https://github.com/Ullrika/quantifying_uncertainty_by_probability

Quantifying uncertainty by probability

This tutorial will cover methods to quantify uncertainty by probability such as Expert Knowledge Elicitation and Bayesian analysis.

We will go through hands on examples on quantifying uncertainty in assessment inputs as well as in assessment output.

All examples will be done using code from the open source program R using the participants own laptops.

It is possible to just sit in and interact without running the code.

Rare events risk analysis – import risk analysis

- K items from a product are imported
- An item is infected with a certain probability: $P(X = 1) = p$, $P(X = 0) = 1-p$
- Items are infected independent of each other
- Number of infected items out of K is $Xall \sim Bin(K, p)$
- The custom make an inspection and randomly select k items:
  - Product A: 0, 0, 0, 0
  - Product B: 0, 1
- What is the probability that more than 10% of the imported items are infected i.e. $P(Xall>K/10)$?

Exceeding a threshold risk analysis - missing the last bus home

- Time for you to walk to the bus station: speed*distance
- P(miss the bus) = P(time < speed*distance)
- The distance to the bus stop is 500 m
- Walking speed is according to Wikipedia about 1.4 m/s = 1.4·60 m/min
- When should I leave to make sure the risk of missing the bus is at an acceptable level, say 5%?

A. Your believe the last bus is expected to arrive in 5 minutes
B. Observations of the time when the bus leaves are 13, 2, 5, 10, 8 min and observations of your walking speed is 1.234, 1.2, 1.4 m/s

Why this workshop?

Why is ”quantification of uncertainty by probability” important

- A desire to communicate uncertainty - creates a need to describe uncertainty
- In 10 years any scientific expert and assessor involved in producing decision support needs to be able to describe and communicate uncertainty! [my belief]
- Principles to quantify uncertainty by probability may enrich your methods for research to better acount for data quality, model complexity, analysis flexibility and make predictions with uncertainty
- Probabilistic treatment of uncertainty is widely applicable and should be part of a standard curriculum

Why communicate uncertainty

- Decision making involves uncertainty
  - in facts – what will happen when we make a choice?
  - in values – what do we want when we cannot have everything?

- Unless uncertainty is known
  - a DM can place too much confidence in experts and face unexpected problems
  - or a DM can place too little confidence in experts and miss opportunities and resources to collect information has been waisted

Fischhoff and Davis, 2014

When to communicate uncertainty

- Dont communicate with more or less detail than the DM need
- May require to communicate more things than are usual within a field, things that are assumed or ignored
- May require to communicate less about details academics like to discuss about
  - i.e. both simplifying and complicating normal scientific discurse
- ALSO – reduce to talk about uncertainty in association to decision-relevant elements
- All uncertainty must be uncovered

Fischhoff and Davis, 2014

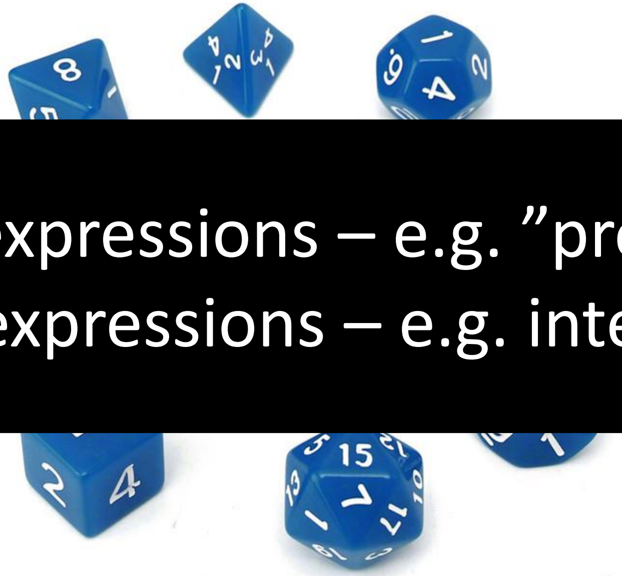
How to communicate

- Characterise uncertainty
- Assess uncertainty
- Convey uncertainty – create messages that afford DM the detail that their choices warrant
- Persuasive
  - when DM wants to change other DMs behaviour
  - shading or hiding uncertainty might be justified
- Non-persuasive
  - Goal to help people to make decisions that serve their own, self-defined best interests
  - Honesty is the only policy

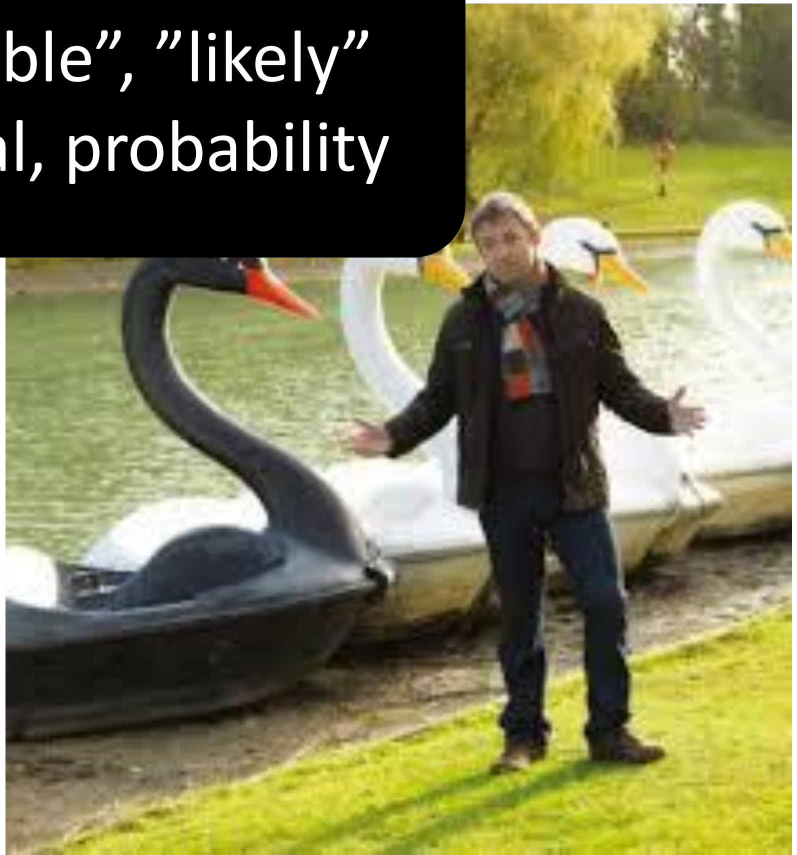
Fischhoff and Davis, 2014

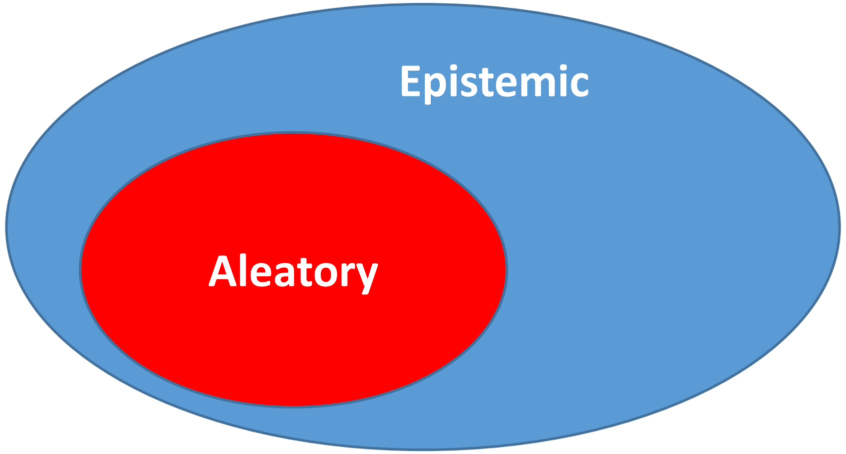
What is uncertainty?

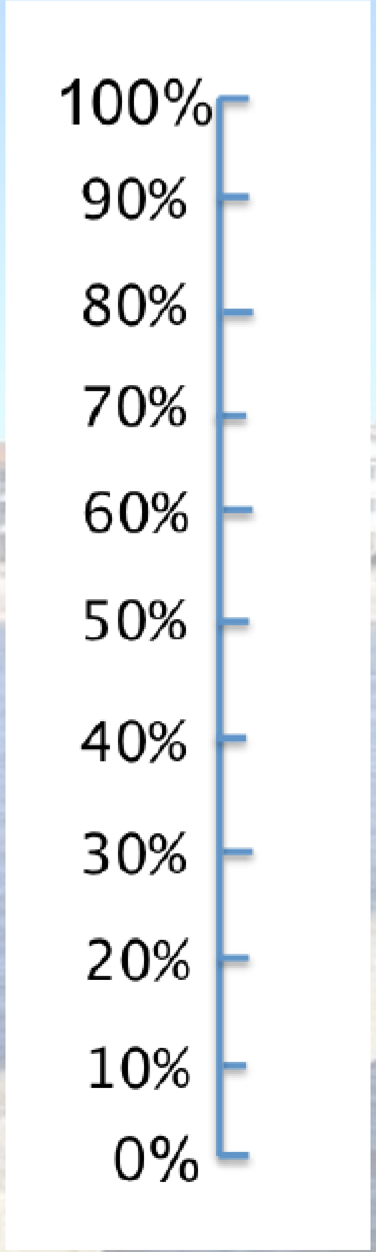
Different kinds of uncertainty



Qualitative expressions – e.g. ”probable”, ”likely”
Quantiative expressions – e.g. interval, probability





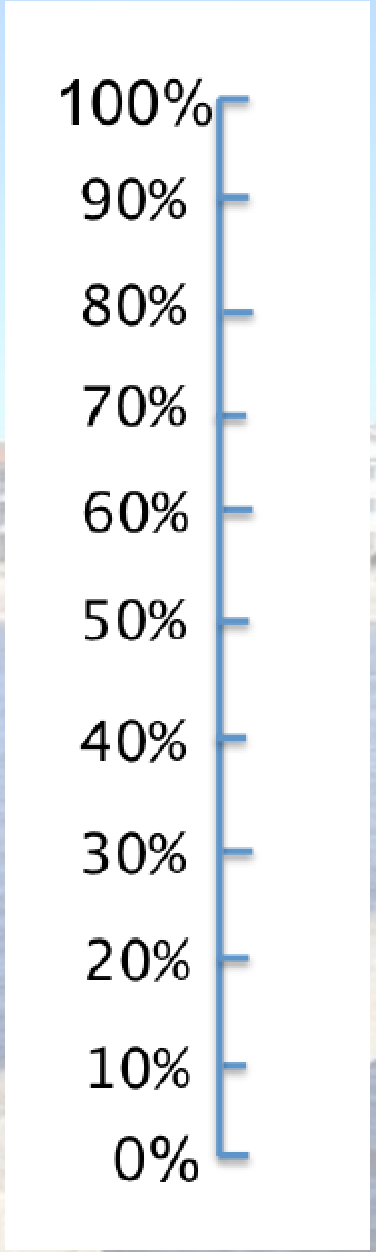


Possible

Likely

Unlikely

Highly probable

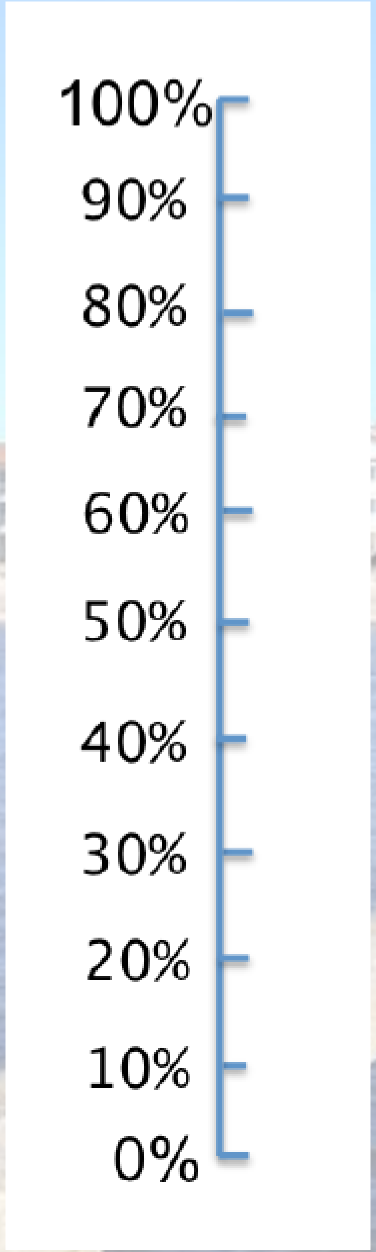


Highly probable

Likely

Possible

Unlikely



Highly probable

Likely

Possible

Unlikely

Table 1. Likelihood Scale	
Term*	**Likelihood of the Outcome**
Virtually certain	99-100% probability
Very likely	90-100% probability
Likely	66-100% probability
About as likely as not	33 to 66% probability
Unlikely	0-33% probability
Very unlikely	0-10% probability
Exceptionally unlikely	0-1% probability

Qualitative - Quantitative

Hur stor är sannolikheten att du rekommenderar oss till en vän/kollega?

Mycket liten — Mycket stor

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

Kommentar: ______________________

____

____

○ Privatgäst

○ Konferensgäst

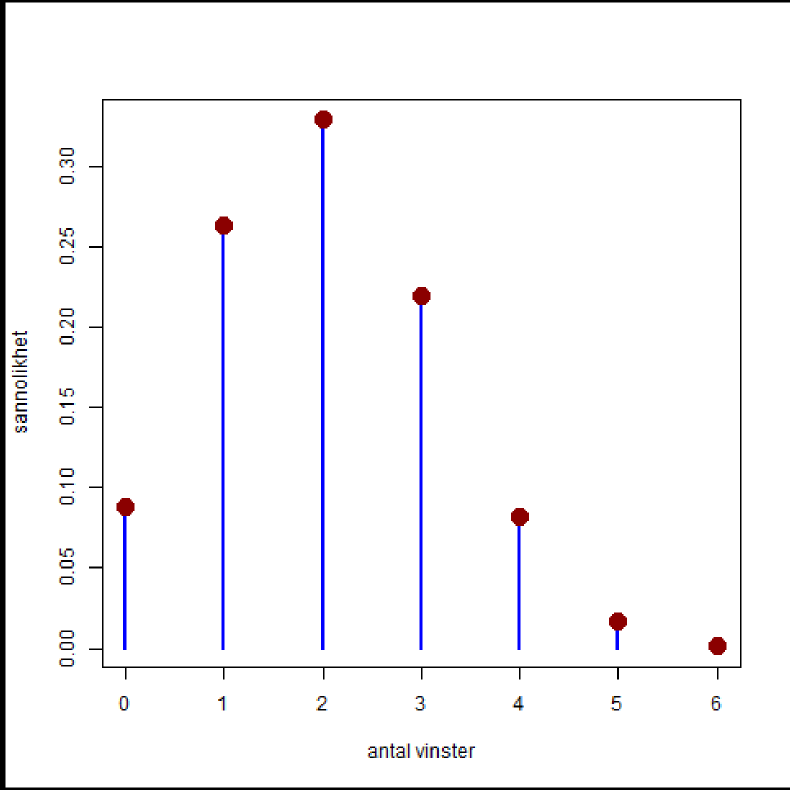
Epost: ______________________

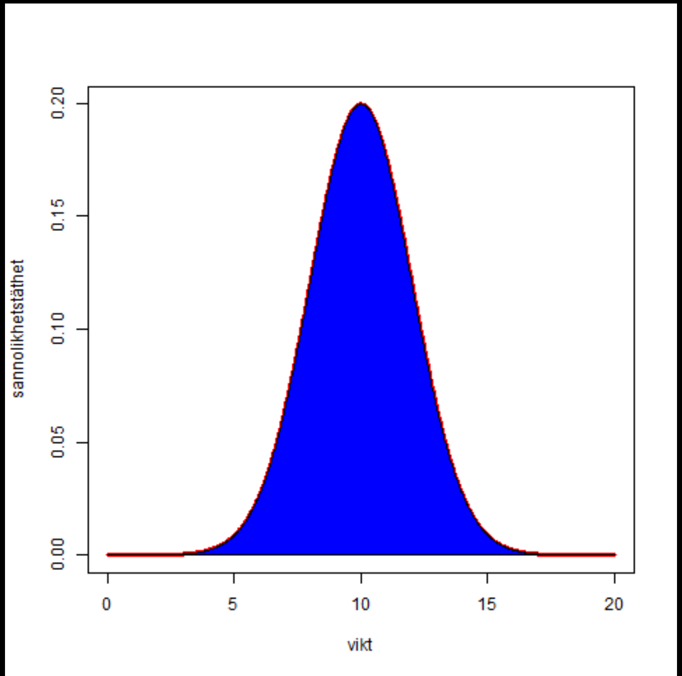
Rumsnummer: ______________________

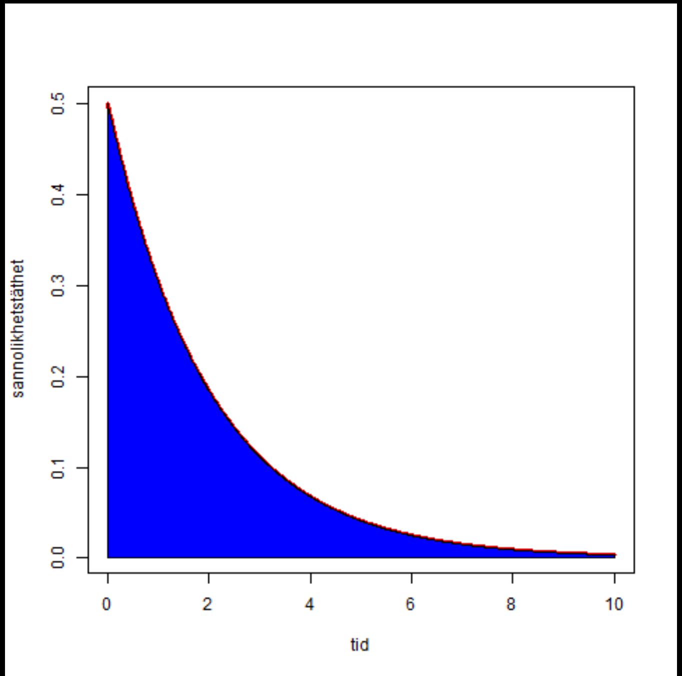
FALKENBERG STRANDBAD

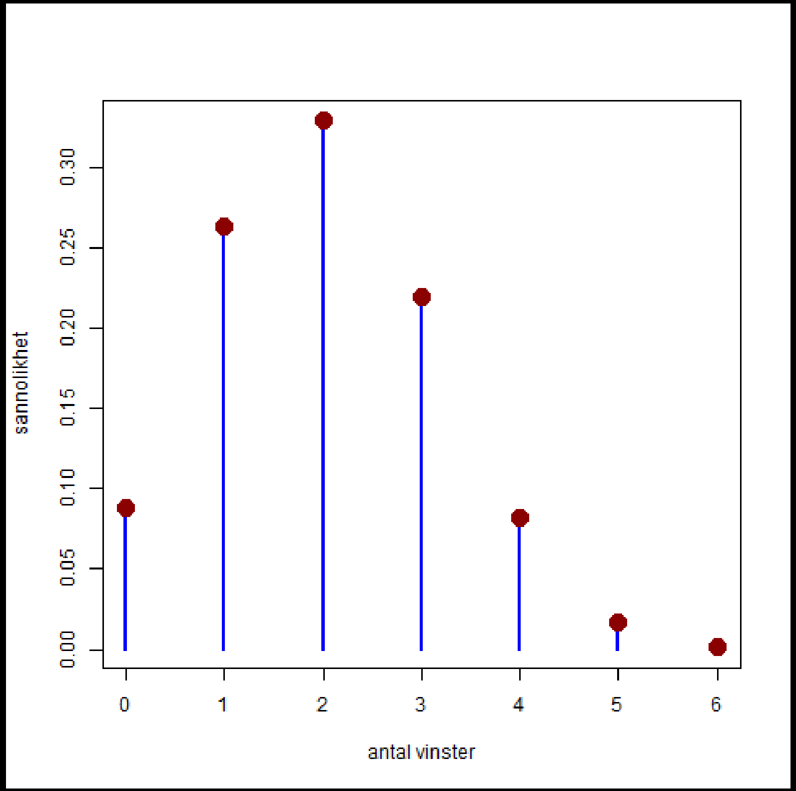
0346-71 49 00 | info@strandbaden.se | www.strandbaden.se

Semi-quantitative


sannolikhet
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0
1
2
3
4
5
6
antal vinster


sannolikhetstäthet
0.00
0.05
0.10
0.15
0.20
0
5
10
15
20
vikt


sannolikhetstäthet
0.0
0.1
0.2
0.3
0.4
0.5
0
2
4
6
8
10
tid


sannolikhet
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0
1
2
3
4
5
6
antal vinster

?

Quiz

- You will get questions with two possible answers: A or B
- You will answer and then state how certain you are in this answer

5	If you have no clue if the answer is A or B
6	If you think you know, but are not fully certain
7	
8	
9	
10	If you are certain that you are right

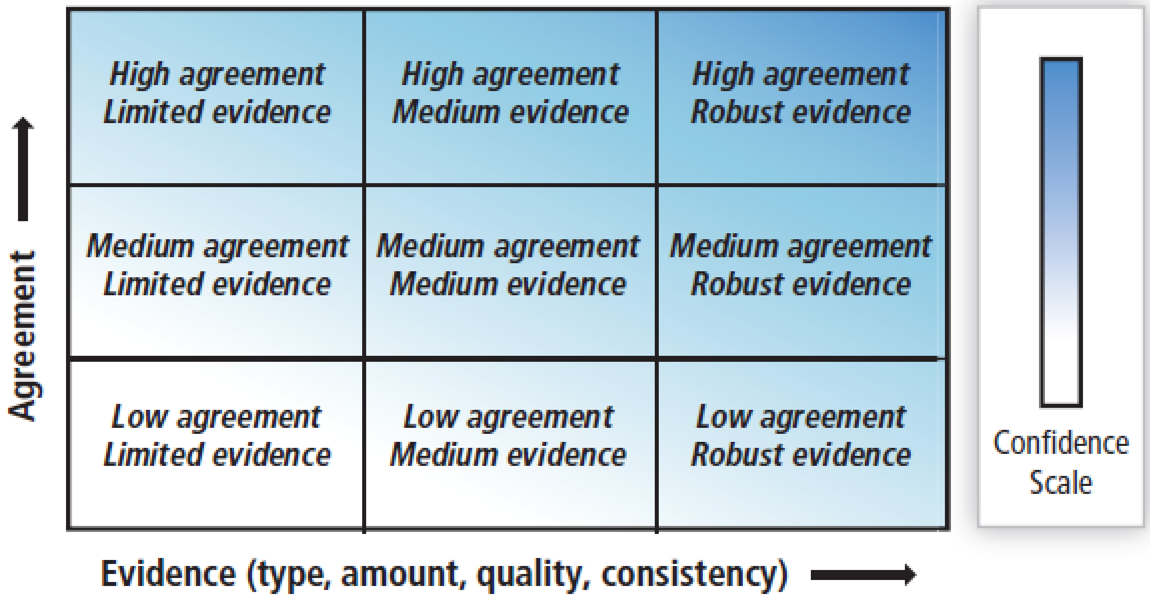
1. Daniel Kahneman was awarded the Nobel price in Economy in A) 1978 or B) 2002
2. The Society for Risk Analysis was founded A) 1980 or B) 1975
3. Accra is the capital of A) Kazakstan or B) Ghana
4. 1 USD is today worth A) less than 0.9 EUR or B) more than 0.9 EUR

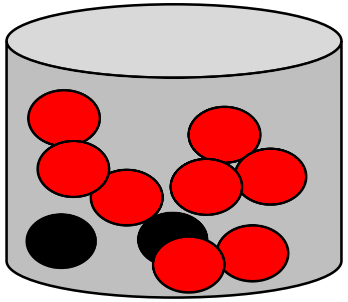
Your confidence	5	6	7	8	9	10
Score if you were right	0	9	16	21	24	25
Score if you were wrong	0	-11	-24	-39	-56	-75

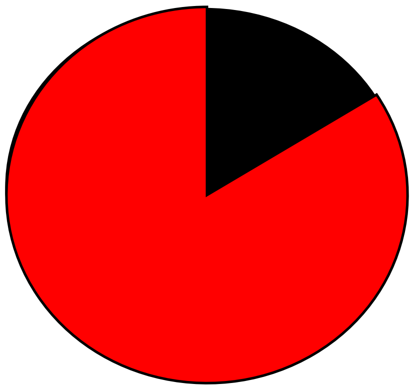
Confidence

Quantitatively calibrated levels of confidence.

Terminology	Degree of confidence in being correct
Very High confidence	At least 9 out of 10 chance of being correct
High confidence	About 8 out of 10 chance
Medium confidence	About 5 out of 10 chance
Low confidence	About 2 out of 10 chance
Very low confidence	Less than 1 out of 10 chance







Spiegelhalter's tweet
https://twitter.com/EFSA_EU/status/1042123576713900033

There is no single 'true' uncertainty

It is my, yours or our uncertainty

Don't apologize for being uncertain

Reclaim uncertainty – be confident about your uncertainty!

Different kinds of uncertainty and probability

P("get a 6")=p6

P(p6 > 1/6) = 0.2

P("our assessment is TRUE")

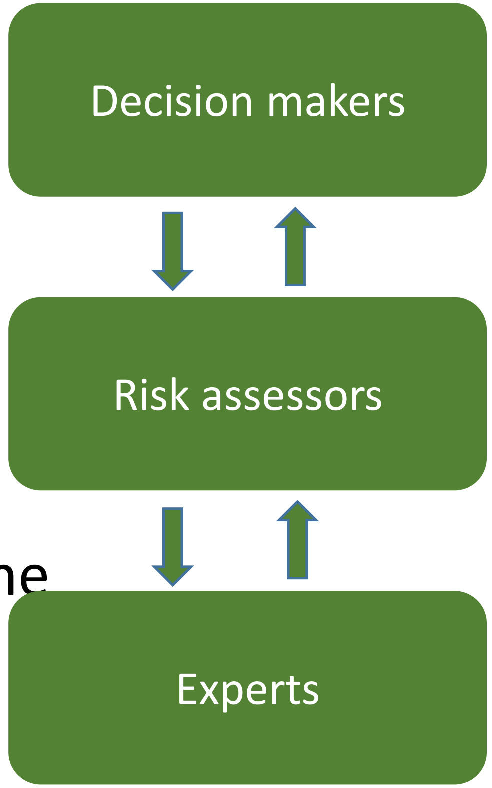
Aleatory and epistemic

Probability can be used as a measure of

Relative frequency, Subjective probability and even Confidence although these describe different types of uncertainty.

Who is responsible for addressing uncertainty?

- Decision-makers are responsible for resolving the impact of uncertainty on decisions – requires weighing the scientific assessment against other considerations (EFSA)
- The decision maker's asks:
  - What is the range of possible answers, and how probable are they?
  - Is further investigation needed?
  - What are the main sources of uncertainty affecting the outcome of the assessment? (helps communication with stakeholders)
- Assessors are responsible for characterizing uncertainty – requires scientific expertise (EFSA)



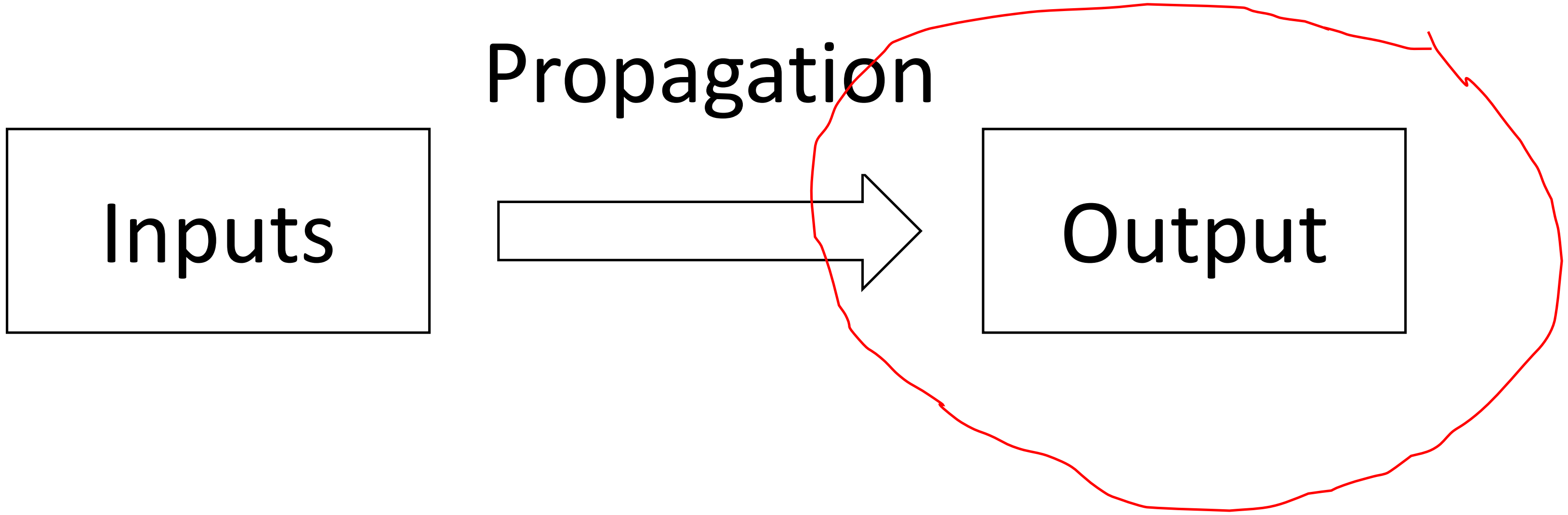
To the assessors

- **Combined uncertainty** should be **quantified**
  - By calculation or expert judgment
- It is **not necessary** to characterise all uncertainties individually
- It is **necessary** to characterise the combined uncertainty
  - This is what matters for decision-making
- Guidance recognises assessors may be unable to quantify some uncertainties
  - Identify and describe these

No need to quantify all sources to uncertainty

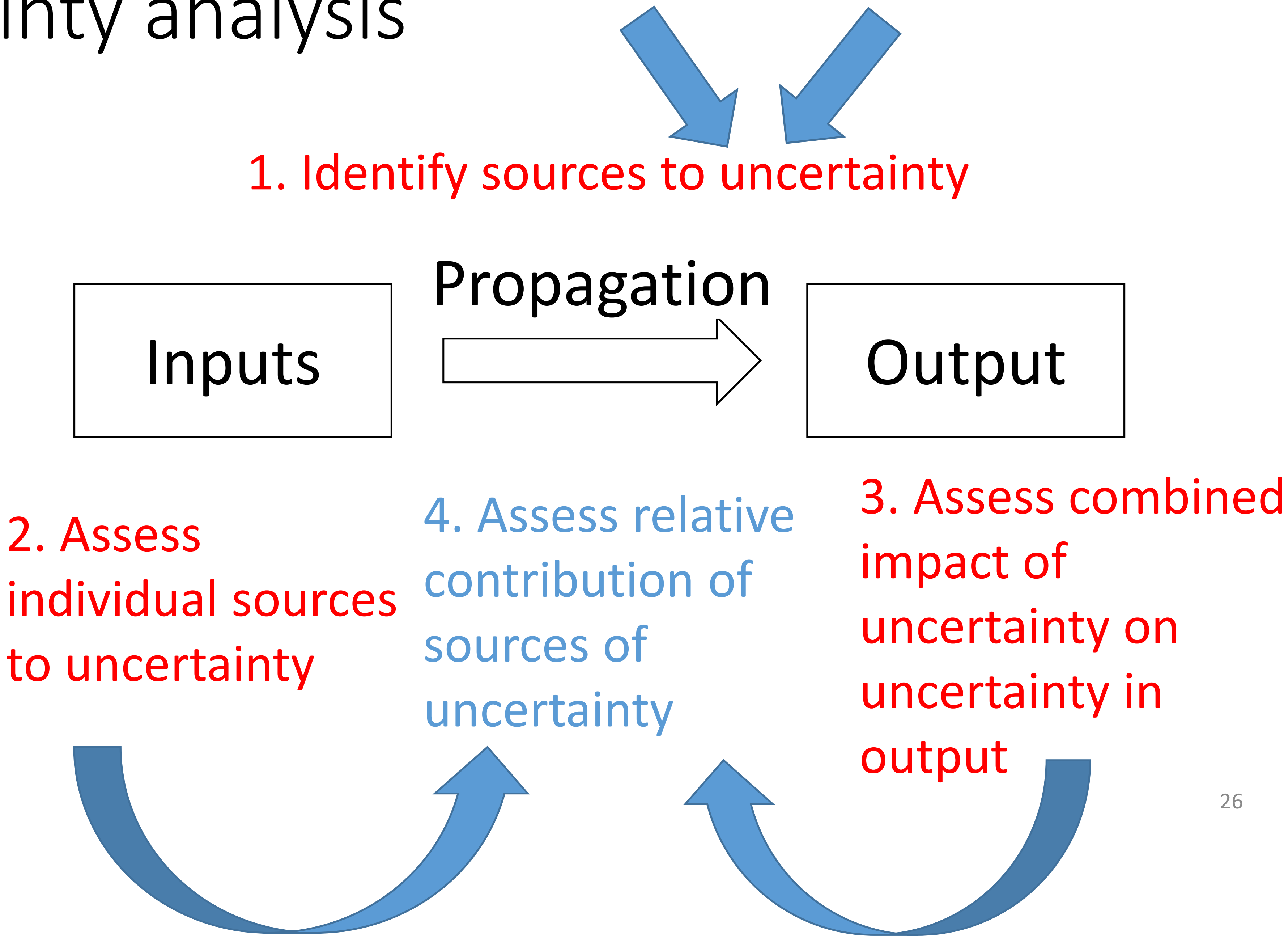
EFSA

Uncertainty analysis



Most important for
decision makers!

Uncertainty analysis



Risk and uncertainty analysis

- Purpose to evalute the impact of uncertainty on objectives (ISO, EFSA)
- Separate risk and uncertainty as concepts and how they are expressed
- There are different methods to express uncertainty
- Advantages with quantitative methods - objective, able to cover both inputs and outputs of an assessment in a transparent way
- Advantages with qualitative methods – all things should not be quantified

Methods for uncertainty analysis

- Descriptive expression
- Ordinal scales
- Matrices
- Conservative assumptions
- Interval Analysis
- Probability Bounds Analysis
- Expert knowledge elicitation
- Monte Carlo simulation (1D, 2D)
- Confidence Intervals
- The Bootstrap
- Bayesian inference (i.e. full probability distributions which can give probability intervals)
- Bayesian modelling
- Sensitivity analysis (one at time, global SA)
- NUSAP (qualitative sensitivity analysis on a complex model)

Methods for uncertainty analysis

Uses probability to quantify epistemic uncertainty

- Descriptive expression
- Ordinal scales
- Matrices
- Conservative assumptions
- Interval Analysis
- Probability Bounds Analysis
- Expert knowledge elicitation
- Monte Carlo simulation (1D, 2D)

Uses bounds on probability to quantify epistemic uncertainty

- Confidence Intervals
- The Bootstrap
- Bayesian inference (i.e. full probability distributions which can give probability intervals)
- Bayesian modelling
- Sensitivity analysis (one at time, global SA)
- NUSAP (qualitative sensitivity analysis on a complex model)

Two methods to quantify uncertainty

- Expert Knowledge Elicitation
  - Extract probability distributions from experts

- Bayesian Analysis
  - Use probability distributions to describe uncertainty
  - Coherent principle of inference to learn from data
  - Revise probability distributions integrating expert knowledge and data

Bayesian analysis

Bayesian data analysis

Bayesian inference

Bayesian updating

Bayesian learning

Bayesian statistics

Bayesian networks

Bayesian modelling

Bayesian emulation

Expert judgement

To put it simple - any use of Bayesian probabilities for:

- Probabilistic causal modelling

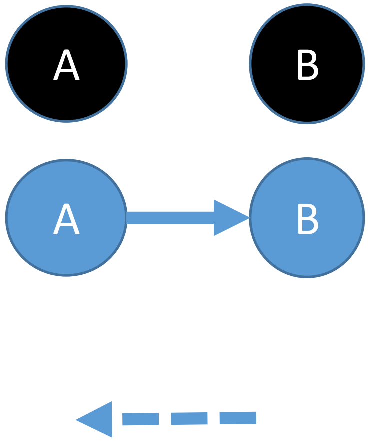
or

- Quantification of uncertainty

$$P(A \,\&\, B) = P(A)P(B)$$

$$P(A \,\&\, B) = P(B|A)P(A)$$

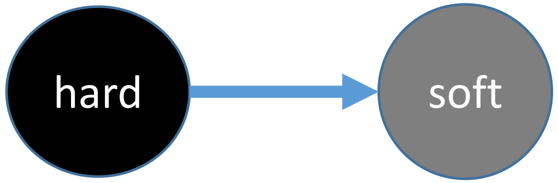
$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

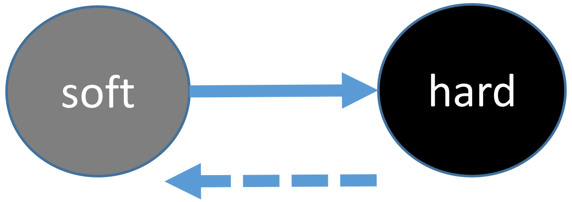


Types of inference

- *Inference* – reasoning from factual knowledge or evidence
- *Statistical inference* – making conclusions about a true probability distribution
- *Parametric inference* – making conclusions about probability distributions of parameters
- *Predictive inference* – making conclusions about probability distributions of future states or future observations

Types of calculations/simulations





- Propagation
- Simple probabilistic calculations

- Forward sampling

e.g. Monte Carlo simulation

MCMC for sampling extremes

- Bayesian updating
- Analytical derivation simple with conjugacy properties (parametric distribution of posterior is known and easy to compute)
- Backward sampling

e.g. Markov Chain Monte Carlo simulation and Approximate Bayesian Computation

I will draw these things on the board

- A is the event of interest
- We want to know $P(A)$
- In general we always condition on what we know $P(A|\text{Knowledge})$
- We can formalise this into a model
- $P(A|\theta)$ We simply the world into a set of parameters θ
- If uncertainty is important we should also consider $P(\theta)$ ”prior”
- We are then interested in $P(A|\theta)P(\theta)$
- We want to incorportate data
- In order to learn from data we need to specify a model linking data to the parameters of the model $P(\text{data}|\theta)$ ”likelihood”
- How to get $P(\theta|\text{data})$? ”Bayesian updating”
- $P(\theta|\text{data})$ ”parametric inference”
- $P(A|\text{data}) = P(A|\theta)\, P(\theta|\text{data})$ ”preditive inference”

EKE

Bayesian analysis

Expert Knowledge Elicitation

Now you are going to be experts in a session

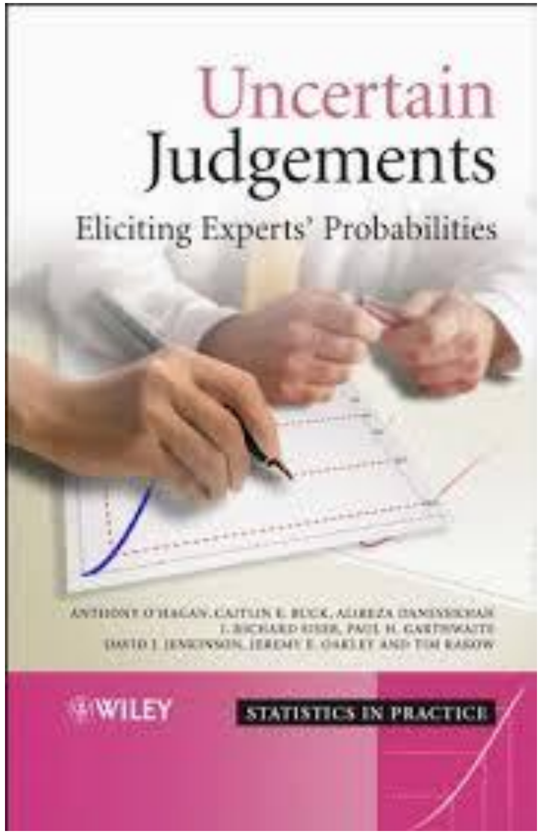
- https://goo.gl/forms/UFoXTLgcAOuRZGX23

- Answer the questions as honestly as possible without looking at the answer.

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



An Expert Knowledge Elicitation

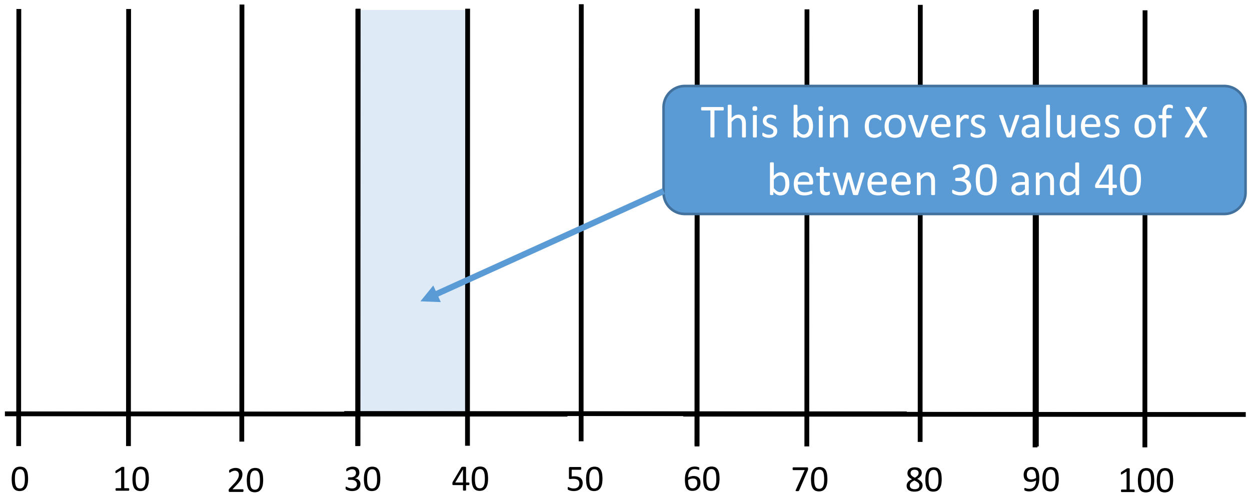
- Formulate the elicitation questions
- Ask experts about
  - Probabilities
  - Quantiles
  - Probability intervals
  - Moments or other descriptions of a probability distribution
- Fit and aggregate into a probability distribution for the uncertain quantity

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…

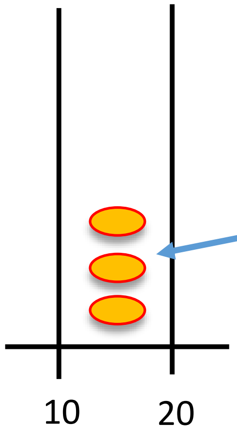
The roulette method – eliciting a pdf

- 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



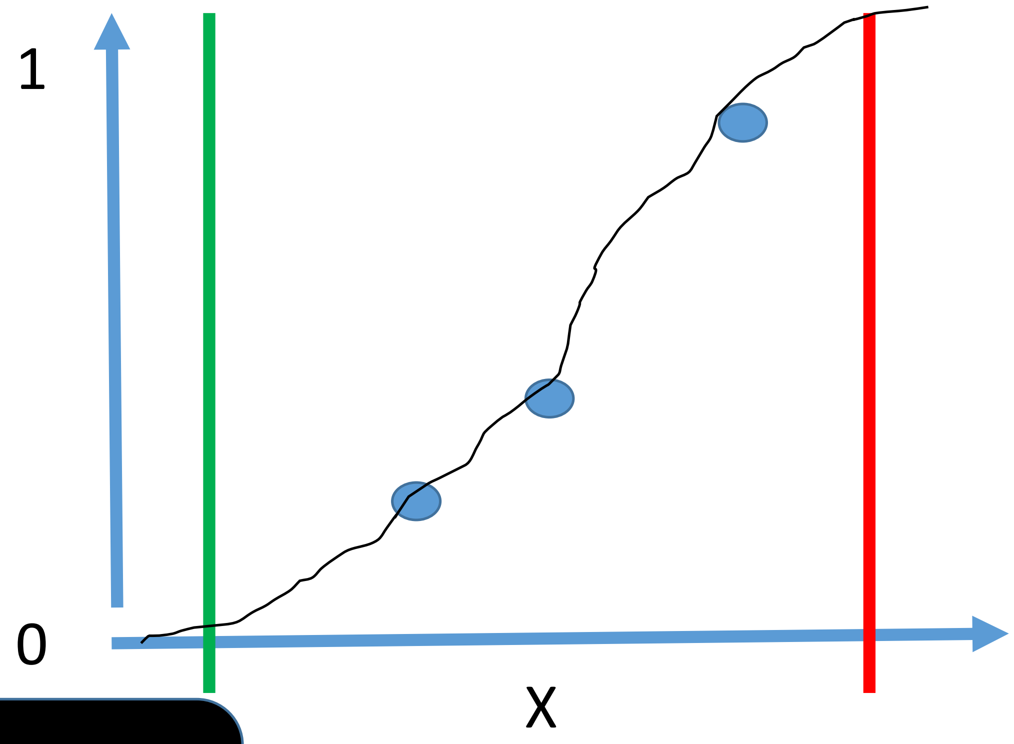
This expert has 25 probs, so each is worth 0.04

With these 3 probs she specifies that her probability that X is between 10 and 20 is 0.12

- The code is at https://github.com/Ullrika/quantifying_uncertainty_by_probability

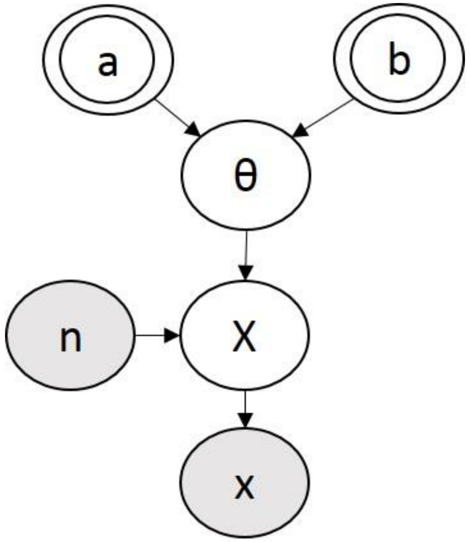
The time it takes to travel by car from Oslo to Stavanger (in hours) – test direct elicitation of the CDF and PDF

- X denotes the time
- Lowest possible value of X is ם
- Highest possible value of X is ם
- Assign a set of points from the CDF-curve: P(X < ם) = ם



Run shelf in R!
CDF: ej_cdf = elicit()
PDF: ej_pdf = roulette(lower = ם , upper = ם)

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?*

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

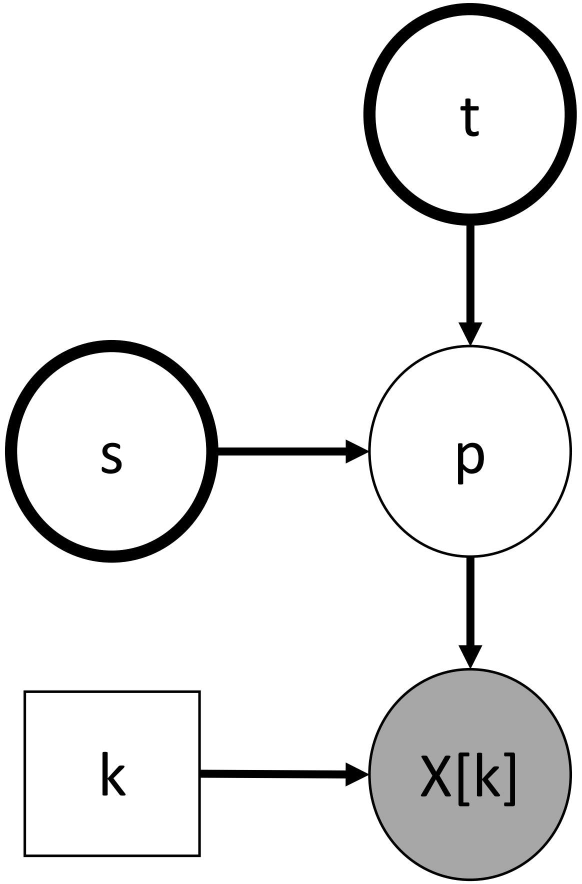
I used an indirect method to elicit your uncertainty in being delayed

Bayesian analysis

Rare event risk analysis

Import risk analysis

Beta-binomial model



prior

~beta(st, s(1-t))

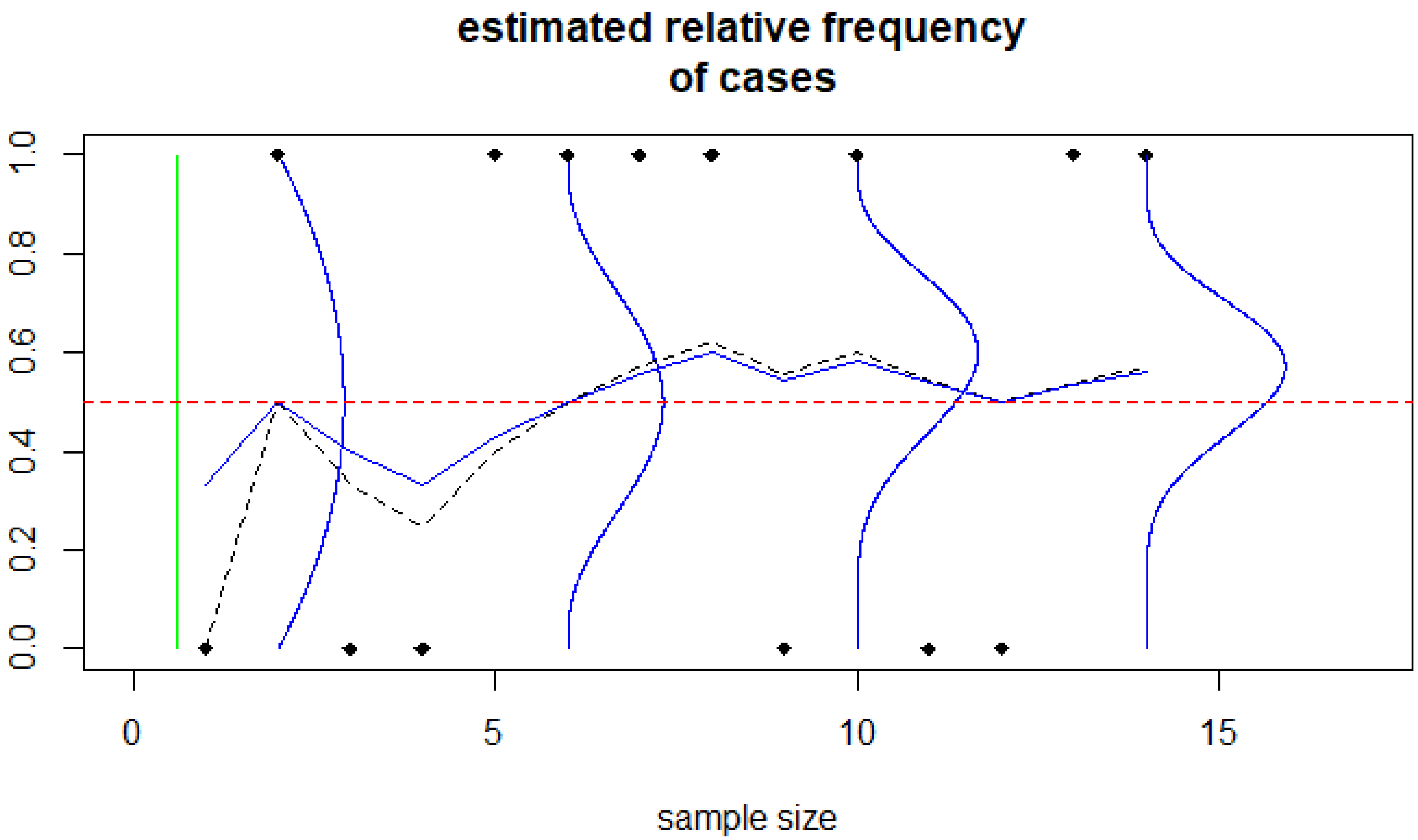
~binomial(k, p)

posterior

~beta(st + X[k], s(1-t) + k – X[k])

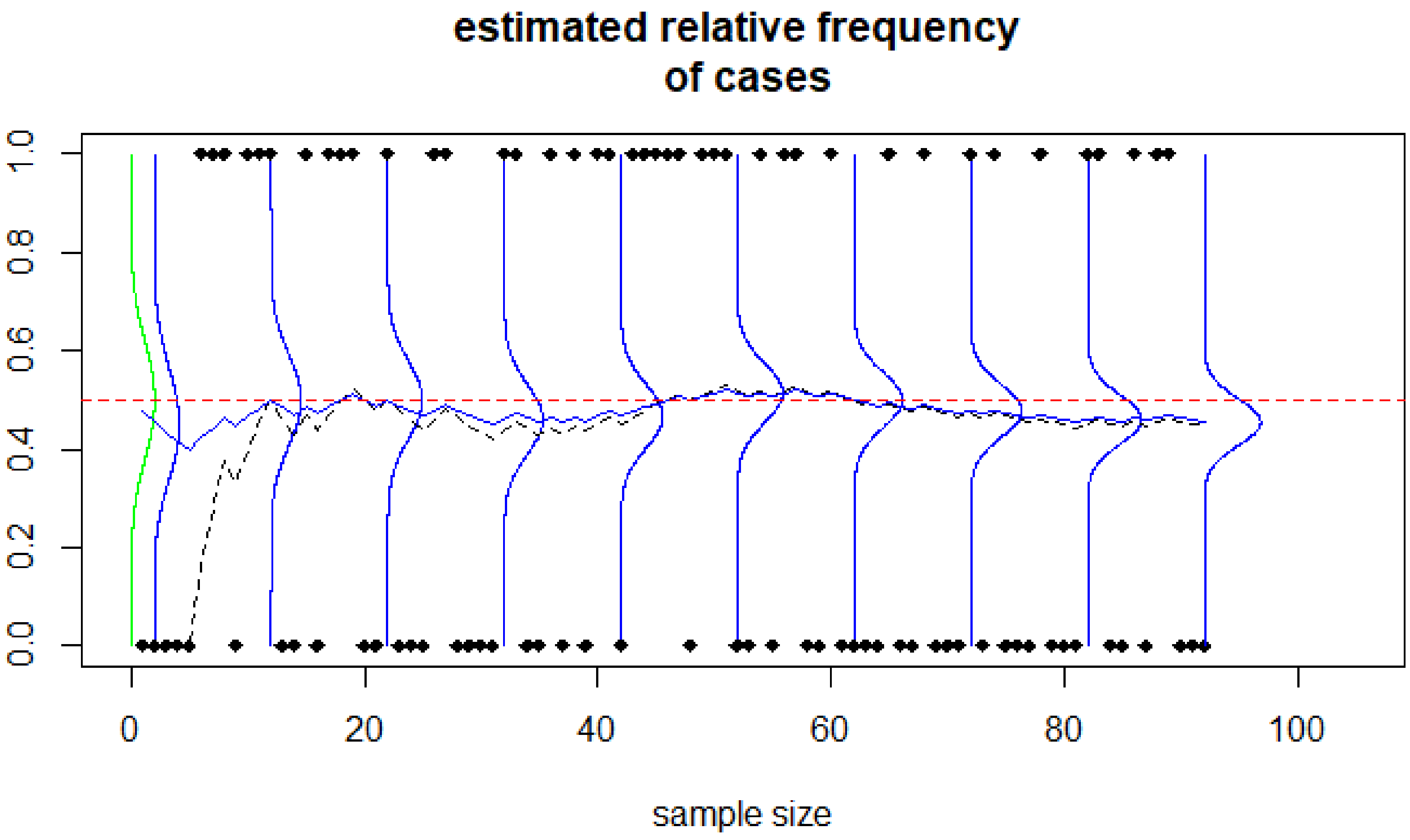
Bayesian updating

Study the effect of choice of prior in combination with the true relative frequency using
beta_binomial_learning.R

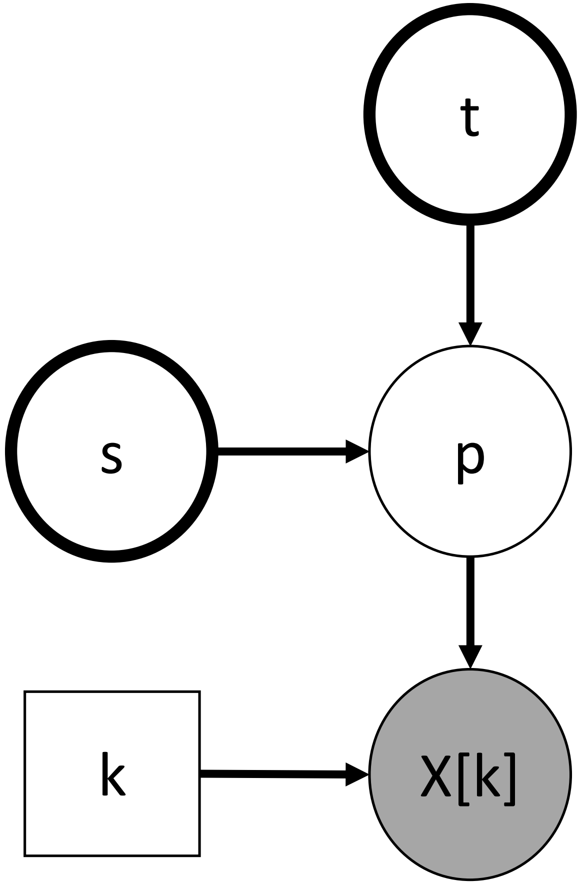


Bayesian updating

Study the effect of choice of prior in combination with the true relative frequency using
beta_binomial_learning.R



Beta-binomial model



prior

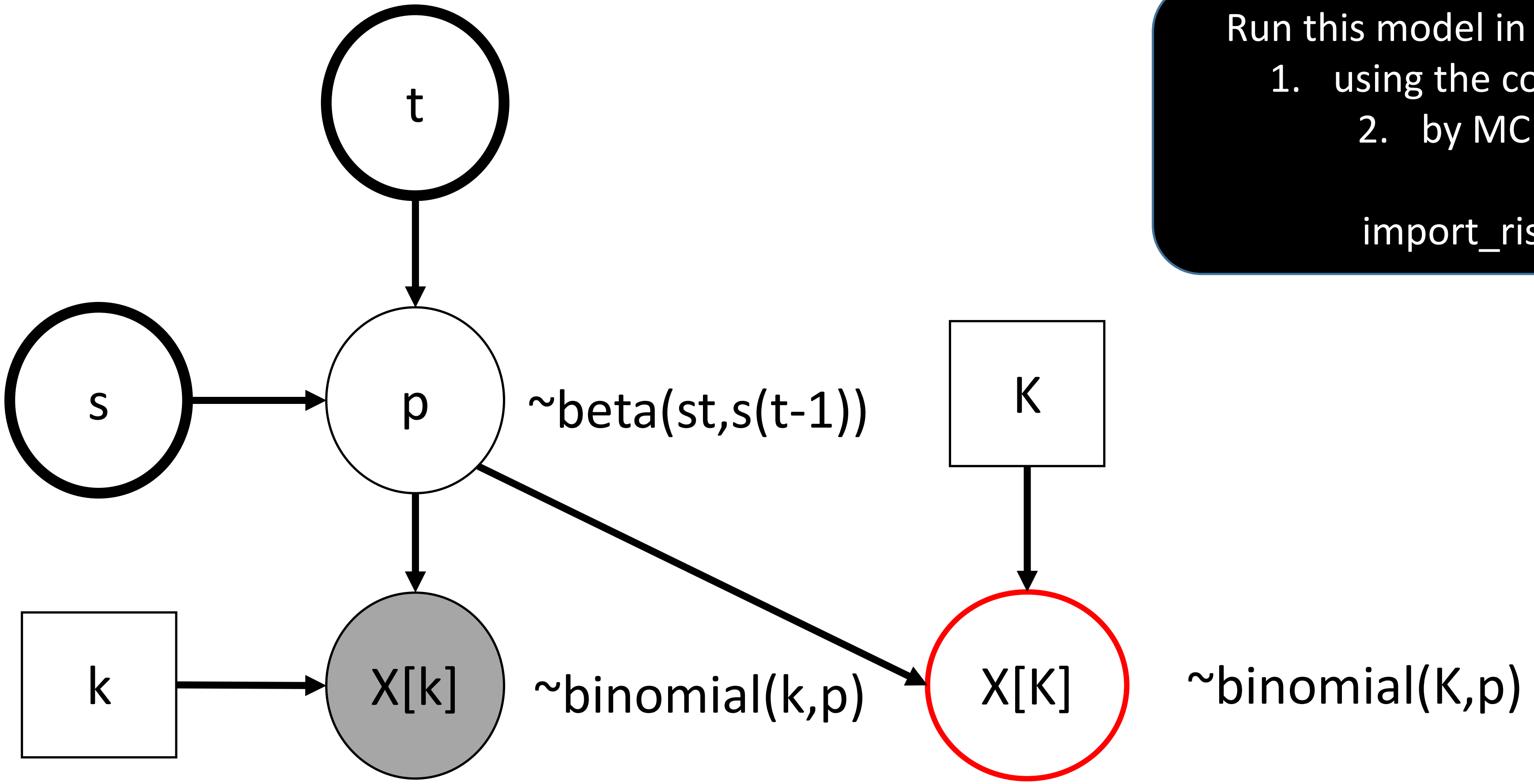
~beta(st, s(1-t))

~binomial(k, p)

posterior

~beta(st + X[k], s(1-t) + k – X[k])

Bayesian updating and making predictions



Run this model in two ways to update:

1. using the conjugacy property
2. by MCMC sampling

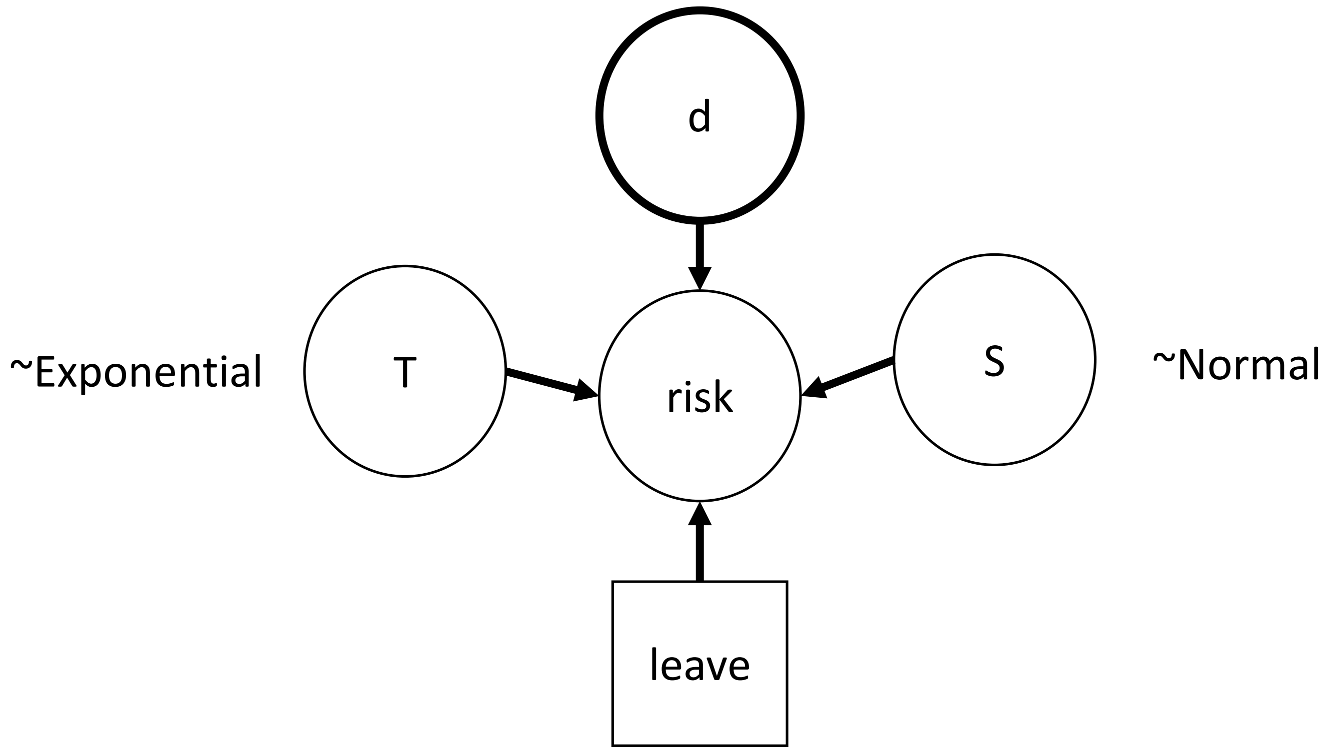
import_risk_analysis.R

Exceeding a threshold risk analysis
- missing the last bus home

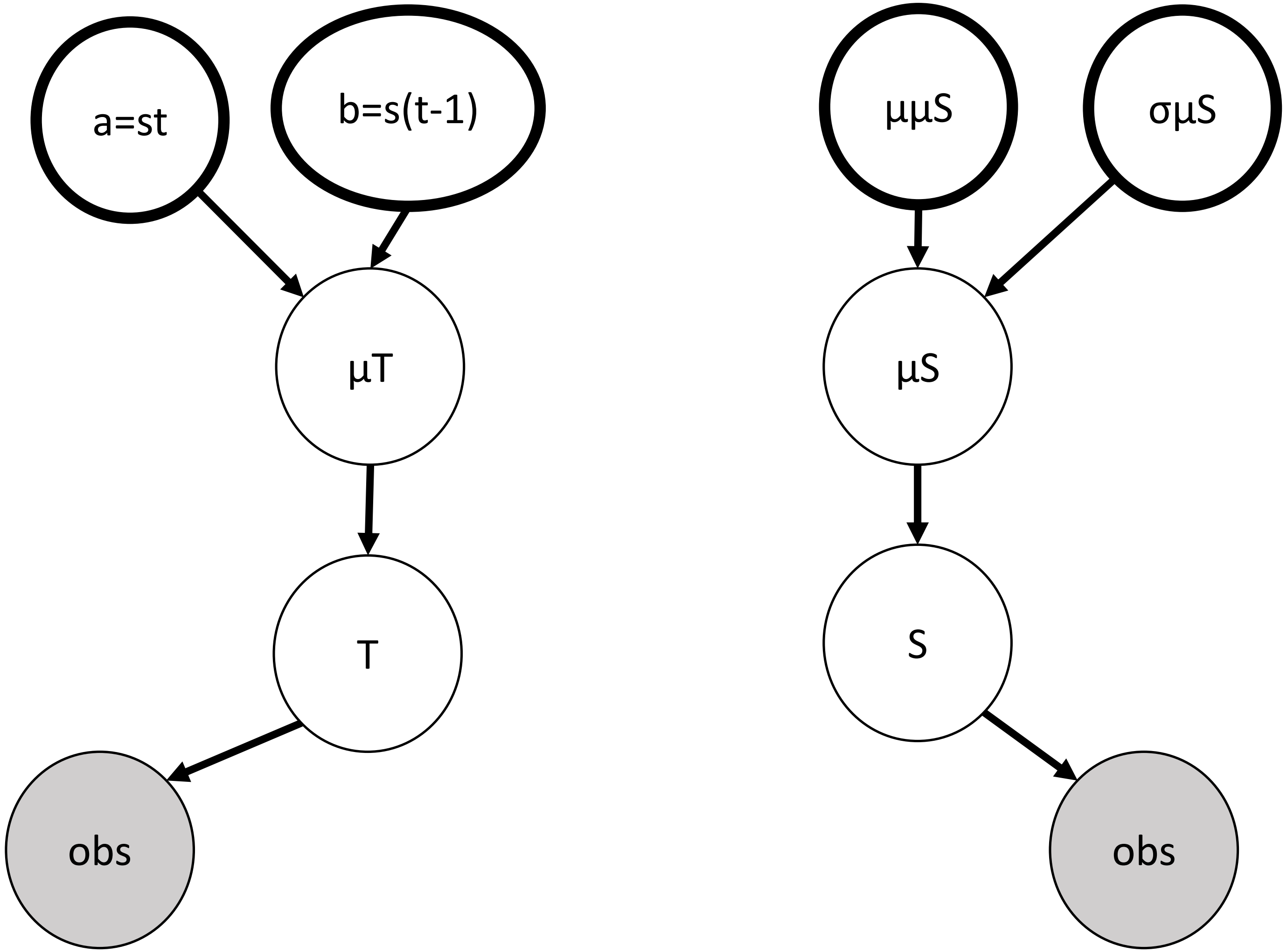
The code is at https://github.com/Ullrika/quantifying_uncertainty_by_probability

System model

bus_part1.R



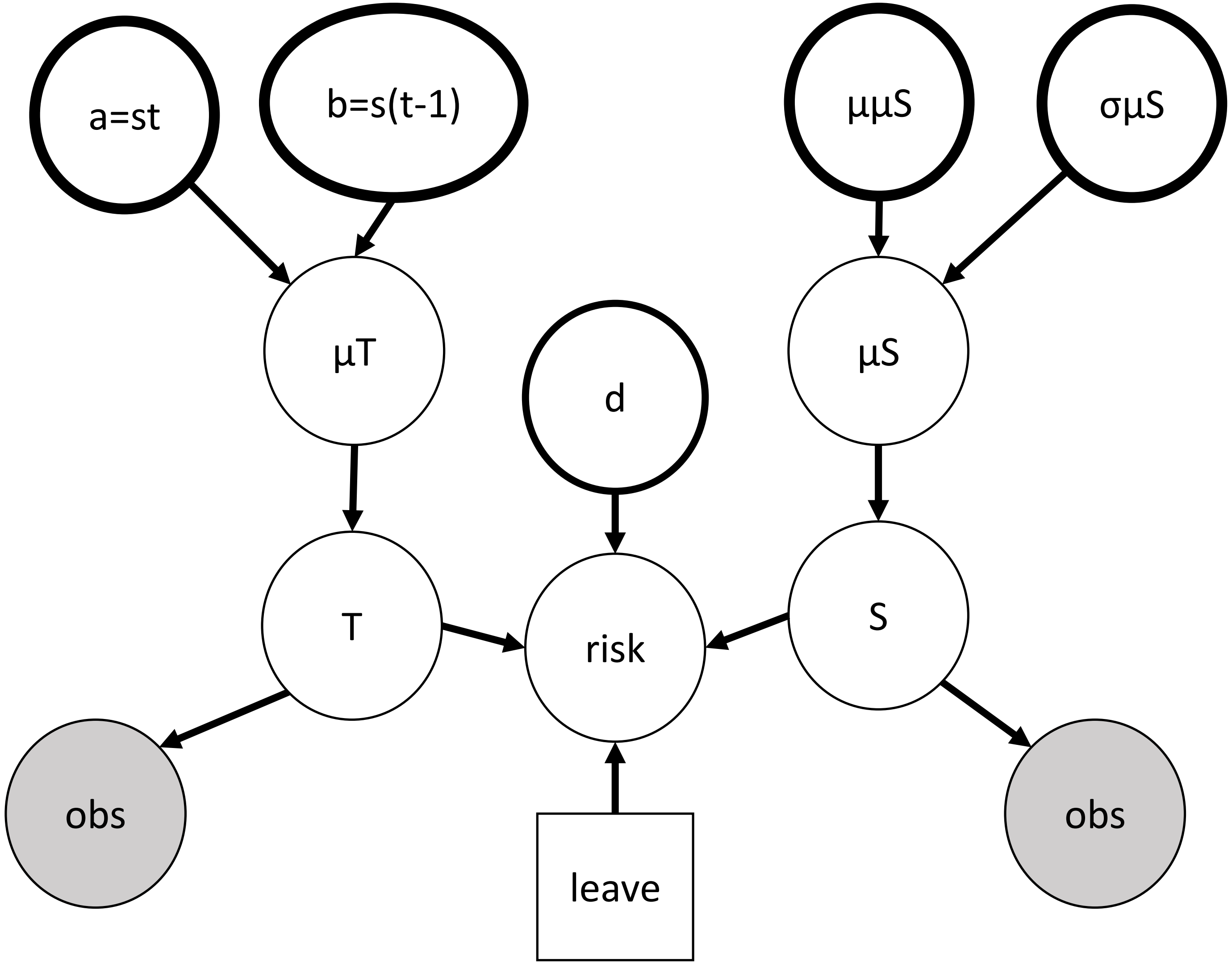
Data generating processes



bus_part2.R

We have data on waiting time and on walking speed.
We build a probabilistic graph so we can derive the likelihood for the data given parameters.

Bayesian Evidence Synthesis



bus_part2.R

When we run this analysis we can run the parametric and predictive inference in the same step.

More on Expert Knowledge Elicitation

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

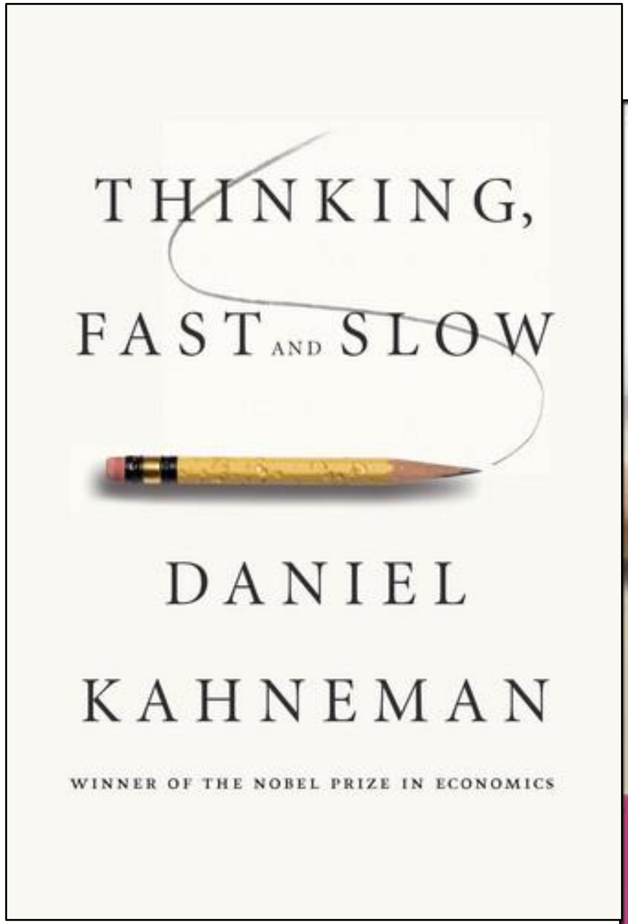
Source: http://www.expertsinuncertainty.net/Software/tabid/4149/Default.aspx

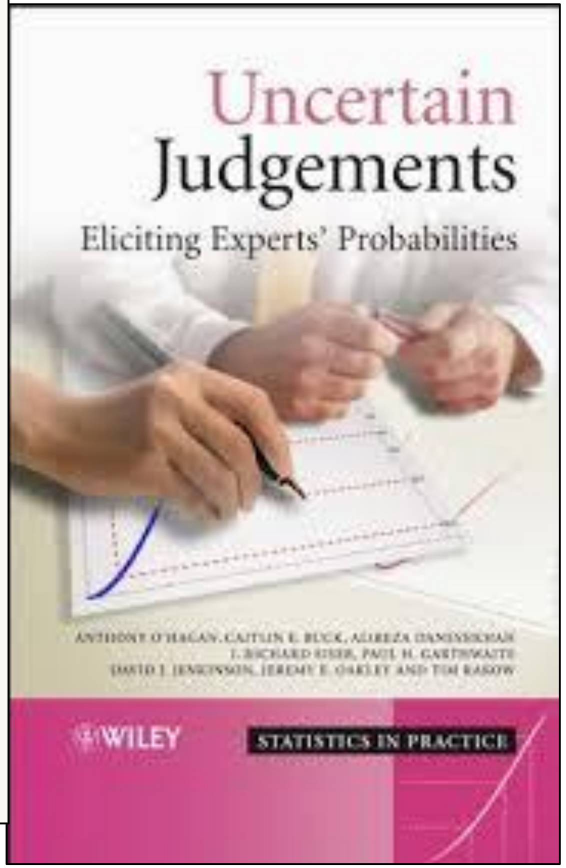
Expert Knowledge Elicitation

- 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





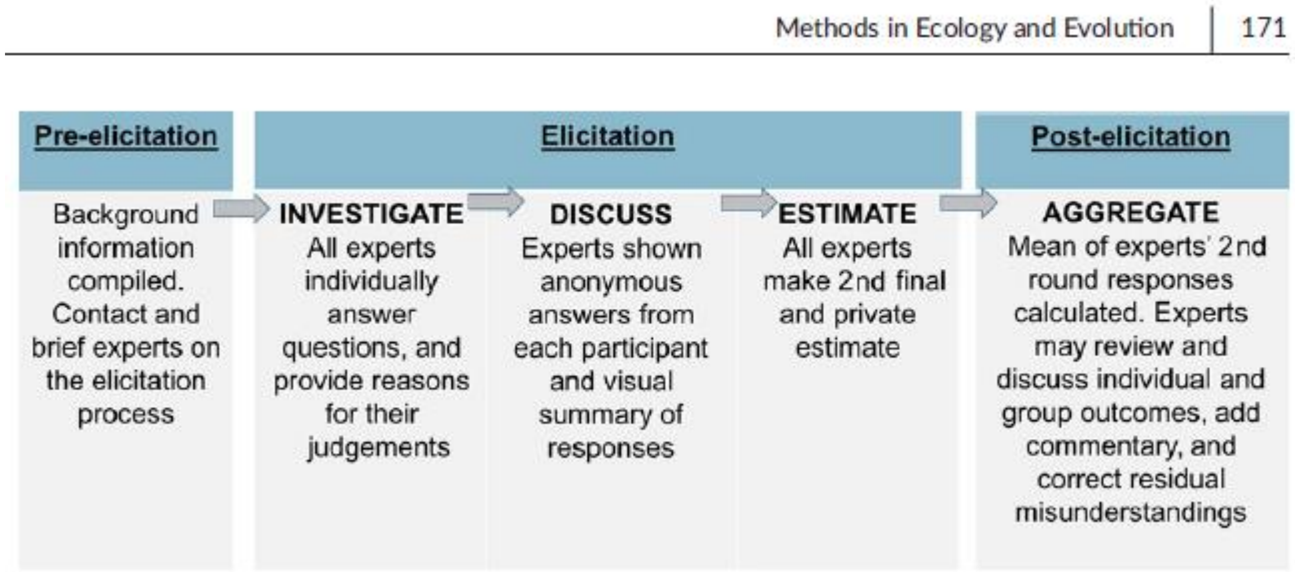
Elicitation with multiple experts

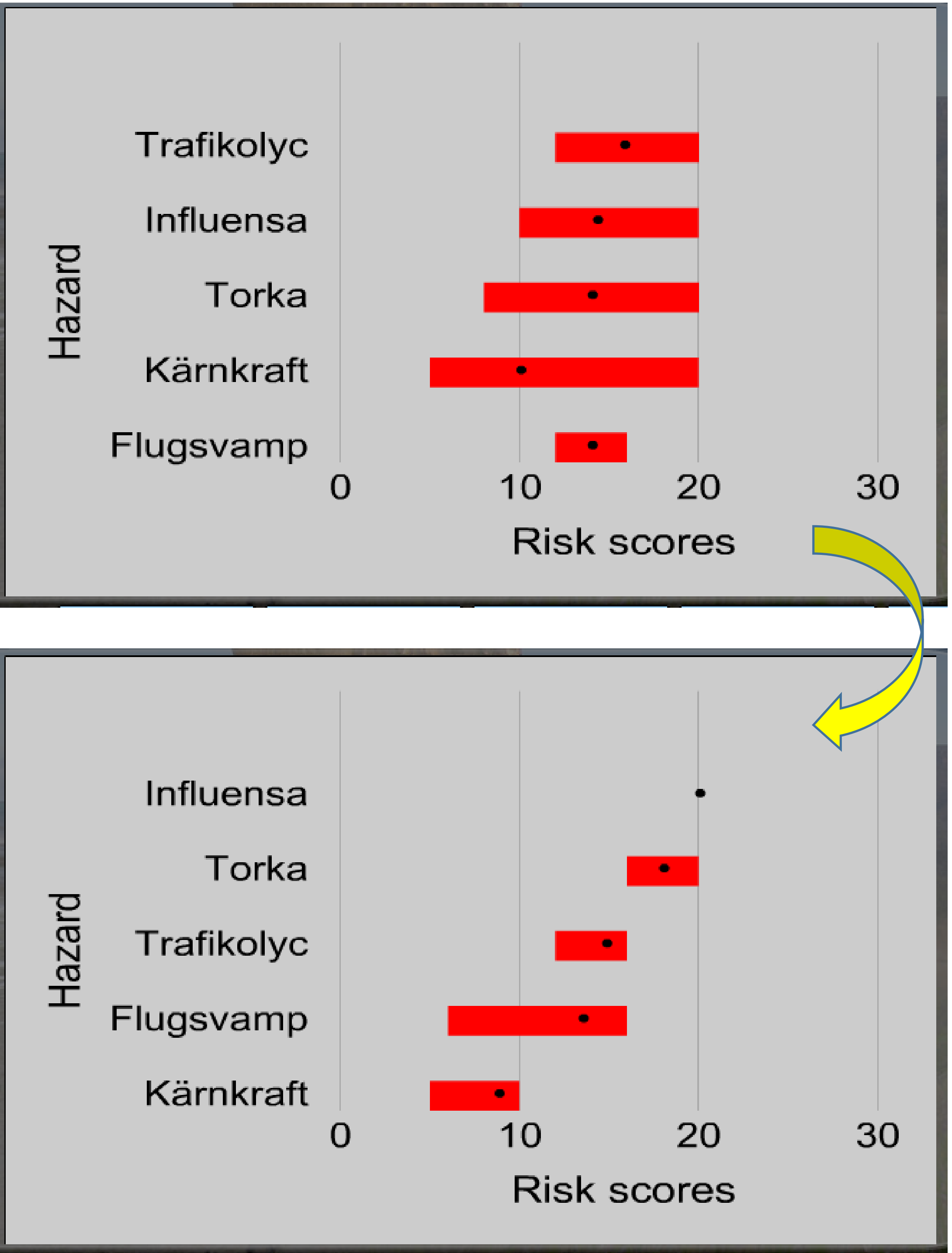
- More psychological factors add added when working with several experts.
- On the other hand, the result seems to be better (Wisdom of the crowds, multiple perspectives)

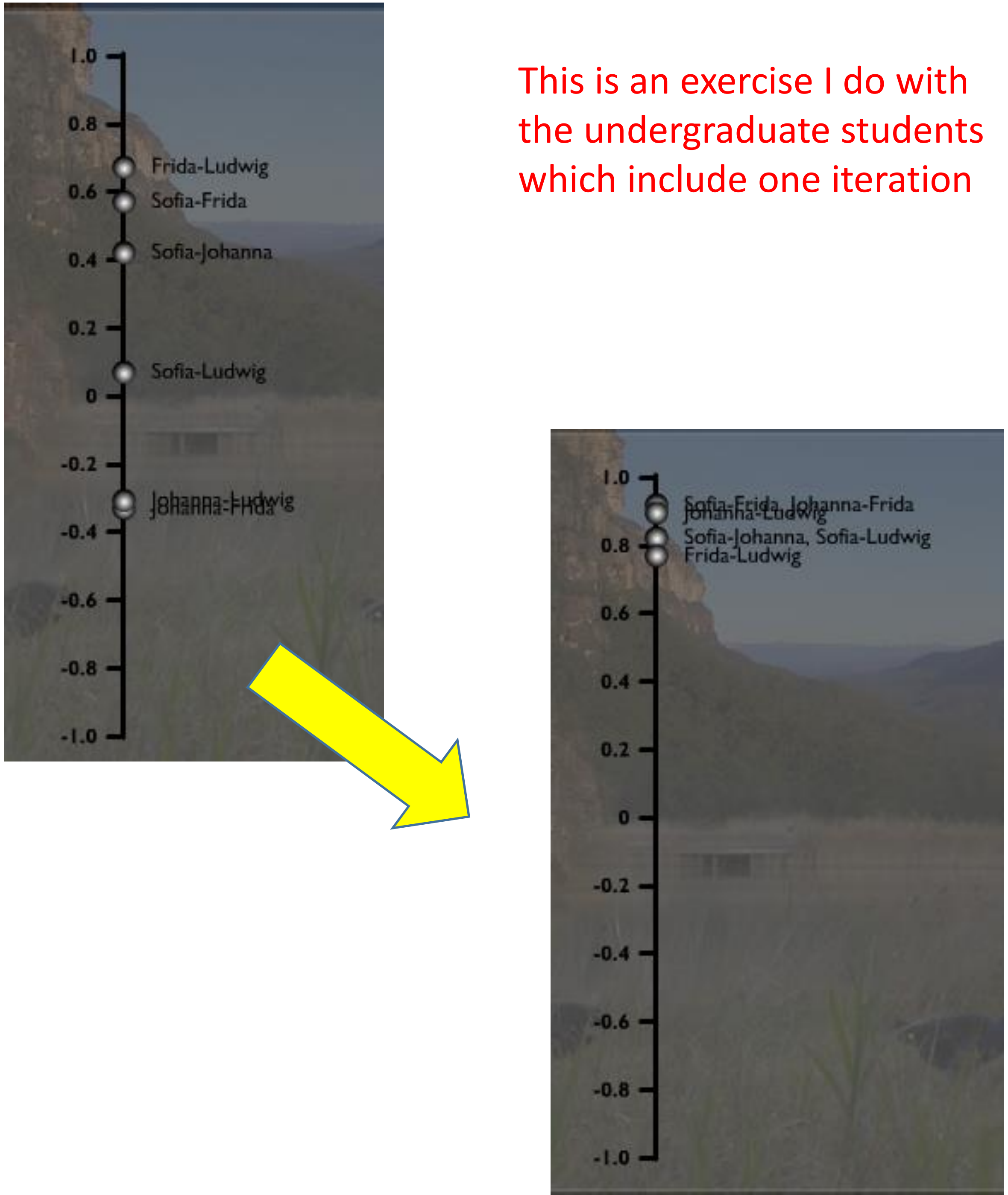
- Multiple experts needs to be aggregated somehow
- Behavioural aggregation
  - Group elicitation
  - One or several iterations, individually and in group
- Mathematical aggregation
  - Treat each expert’s distribution as data and update the decision maker’s belief
  - Pooled opinions – linear or logarithmic 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
- The IDEA protocol individual judgements, interaction of experts then individual judgements at the end

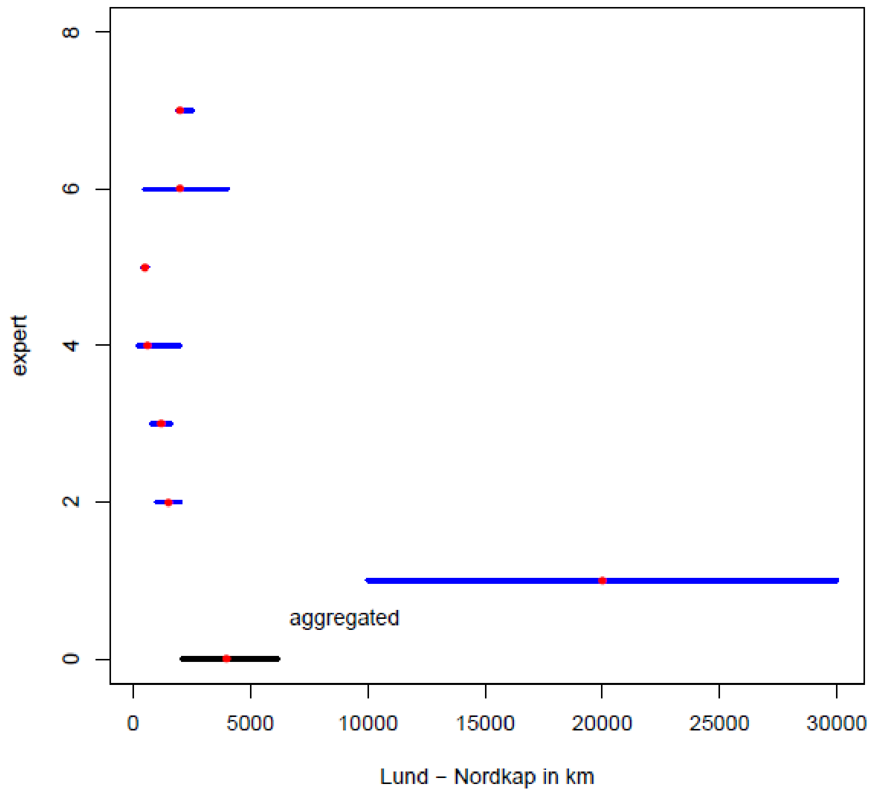


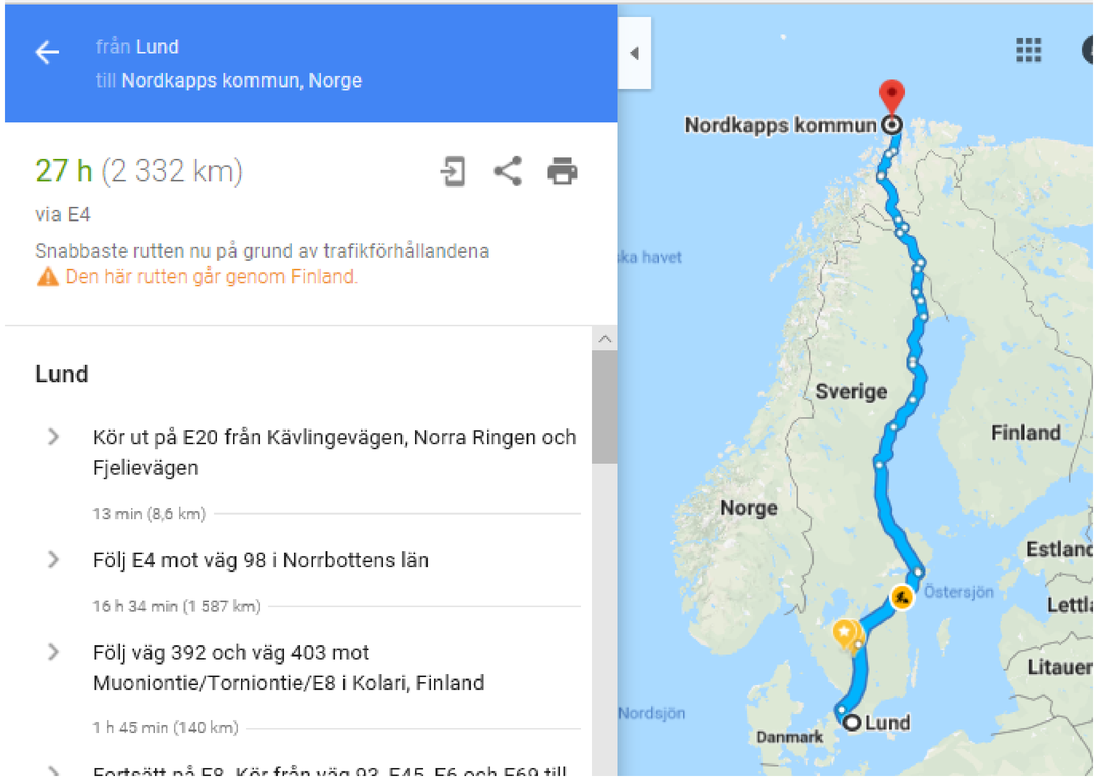




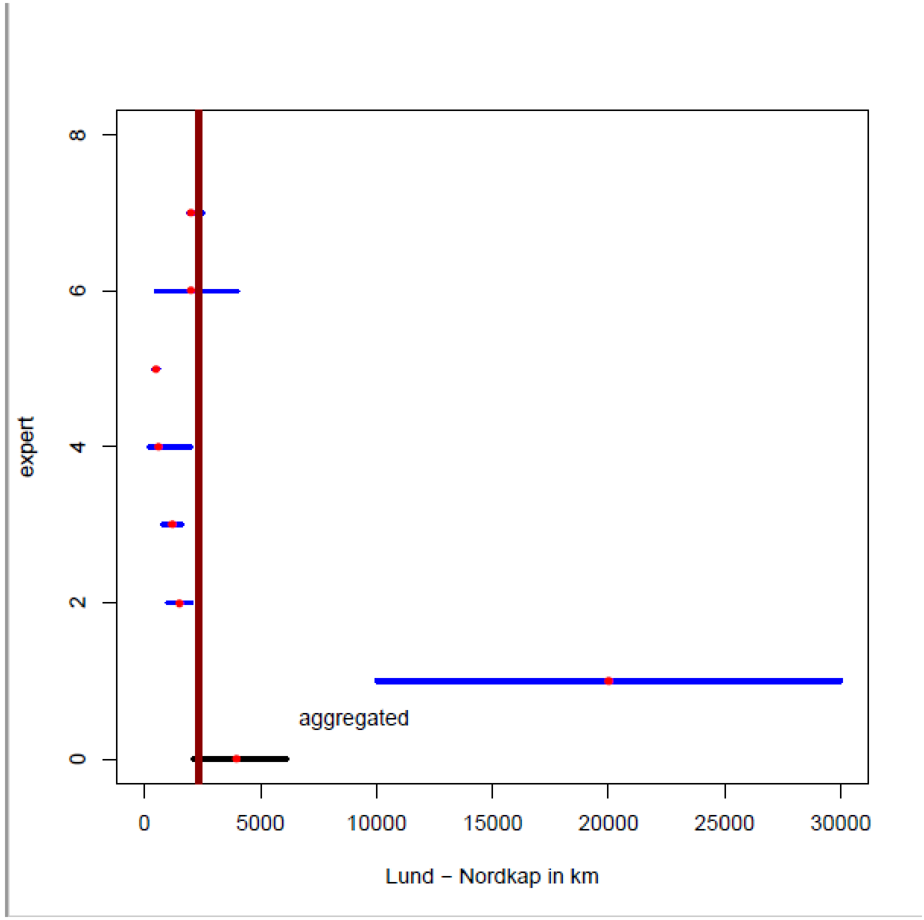
This is an exercise I do with the undergraduate students which include one iteration

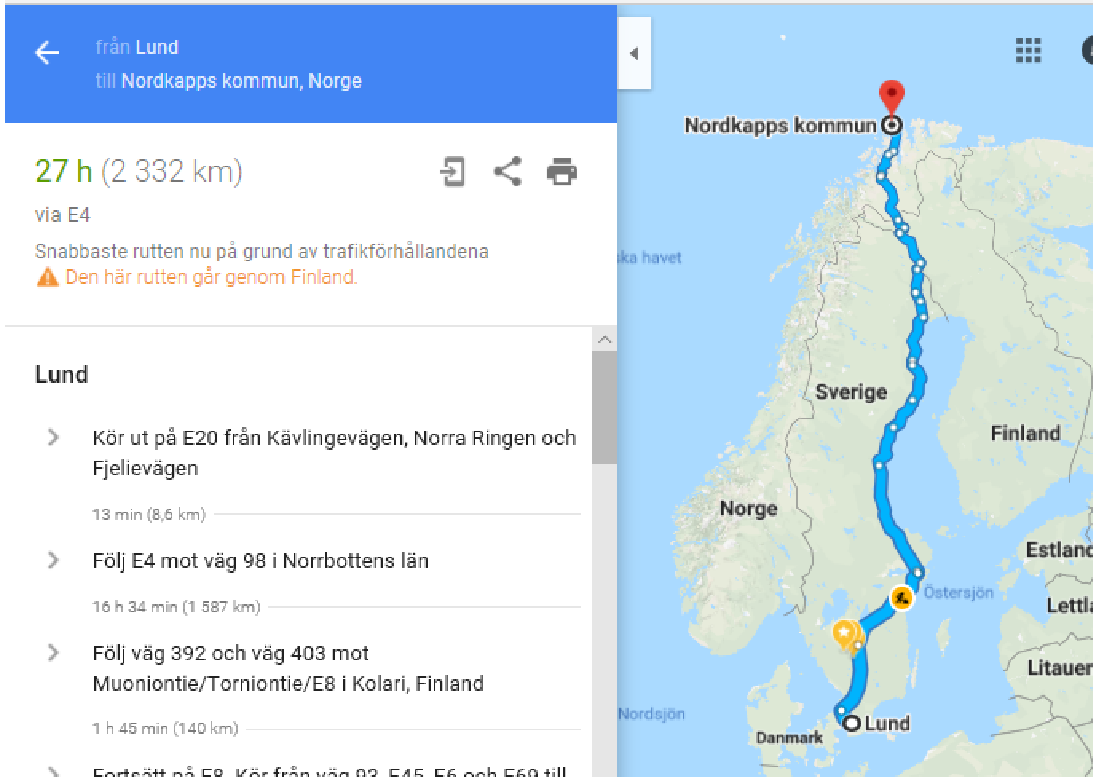
One by one, most of you could not hit the correct distance with a probablity interval





....but the aggregated interval is rather precise and covers the true value.





Recommended literature/courses/meetings

Probabilistic judgements by Anthony O'Hagan
http://www.tonyohagan.co.uk/shelf/ecourse.html

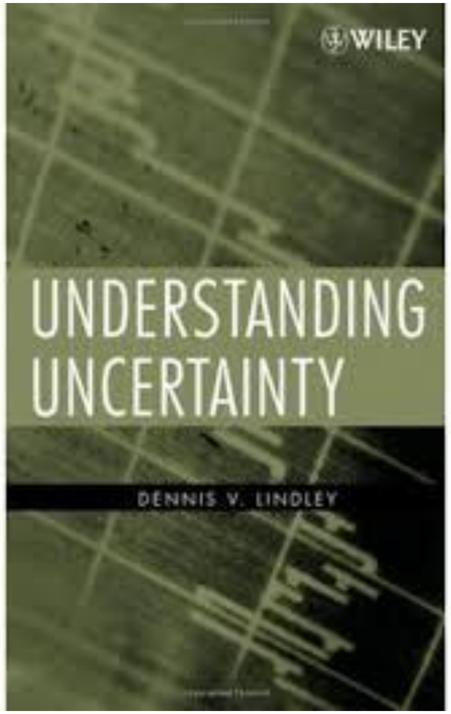
DataCamp Fundamentals of Bayesian Data Analysis in R by Rasmus Bååth
https://www.datacamp.com/courses/fundamentals-of-bayesian-data-analysis-in-r

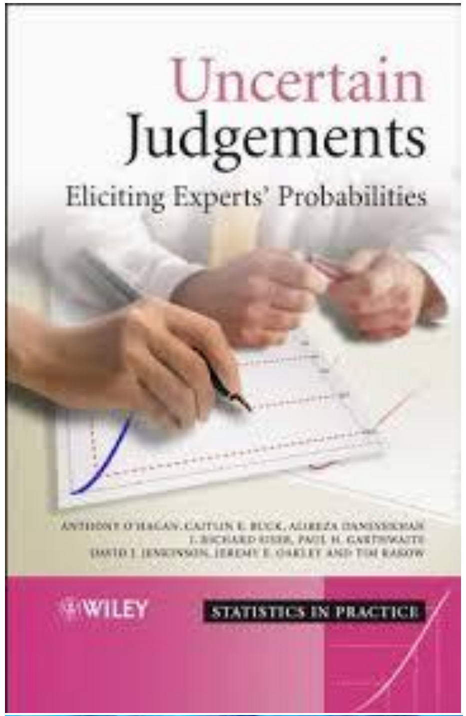
Bayesian methods course at the Department of Statistics, LU, Jonas Wallin
https://www.stat.lu.se/en/education/courses/stae02_bayesian_methods

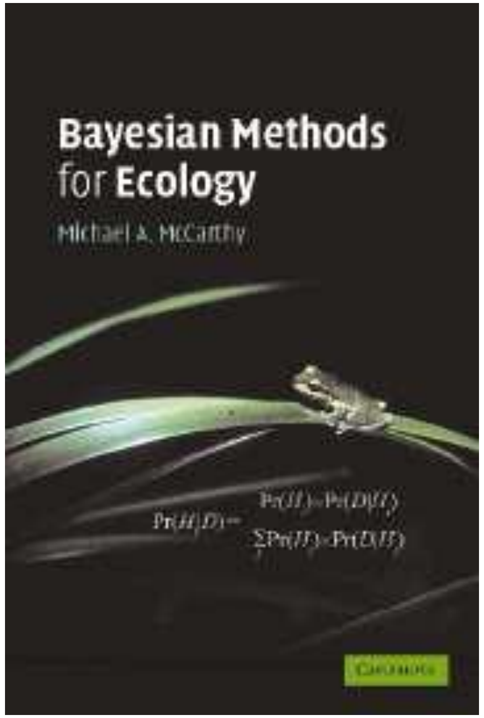
Bayesian Analysis and Decision Theory - Graduate course, CEC, Lund, Ullrika Sahlin

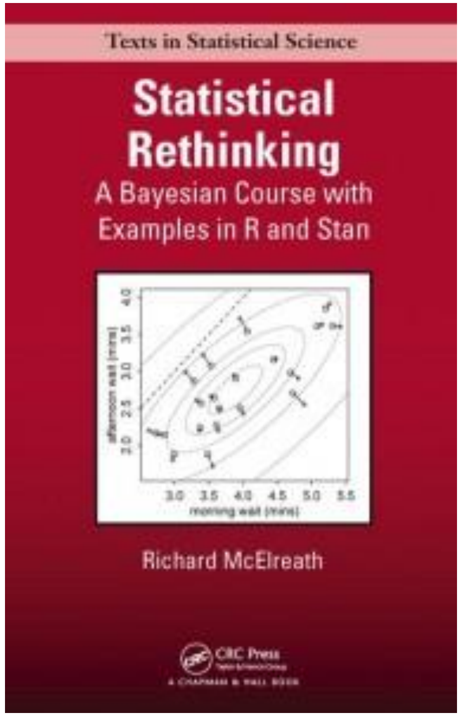
Rasmus blog http://www.sumsar.net/blog/2015/12/bayes-js-a-small-library-for-doing-mcmc-in-the-browser/

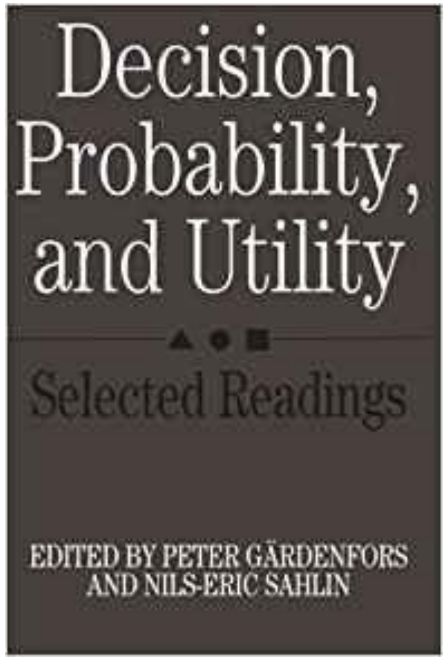
Bayes@Lund – annual and accessible meeting in Lund (sign up to the Bayesian email list at Lund University to get early information http://www.lucs.lu.se/bayes/)

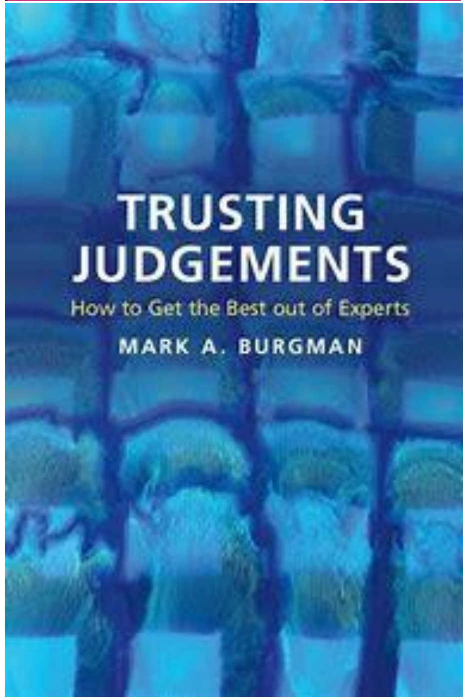

WILEY
UNDERSTANDING UNCERTAINTY
DENNIS V. LINDLEY

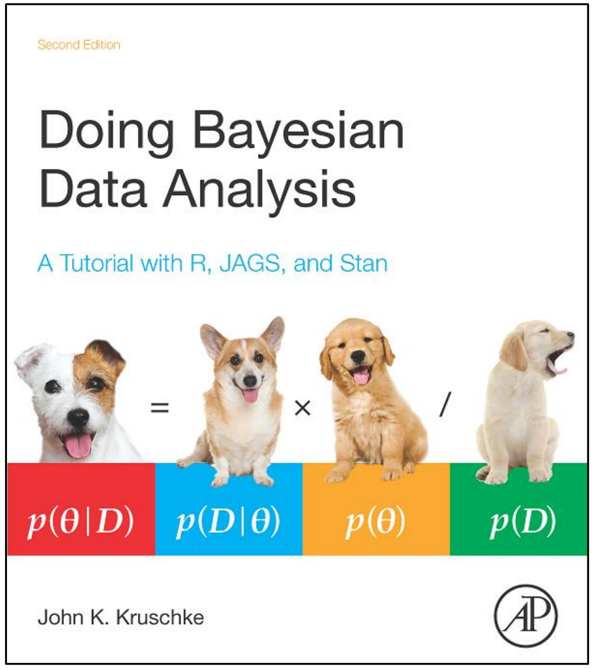

Uncertain Judgements
Eliciting Experts' Probabilities
WILEY
STATISTICS IN PRACTICE

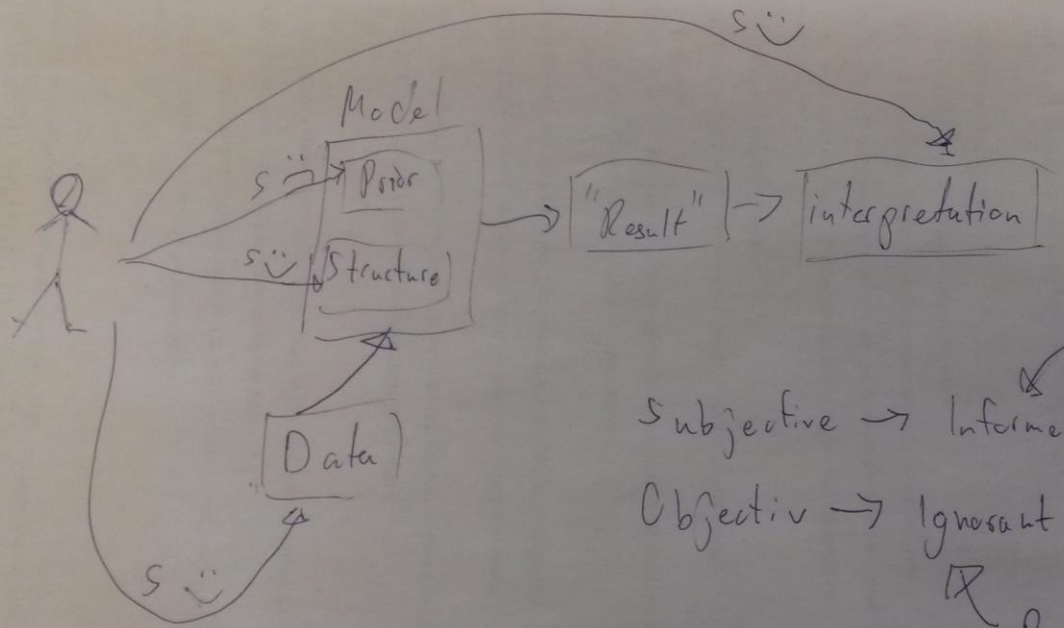

Bayesian Methods for Ecology
Michael A. McCarthy


Texts in Statistical Science
Statistical Rethinking
A Bayesian Course with Examples in R and Stan
Richard McElreath
CRC Press


Decision, Probability, and Utility
Selected Readings
EDITED BY PETER GÄRDENFORS AND NILS-ERIC SAHLIN


TRUSTING JUDGEMENTS
How to Get the Best out of Experts
MARK A. BURGMAN


Second Edition
Doing Bayesian Data Analysis
A Tutorial with R, JAGS, and Stan
=
×
/
$p(\theta|D)$
$p(D|\theta)$
$p(\theta)$
$p(D)$
John K. Kruschke



Depends

Good information?

Subjective → Informed

Objectiv → Ignorant

Best you can do in a bad situation

Who's degree of belief?

A protocol for summarising scientific uncertainty

1. Identify key outcomes for decision makers and how to measure them
2. Summarise variability
3. Summarise internal validity
4. Summarise external validity
5. Summarise the strenght of the basic science (e.g. NUSAP)
6. Summarise uncertainty (e.g. probability distributions or credibel intervals)

Recommendations in the Spiegelhalter and Riesch paper

1. Use quantitative models with aleatory and epistemic uncertainty expressed as Bayesian probability distributions.
2. Conduct sensitivity analysis to alternative model forms and assess evidential support for alternative structures, without putting probabilities on models.
3. Provide a list of known model limitations and a judgement of their qualitative or quantitative influence, possibly along the lines shown in table 6, and ensuring there has been a fully imaginative consideration of possible futures.
4. Provide a qualitative expression of confidence, or lack of it, in any analysis based on the quality of the underlying evidence, possibly expressed using an adapted GRADE scale or the IPCC guidance [46].
5. In situations of low confidence, use deliberately imprecise expressions of uncertainty about quantities, such as their orders-of-magnitude, whether they are positive or negative, or even refuse to give any judgement at all; the IPCC guidance suggests a calibrated scale for these expressions.
6. When exploring possible actions, look for robustness to error, resilience to the unforeseen, and potential for adaptivity in the face of the unexpected [10].
7. Seek transparency and ease of interrogation of any model, with clear expression of the provenance of assumptions.
8. Communicate the estimates with humility, communicate the uncertainty with confidence.
9. Fully acknowledge the role of judgement: this ‘..means engaging in policy making by fully accepting the constructive, participatory, ultimately open-ended and untamed nature of judgements under uncertainty’

Approach uncertainty sceptisism (Fischoof and Davis again)

Table 3. Frequently asked questions addressing four concerns of scientists reluctant to express their uncertainty in credible-interval form

Concern 1	If I give credible intervals, people will misinterpret them, inferring greater precision than I intended.
Response	Behavioral research has found that most people (*i*) like receiving explicit quantitative expressions of uncertainty (such as credible intervals), (*ii*) can interpret them well enough to extract their main message, and (*iii*) misinterpret verbal expressions of uncertainty (e.g., "good" evidence, "rare" side effect). For audiences that receive the reports created with the protocol (Table 2), understanding should be greater if they receive credible intervals than if they have to infer them (63).
Concern 2	People cannot use probabilities.
Response	Behavioral research has found that laypeople can often provide reasonably consistent probability judgments if asked clear questions and extract needed information if provided with well-designed displays (41, 60, 74). Whether they do so well enough to satisfy their decision-making needs is an empirical question, which should be answered with evidence rather than speculation.
Concern 3	My credible intervals will be used unfairly in performance evaluations.
Response	Such judgments can protect experts from unfair evaluations, unjustly accusing them of having conveyed too much or too little confidence, especially when supported by the rationale for those judgments. The protocol provides such protection—if the experts' management stands behind it.
Concern 4	People do not need such judgments.
Response	Decision makers must act with some degree of uncertainty. Not helping them means preferring the risk of having them guess incorrectly over the risk of expressing oneself poorly.