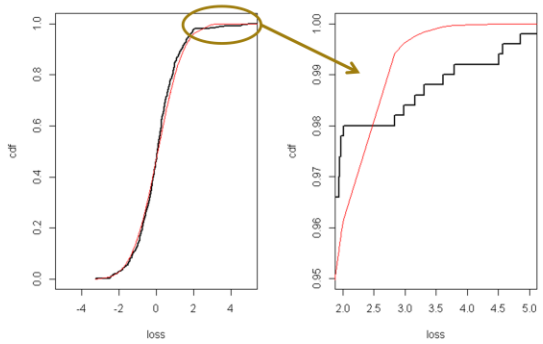

This chart shows what can happen with model fitting for different sized data sets.

We know that fitting complex models to small data sets is a recipe for instability and a whole host of other problems. The only practical way forward is to *use* simple models, even though its obviously possible that a complex model might have generated the data.

There's a different problem with very large data sets. Whatever model we try to fit, we can always find some pattern in the data which the model doesn't capture. There's an apparent sweet spot in the middle where the data is just right, but we won't find ourselves here except by luck.

What then can we do about model failure for large data sets? This question is not so well addressed in the statistical literature, but we think the concept of accounting materiality is helpful here. No model is correct, and we can find tests that fail, but the model might still be useful for the intended purpose. For example, if we want to set capital at a high level of confidence, we might focus on tests in the tail and be prepared to overlook failure of other, less relevant, tests.

# Example of Normal Model not Rejected

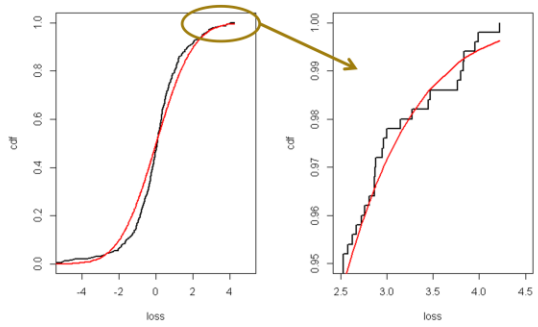



| Estimation method | 99% VaR | 99.5% VaR |
|---|---|---|
| Fitted model | 2.62 | 2.89 |
| Empirical | 3.62 | 4.53 |
| Theoretical | 3.36 | 4.03 |

We **accept** a model that cannot be shown to be wrong, but **is not useful** (for capital calculations, because the errors are in the tail where the impact is most material).

Institute and Faculty of Actuaries

# Example of a Rejected Model




| Estimation method | 99% VaR | 99.5% VaR |
|---|---|---|
| Fitted model | 3.66 | 4.05 |
| Empirical | 3.78 | 3.90 |
| Theoretical | 2.32 | 2.58 |

**We reject** a model that appears to be wrong, but **is useful** (errors have an immaterial impact because we capture tail behaviour well)

Institute and Faculty of Actuaries

## Some new ideas for actuarial statistics

*Familiar Concepts:*

Fitting set of models

Unbiased and asymptotically efficient parameter estimates.

Type I and Type II errors of statistical tests

*Newer Concepts:*

Ambiguity set of models that might have generated the data.

Quality criteria for VaR estimates

Robustness of a methodology to model mis-specification.

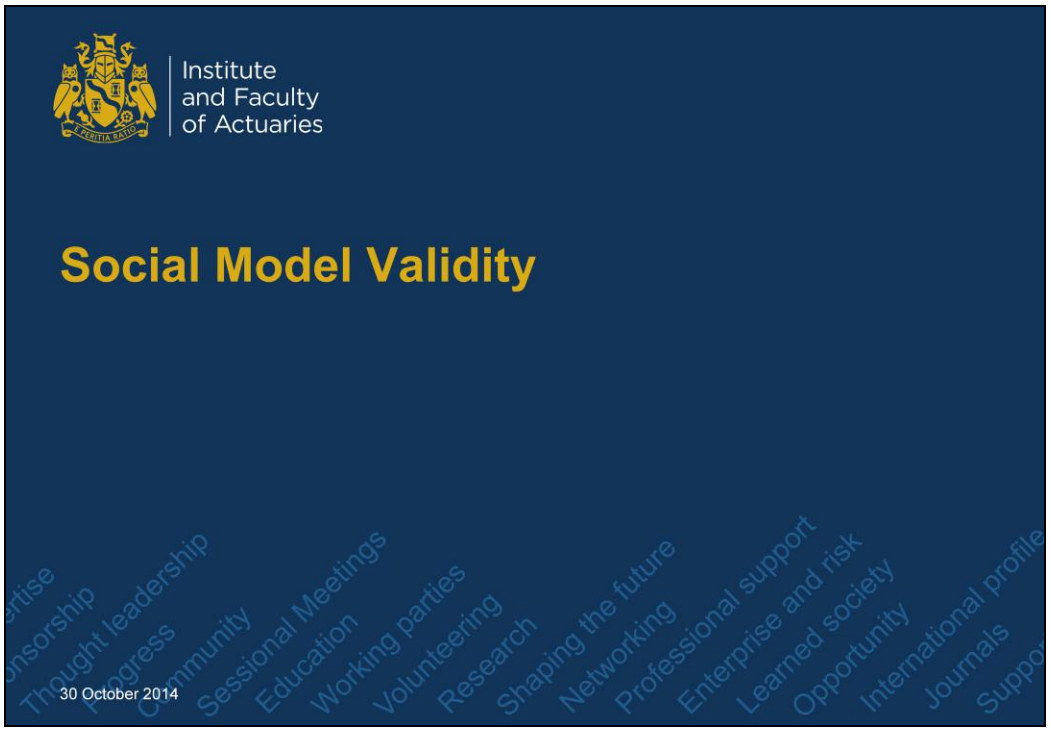
Statistical test materiality

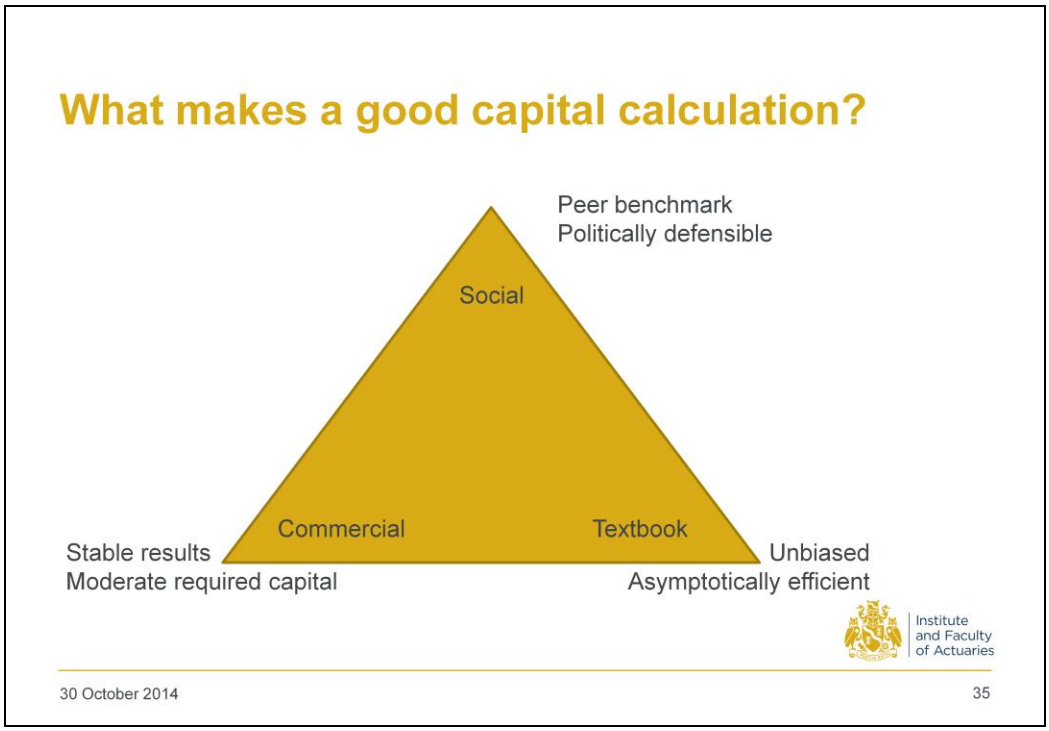
Institute and Faculty of Actuaries



I guess that many of you, like me, would have learned statistics according to the methods of great 20th century frequentists: Ronald Fisher, Jerzy Neyman and Egon Pearson. We try to do something sensible and logical in our day job, but we struggle to articulate what we're doing in formal statistical terms.

This struggle is real; the statistical tools we think we're using were not designed for our purposes – Ronald Fisher was mostly interested in improving agricultural productivity. We should not be afraid to rebuild foundations that work for us, and I expect ambiguity, robustness and test materiality to be among the concepts we need.


Institute
and Faculty
of Actuaries
Social Model Validity
30 October 2014
Thought leadership
Progress
Community
Sessional Meetings
Education
Working parties
Volunteering
Research
Shaping the future
Networking
Professional support
Enterprise and risk
Learned society
Opportunity
International profile
Journals




Technical validity is not sufficient for a model to be used, or to be useful.

In many cases, we also have to comply with social constraints. When a textbook states that an estimate is "best", this assertion is in respect of some, perhaps implicit, loss function. There is nothing wrong with replacing that loss function by something more relevant for our business. For example we might focus on the average amount of capital required, or on the stability of capital estimates, rather than on the expected squared deviation of a parameter from its true value.

Some of our choices are more socially driven than is often acknowledged. Benchmarking is rife in our industry; when we don't know what to do, we take comfort from doing the same things as everyone else. We can defend methods as consistent with industry best practice even if we can't prove they work.

The problem here is not one of ineffective methods, as we have to do the best we can, but rather in misrepresenting a social convention as a mathematical theorem. For example, we might decide that insurers should hold sufficient capital to withstand a mass lapse of 40% of their policies. I can no more prove this is a 99.5%-ile than I can prove you should wear a tie to meetings in Staple Inn, or that you should eat with a fork in your left hand. These are all social conventions and we do ourselves no favours with spurious quasi-statistical justifications. Likewise, the size of the ambiguity set is part of the social contract between policyholders and financial services firms. You won't find a theorem proving how big that set should be.

The terms "technical" and "social validity" were introduced by J. G. March (1994), *A Primer on Decision Making: How Decisions Happen*. A discussion of these concepts in the context of capital modelling can be found in Tsanakas (2012), 'Modelling: The Elephant in the Room, *Actuary Magazine*.

# Transparency

- We need to accept the legitimacy of social and commercial constraints on modelling decisions
  - See forthcoming paper from Model Risk Working Party
- Pretending modelling decisions are purely technical inhibits careful thinking
  - Disincentive for stating uncertainty
  - Fear of the "wrong answer"
- If it is "not about the number" what is it *really* about?

Institute
and Faculty
of Actuaries
Conclusions
30 October 2014
Thought leadership
Progress
Community
Sessional Meetings
Education
Working parties
Volunteering
Research
Shaping the future
Networking
Professional support
Enterprise and risk
Learned society
Opportunity
International profile
Journals

## Concluding Remarks

- You can never know that a model is right. You can know that
  - a series of tests have not rejected a model
  - the way our estimates were worked was good
- Given the purpose of solvency regulation, statistical bias is best considered at the end outcome (capital shortfalls) rather than on intermediate steps such as parameter estimates
- Robust methodologies should avoid downward capital bias under a wide range (an ambiguity set) of possible underlying models

Institute and Faculty of Actuaries

## Concluding Remarks (continued)

- Models do not need to be perfect; they need to be fit for purpose. Some statistical tests will fail on any large data set, but a model can still be useful even if it is wrong.
- There will always be social and commercial constraints in model choice. Validation should understand the impact of these constraints.

# References


- Balter and Pelsser (2014), *Pricing and Hedging in Incomplete Markets with Model Ambiguity*. http://ssrn.com/abstract=2459405
- Barndorff-Nielsen and Cox (1996), Prediction and Asymptotics, *Bernoulli.*
- Basel Committee (1996), http://www.bis.org/publ/bcbs22.pdf
- Bignozzi and Tsanakas (2013), *Model Uncertainty in Risk Capital Measurement*. http://ssrn.com/abstract=2334797
- Bignozzi and Tsanakas (2014), *Parameter Uncertainty and Residual Estimation Risk.* http://ssrn.com/abstract=2158779
- Eshun, Machin, Sharpe and Smith (2011), *Extreme Value Theory for a 1-in-200 Event.* http://www.actuaries.org.uk/sites/all/files/documents/pdf/c4extremeworkingparty.pdf

# References (continued)


- European Insurance and Occupational Pensions Authority (2010) Advice for Level 2 Implementing Measures on Solvency II: Calibration of Market Risk Module. https://eiopa.europa.eu/fileadmin/tx_dam/files/consultations/consultationpapers/CP70/CEIOPS-L2-Advice-Market-risk-calibration.pdf
- Frankland et al. (2013), *Difficult Risks and Capital Models*. http://www.actuaries.org.uk/events/one-day/sessional-research-event-extreme-events-working-party-paper-0
- Gerrard and Tsanakas (2011), *Failure Probability Under Parameter Uncertainty*. http://ssrn.com/abstract=1486317
- Nicholson and Smith (2013). *GI: Prepare for the Worst* http://www.theactuary.com/features/2013/10/gi-prepare-for-the-worst
- Smith (2014) Overlapping Intervals for Predicting Insurance Risk, http://www.turing-gateway.cam.ac.uk/financial_risk2014-program


Institute and Faculty of Actuaries

Expressions of individual views by members of the Institute and Faculty of Actuaries and its staff are encouraged.

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

Institute and Faculty of Actuaries