Notes on Model Risk
(Or: Black Swans and Red Herrings)

CAS Annual Meeting
Boston
November, 2009

Jim Guszcza, FCAS, MAAA
Deloitte Consulting LLP
jguszcza@deloitte.com

A Question

<u>Interview with John Kay (Financial Times)</u>:

Q: Tell me, why did most investment models, built by Harvard, Yale and Cambridge Mathematics PhDs, appear to fail?

A: Put simply, people made the mistake of believing the model. The people who built them – the mathematics PhDs – didn't know very much about the world. The people who knew about the world didn't understand the mathematics. Both groups had inappropriate confidence in the value of these models. They aren't useless – but models can only illuminate the world, never be a substitute for judgment.

- Keep this question in mind during the talk.

Motivation

> "It's about time that actuaries got more involved in quantitative finance and brought some common sense back into this field. We need models people can understand and a greater respect for risk… what high finance needs now are precisely the skills that actuaries have, a deep understanding of statistics, an historical perspective, and a willingness to work with data."
>
> -- Paul Wilmott, "Actuaries vs Quants", *The Actuary* (UK), November 2008

- The fact that the topic of "model risk" regularly shows up in actuarial papers and presentations supports Wilmott's sentiment.
    - So let's explore the concept of "model risk" a little more deeply.

Themes

The Three Faces of Risk

Model Risk: modeling the data

Model Risk: modeling reality

Model Risk: applying models

Deloitte.

The Three Faces of Risk

Process Risk
Parameter Risk
Model Risk

The Three Faces of Risk

Sources of uncertainty (in order of insidiousness):

- Process risk: uncertainty due to the stochastic nature of the model.
- Parameter risk: uncertainty in the values of the model parameters.
- Model risk: uncertainty in the appropriateness of our model.

A Discussion of
Parameter and Model Uncertainty in Insurance

Andrew J.G. Cairns
Department of Actuarial Mathematics and Statistics,
Heriot-Watt University,
Edinburgh, EH14 4AS, United Kingdom

E-mail: A.Cairns@ma.hw.ac.uk
WWW: http://www.ma.hw.ac.uk/~andrewc/

The Three Faces of Risk

<u>Sources of uncertainty (in order of insidiousness)</u>:

- Process risk: uncertainty due to the stochastic nature of the model.
- Parameter risk: uncertainty in the values of the model parameters.
- Model risk: uncertainty in the appropriateness of our model.

$$\log(severity) = \alpha + \beta_1 priorClaim + \beta_2 tenure + \beta_3 credit + \ldots + \varepsilon$$

The Three Faces of Risk

<u>Sources of uncertainty (in order of insidiousness)</u>:

- Process risk: uncertainty due to the stochastic nature of the model.
- Parameter risk: uncertainty in the values of the model parameters.
- Model risk: uncertainty in the appropriateness of our model.

$$\log(severity) = \alpha + \beta_1 priorClaim + \beta_2 tenure + \beta_3 credit + \ldots + \varepsilon$$

The Three Faces of Risk

Sources of uncertainty (in order of insidiousness):

- Process risk: uncertainty due to the stochastic nature of the model.
- Parameter risk: uncertainty in the values of the model parameters.
- Model risk: uncertainty in the appropriateness of our model.

$$\log(severity) = \alpha + \beta_1 priorClaim + \beta_2 tenure + \beta_3 credit + \ldots + \varepsilon$$

Illustration of Model Risk: Specification Error

- These four datasets were introduced by F.J. Anscombe in 1973.
  - See also *The Visual Display of Quantitative Information* by Edward Tufte.

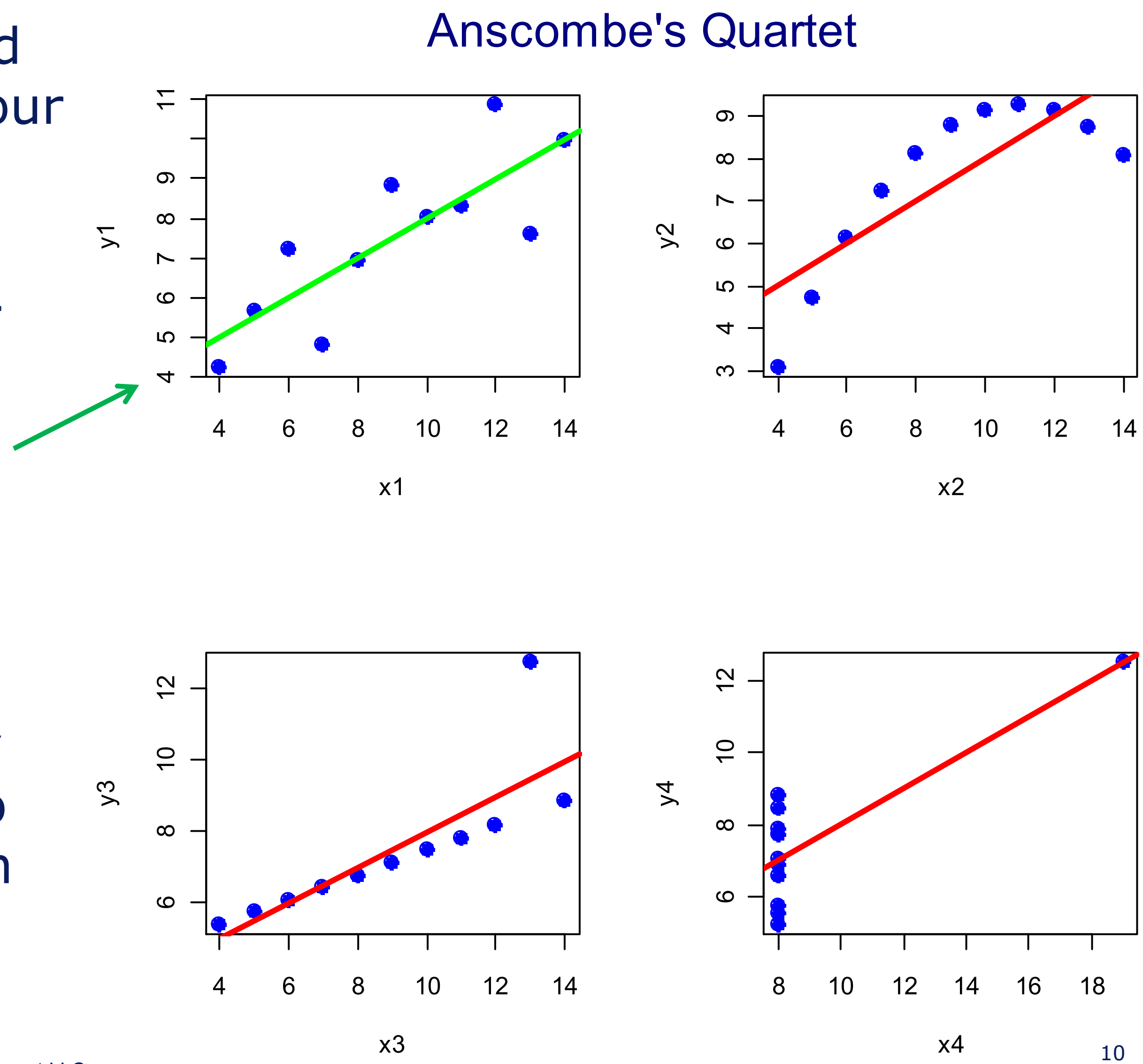
A		B		C		D	
x1	y1	x2	y2	x3	y3	x4	y4
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.1	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.1	4.0	5.39	19.0	12.5
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

- The regression models fit to each of these four datasets are identical.
  - Does the regression output tell the whole story?

```
> coef(lm(y1~x1, data=anscombe))
(Intercept)          x1
  3.0000909   0.5000909
> coef(lm(y2~x2, data=anscombe))
(Intercept)          x2
   3.000909    0.500000
> coef(lm(y3~x3, data=anscombe))
(Intercept)          x3
  3.0024545   0.4997273
> coef(lm(y4~x4, data=anscombe))
(Intercept)          x4
  3.0017273   0.4999091
```

Anscombe's Quartet

- The parameters and statistics from all four models are nearly identical.
  - α, β, σ^2, t, F, R^2, …
- But the model is appropriate only in the 1st case.
- Looking at model output is necessary but not sufficient to control specification error.



Model Specification Error: the Basics

- Some examples of what can go wrong:
  - Improper assumption of additivity / linearity
  - Inappropriate distributional assumptions
    - Exponential family?
    - Constant dispersion parameter?
  - Missing or incorrect weight/offset
  - Non-iid data points
  - Missing variables (omitted variable bias)
  - Too many variables: (irrelevant variables, multicollinearity)
  - Overparameterization
  - Variables with unintended proxy relationship
  - Bias due to improper missing data imputation methods
  - Extreme leverage of outlier data points
  - …

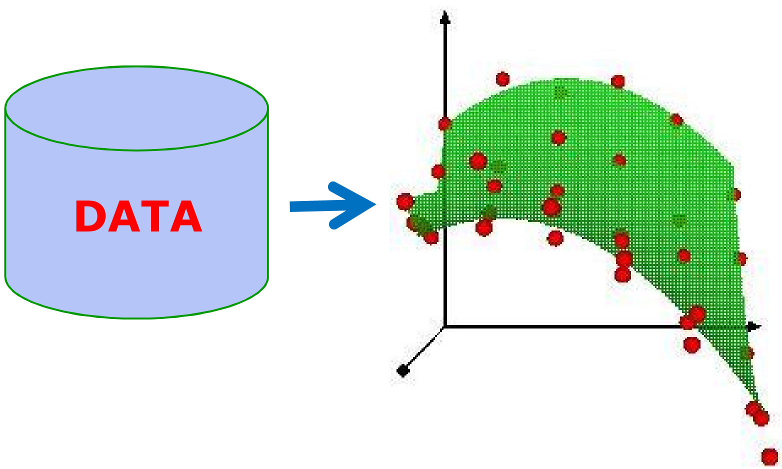
… And this is only the beginning of our discussion of model risk!

Beyond the Data

Using the past to predict the future
The bias-variance tradeoff
Black Swans and Knightian Uncertainty

What We Talk about When We Talk About Model Risk

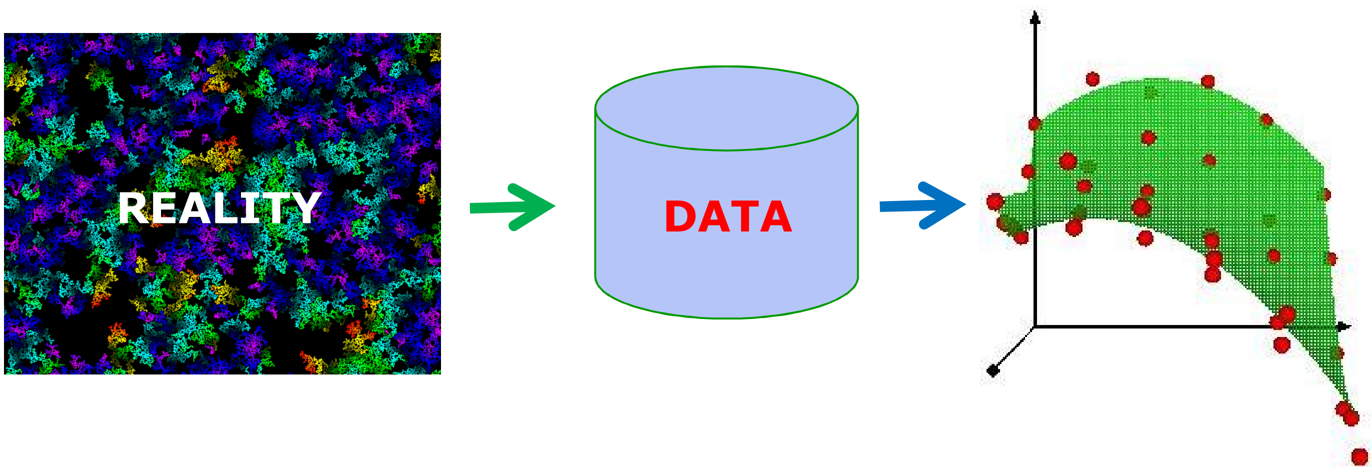
- The Anscombe Quartet illustrates forms of model risk that are largely under the analyst's control, sitting in front of a computer screen, analyzing data.



- This is model risk in the sense of **model specification error**.
- But this shouldn't be the last word on model risk.
- The goal of a statistics project is not to fit a model to data….

What We Talk about When We Talk About Model Risk

- The goal of a statistics project isn't to fit a model to data…



- Closer to the mark: the goal is to create a model that captures the <u>reality</u> underlying the data.

We need to be aware of possible ways in which the data at hand do not capture the underlying reality we wish to model.

Modeling Reality Rather than Modeling Data

- Traditional actuarial techniques
    - Accounting for historical changes in a book of business
    - On-leveling premium, trending and developing losses
    - Credibility theory to account for data limitations

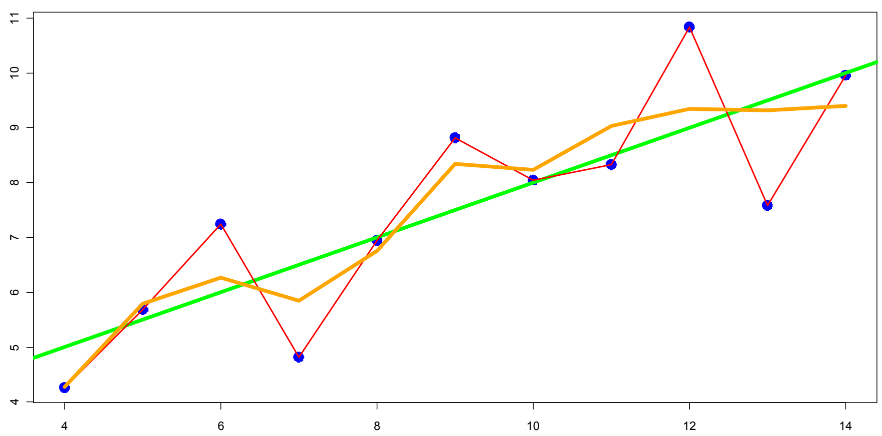
- Modern techniques to avoiding fitting *all* patterns in data
    - Diagnostic plots (QQ, etc)
    - Blind-test validation using holdout data
    - Cross-validation to manage the bias-variance tradeoff
    - Shrinkage techniques
    - Bayesian and Hierarchical modeling techniques
    - General theme: not all patterns in the data will generalize to the future

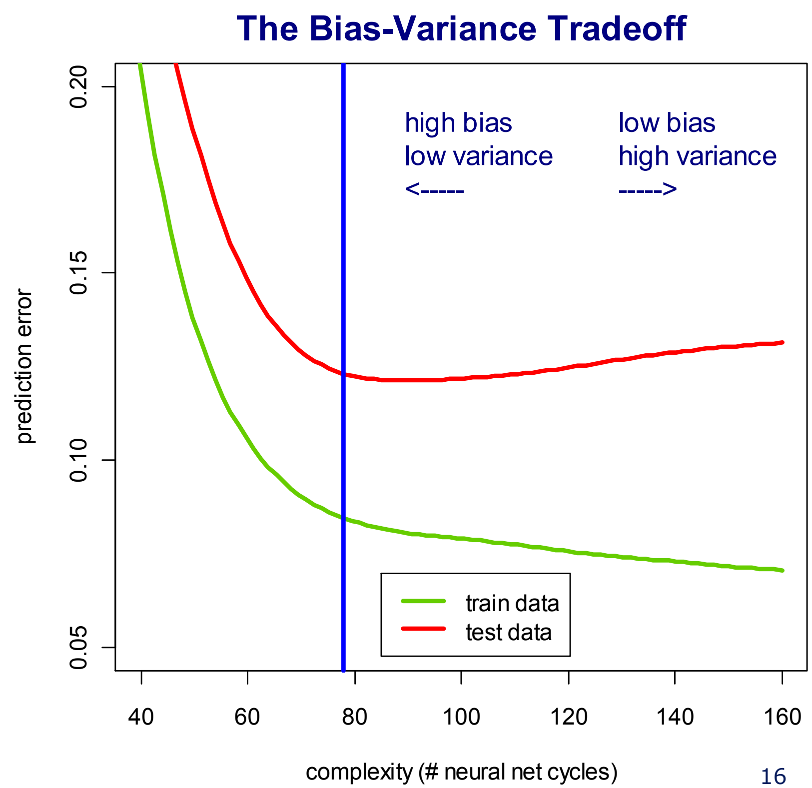
- Recognition that some patterns might be absent from the data
    - Tail risk, high-severity/low-frequency events
    - Need for cat modeling, extreme value theory
    - Knightian Uncertainty, Black Swans, etc

The Data Contains too Many Patterns

- A type of model risk: fitting too much of the random noise in the data and/or missing the true signal.

"Everything should be made as simple as possible, but not simpler"
– Albert Einstein





- Cross-validation has become a widely used technique to avoid building models that over-fit the data.
  - Manage the bias-variance tradeoff

The Data Contains too Few Patterns

- David Hume: the problem of induction
  - Also Sextus Empiricus, Karl Popper, Nelson Goodman, Nassim Taleb, …
  - We can never be certain when making inferences from observed facts to unobserved facts.
  - All observed swans have been white; *x* is a swan $\not\Rightarrow$ *x* is white
  - Before black swans were discovered, what was the probability of a non-white swan? Could it even be measured?

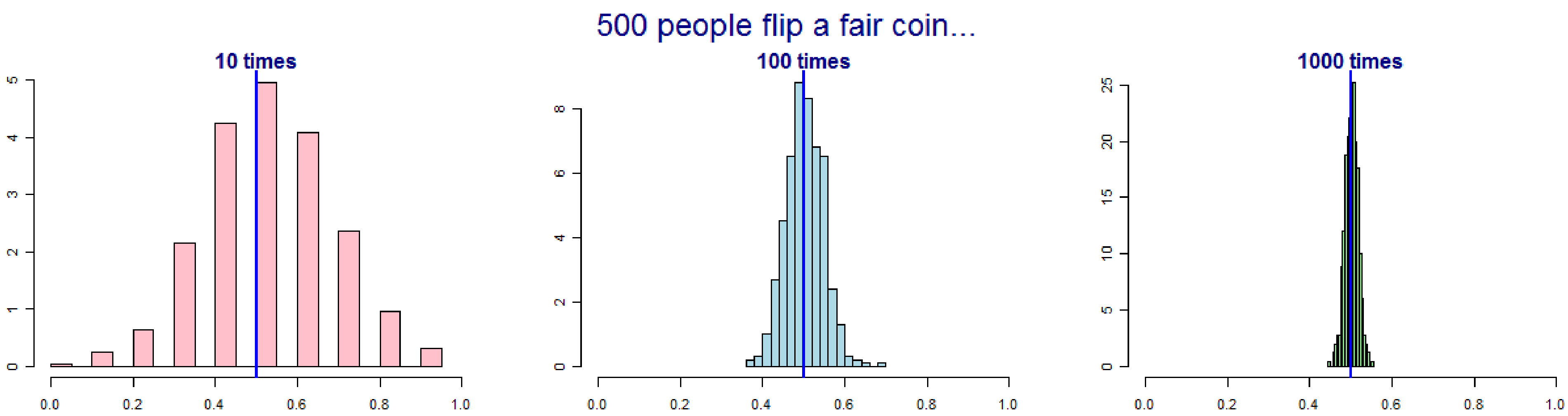
Frank Knight's Distinction between Risk and Uncertainty

"Uncertainty must be taken in a sense radically distinct from the familiar notion of Risk, from which it has never been properly separated. The term "risk," as loosely used in everyday speech and in economic discussion, really covers two things which… are categorically different… "Risk" means in some cases a quantity susceptible of measurement… **A measurable uncertainty, or "risk" proper, as we shall use the term, is so far different from an unmeasurable one that it is not in effect an uncertainty at all. We shall accordingly restrict the term "uncertainty" to cases of the non-quantitative type.** It is this "true" uncertainty, and not risk, as has been argued, which forms the basis of a valid theory of profit and accounts for the divergence between actual and theoretical competition."

--*Risk, Uncertainty, and Profit* (1921)

Simple Illustration of Knightian Uncertainty

- Consider a bunch of people tossing coins.
- Let's take a step back and consider process risk, parameter risk, and model risk in turn.
- **Process risk:** the process is stochastic… but the more flips, the more confident we can be about the outcome.



Simple Illustration of Knightian Uncertainty

- **Parameter risk:** even if we are 100% sure that the "true model" is Bernoulli, we don't know the "true probability".
  - We must infer it from the data. The more data, the more confident we are in our inference.

- Bayesian modeling: rather than adopt a model with a parameter that we are uncertain about, we integrate over that parameter.

$$\boxed{\binom{n}{k}\theta^k(1-\theta)^{n-k} \quad \Rightarrow \quad \int_0^1 \binom{n}{k}\theta^k(1-\theta)^{n-k}\,d\mu(\theta)}$$

- **De Finetti's Representation Theorem:** an <u>exchangeable</u> sequence of random variables can be represented as a mixture of iid random variables.
  - See appendix: exchangeability ≈ "the future will resemble the past"

Simple Illustration of Knightian Uncertainty

- **Model risk #1 (specification error):** some magicians can influence the outcome of a coin toss so that the probability of heads is greater than the probability of tails.

Suppose a coin landed heads 507 times in 1000 tosses. What is the probability of heads on the 1001st toss?

Simple Illustration of Knightian Uncertainty

- **Model risk #1 (specification error):** some magicians can influence the outcome of a coin toss so that the probability of heads is greater than the probability of tails.

Suppose a coin landed heads 507 times in 1000 tosses.
What is the probability of heads on the 1001st toss?

Closer inspection reveals that:

Freq(Heads | Heads) ≈ .405
Freq(Heads | Tails) ≈ .613

- In this case Persi the Magician influenced the flips differently depending on the outcome of the previous flip.

➔ Using a simple Binomial model would be specification error.

  - (We need to introduce transition probabilities.)
  - This is still not Knightian uncertainty! … Just garden variety model risk.

Simple Illustration of Knightian Uncertainty

- **Model risk #2 (a Black Swan):** Consider a different sequence of coin tosses.

This time Persi tossed 505 heads in 1000 tosses.

Close inspection shows that the frequency of heads was ≈50% regardless of the outcome of the previous toss.

So now what is the probability of heads in the next toss?

Simple Illustration of Knightian Uncertainty

- **Model risk #2 (a Black Swan):** Consider a different sequence of coin tosses.

> This time Persi tossed 505 heads in 1000 tosses.
>
> Close inspection shows that the frequency of heads was ≈50% regardless of the outcome of the previous toss.
>
> So now what is the probability of heads in the next toss?

- 50%, right?

Simple Illustration of Knightian Uncertainty

- **Model risk #2 (a Black Swan):** Consider a different sequence of coin tosses.

> This time Persi tossed 505 heads in 1000 tosses.
>
> Close inspection shows that the frequency of heads was ≈50% regardless of the outcome of the previous toss.
>
> So now what is the probability of heads in the next toss?

- 50%, right?

- Suppose someone just sent Persi a secret message, offering him $1 for every toss that lands heads.

- Your model will fail <u>even though it was correctly specified</u>.
  - Back to Hume: no guarantees that the future must resemble the past.

Real World Knightian Uncertainty

- This is all obvious, right? Maybe not.

Comment by Edmund Phelps in the *Financial Times*:

"But why did big shareholders not move to stop over-leveraging before it reached dangerous levels? Why did legislators not demand regulatory intervention? The answer, I believe, is that they had no sense of the existing **Knightian uncertainty**. So they had no sense of the possibility of a huge break in housing prices and no sense of the fundamental inapplicability of the risk management models used in the banks. "Risk" came to mean volatility over some recent past. The volatility of the price as it vibrates around some path was considered but not the uncertainty of the path itself: the risk that it would shift down. The banks' chief executives, too, had little grasp of uncertainty. Some had the instinct to buy insurance but did not see the uncertainty of the insurer's solvency."

Financial Times, April 15, 2009

http://www.ft.com/cms/s/0/41f536ee-2954-11de-bc5e-00144feabdc0.html?nclick_check=1

Examples of Knightian Uncertainty

- 9/11
- Superfund: retroactive liability
- The banking crisis
- The emergence of the internet (a positive black swan)
- Possible emerging black swan: a scientific breakthrough that would dramatically increase the human lifespan (also positive, but with financial implications for insurers).

*Knightian uncertainty can't be **modeled**… but it must be **managed**.*

Risk Management Implications

From the Turner Review (the FSA's report on the banking crisis):

"**Non-independence of future events; distinguishing risk and uncertainty.** More fundamentally, however, it is important to realize that the assumption that past distribution patterns carry robust inferences for the probability of future patterns is methodologically insecure. It involves applying to the world of social and economic relationships a technique drawn from the world of physics, in which a random sample of a definitively existing universe of possible events is used to determine the probability characteristics which govern future random samples. But it is unclear whether this analogy is valid when applied to economic and social relationships, or whether instead, we need to recognise that we are dealing not with mathematically modellable risk, but with inherent 'Knightian' uncertainty. This would further reinforce the need for a macro-prudential approach to regulation.

But it would also suggest that no system of regulation could ever guard against all risks/uncertainties, and that there may be extreme circumstances in which the backup of risk socialization (e.g. of the sort of government intervention now being put in place) is the optimal and the only defence against system failure."

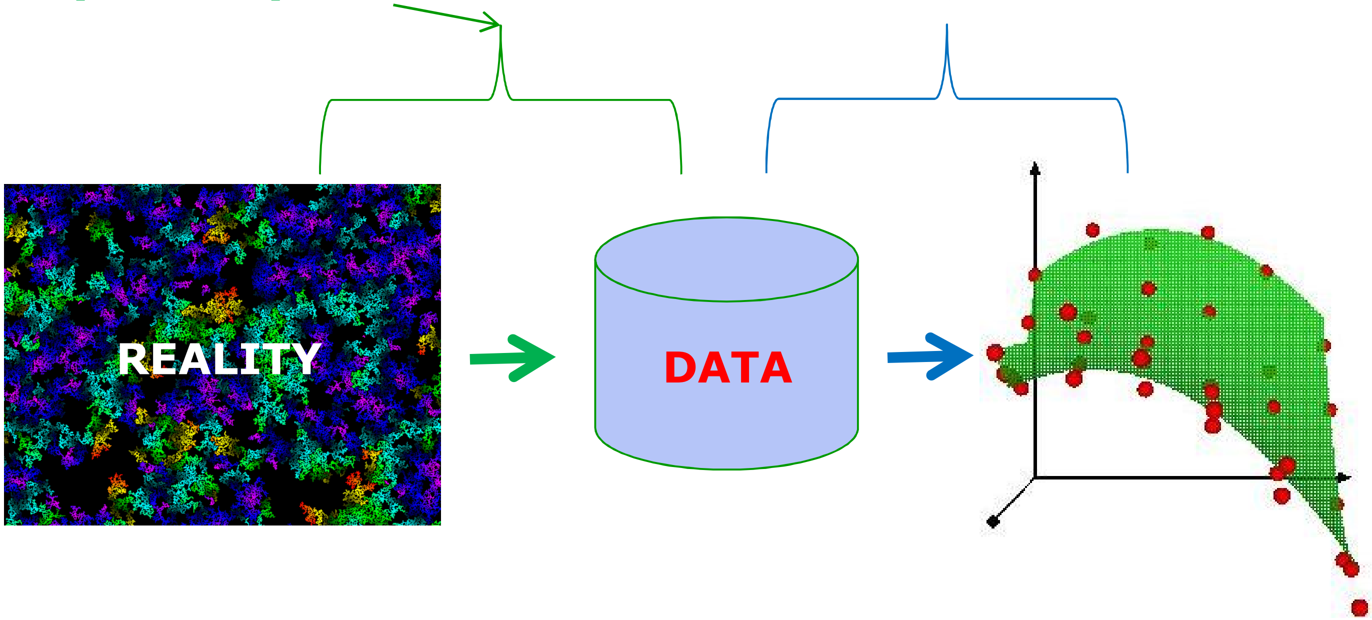
http://www.fsa.gov.uk/pubs/other/turner_review.pdf

“Enterprise” Model Risk

Practical Risks
The Computer Factor
The Human Factor

Stretching the Concept Still Further

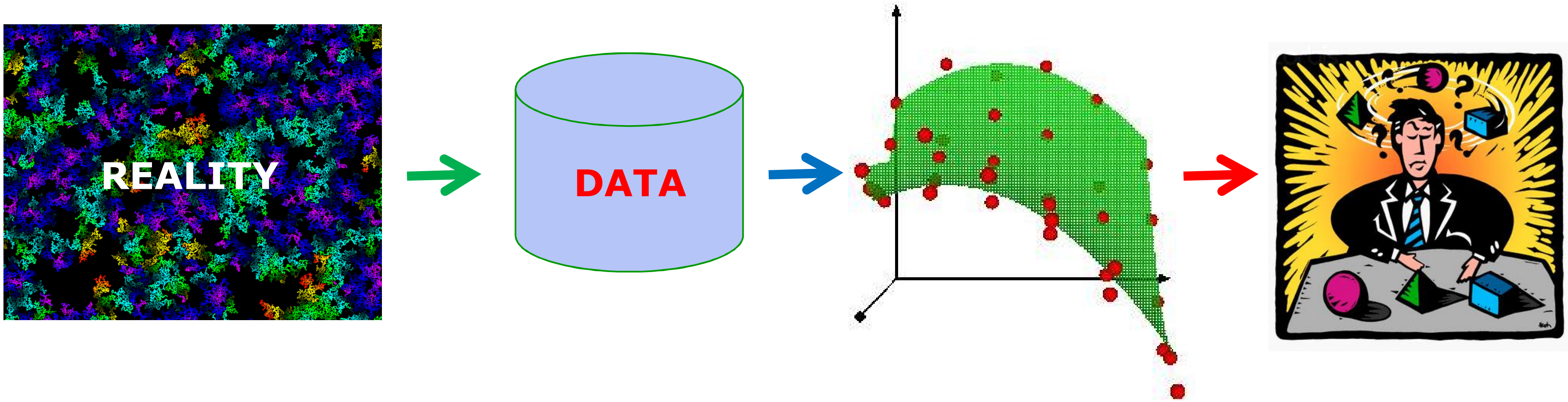
- Our discussion started with the **theoretical** and veered towards the **philosophical**.



Stretching the Concept Still Further

- Our discussion started with the **theoretical** and veered towards the **philosophical**.

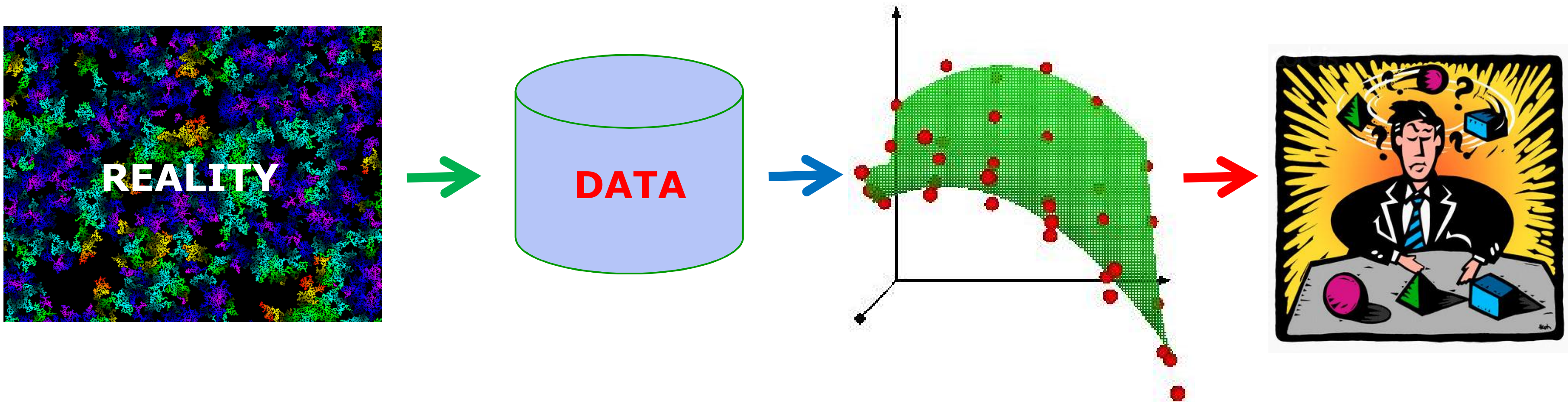
- But “model risk” is a highly **practical** topic.



Stretching the Concept Still Further

- Our discussion started with the **theoretical** and veered towards the **philosophical**.

- But “model risk” is a highly **practical** topic.



- Statistical models are built to help people make decisions
  - or otherwise direct business actions

- Any potential error or omission that impedes this goal is a risk that should be managed.

The Computer Factor

- Emanuel Derman is a prominent physicist / financier who has written about model risk from a broad perspective.

- Models are **software solutions**.

- Possible sources of problems.
  - Coding errors
  - Logic errors
  - Rounding errors
  - Too much complexity
  - Poor metadata
  - Poor documentation
  - Mismatch between data used to fit the model and data fed into the model upon implementation
  - Handling of missing data elements
  - Version control
  - …

Goldman Sachs

Quantitative Strategies Research Notes

April 1996

Model Risk

Emanuel Derman

"Many of the worst risks center around **implementation**. These days, models are sophisticated programs, thousands of lines long, with rich data structures that are used to perform detailed computation. Models undergo revisions by people who were not the original authors. Equally important in making them useful, models need user interfaces, position databases, trade entry screens and electronic price feeds..."

The Human Factor

- Models are also **decision support tools**.

- The human factor is also a source of model risk
  - Lack of executive sponsorship
  - Poor communication between modelers and
    - IT
    - Management
    - End-users
  - Lack of faith in the model by management and/or end users
  - Poor (or non-existent) training and change management
  - Poor business implementation
  - Inconsistent use of the model
  - Poorly conceived business rules
  - Gaming the system by end-users
  - "Red Herrings": excessively focusing on inconsequential technical issues rather than key business decisions

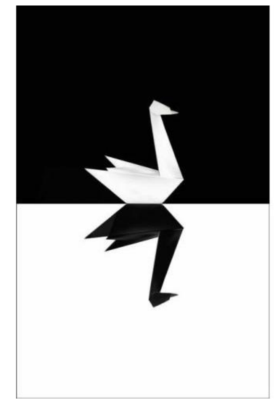
As with "philosophical" model risk, these forms of model risk must be recognized and managed.

Modeling is Only Part of a Modeling Project

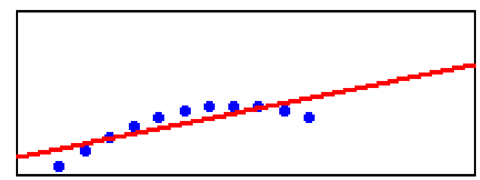
- The modeling project doesn't <u>begin</u> with data; it begins with a:
    - Executive-level initiative (hopefully)
    - Question
    - Strategic goal
- The modeling project doesn't <u>end</u> with the model; it ends with:
    - IT implementation
    - Business rule creation
    - Training
    - Change management
    - Monitoring

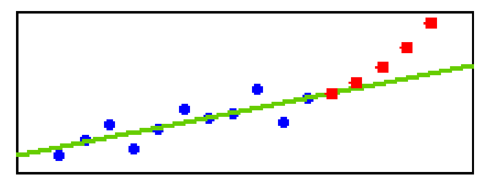
"Model risk": any potential shortcoming that could prevent the model from addressing the ultimate strategic goal.

Summing Up: Varieties of Model Risk

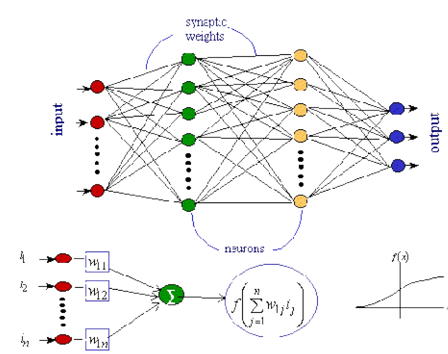


- **Philosophical:** will the future reflect the past?
  - The problem of induction (David Hume)
  - Knightian Uncertainty (Frank Knight)
  - Black Swans (Karl Popper, Nassim Taleb)





- **Theoretical:**
  - Model misspecification
  - Problems with extrapolating beyond the data
  - …



- **Practical:**
  - Quality control
  - Communication, misunderstanding
  - Willful misuse (gaming the system)
  - Documentation
  - Complexity
  - Unintended consequences

Appendix: De Finetti's Representation Theorem

- If our beliefs satisfy a certain symmetry condition (exchangeability) it follows that our posterior probability distribution can be represented as a mixture of Binomial distributions.
  - **Exchangeability:** the order of a finite set of random variables does not affect the joint probability. For all n and all permutations σ:

$$\boxed{\Pr(X_1 = e_1, X_2 = e_2, \ldots, X_n = e_n) = \Pr(X_1 = e_{\sigma(1)}, X_2 = e_{\sigma(2)}, \ldots, X_n = e_{\sigma(n)})}$$

  - The order doesn't matter ➔ "the future will be like the past"

- **De Finetti's Representation Theorem:** If $\{X_i\}$ is exchangeable then the limiting relative frequency $\lim_{n\to\infty}(\frac{1}{n}\Sigma X_i)$ exists with probability 1 and:

$$\boxed{\Pr\left(\sum_{i=1}^{n} X_i = k\right) = \int_0^1 \binom{n}{k} \theta^k (1-\theta)^{n-k}\, d\mu(\theta)}$$

In a phrase: an exchangeable sequence is a mixture of iid sequences.

Back to Our Question

Interview with John Kay (Financial Times):

Q: Tell me, why did most investment models, built by Harvard, Yale and Cambridge Mathematics PhDs, appear to fail?

A: Put simply, people made the mistake of believing the model. The people who built them – the mathematics PhDs – didn't know very much about the world. The people who knew about the world didn't understand the mathematics. Both groups had inappropriate confidence in the value of these models. They aren't useless – but **models can only illuminate the world, never be a substitute for judgment**.