Tutorial: Quantifying uncertainty using data and experts

Ullrika Sahlin

Lund University, Sweden

Overview of this tutorial

- Background
  - Uncertainty
  - Unceratinty in Bayesian models
  - Uncertainty analysis
- Tools
  - Bayesian updating in JAGS
  - Expert elicitation - SHELF
- Hands on
  - Exercise – Missing the bus
  - Exercise – Is there a problem out there?

”If you dont have enough data you can use Bayesian methods….”

a random book in quantitative risk assessment

- Classical vs Bayesian statistics
- Depends on the purpose of the analysis
- The focus here is to quantify uncertainty

Uncertainty in risk assessment

This paper is concerned about how to define and describe risk in an engineering context. There exist many definitions of risk in such a setting, but most of them include the following three components:

A: what can go wrong (the initiating events).
C: the consequences of these events if they should occur.
P: the probabilities of *A* and *C*.

In short we write Risk = (*A*, *C*, *P*). There are basically two ways of interpreting the probability *P*:

(a) as a relative frequency, i.e. the relative fraction of times the event occurs if the situation analyzed were hypothetically "repeated" an infinite number of times.
(b) as a subjective measure of uncertainty, conditional on the background knowledge (the Bayesian perspective).

The former interpretation means that probability is used to reflect variation (i.e. what is commonly referred to as stochastic or aleatory uncertainty, Apostolakis, 1990), whereas the latter inter-

Aven 2015

Safety assessments of technological systems, such as nuclear power plants, chemical process facilities, and hazardous waste repositories, require the investigation of the occurrence and consequences of rare events. The subjectivistic (Bayesian) theory of probability is the appropriate framework within which expert opinions, which are essential to the quantification process, can be combined with experimental results and statistical observations to produce quantitative measures of the risks from these systems. A distinction is made between uncertainties in physical models and state-of-knowledge uncertainties about the parameters and assumptions of these models. The proper role of past and future relative frequencies and several issues associated with the elicitation and use of expert opinions are discussed.

PROBABILISTIC RISK ASSESSMENT (PRA) OR PROBABILISTIC safety assessment (PSA) is a method that has evolved during the last 20 years. Its aim is to produce quantitative estimates of the risks associated with complex engineering systems such as nuclear plants, chemical process facilities, waste repositories, and

Apostolakis 1990

Uncertainty in Scientific Assessment

EFSA Journal

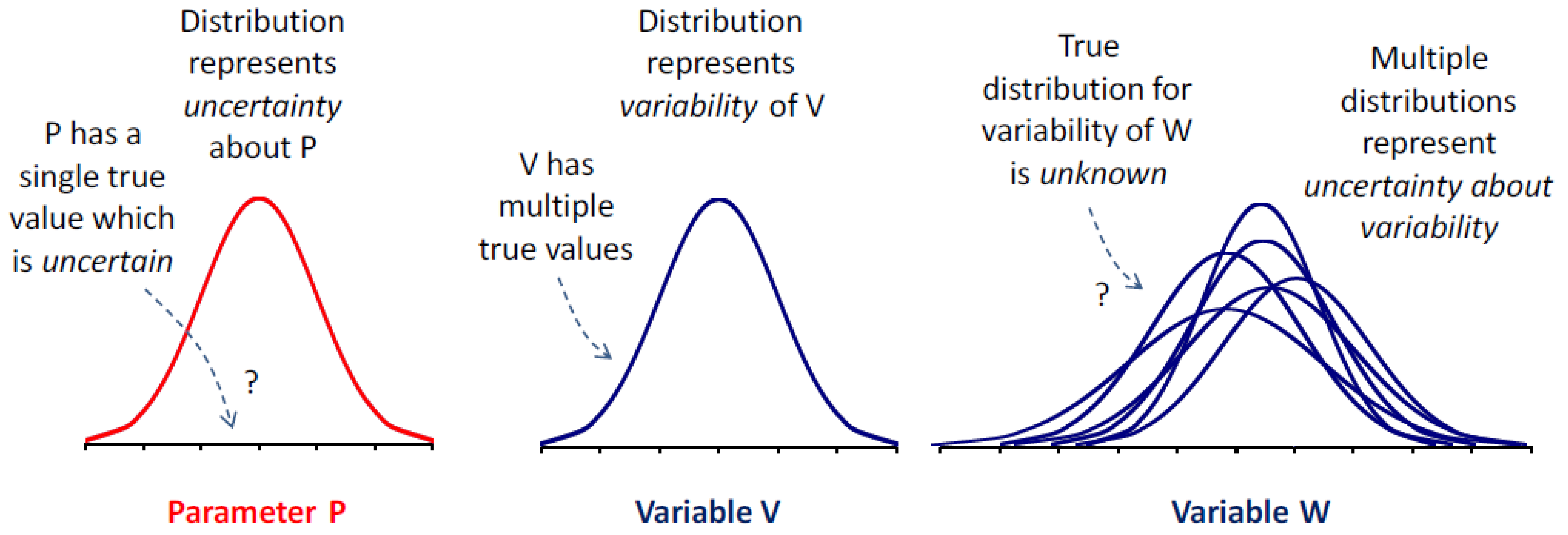
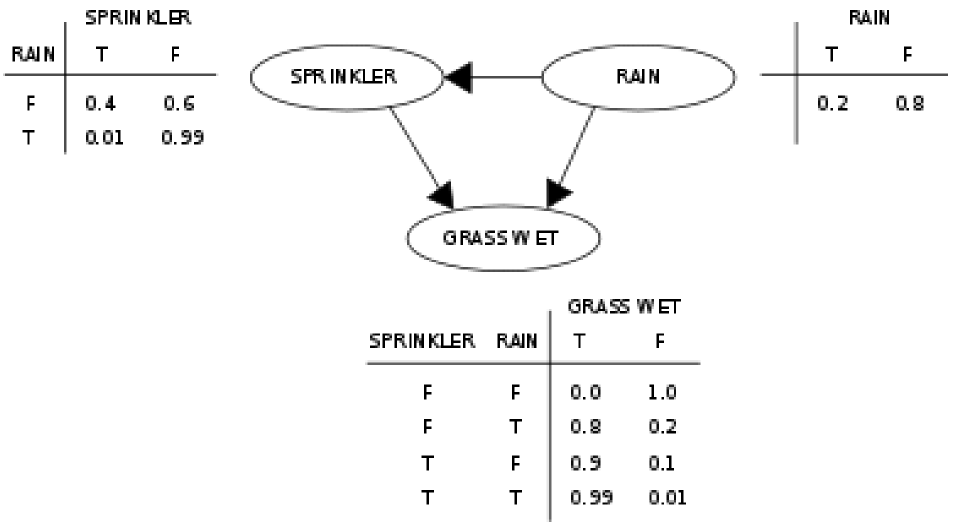


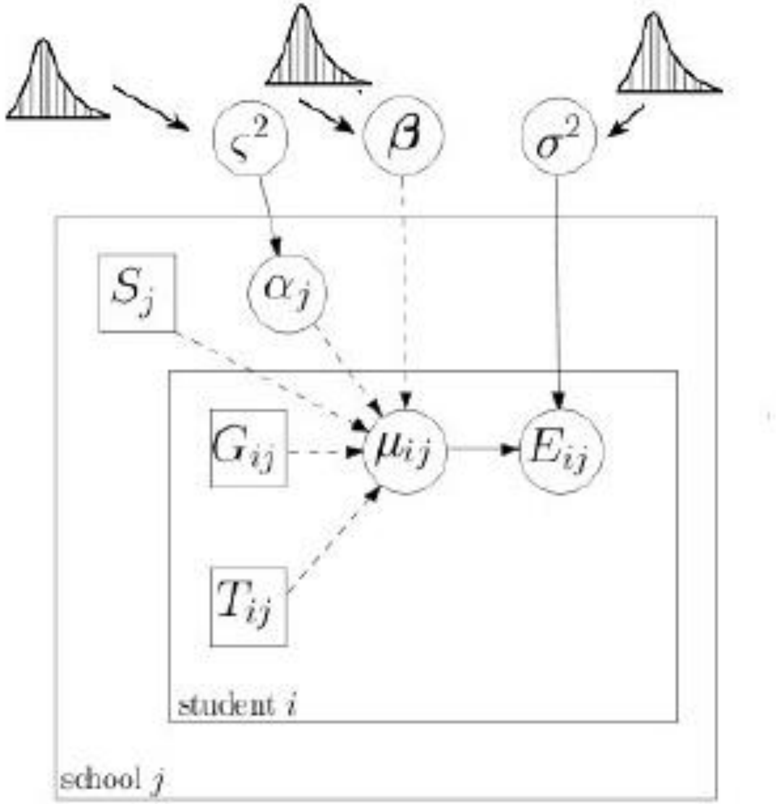
Figure 2: Illustration of the distinction between uncertainty and variability (left and central graphs), and that both can affect the same quantity (right hand graph).

Relative frequency, subjective probability or both?

Bayesian Network

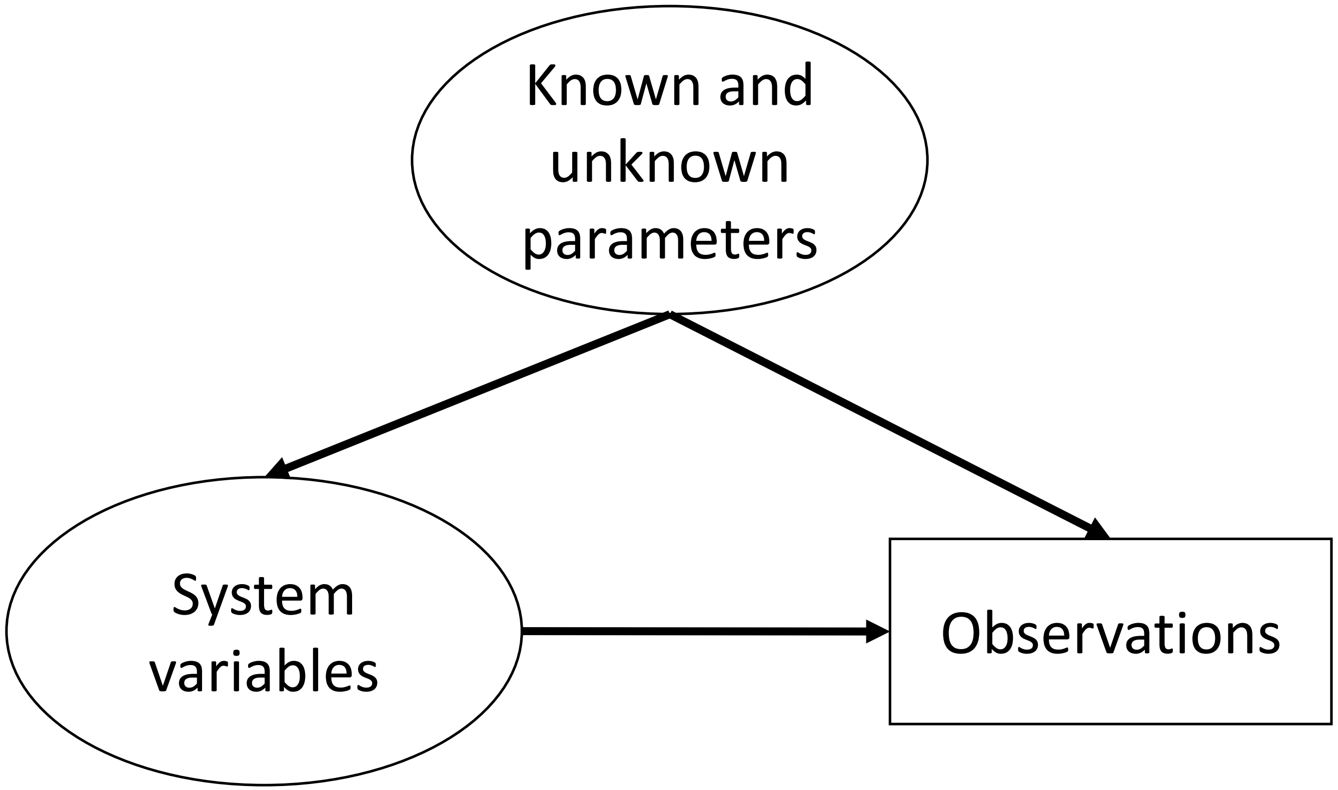


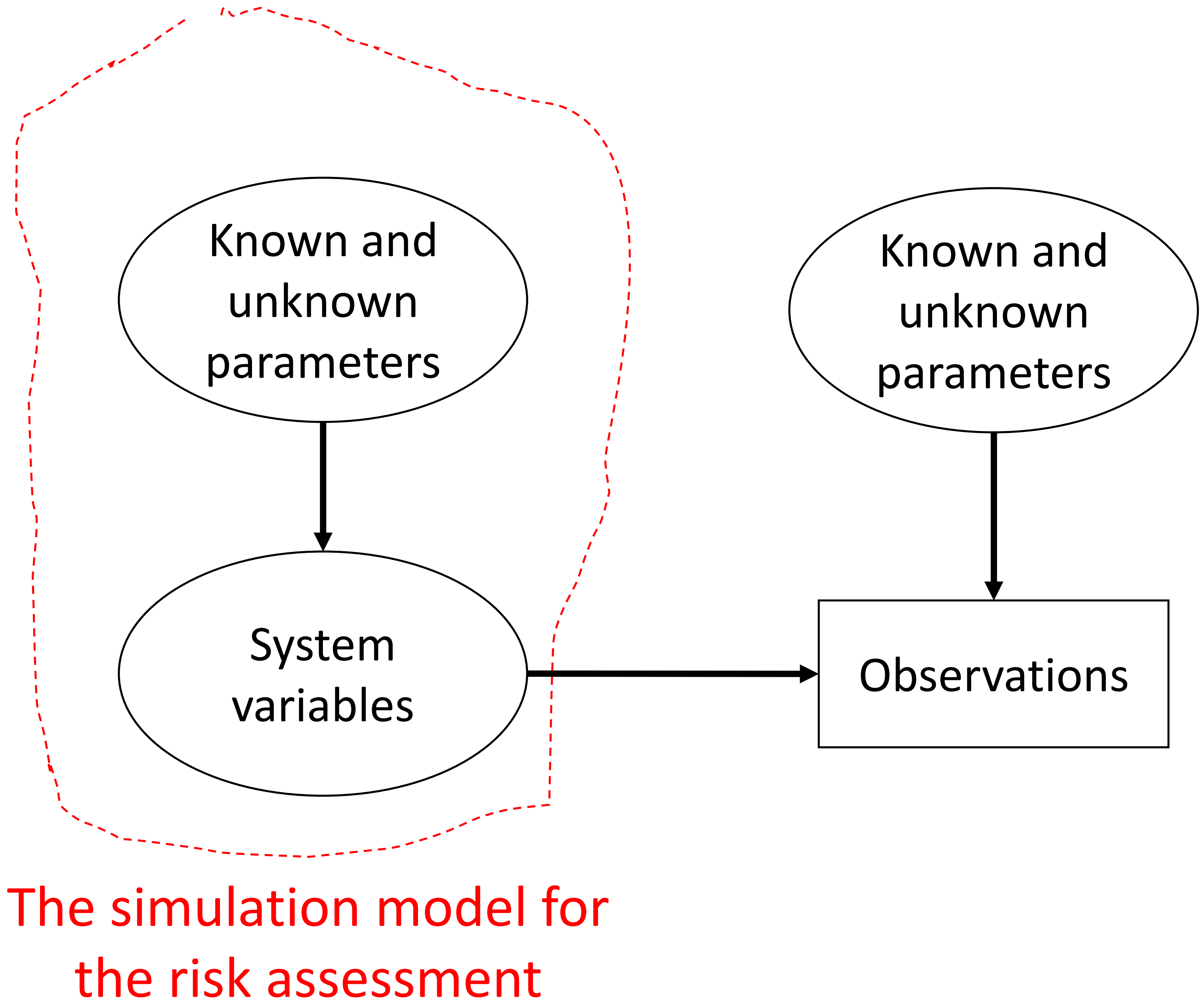
Bayesian Hierarhical Model



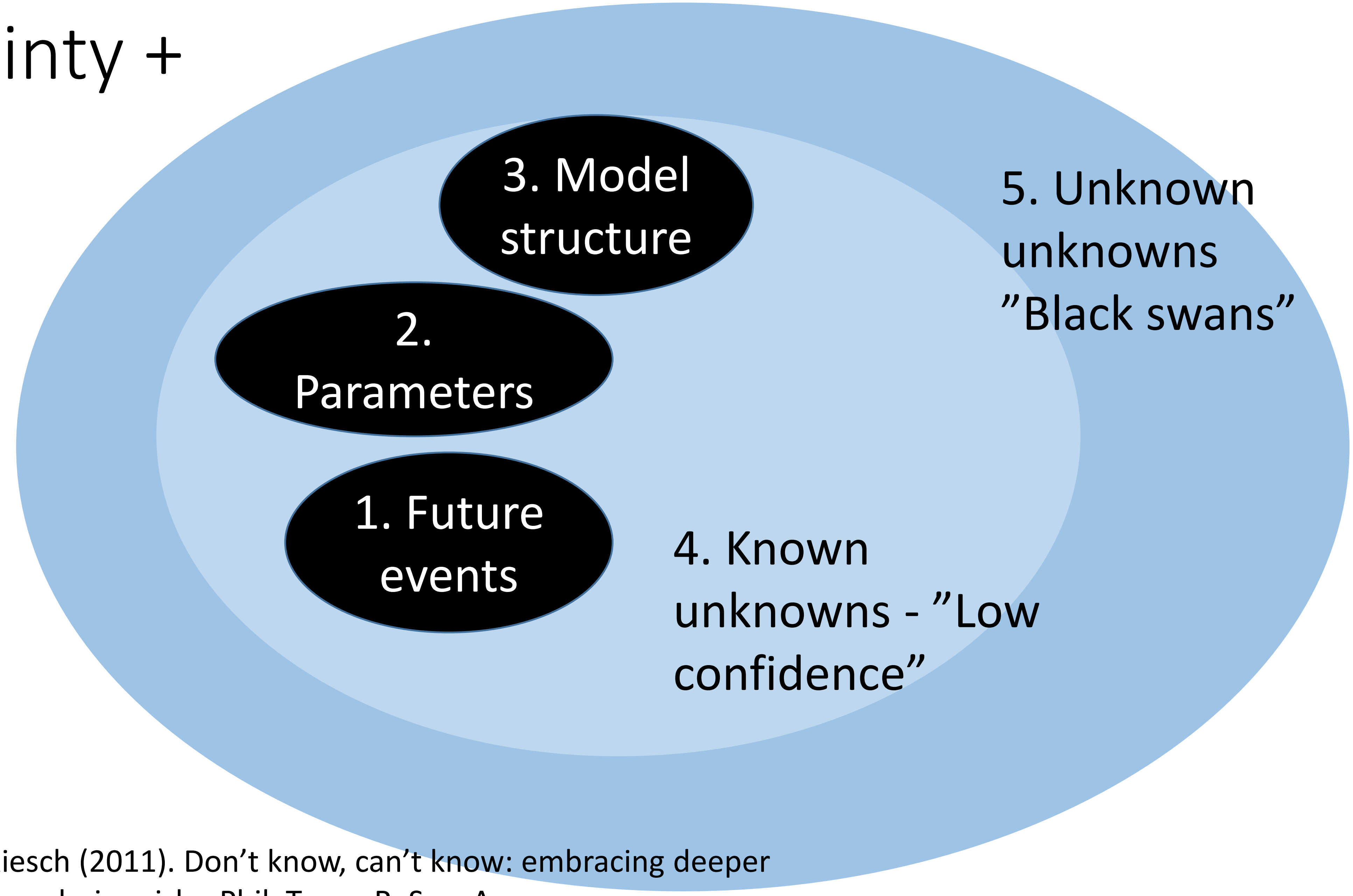
http://www.bias-project.org.uk

Uncertainty in Bayesian models




Known and unknown parameters
System variables
Known and unknown parameters
Observations
The simulation model for the risk assessment

Uncertainty +



Spiegelhalter and Riesch (2011). Don’t know, can’t know: embracing deeper uncertainties when analysing risks. Phil. Trans. R. Soc. A

Uncertainty analysis

https://www.efsa.europa.eu/en/topics/topic/uncertainty-assessment

Uncertainty analysis

1. Identify sources to uncertainty

Inputs

Propagation

Output

2. Assess individual sources to uncertainty

4. Assess relative contribution of sources of uncertainty

3. Assess combined impact of uncertainty on uncertainty in output

Methods to quantify uncertainty

Table 3: Summary evaluation of which methods can be applied to which types of assessment question (defined in Section 6), and provide which forms of uncertainty expression (defined in Section 4.1).

Method	Types of assessment question	Forms of uncertainty expression provided
Expert discussion	Quantitative and categorical	All
Expert Knowledge Elicitation (EKE)	Quantitative and categorical	All
Descriptive expression	Quantitative and categorical	Descriptive
Ordinal scales	Quantitative and categorical	Ordinal
Matrices	Quantitative and categorical	Ordinal
NUSAP	Quantitative and categorical	Ordinal
Uncertainty table for quantitative questions	Quantitative	Ordinal, range or range with probability
Uncertainty table for categorical questions	Categorical	Ordinal and distribution
Interval Analysis	Quantitative	Range
Confidence Intervals	Quantitative	Range with probability
The Bootstrap	Quantitative	Distribution
Bayesian Inference	Quantitative and categorical	Distribution
Probability Bounds Analysis	Quantitative and categorical	Bound with probability
Monte Carlo	Quantitative and categorical	Distribution
Approximate probability calculations	Quantitative	Distribution
Conservative assumptions	Quantitative	Bound or bound with probability
Sensitivity Analysis	Quantitative and categorical	Sensitivity of output to input uncertainty

Methods contribution to uncertainty analysis

Table 4: Summary evaluation of which methods can contribute to which steps of uncertainty analysis. Yes/No = yes, with limitations, No/Yes = no, but some indirect or partial contribution. Blank = no. Grey shading highlights the primary purpose(s) of each method. See Annex B for detailed evaluations.

Methods		**Steps of uncertainty analysis**					
	Qualitative or Quantitative	**Identify and list uncertainties**	**Select which to assess individually**	**Assess individual uncertainties**	**Combine uncertainties**	**Investigate influence**	**Describe unquantified uncertainties**
Expert discussion	both	Yes	Yes	Yes	Yes	Yes	Yes
Semi-formal Expert Knowledge Elicitation	both	Yes	Yes	Yes	Yes	Yes	
Formal Expert Knowledge Elicitation	both	Yes	Yes	Yes	Yes		
Typology	both	Yes					Yes
Descriptive expression	Quali			Yes	Yes	Yes	Yes
Ordinal scales	Quali			Yes	Yes	No/Yes	Yes
Matrices	Quali				Yes	Yes/No	
NUSAP	Quali	Yes		Yes		Yes	Yes
Uncertainty table for quantitative questions	both			Yes	Yes	Yes	Yes
Uncertainty table for categorical questions	both	Yes		Yes	Yes	Yes	Yes
Interval Analysis	Quanti			Yes	Yes		
Confidence Intervals	Quanti			Yes			
The Bootstrap	Quanti			Yes	No/Yes		
Bayesian Inference	Quanti			Yes			
Probability Bounds Analysis	Quanti				Yes		
Monte Carlo	Quanti				Yes	Yes	
Approximate probability calculations	Quanti				Yes		
Conservative assumptions	Quanti			Yes	Yes		
Sensitivity Analysis	Quanti					Yes	

Performance criteria on the method to assess uncertainty

- Evidence of current acceptance
- Expertise needed to conduct
- Time needed
- Theoretical basis
- Degree/ extent of subjectivity
- Method of propagation
- Treatment of uncertainty and variability
- Meaning of output
- Transparency and reproducibility
- Ease of understanding for non-specialist

EFSA's uncertainty guidance (draft 2016)

Tools

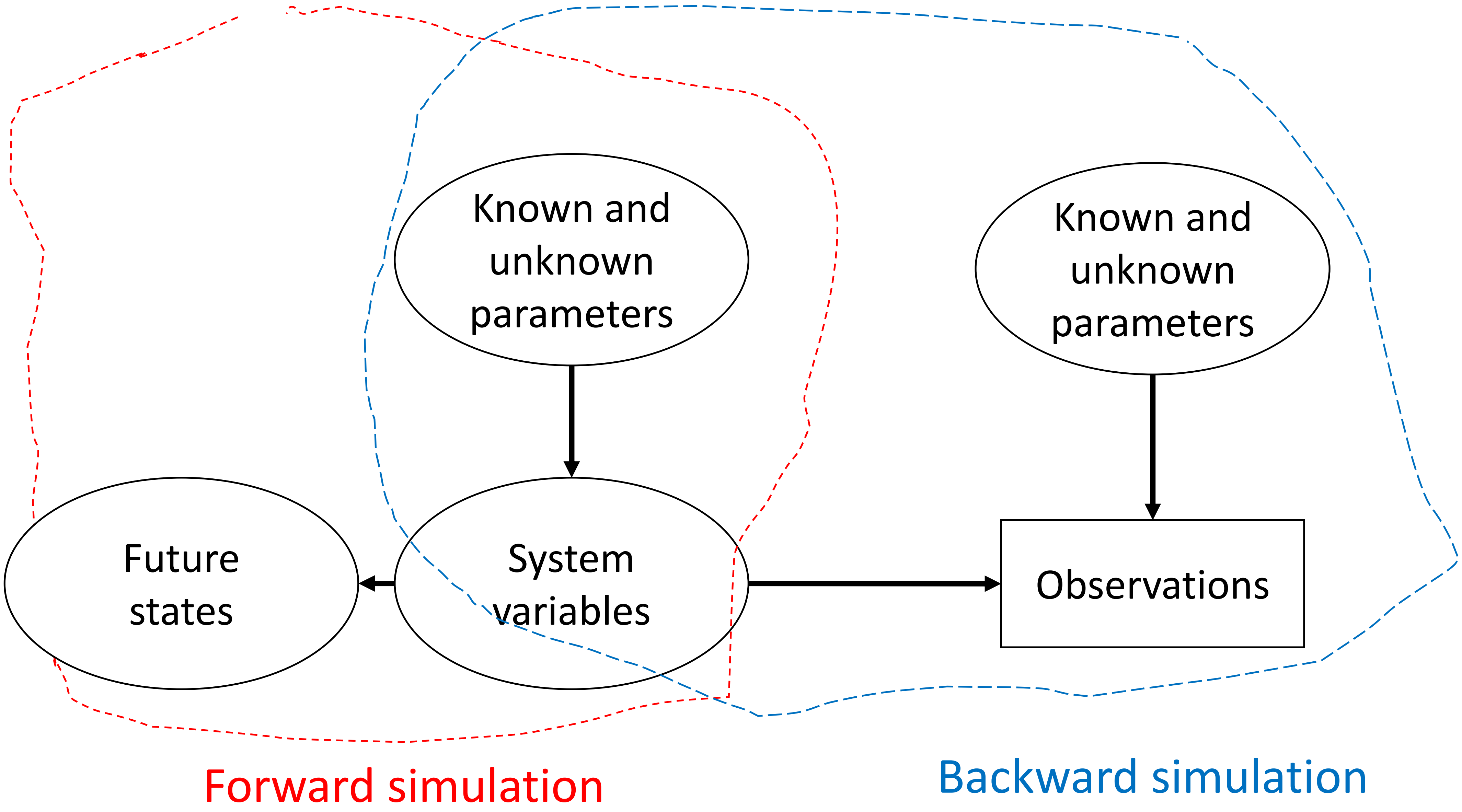
NUSAP – a semi-qualitative unc analysis

Table B.4.1: Example of NUSAP pedigree matrix for scoring parameter strength and influence.

	Strength				Effect
Score	**Proxy**	**Empirical basis**	**Methodological rigor**	**Validation**	**Influence on results**
4	Exact measure of the desired quantity (e.g. from the same geographical area)	Large sample, direct measurements (recent data, controlled experiments)	Best available practice (accredited method for sampling / diagnostic test)	Compared with independent measurements of the same variable (long domain, rigorous correction of errors)	
3	Good fit or measure (e.g. from another but representative area)	Small sample, direct measurements (less recent data, uncontrolled experiments, low non-response)	Reliable method (common within established discipline)	Compared with independent measurements of closely related variable (shorter time periods)	No or negligible impact on the results
2	Well correlated (e.g. large geographical differences , less representative)	Very small sample, modelled/derived data (indirect measurements, structured expert opinion)	Acceptable method (limited consensus on reliability)	Compared with measurements of non-independent variable (proxy variable, limited domain)	Little impact on the results
1	Weak correlation (e.g. very large geographical differences, low representativity)	One expert opinion, rule of thumb	Preliminary method (unknown reliability)	Weak, indirect validation	Moderate impact on the end result
0	Not clearly correlated	Crude speculation	No discernible rigor	No validation	Important impact on the end result

EFSA's uncertainty guidance (draft 2016)

Uncertainty in Bayesian models



MCMC sampling

- JAGS: Just Another Gibbs Sampler
- http://mcmc-jags.sourceforge.net/

- Can do ”Backward simulation” and forward simulation*

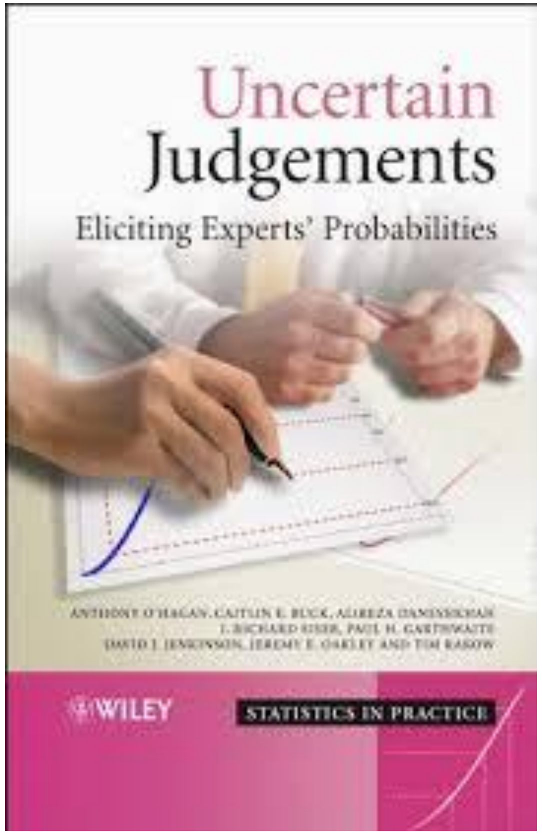
*Inside the MCMC sampling:
- Consider autocorrelation in samples, but takes longer time to run
- Not possible if the system model is very complex and
cannot be programmed in the JAGS language

- The exercises will be run from R using rjags

Expert's Knowledge Elicitation

- Aim to describe the Expert's Knowledge about one or more uncertain quantities in probabilistic form
- i.e. a joint probability distribution for the random variable in question

- EKE can be used to build priors distributions or prior predictive distributions



European Food Safety Authority

An Expert Knowledge Elicitation

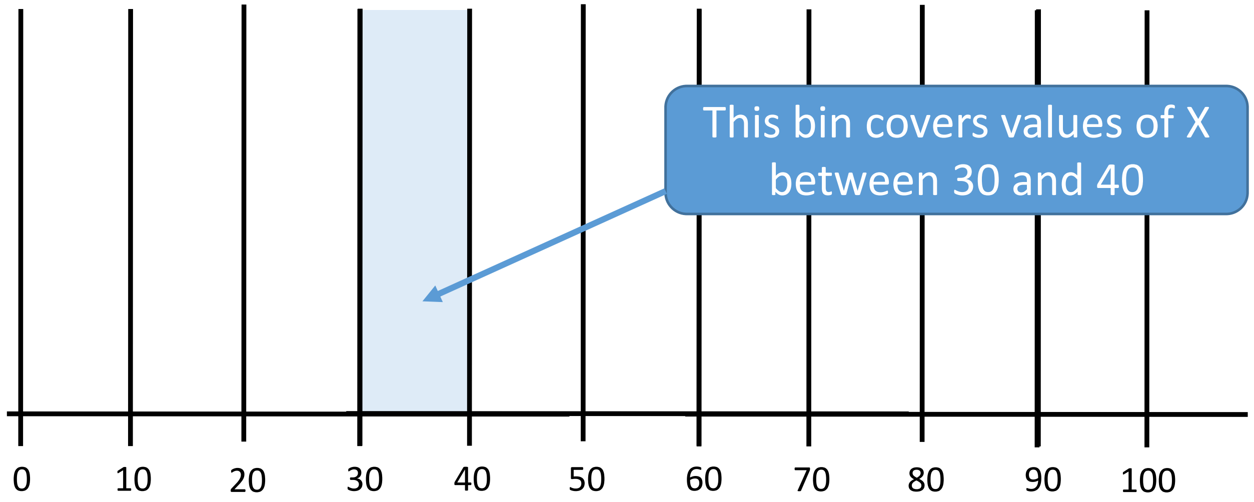
- Formulate the elicitation questions
- Ask experts about
  - Probabilities
  - Quantiles
  - Probability intervals
  - Moments or other descriptions of a probability distribution
- Fit and aggregate into a probability distribution for the uncertain quantity

Direct methods for EKE

- Simple and a bit crude
  - *Intervals* – Lower and Upper limits, then a Uniform distribution
  - *Triangular distributions* – Mode, Lower and Upper limits
- Cumulative Density Function (CDF)
  - *Quartiles* – 4 intervals, median and 25th and 75th percentiles
  - *Tertiles* – 3 intervals with equal probability
  - *Probabilities/Hybrid* – Choose probabilites and intervals
- Probability Density Function (PDF)
  - Mode/Mean, percentiles, shape,…
  - Place chips, draw it by hand…

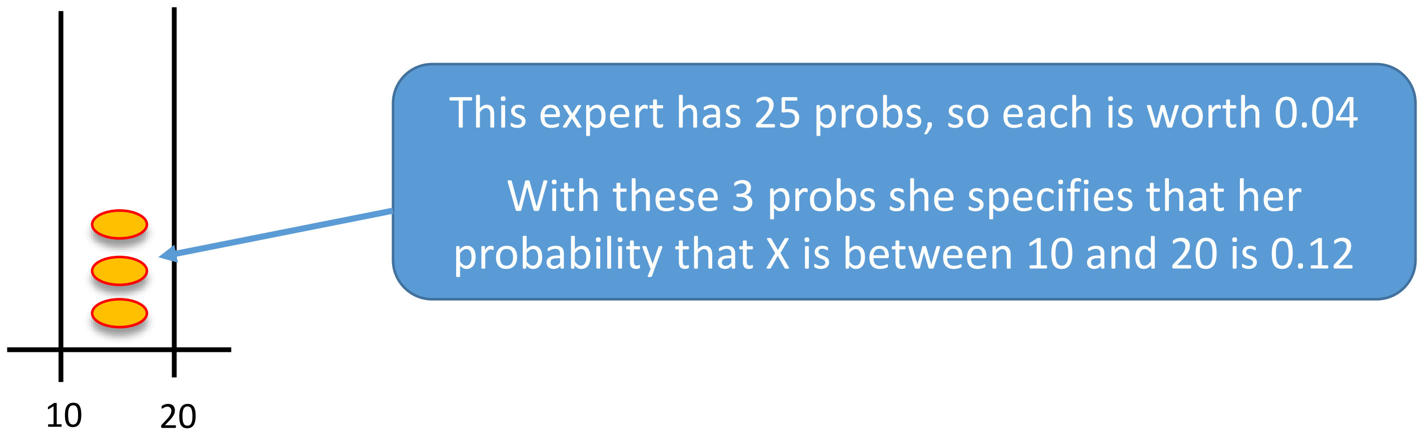
The roulette method – elicting a pdf

- You have a grid comprising a number of columns
  - These represent ranges of possible values of X
  - Called *bins*
- You have labelled the bin boundaries
  - Showing the range of values of X covered by each bin

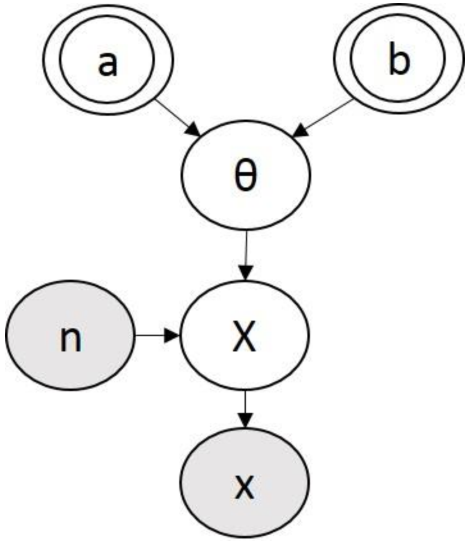


The roulette method – eliciting a pdf

- You also have a number of counters called *probs*
  - Because each one represents an amount of probability
  - For instance, if you have 20 probs, each represents a probability of 0.05 (5%)
- You are asked to place the probs in the bins on your grid, to specify your knowledge and beliefs about X



Indirect methods for EKE



- Equivalent Prior Sample (EPS)
  - *What is the expected frequency of the event?*
  - *What is the size a sample that you imagine to have behind this estimate?*

$$\frac{x}{n} = ? \qquad\qquad n = ?$$

- Hypothetical Future Sample (HFS)
  - *In a future sample of size 100 – in how many times has the event occured?*

$$n = 100 \qquad\qquad x = ?$$

We want the expert's uncertainty!

Wrong	*Right*
Your judgement of the probability that …	Your probability that …
Write down the median for …	Write down your median value for …
The probability of the quantity being below the median is 0.5	Your personal probability of the quantity being below your median value is 0.5
There should be a 50% chance that the quantity lies between the lower and upper quartiles	You should give 50% probability to the quantity lying between your lower and upper quartiles

- Never refer to a relative frequency or a proportion as a probability.

Selection of Structured EKE Software

- EXCALIBUR (EXpert CALIBration): www.lighttwist.net/wp/excalibur
- ElicitN: www.downloadcollection.com/elicitn.htm
- SHELF (The SHeffield ELicitation Framework): www.tonyohagan.co.uk/shelf/
- MATCH Uncertainty Elicitation Tool: optics.eee.nottingham.ac.uk/match/uncertainty.php#
- UncertWeb - The Elicitator: http://elicitator.uncertweb.org/
- Variogram elicitation: www.variogramelicitation.org
- Unicorn: www.lighttwist.net/wp/unicorn-download

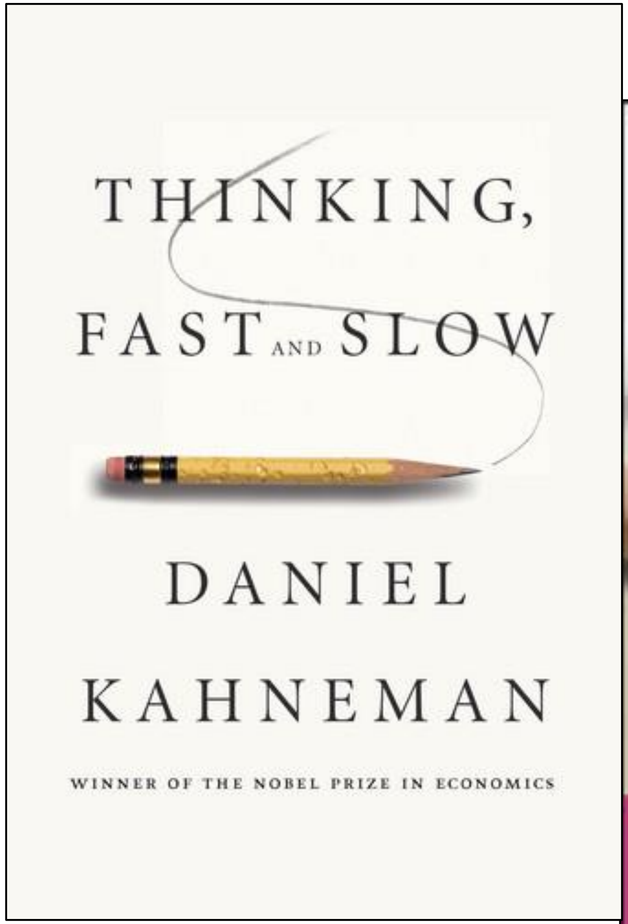
Source: http://www.expertsinuncertainty.net/Software/tabid/4149/Default.aspx

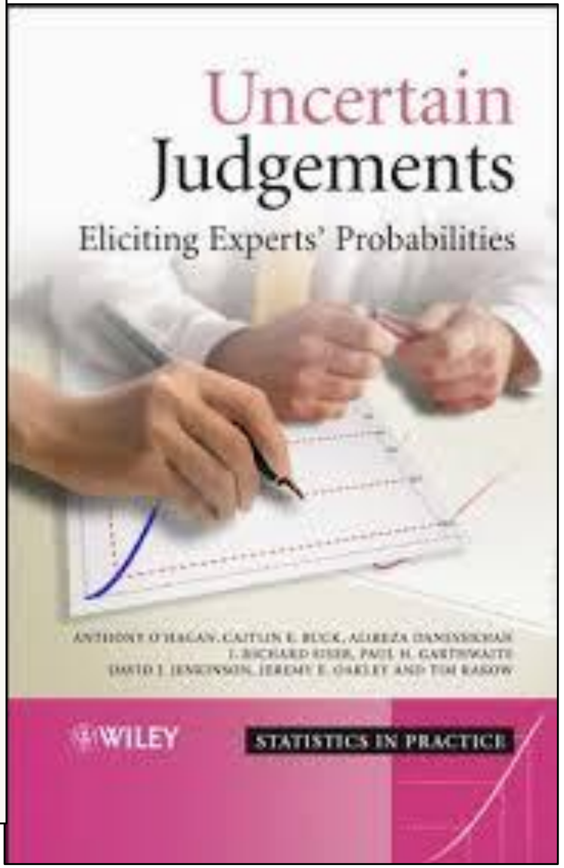
Expert Knowledge Elicitation

- O' Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. E., Garthwaite, P. H., Jenkinson, D. J., Oakley, J. E. and Rakow, T. (2006). Uncertain Judgements: Eliciting Expert Probabilities. Chichester: Wiley
- https://www.efsa.europa.eu/en/efsajournal/pub/3734
- SHELF package: **Tools to Support the Sheffield Elicitation Framework**
- http://optics.eee.nottingham.ac.uk/match/uncertainty.php
- David E. Morris, Jeremy E. Oakley, John A. Crowe, A web-based tool for eliciting probability distributions from experts, Environmental Modelling & Software, Volume 52, February 2014, Pages 1-4, ISSN 1364-8152, http://dx.doi.org/10.1016/j.envsoft.2013.10.010.

Pshycological factors and elicitation

- Anchoring and adjustment
- Availability
- Range–frequency compromise
- Representativeness and baseline neglect
- Conjuction fallacy
- The law of small numbers
- Overconfidence





Elicitation with multiple experts

- More psychological factors when working with several experts

- Behavioural aggregation
  - Group elicitation
  - One or several iterations, individually and in group
- Mathematical aggregation
  - Treat each expert’s distribution as data and update the decision maker’s belief
  - Pooled opinions – linear or logarithmic pooling
  - Calibrate experts and weight according to their performance

Alternative protocols for EKE

- the Sheffield protocol with group interaction of experts, consensus distributions
- the Cooke protocol with use of seed questions for the calibration of experts, no interaction
- the Delphi protocol on written expert elicitation with feedback loops, anonymous sharing of the results between iterations

Exercises

- Probabilistic risk assessment where we quantify uncertainty in input parameters of an assessment model using different sources of information and where uncertainty in assessment output matters
- thebus.R
- A simple risk classification problem under sparse information and several experts with differing judgements
- riskornot.R
- https://github.com/Ullrika/TBDL2017

The probability of missing the bus

- I believe the last bus is expected to arrive in 10 minutes (B time the bus leaves in minutes)
- The distance to the bus stop is 1 km (d km)
- Walking speed (SP minutes per km)
- When should I leave to make sure the risk of missing the bus is at an acceptable level, say 10%?

We will use:

- **Expert judgement on your walking speed**
- **Observations on your walking speed**

$$Risk = P(SP \cdot d > B)$$

$$B \sim Exp\left(\frac{1}{10}\right)$$

Is there something out there?

- You, me, experts have different judgments if there is something out there
- We make two observations, but these are not perfect
- We did not see anything
- Is there something out there?