Capturing uncertainty about model inputs using expert elicitation

John Paul Gosling (University of Leeds)

Overview

UNIVERSITY OF LEEDS

1. Principles of expert elicitation
2. Elicitation of input parameters
3. Eliciting input parameters for a immunology model

Bayes theorem and computer models

UNIVERSITY OF LEEDS

Calculate the posterior using

$$\pi(\theta|x) \propto \pi(\theta)\ \pi(x|\theta).$$

Consider $y = \eta(x)$ for some $x \in X$.

Calculate $E(\eta(x))$.

Expert elicitation

UNIVERSITY OF LEEDS

"An elicitation method forms a bridge between an expert's opinions and an expression of these opinions in a statistically **useful** form."

Garthwaite *et al.* (2005). Statistical methods for eliciting probability distributions. *Journal of the American Statistical Association*, **100**, 680-701.

"By the term expert elicitation we mean those techniques used to inform decisions, forecasts or predictions based on a **formalized treatment** of the judgments or opinions of experts."

Flandoli *et al.* (2010). Comparison of a new expert elicitation model... *Reliability Engineering & System Safety*, **96**,1292–1310.

Subjective, but scientific (1)

You want to use subjective probability judgements? Isn't that totally unscientific? Science is supposed to be objective.

Yes, objectivity is the goal of science, but scientists still have to make judgements. These judgements include theories, insights, interpretations of data. Science progresses by other scientists debating and testing those judgements. Making good judgements of this kind is what distinguishes a top scientist.

Slides stolen from Tony O'Hagan

Subjective, but scientific (2)

But subjective judgements are open to bias, prejudice, sloppy thinking …

Subjective probabilities are judgements but they should be careful, honest, informed judgements. In science we must always be as objective as possible. Probability judgements are like all the other judgements that a scientist necessarily makes, and should be argued for in the same careful, honest, informed way.

Slides stolen from Tony O'Hagan

Expert elicitation

UNIVERSITY OF LEEDS

Elicitation of an expert's beliefs is not a trivial task – can take many hours of effort (and arguing).

Expert elicitation

UNIVERSITY OF LEEDS

Elicitation of an expert's beliefs is not a trivial task – can take many hours of effort (and arguing).

Elicitation can help us to take stock of the uncertainty about quantities of interest without the cost of data collection.

Expert elicitation

UNIVERSITY OF LEEDS

Elicitation of an expert's beliefs is not a trivial task – can take many hours of effort (and arguing).

Elicitation can help us to take stock of the uncertainty about quantities of interest without the cost of data collection.

Elicitation is far from being a precise science:
it can be difficult for the experts to articulate their beliefs, and there are other complications due to the biases of experts and the biases created by the questioning process.

Expert elicitation

UNIVERSITY OF LEEDS

Key principles:

Clear definitions of the variables of interest.

Well-structured questions for capturing judgements.

Transparency in the process.

Opportunities for **feedback and revision**.

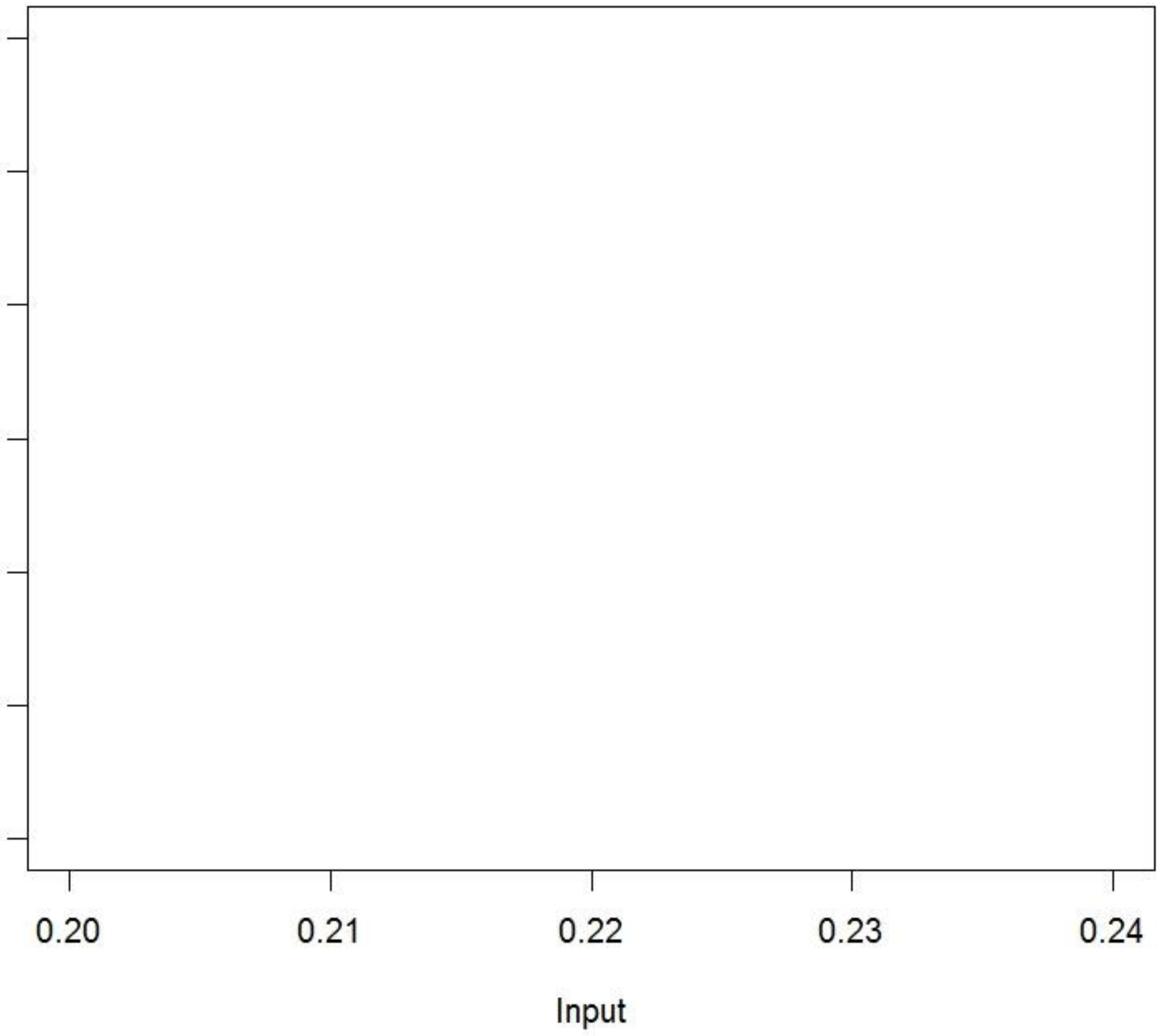
Elicitation for input parameters

UNIVERSITY OF LEEDS

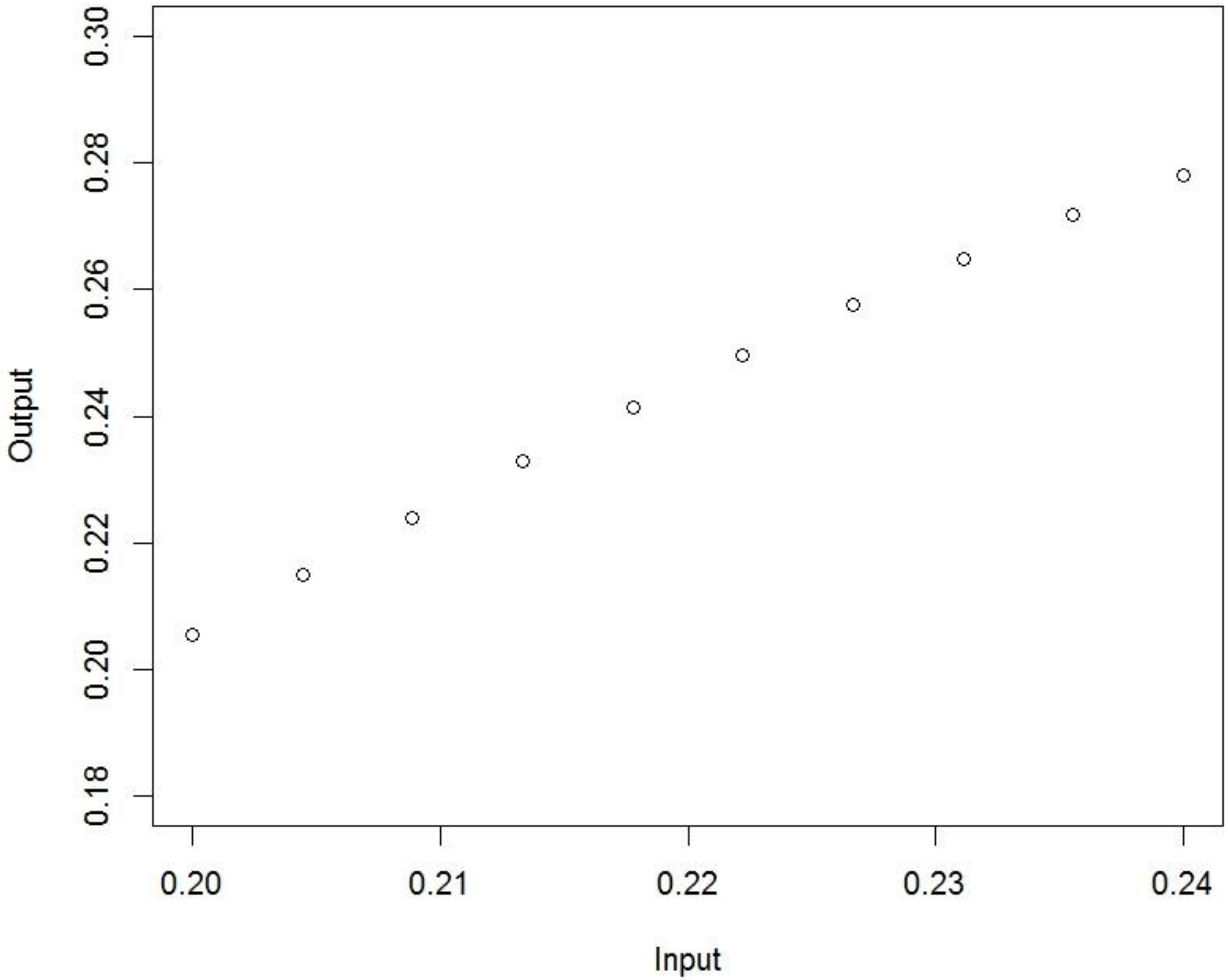
Determining where we want to explore

Careful characterisation of uncertainty for a Bayesian analysis

Getting the initial range right

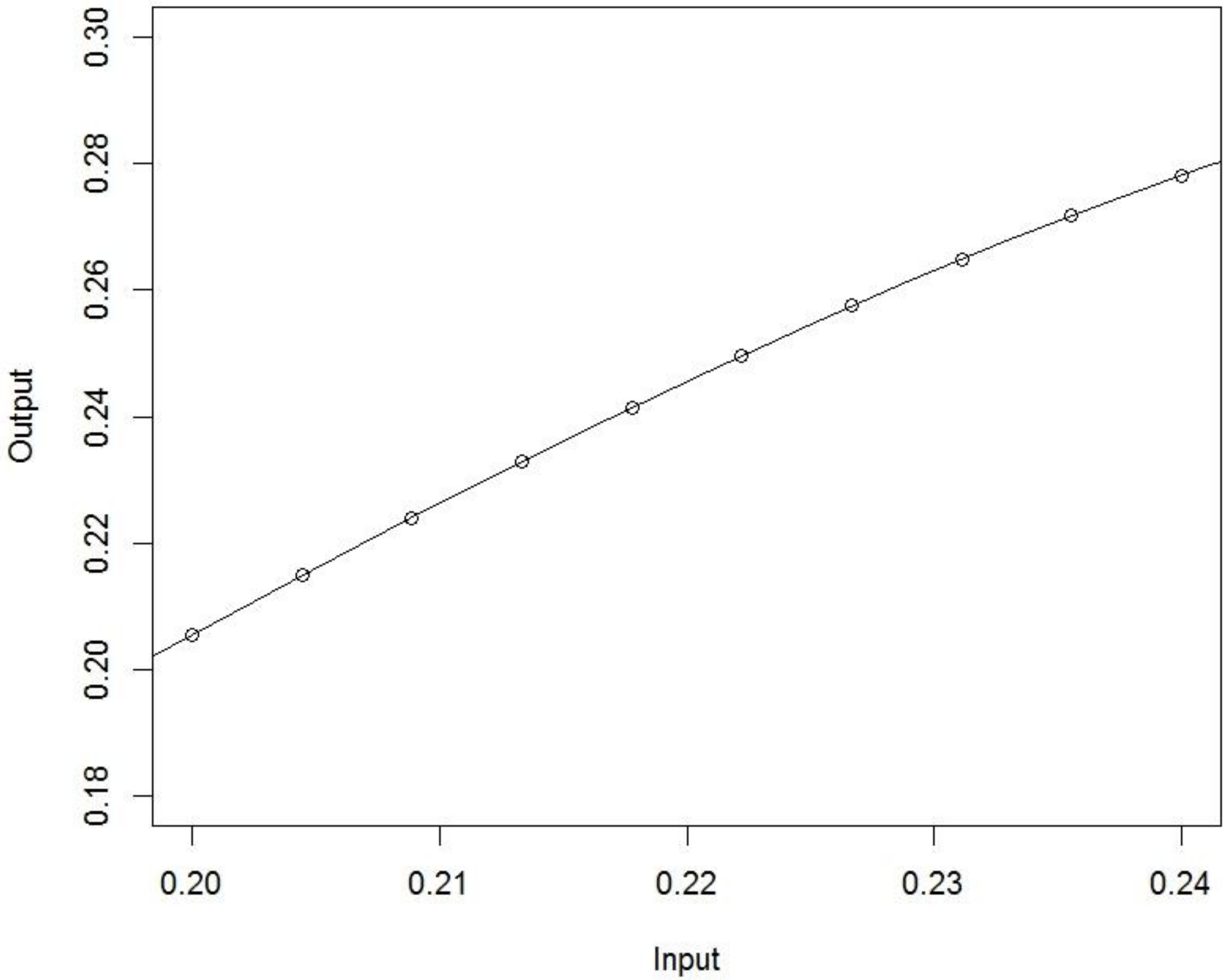


Getting the initial range right

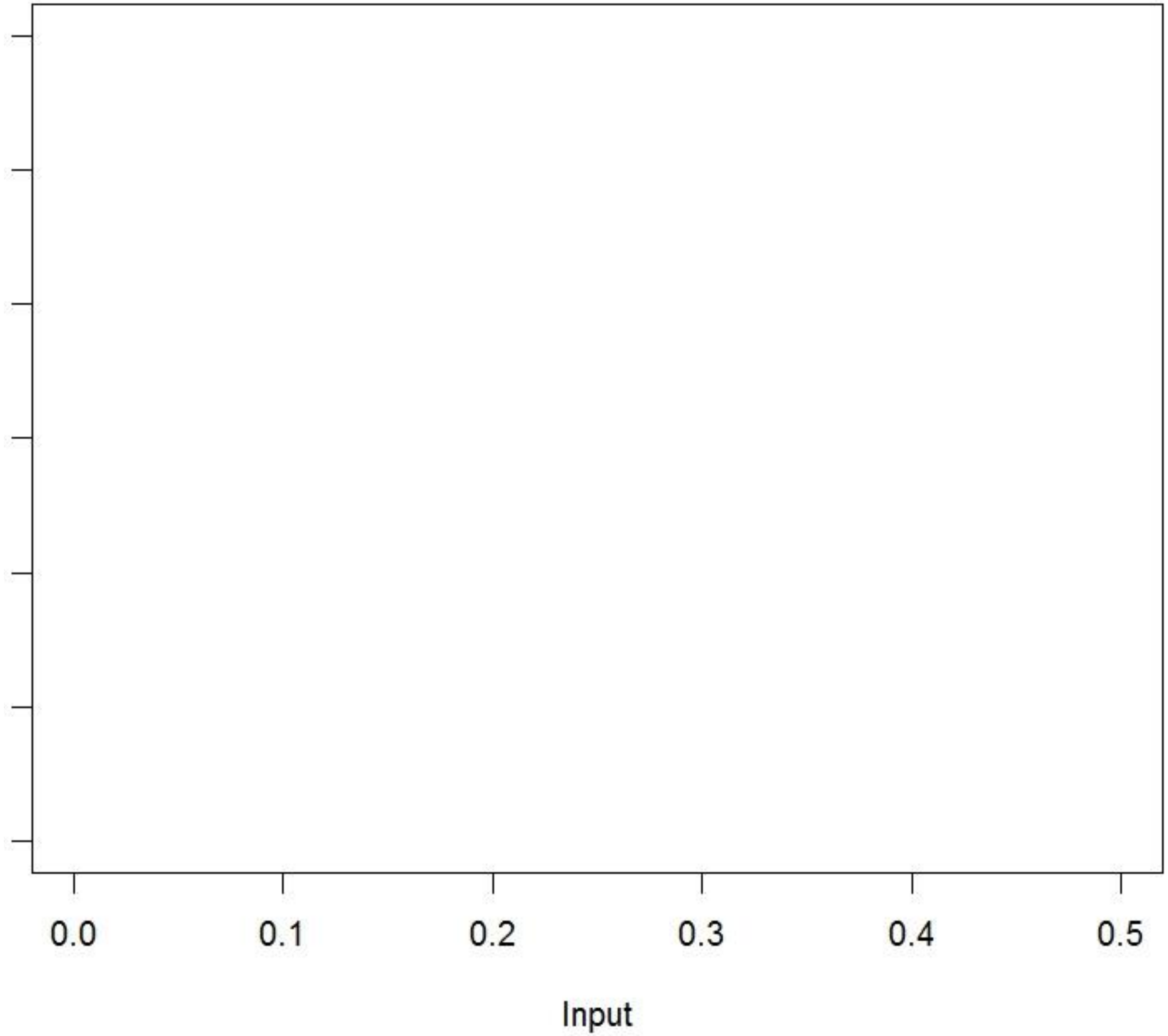


Getting the initial range right

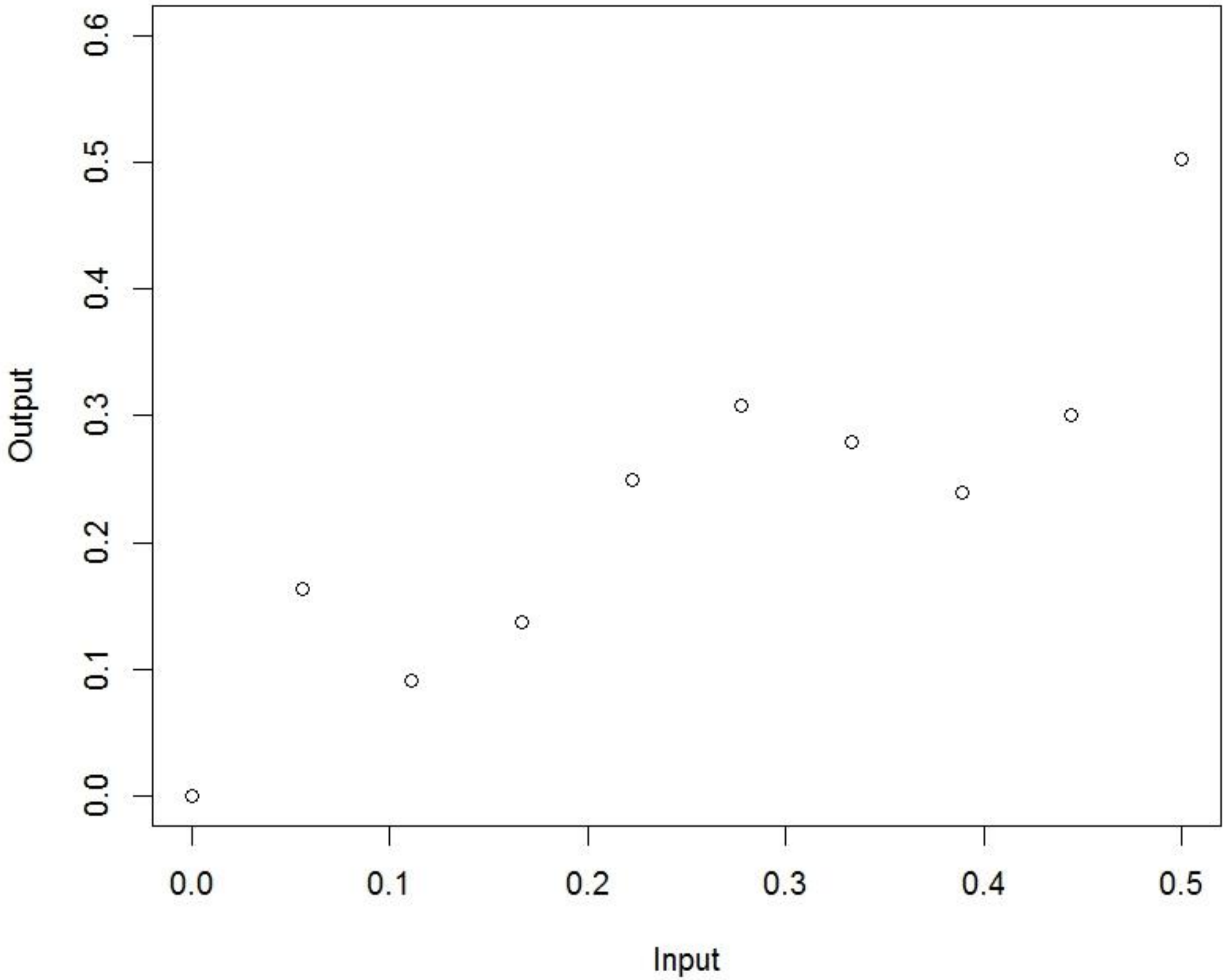
UNIVERSITY OF LEEDS



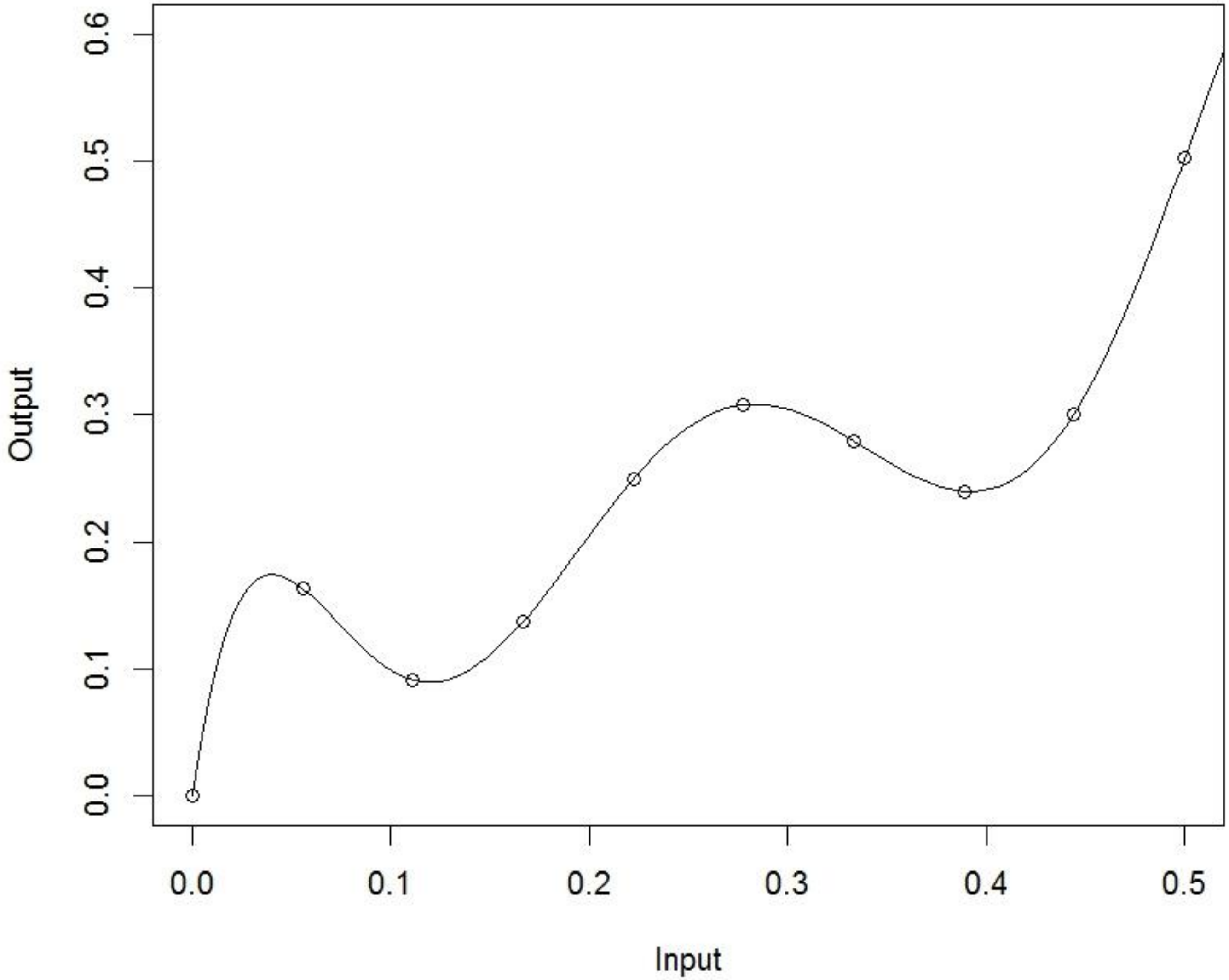
Getting the initial range right



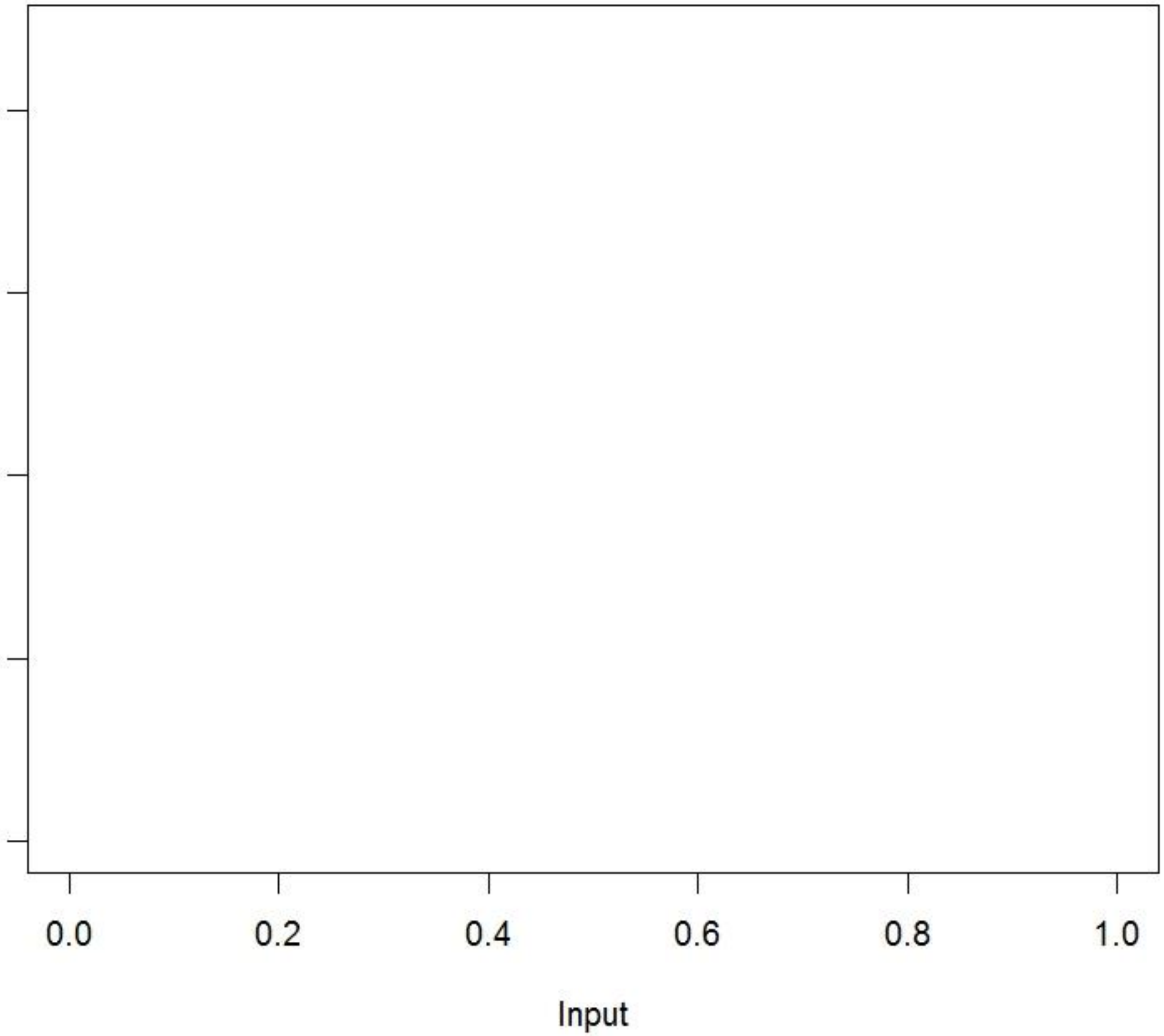
Getting the initial range right



Getting the initial range right

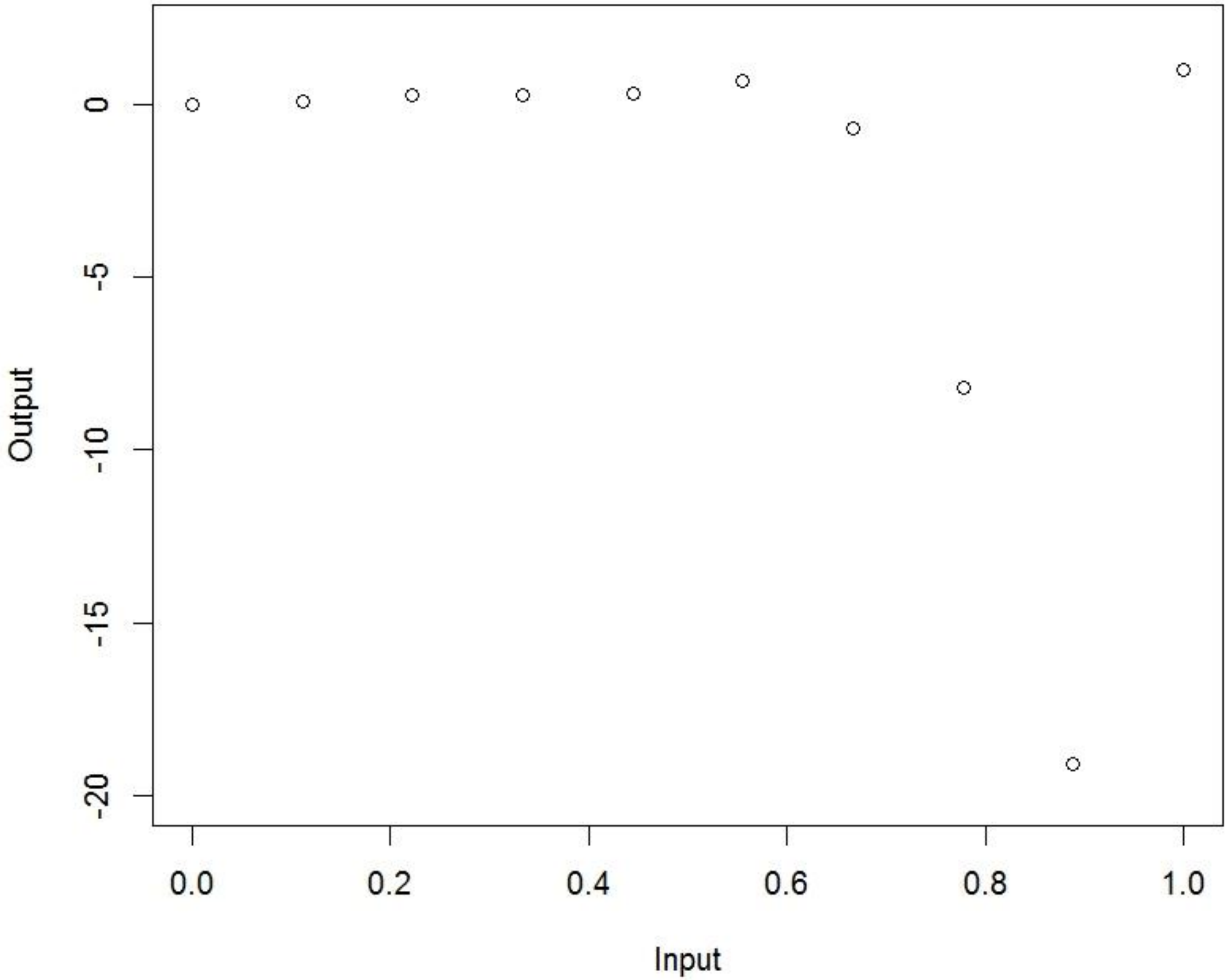


Getting the initial range right



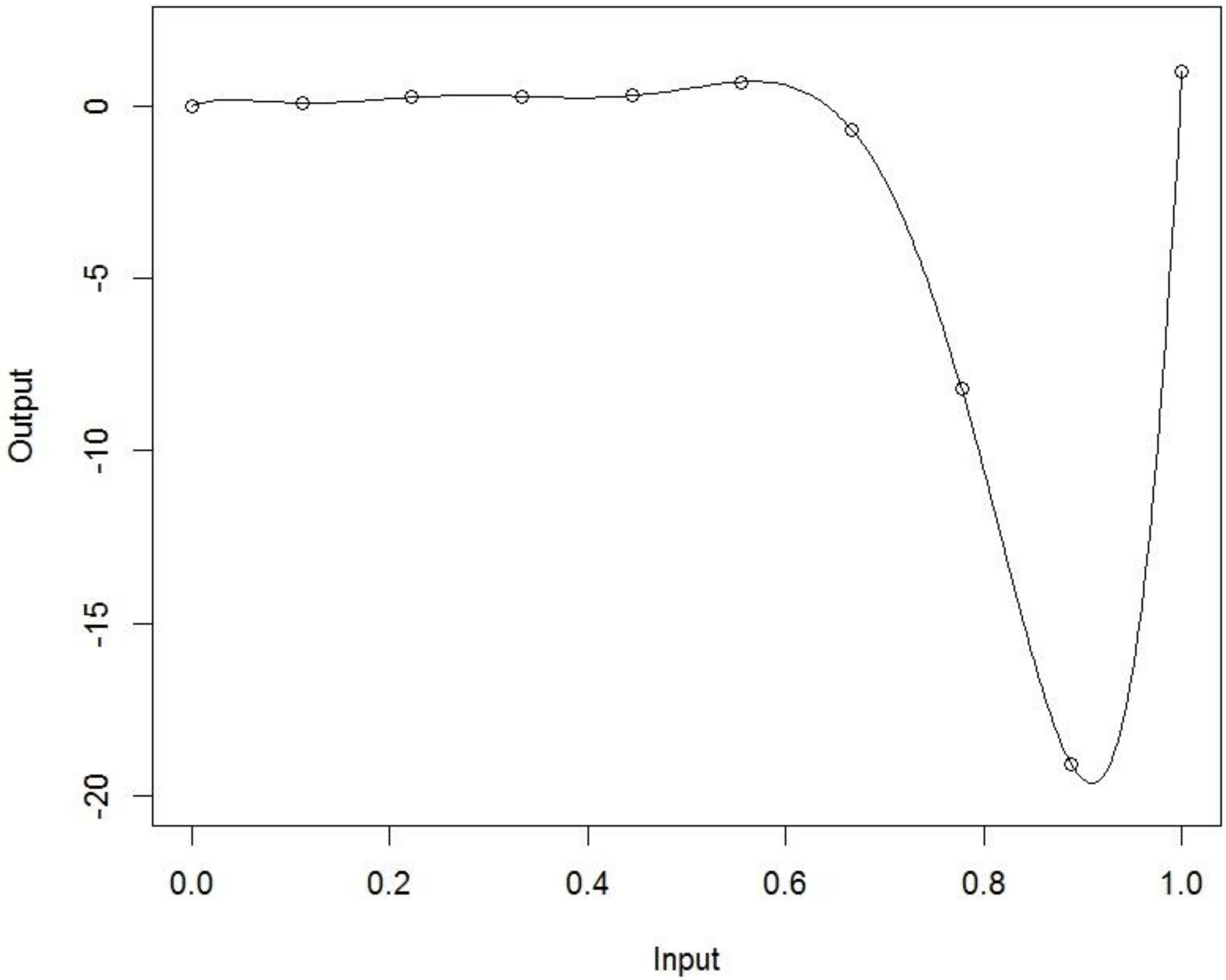
Getting the initial range right

UNIVERSITY OF LEEDS



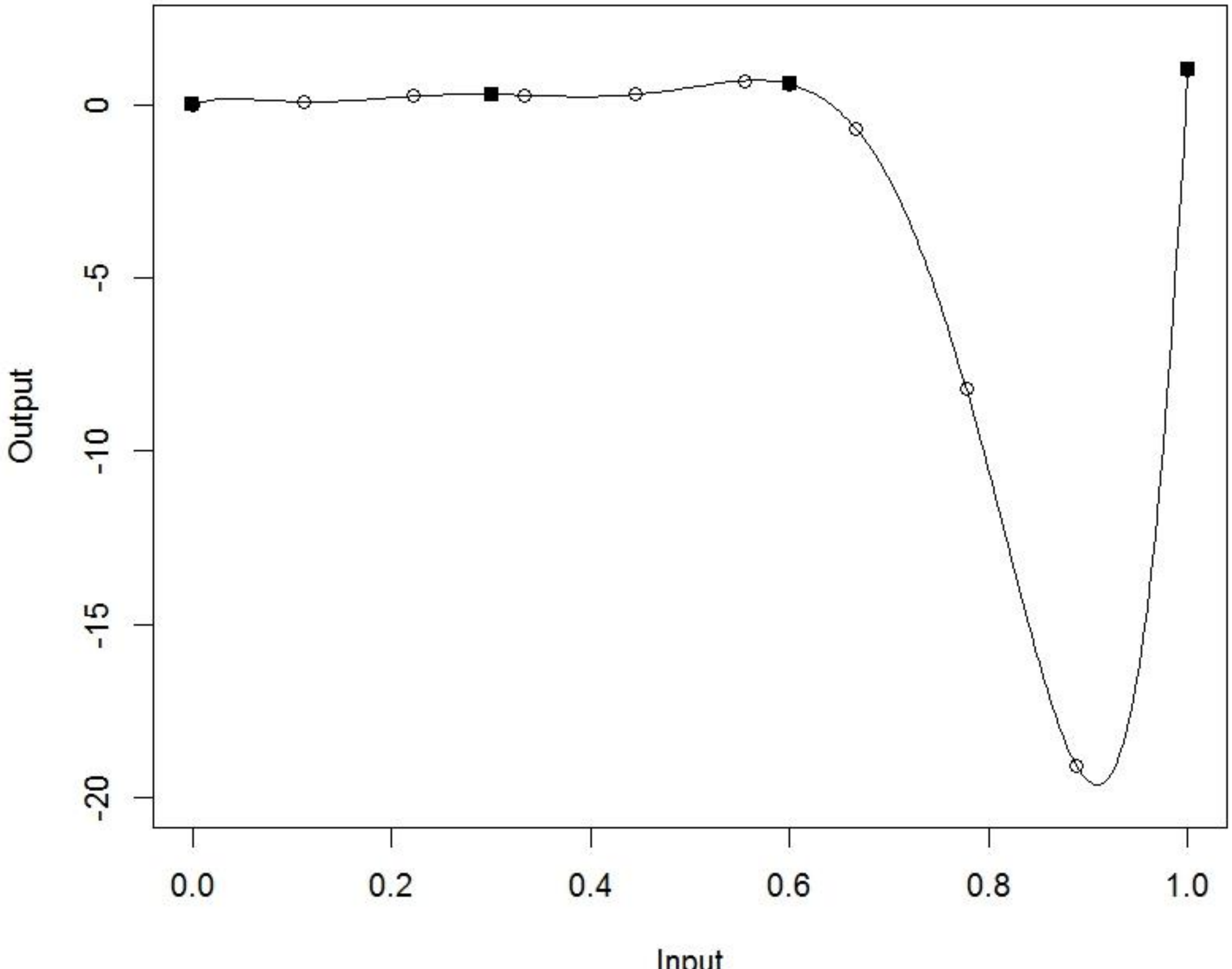
Getting the initial range right

UNIVERSITY OF LEEDS



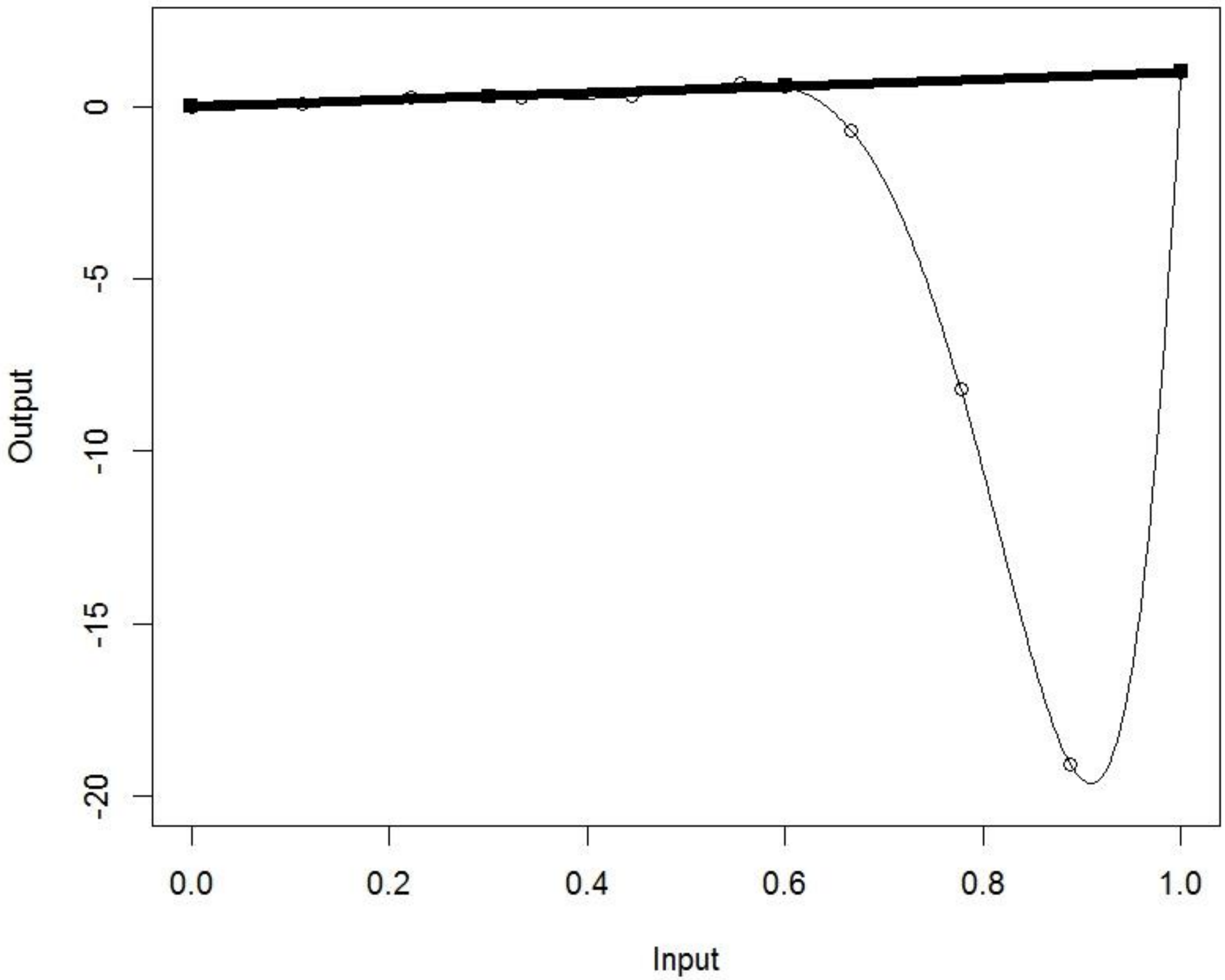
Getting the initial range right

UNIVERSITY OF LEEDS



Getting the initial range right

UNIVERSITY OF LEEDS



Eliciting ranges

UNIVERSITY OF LEEDS

For some parameters, the specification of parameter ranges could be simple due to the input being interesting over its full range.

For others, a 95-99% credible interval will usually suffice.

99% credible interval:

Can you specify a range for this input such that you believe there is a 1% chance that the true value for the input falls outside the range?

Cromwell's rule

In a letter to the General Assembly of the Church of Scotland on 5 August 1650:

"I beseech you, in the bowels of Christ, think it possible that you may be mistaken."

What now?

Once the ranges for all parameters have been elicited, we can

- specify a training design over that part of input space,
- build an emulator using that training design,
- run basic screening and exploratory sensitivity analyses.

What now?

UNIVERSITY OF LEEDS

Once the ranges for all parameters have been elicited, we can

- specify a training design over that part of input space,
- build an emulator using that training design,
- run basic screening and exploratory sensitivity analyses.

BUT ranges aren't enough for a careful Bayesian treatment when we want to do probabilistic sensitivity analyses or calibration.

Eliciting a distribution

UNIVERSITY OF LEEDS

We can ask an expert for judgements about individual probabilities, but what about a probability distribution?

If the quantity of interest is continuous, the expert is effectively making **infinitely** many judgements.

Of course, we are only interested in finding a distribution that is **good enough**.

Heuristics and biases

Adjustment and anchoring:

judgements are anchored at some starting value
and adjust outwards, usually insufficiently.

There is some evidence to suggest that, if I asked you to write down the last few digits of your phone number, then your subsequent judgements would be influenced by this.

Kadane and Wolfson (1998). Experiences in elicitation. *The Statistician*, **47**, 3-19.

The bisection method

Specify a value such that you think it is equally likely the parameter falls below or above it.

Assessing quantities

UNIVERSITY OF LEEDS

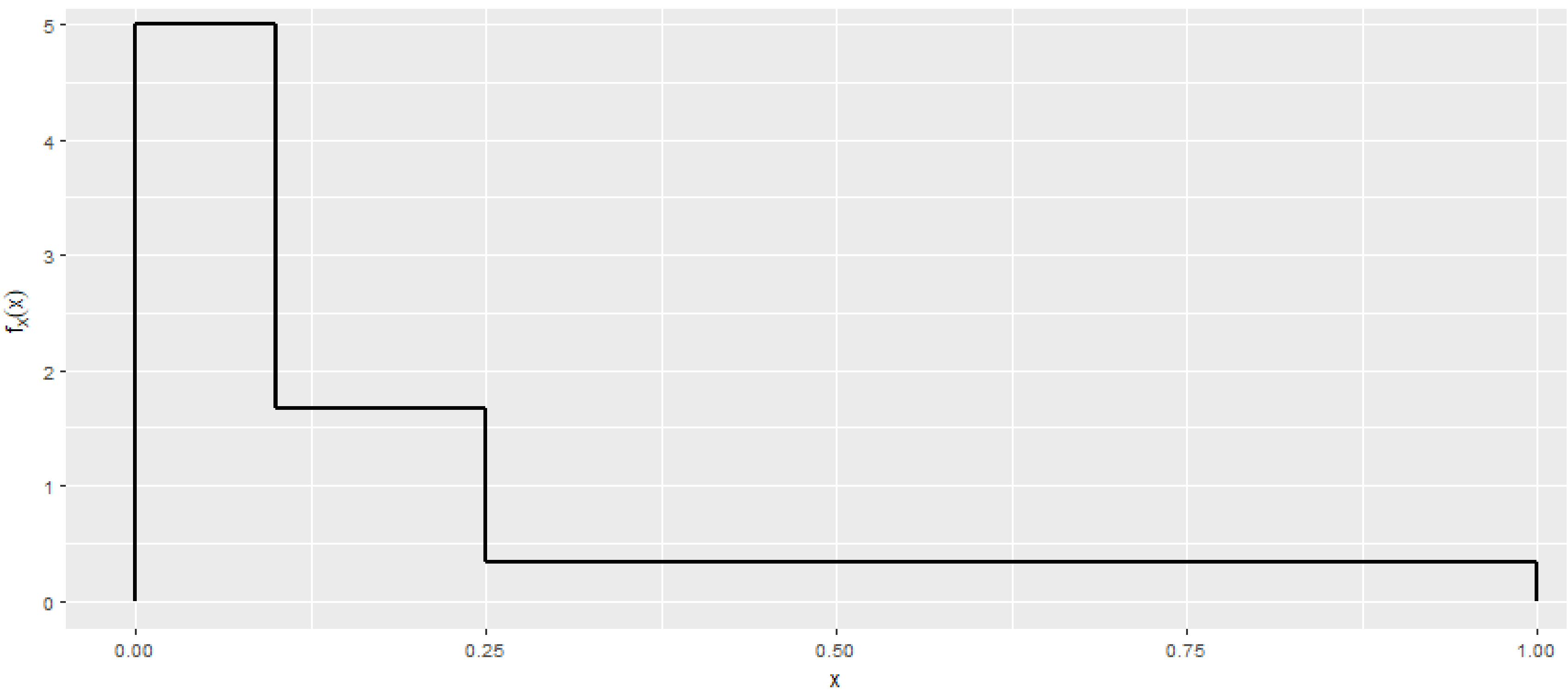
The bisection method

Specify a value such that you think it is equally likely the parameter falls between 0.1 and that number or above it.

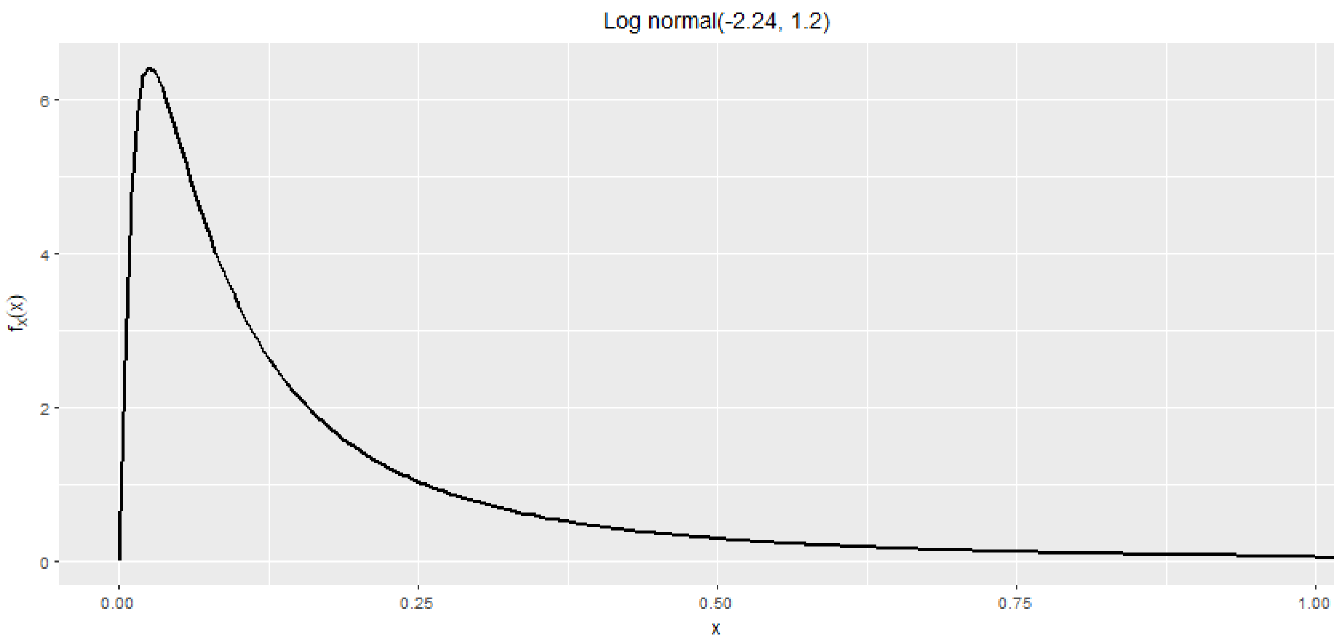
Assessing quantities

UNIVERSITY OF LEEDS

The bisection method



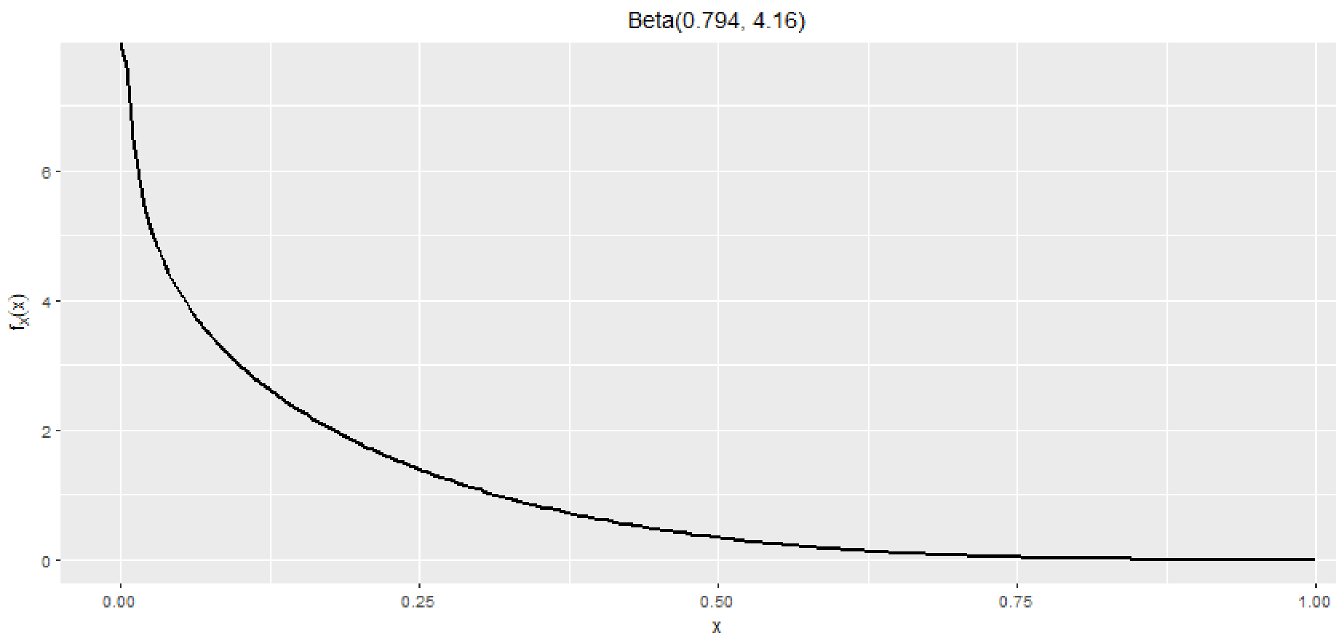
Assessing quantities



We can then fit a distribution by choosing an appropriate parametric family and using a least squares approach to determine its parameters:

$$\text{Minimise} \sum (\text{Judged statistic} - \text{fitted statistic})^2$$

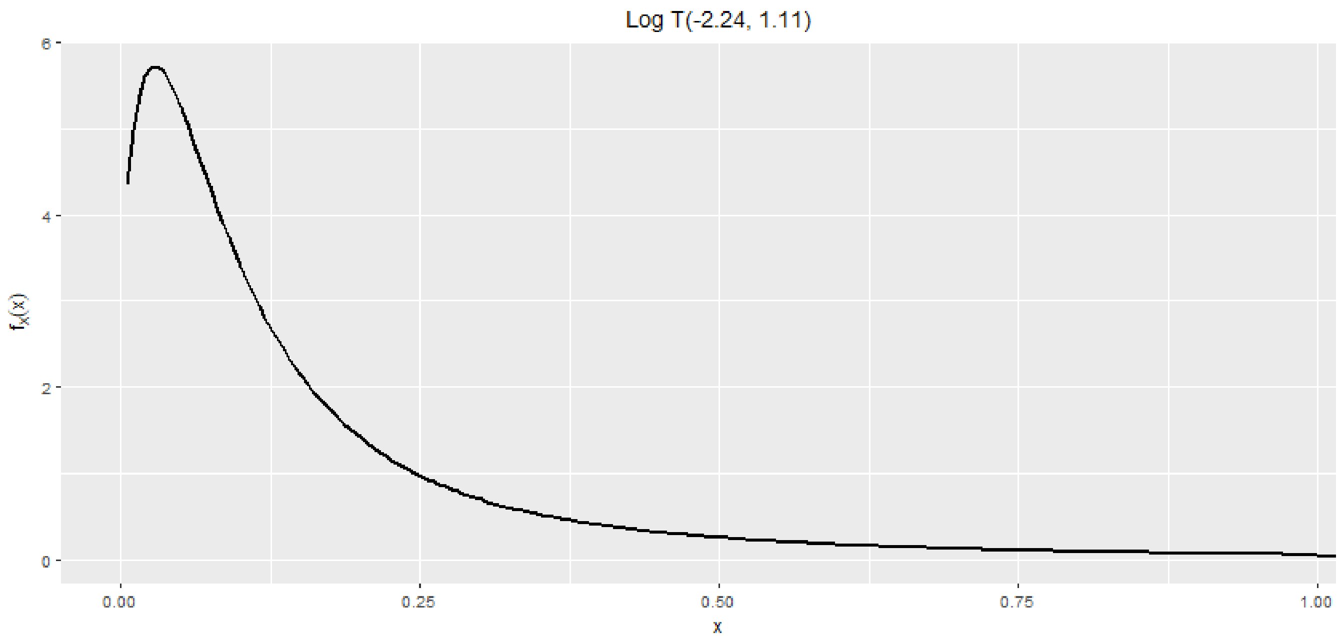
Assessing quantities



We can then fit a distribution by choosing an appropriate parametric family and using a least squares approach to determine its parameters:

$$\text{Minimise} \sum (\text{Judged statistic} - \text{fitted statistic})^2$$

Assessing quantities



We can then fit a distribution by choosing an appropriate parametric family and using a least squares approach to determine its parameters:

$$\text{Minimise} \sum (\text{Judged statistic} - \text{fitted statistic})^2$$

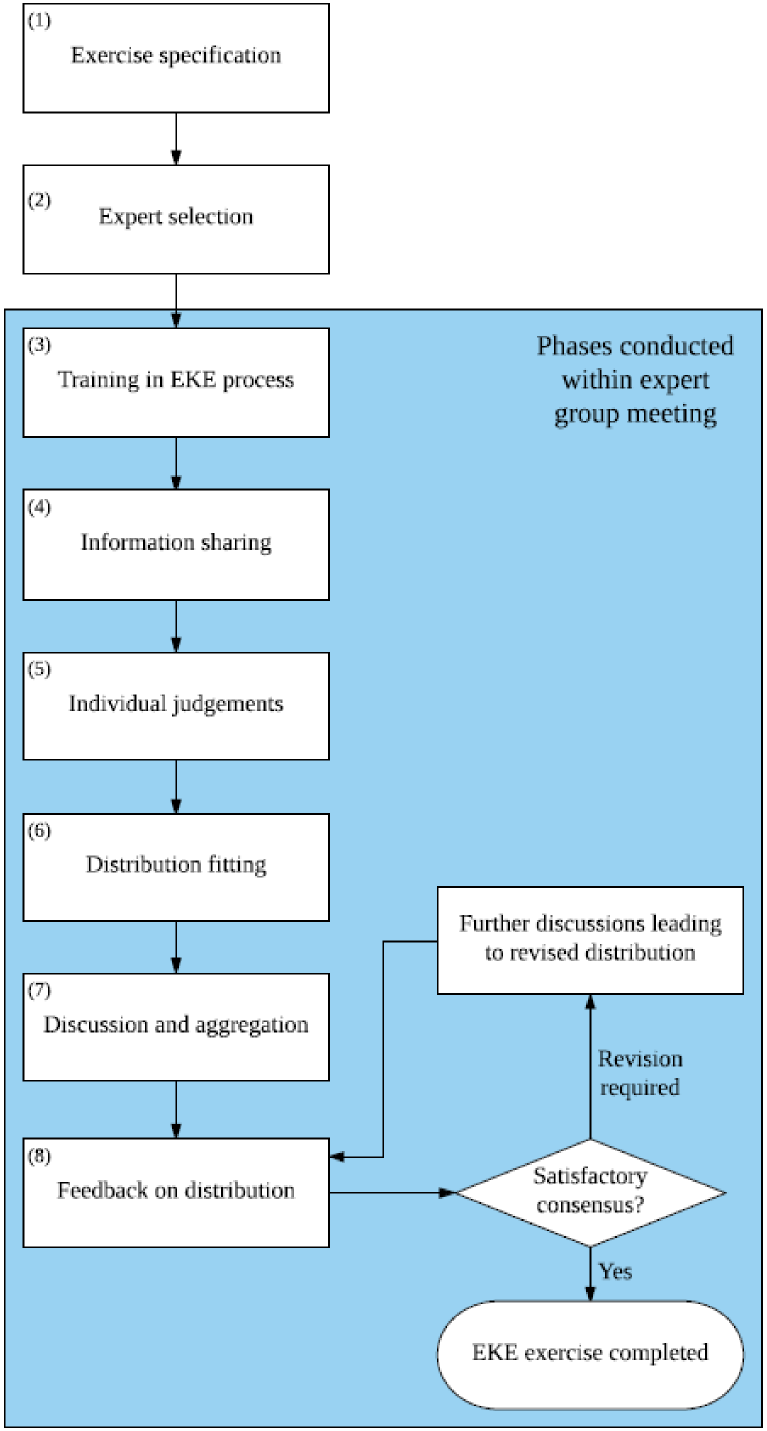
Heuristics and biases

Availability: assessors link their probabilities to the frequency with which they can recall an event.

Things that are more memorable are deemed more probable.

For example, high profile train accidents lead people to imagine rail travel is more risky than it really is.

Kadane and Wolfson (1998). Experiences in elicitation. *The Statistician*, **47**, 3-19.



This is the Sheffield elicitation framework (SHELF) approach to eliciting judgements about single continuous variables.

There are two groups involved: experts and facilitators.

The opportunities for **feedback** and **revision** are very important.

UNIVERSITY OF LEEDS

Roles in an elicitation exercise

The expert(s):

often chosen by the problem owner;

their knowledge encapsulates years of experience and the synthesis of many strands of evidence;

able to communicate information about their knowledge.

Roles in an elicitation exercise

UNIVERSITY OF LEEDS

The facilitator(s):

utilise a protocol that will enable the experts to make judgements about the quantities of interest;

facilitates discussions and records discussions;

provides meaningful feedback.

An extra issue

UNIVERSITY OF LEEDS

My input is called:

"The rate at which cold gas is converted into stars"*

* Taken from Vernon et al. (2014), Galaxy Formation: Bayesian History Matching for the Observable Universe.

An extra issue

UNIVERSITY OF LEEDS

My input is called:

"The rate at which cold gas is converted into stars"*

Is this something that exists in reality?

Is its role in reality the same as its role in the model?

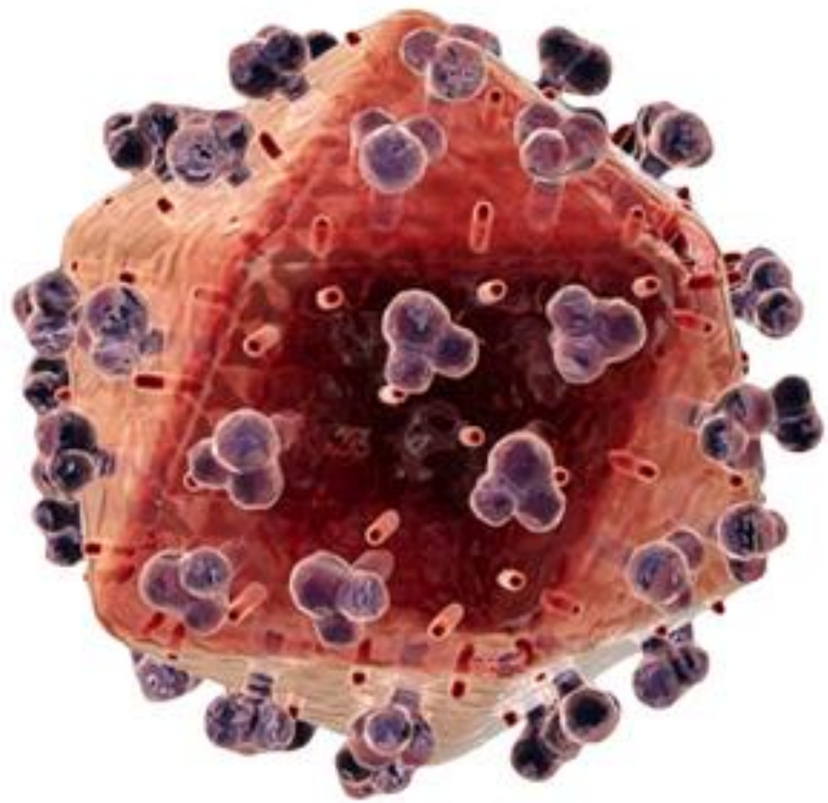
If we knew the true value of this parameter in reality, would it be the best value to use in our imperfect model?

* Taken from Vernon et al. (2014), Galaxy Formation: Bayesian History Matching for the Observable Universe.

Yellow fever

UNIVERSITY OF LEEDS

Yellow fever is an acute viral disease transmitted by mosquitoes.

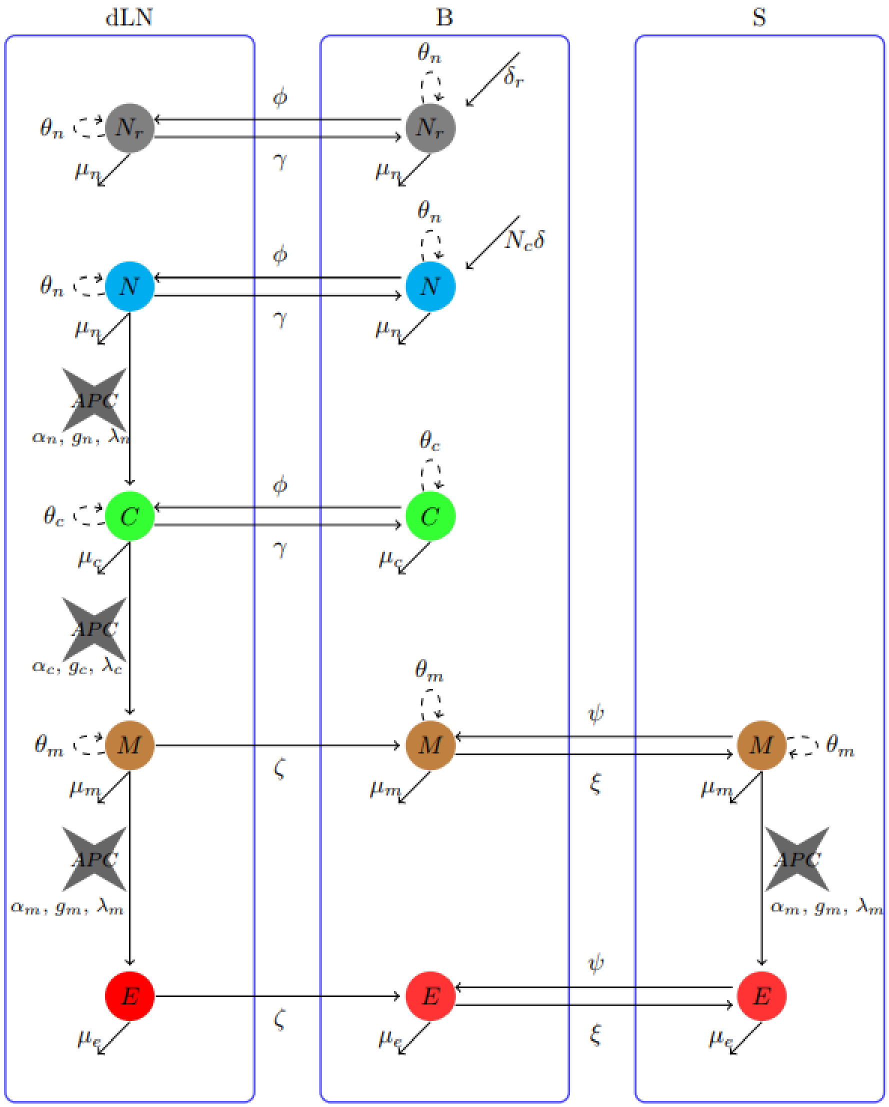


In 2013 alone, it was estimated to have caused over 30,000 deaths.

Fortunately, there is an extremely effective vaccination, but the immune system response is not well understood.

Mathematical model of infection

UNIVERSITY OF LEEDS



Mathematical model of infection

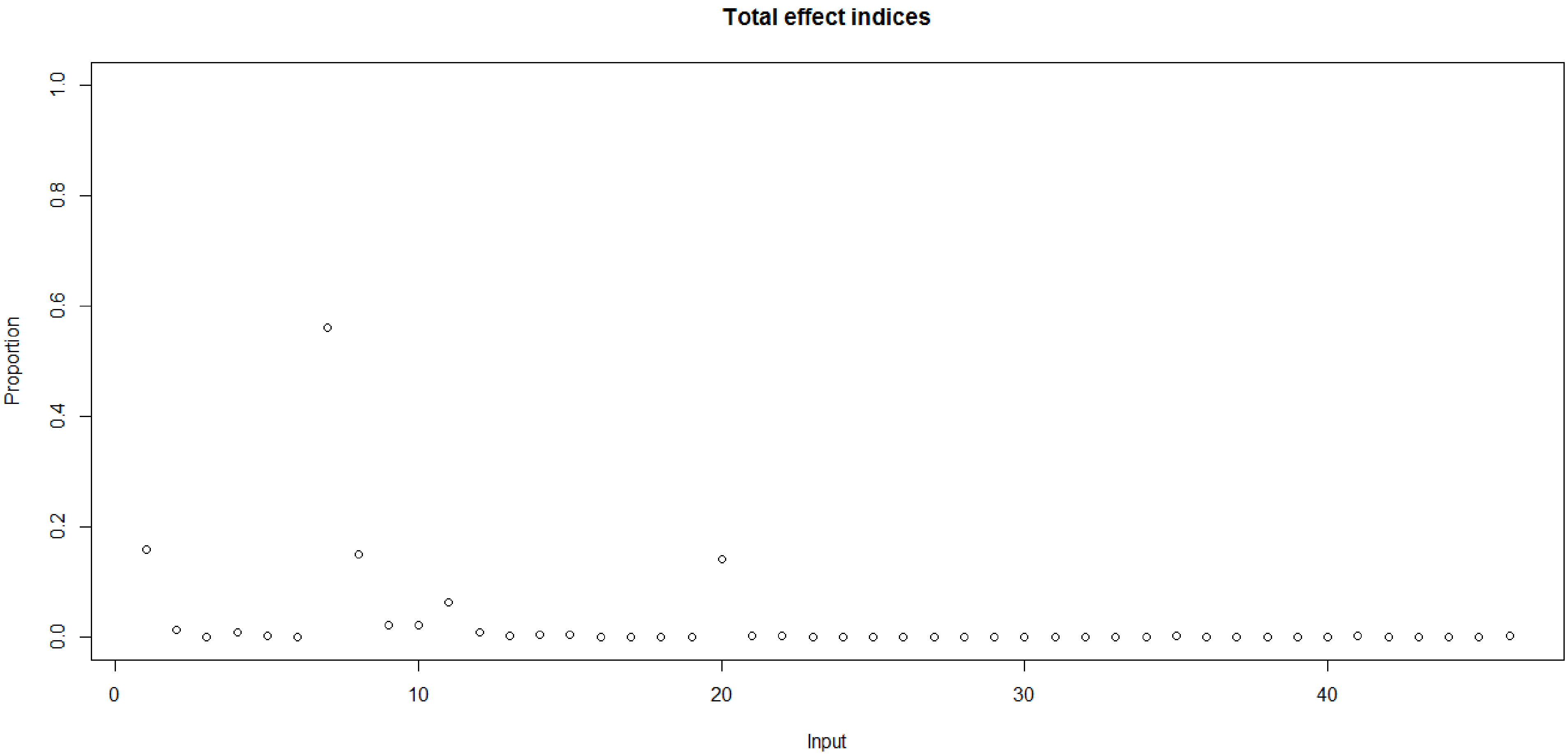
UNIVERSITY OF LEEDS

Parameter	Definition	Value
κ_n	Carrying capacity per clonotype of naive cells	1,940 cells
	Carrying capacity of specific naive cells	$N_c \times \kappa_n$
κ_r	Carrying capacity of non-specific naive cells	$(N_R - N_c) \times \kappa_n$
κ_c	Carrying capacity per clonotype of specific central memory cells	87,750 cells
κ_m	Carrying capacity per clonotype of specific effector memory cells	87,750 cells
θ_n	Division rate of naive cells	5.63×10^{-4} per day
θ_c	Division rate of central memory cells	6.49×10^{-3} per day
θ_m	Division rate of effector memory cells	6.49×10^{-3} per day
μ_n	Death rate of naive cells	4.46×10^{-4} per day
μ_c	Death rate of central memory cells	3.67×10^{-3} per day
μ_m	Death rate of effector memory cells	3.67×10^{-3} per day
μ_e	Death rate of effector cells	3.57×10^{-2} per day

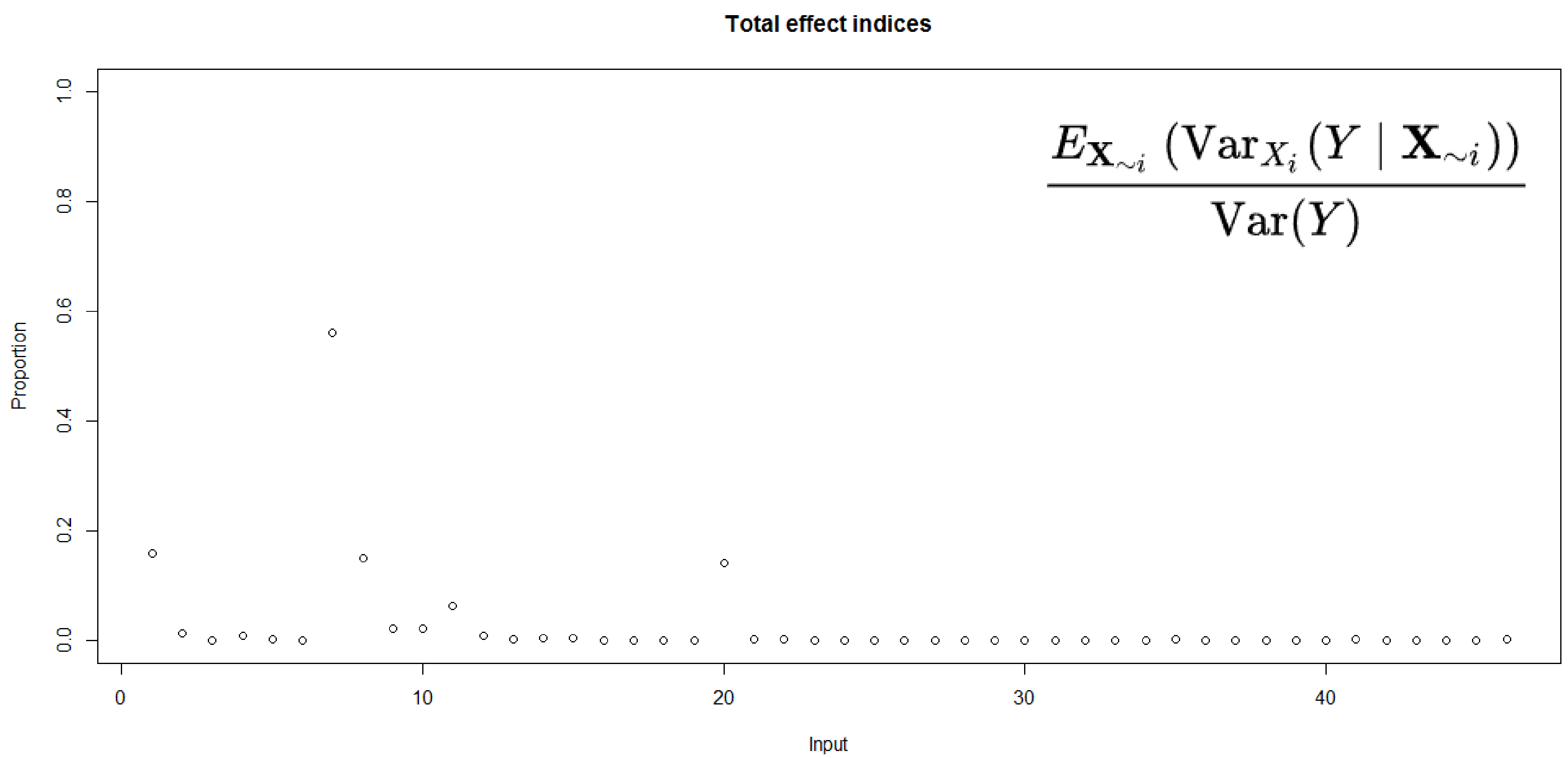
The modellers had started with just values taken from the literature.

We first worked to get plausible ranges.

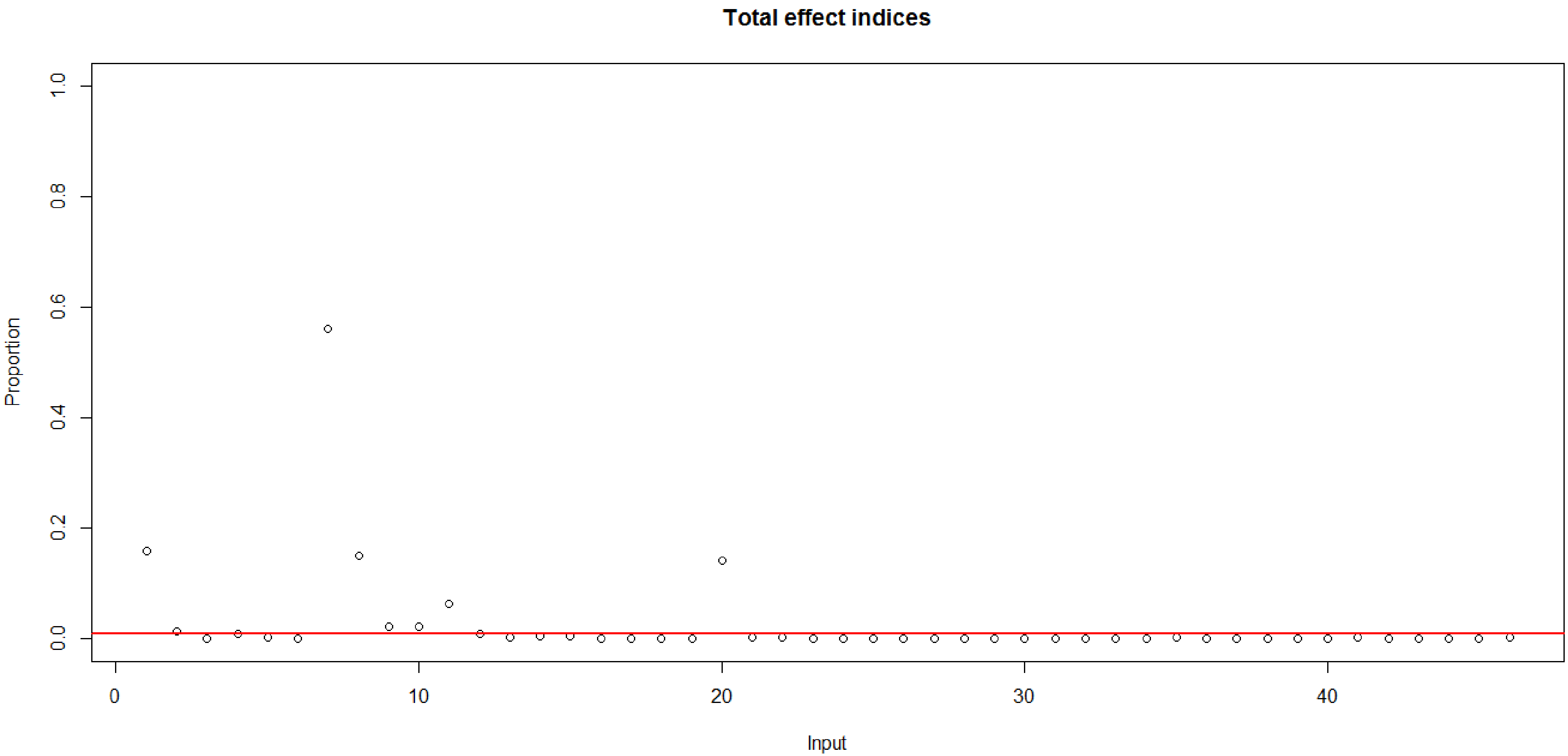
Sensitivity indices



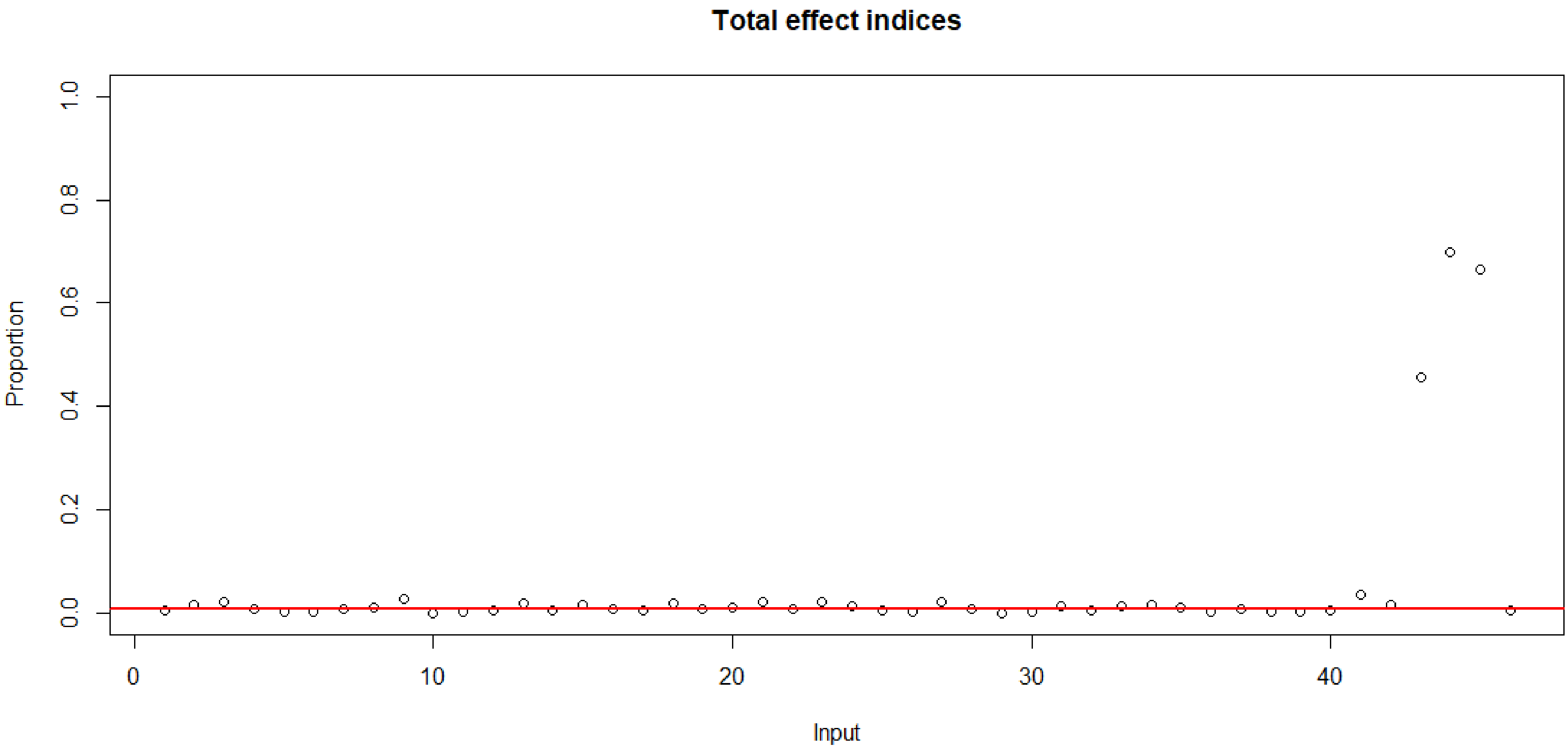
Sensitivity indices



Sensitivity indices

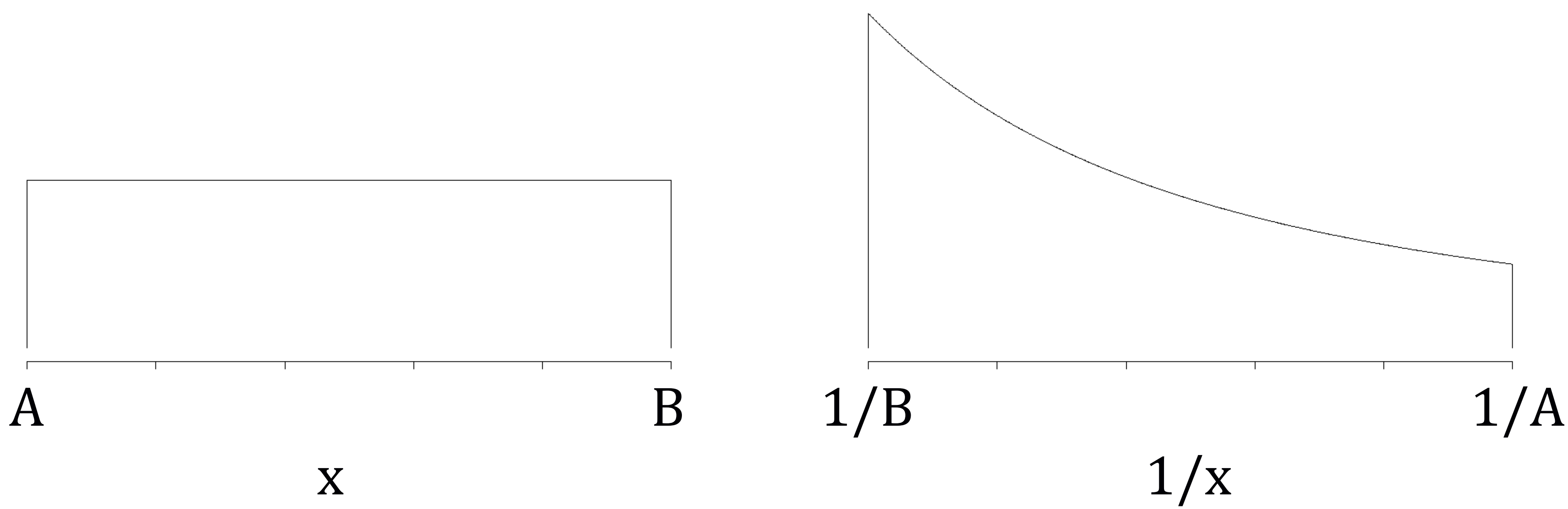


Sensitivity indices



Lazy uniforms

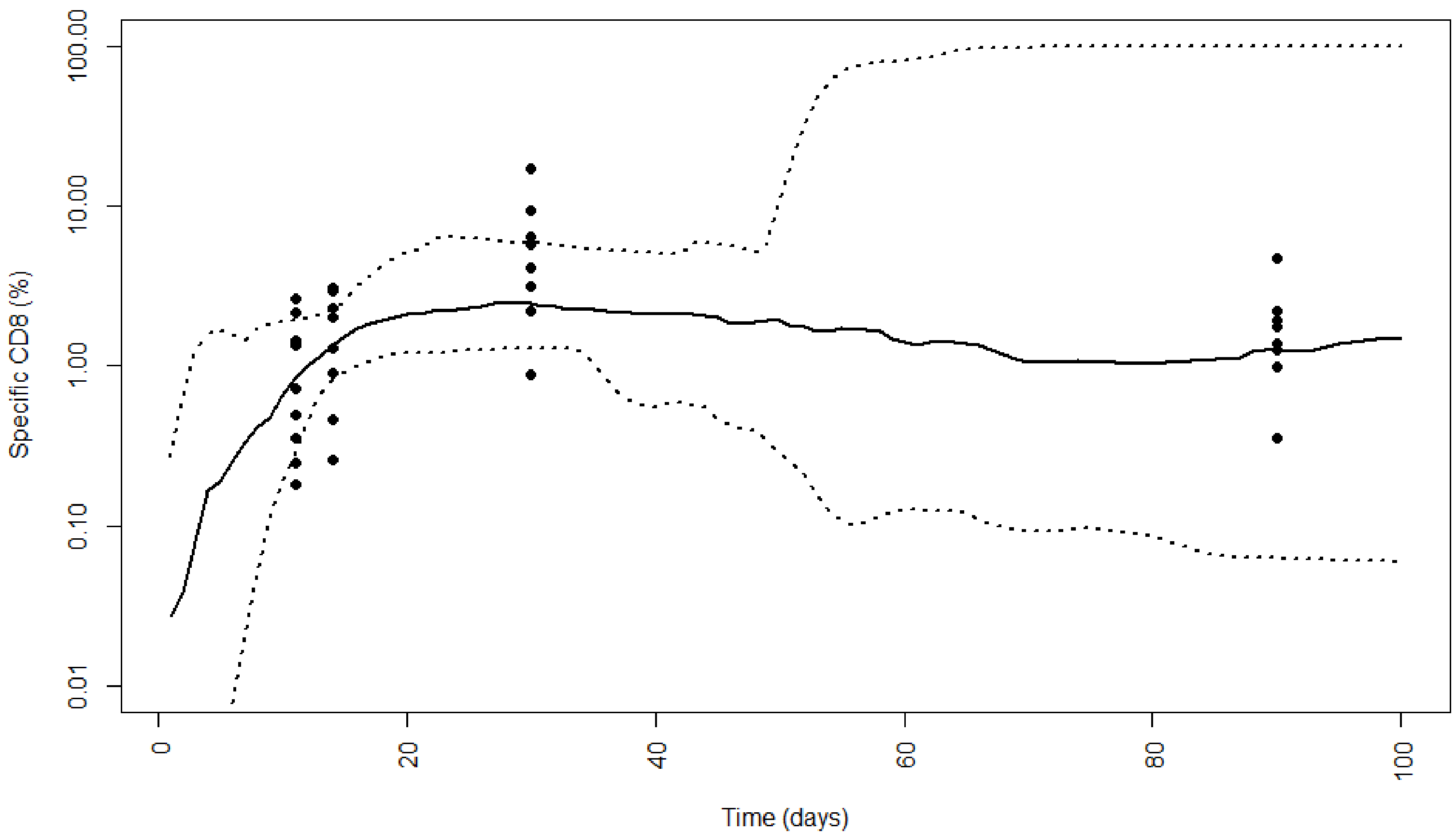
$X \sim \text{Uniform}(A,B)$



Lazy uniforms

UNIVERSITY OF LEEDS

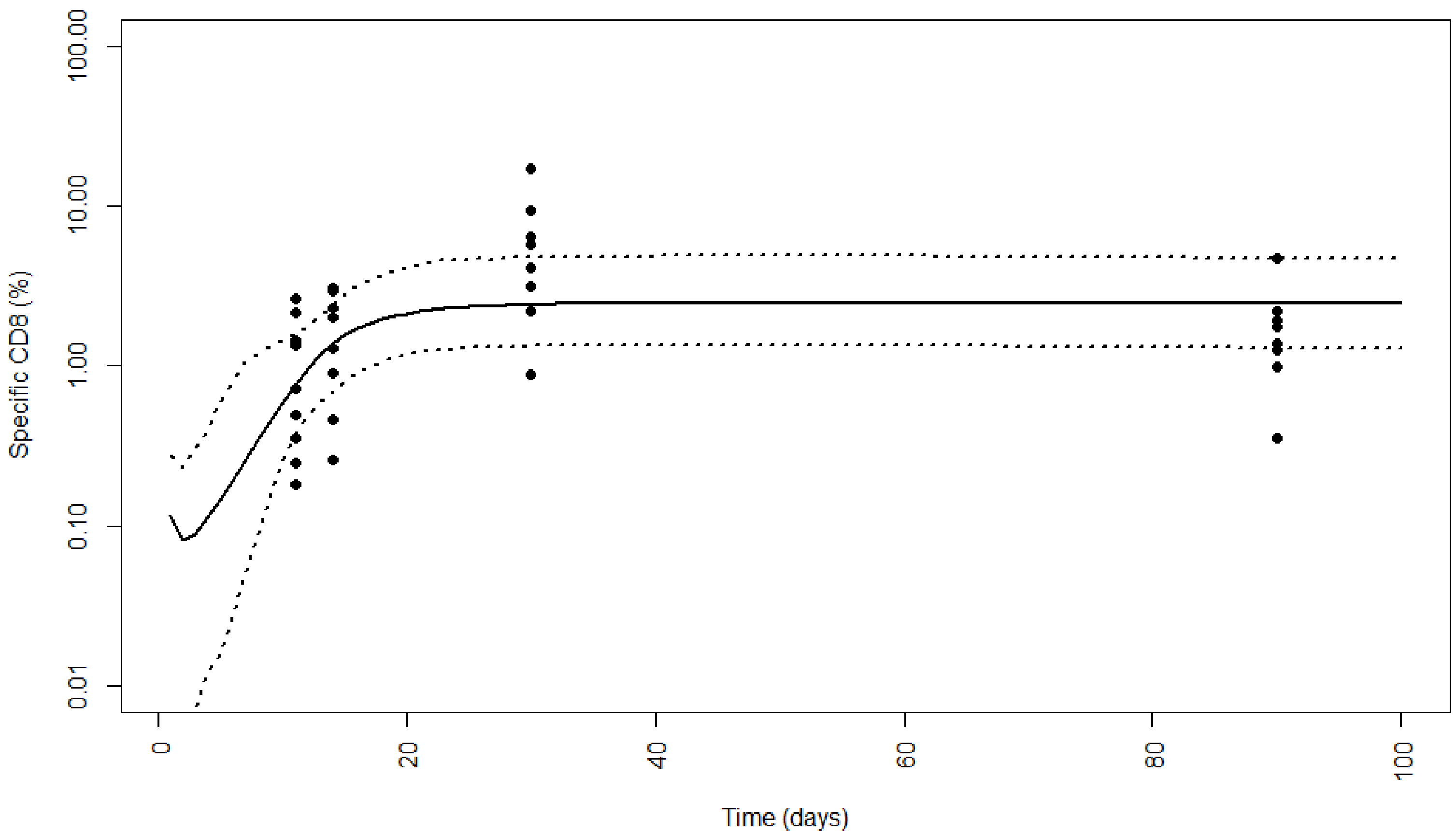
Uniform over average rates



Lazy uniforms

UNIVERSITY OF LEEDS

Uniform over average times



Final remarks

UNIVERSITY OF LEEDS

Carefully performed selection of parameter ranges can save time and effort.

Expert elicitation should be a key part of any Bayesian analysis (especially, if the prior is influential).

Many examples of expert elicitation in the literature for many types of input (and covering Sheffield approach).

References

UNIVERSITY OF LEEDS

Modern treatments:

O' Hagan, A., *et al.* (2006). *Uncertain Judgements: Eliciting Expert Probabilities*. John Wiley and Sons.

Dias, L.C., Morton, A. and Quigley, J., 2018. *Elicitation: the science and art of structuring judgement*. Springer.

Tools:

SHELF documentation available from tonyohagan.co.uk and SHELF package available for R.