Bayesian inference

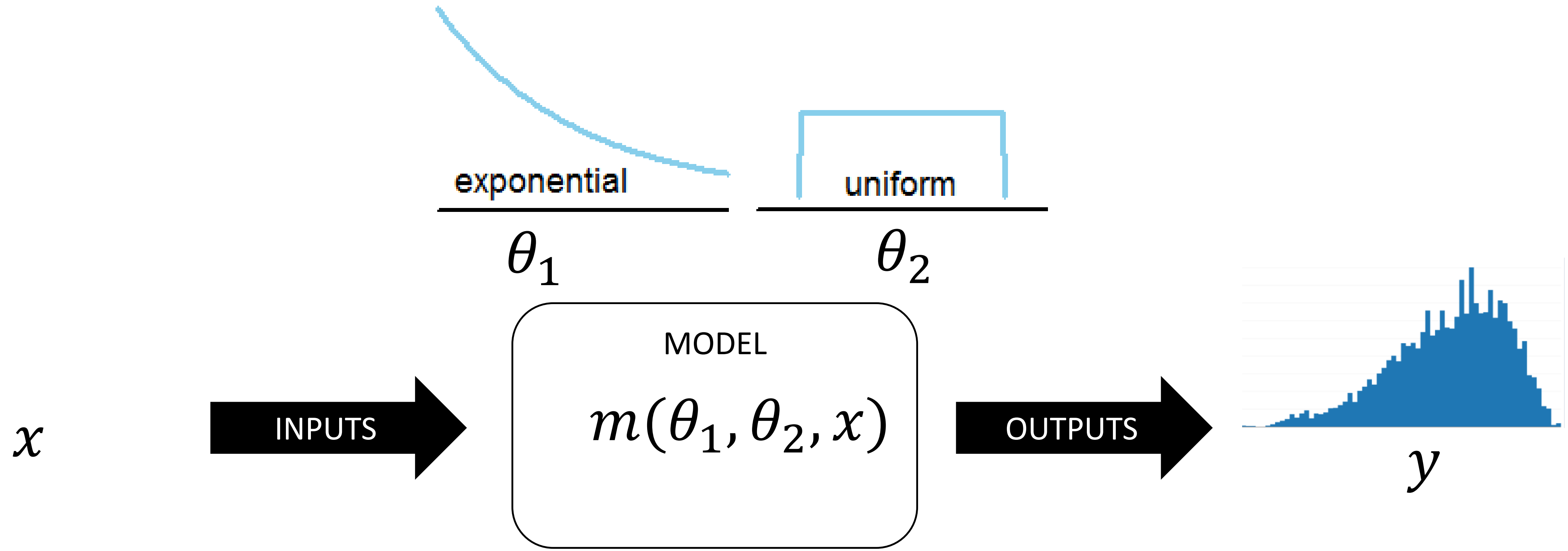
REACH seminar on data analysis 27 Oct 2017

Ullrika Sahlin ullrika.sahlin@cec.lu.se

Lund University Centre for Environmental and Climate Research

Questions

1. Have you ever done a Monte Carlo simulation? How did you interpret the probability distributions?



Questions

1. Have you ever done a Monte Carlo simulation? How did you interpret the probability distributions?

2. What type of data analysis do you associate with Bayesian inference?

Data analysis - for what purpose

Purpose	**Bayesian inference**
Hypothesis testing	Yes
Estimation	Yes
Assessment	Yes
Quantification of uncertainty	Yes
Decision analysis	Yes

Data analysis – different situations

$$m(\theta, x) \to y$$

	Parametric inference	Predictive inference
Data rich		
Data sparse		

	Parametric inference	Predictive inference
Data rich	Bayesian Networks in supervised learning	
Data sparse		Bayesian Network as an influence diagram

	Parametric inference	Predictive inference
Data rich	Bayesian Analysis	Bayesian Cost-Benefit analysis
Data sparse	Bayesian analysis with informed priors	Bayesian Decision analysis

	Parametric inference	Predictive inference
Data rich	Bayesian Analysis	Bayesian Cost-Benefit analysis
Data sparse	Bayesian analysis with informed priors	Bayesian Decision analysis
No data	Expert elicitation*	Forward Monte Carlo simulation**

Uncertainty

Aleatory and epistemic

- Randomness vs lack of knowledge
- Non-reducible vs reducible

- Epistemic uncertainty measured by subjective probability
- Aleatory uncertainty measured by relative frequency

$$m(\theta, x) \rightarrow y$$

- Parameters θ are uncertain
- The model structure m is fixed*
- The model can express aleatory uncertainty

Bayesian analysis

A **probability model** of new data given the system and the data generating processes

Prior – uncertainty in parameters before considering available data

Likelihood – the probability mass of available data given a specific combination of parameter values

Posterior – uncertainty in parameters after learning from data using Bayesian updating

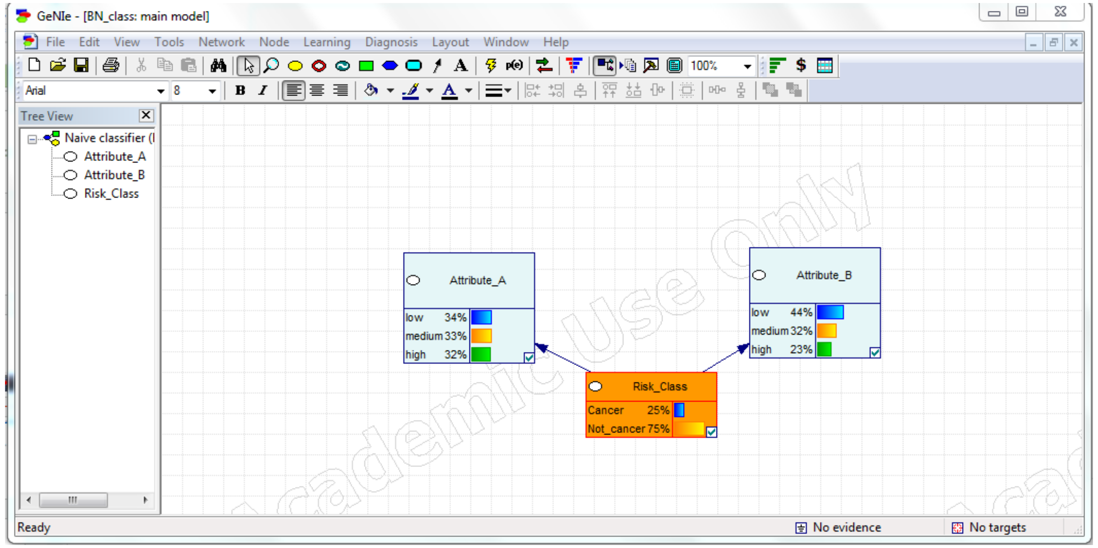
Predictive posterior – uncertainty in future system states or observables

Bayesian Networks

- Probabilistic graphs
- A model to derive the likelihood to learn a network structure
- Decision analysis (influence diagram) – GENIE, HUGIN, NETICA
- Usually categorical or discrete nodes
- Gaussian networks, mixture of Gaussian and categorical nodes
- Expert driven, data driven or both
- Uncertainty in either state variables or parameters

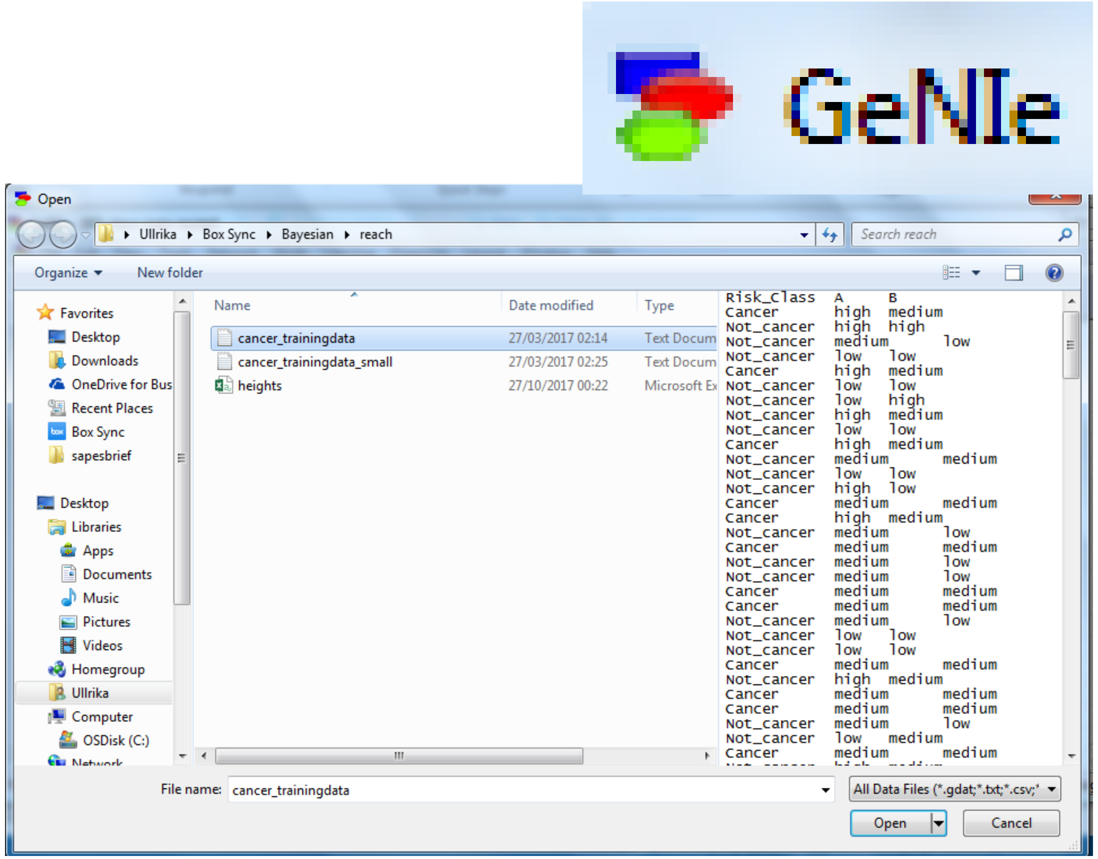
Example of Bayesian Network

- Bayes Naive Classifier
- Parameters assigned by experts



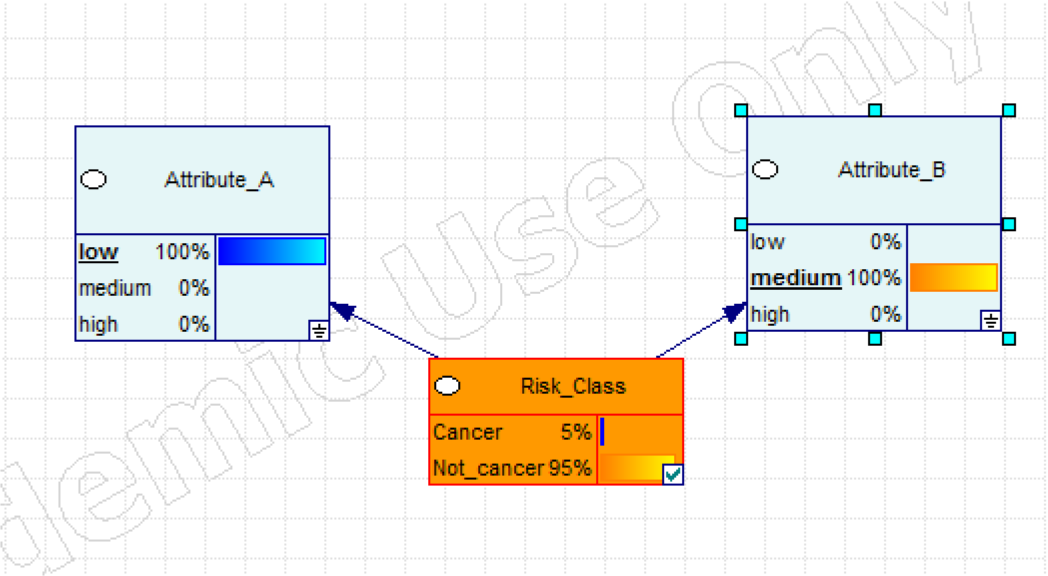
Example of Bayesian Network

- Load training data
- Learn parameters
- One can choose to rely on priors (good when data is small)



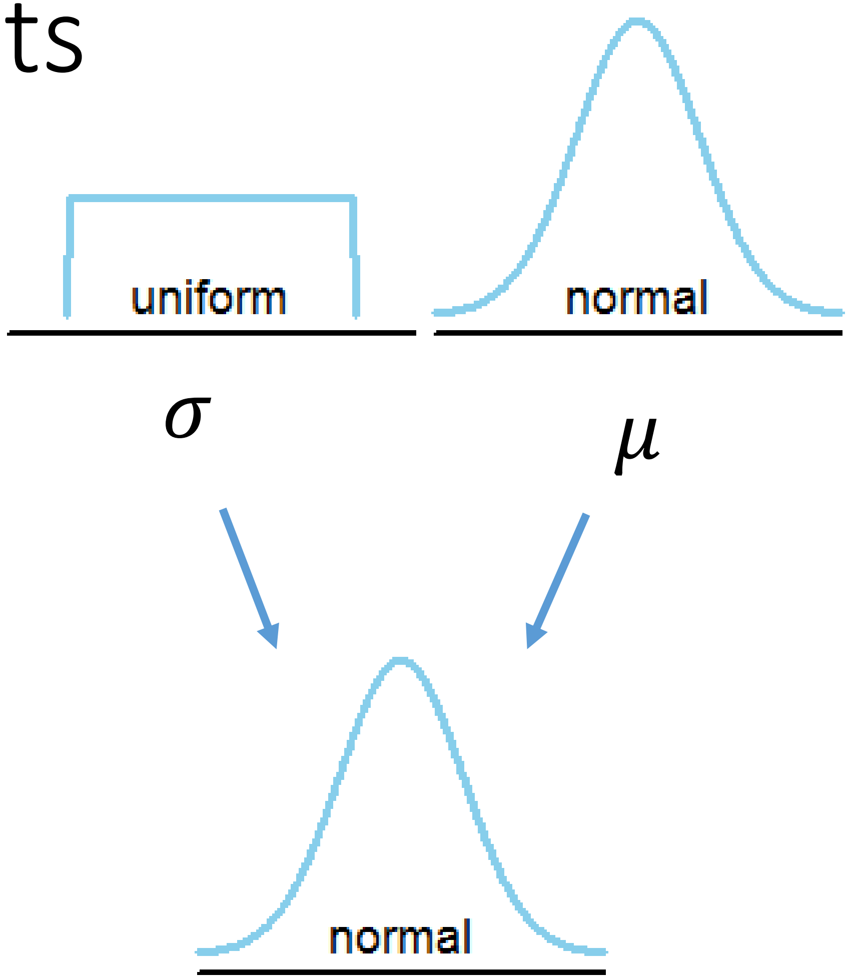
Example of Bayesian Network

- The model can be used to assess the probability to have cancer given known values on an attribute
- Decison support tool



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

The height of US presidents



$$h \sim N(\mu, \sigma)$$

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

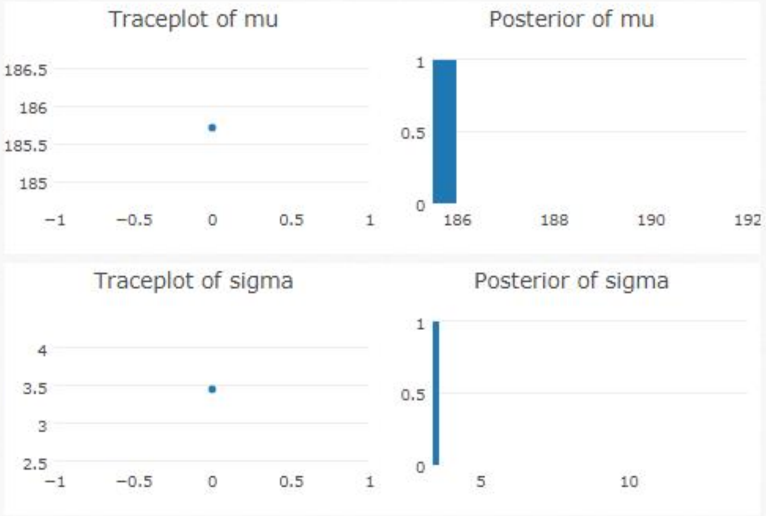
Sampling from the posterior of the height model

Here is a motivating example: Say that you have the heights of the last ten American presidents...

```
// The heights of the last ten American presidents in cm, from Kennedy to Obama
var heights = [183, 192, 182, 183, 177, 185, 188, 188, 182, 185];
```

... and that you would like to fit a Bayesian model assuming a Normal distribution to this data. Well, you can do that **right now** by clicking “Start sampling” below! This will run an MCMC sampler in your browser implemented in JavaScript.

Start sampling | Stop sampling and plot histograms | Clear samples



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

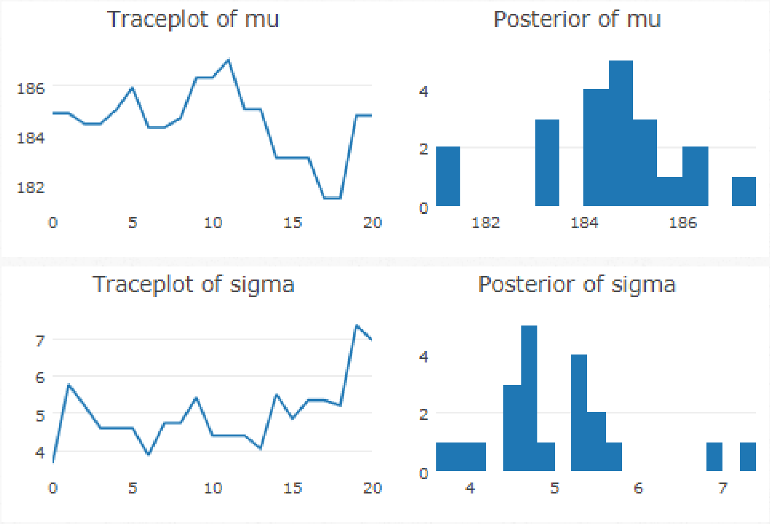
Sampling from the posterior of the height model

Here is a motivating example: Say that you have the heights of the last ten American presidents...

```
// The heights of the last ten American presidents in cm, from Kennedy to Obama
var heights = [183, 192, 182, 183, 177, 185, 188, 188, 182, 185];
```

... and that you would like to fit a Bayesian model assuming a Normal distribution to this data. Well, you can do that **right now** by clicking "Start sampling" below! This will run an MCMC sampler in your browser implemented in JavaScript.

Start sampling | Stop sampling and plot histograms | Clear samples



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

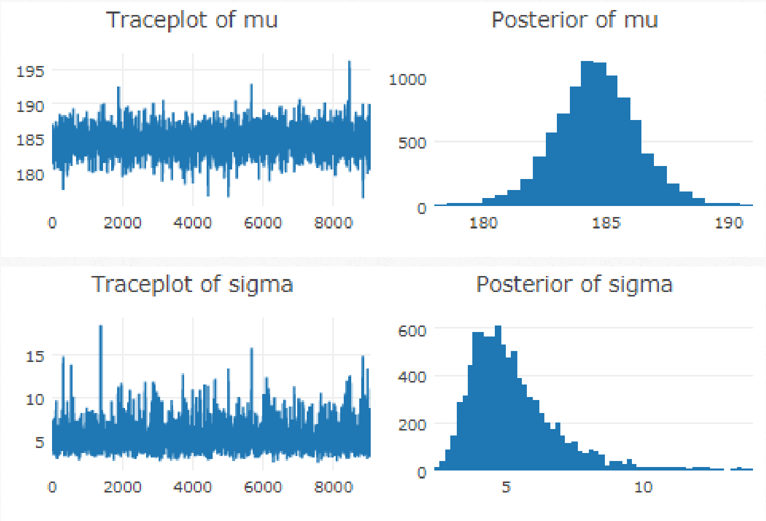
Sampling from the posterior of the height model

Here is a motivating example: Say that you have the heights of the last ten American presidents...

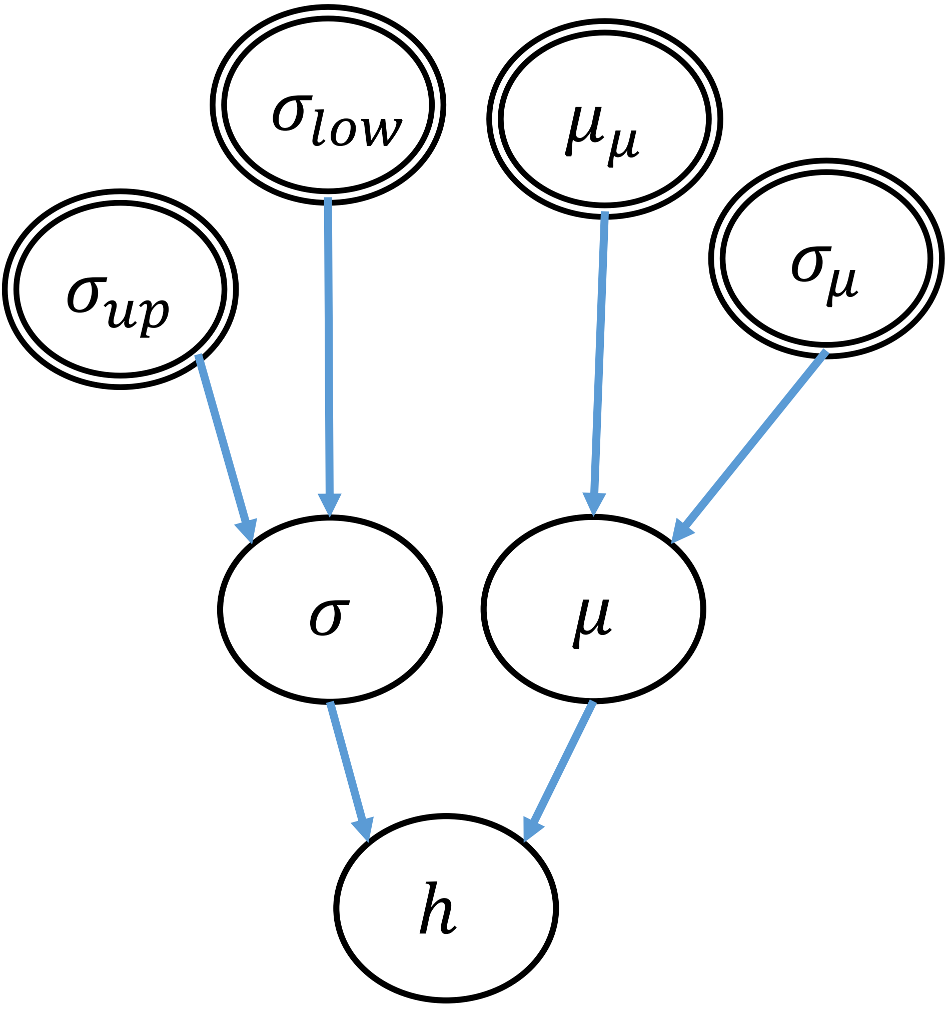
```
// The heights of the last ten American presidents in cm, from Kennedy to Obama
var heights = [183, 192, 182, 183, 177, 185, 188, 188, 182, 185];
```

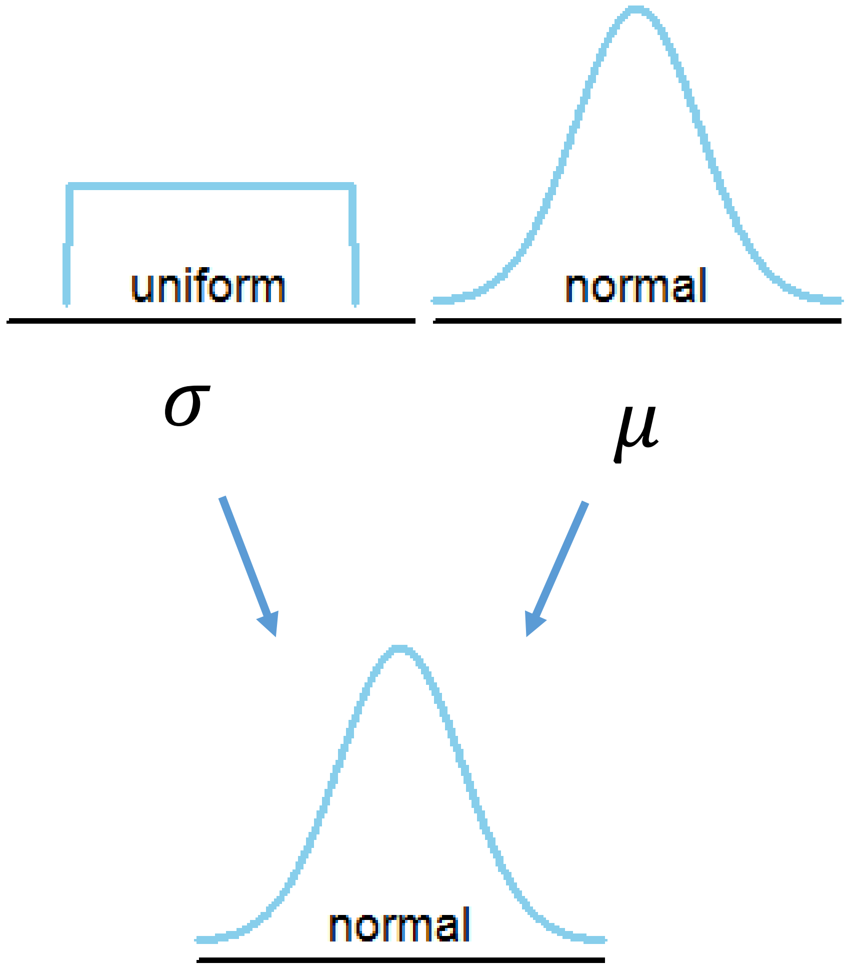
... and that you would like to fit a Bayesian model assuming a Normal distribution to this data. Well, you can do that **right now** by clicking “Start sampling” below! This will run an MCMC sampler in your browser implemented in JavaScript.

Start sampling | Stop sampling and plot histograms | Clear samples



Alternative ways to express the model





$$h \sim N(\mu, \sigma)$$

Advantage: likelihood can be built by conditional probaility models

Hypothesis testing

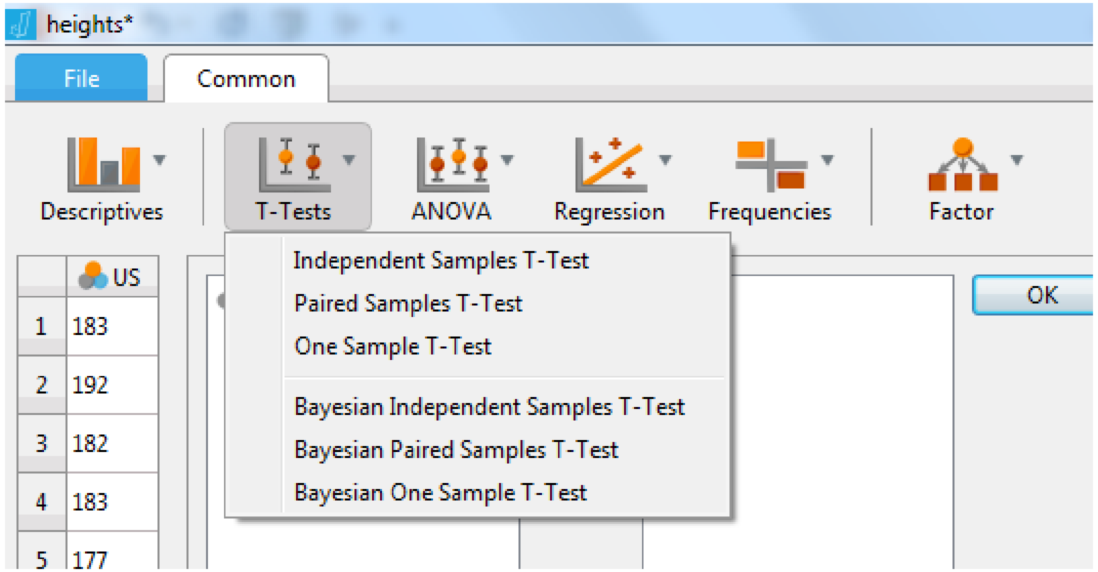
Bayes factor

Bayesian p-value

> 185

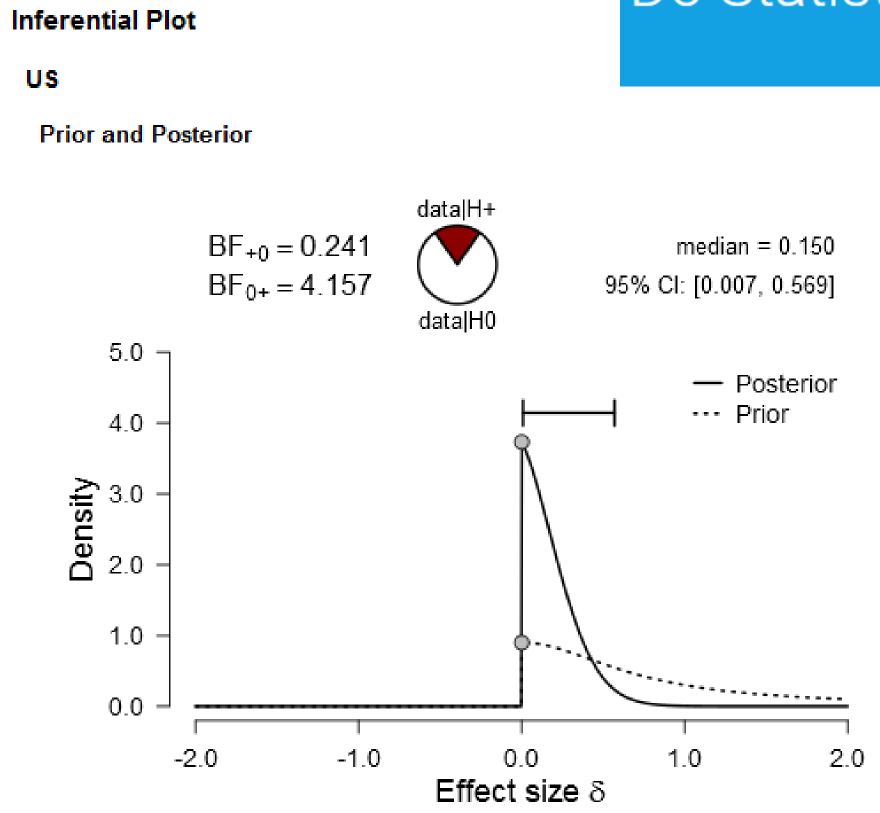
Hypothesis testing

- SPSS like environment
- Classical frequentic tests and Bayesian tests



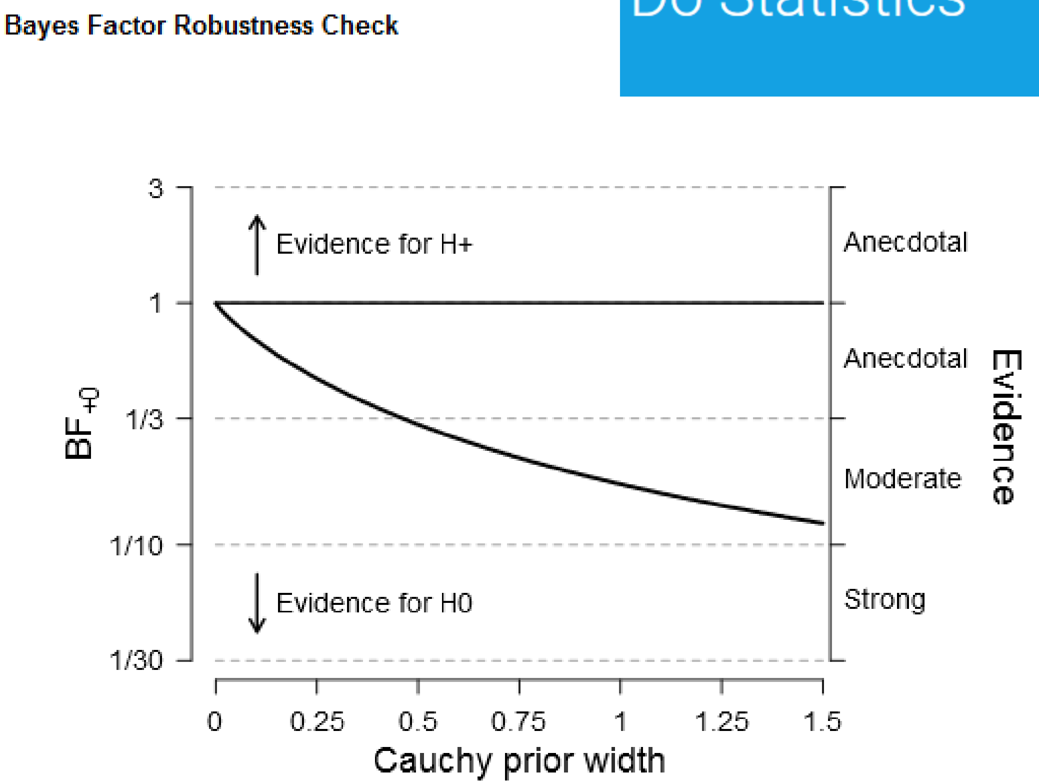
Hypothesis testing

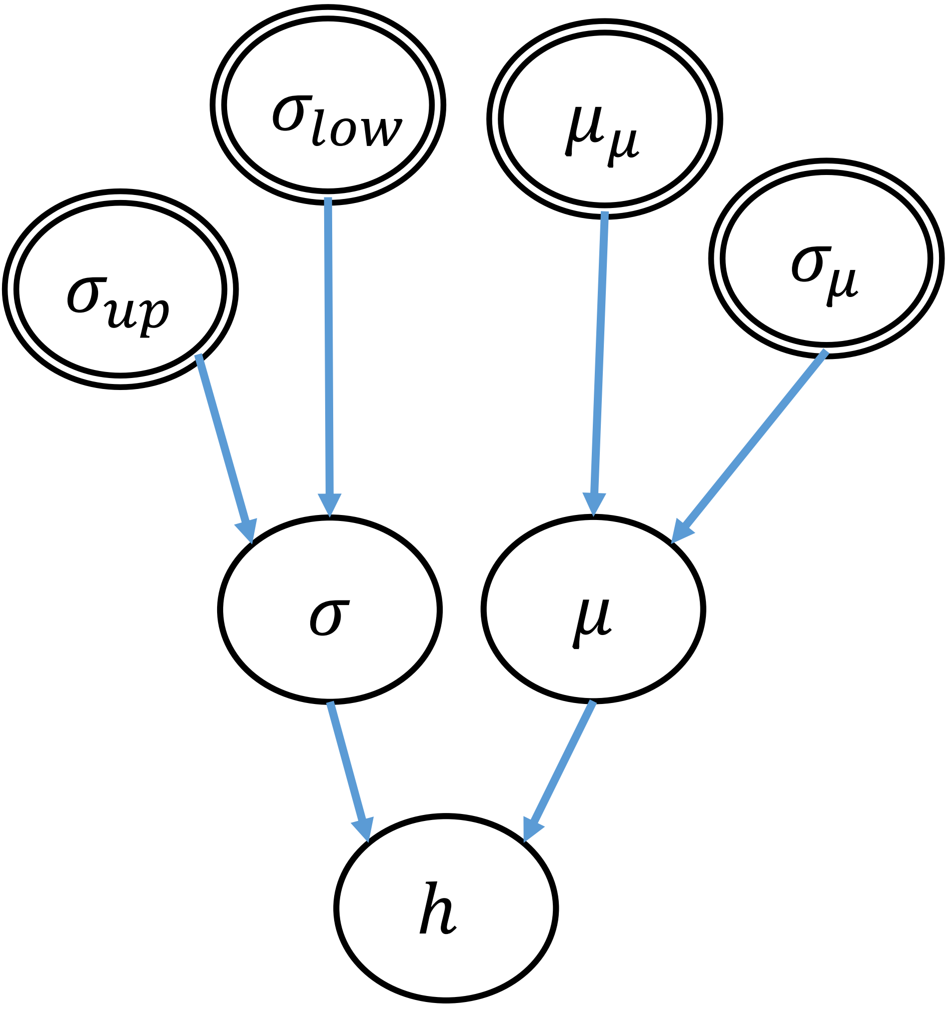
- Prior and posterior
- Credible interval
- Bayes factor
- Effect size (estimation)



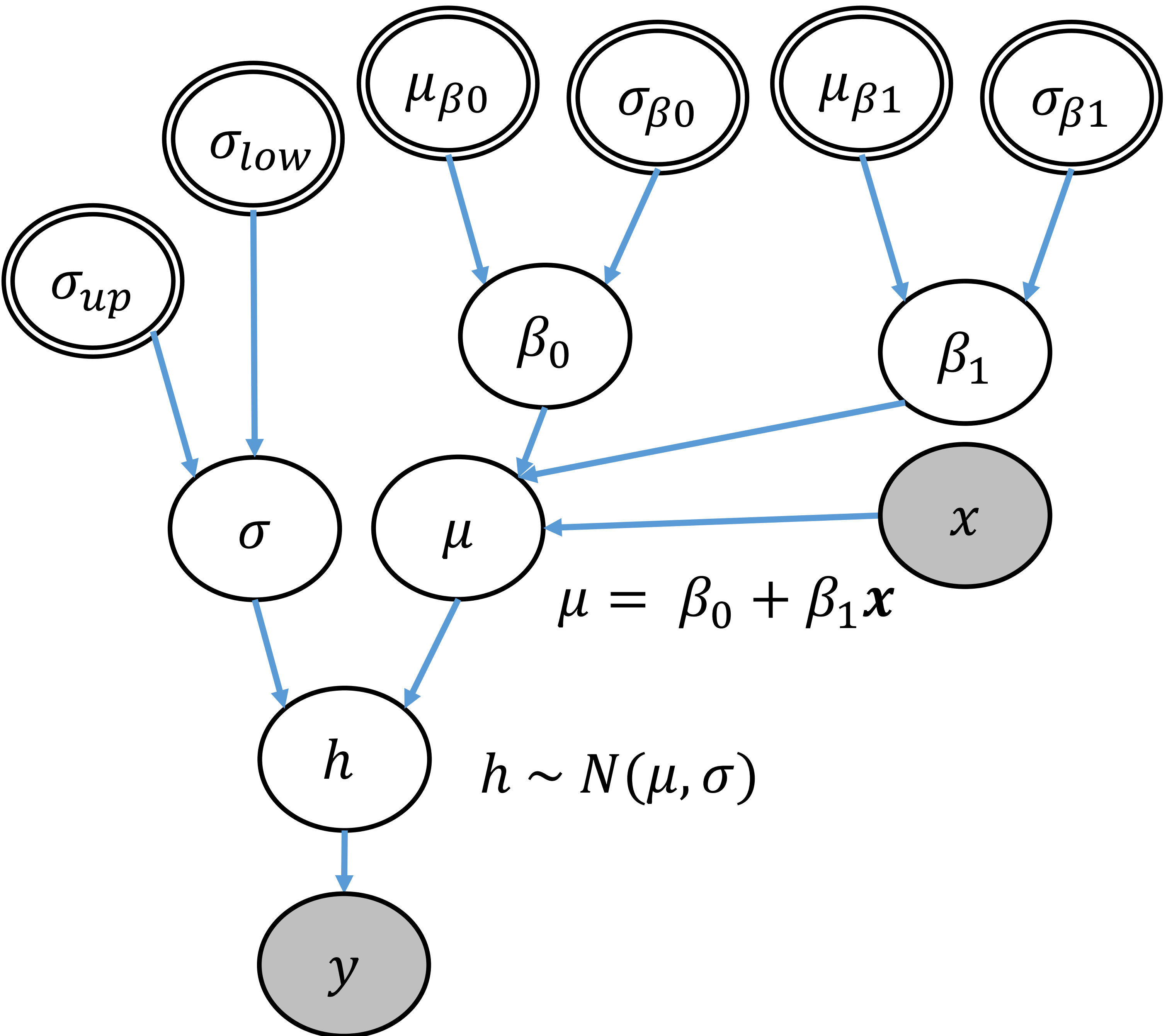
Hypotesis testing

- Robustness to the choice of prior





Let us complicate this model a bit !



Alternative ways to specify a model:
Left – as a graph
Right – by equations

BN, BUGS and JAGS relies on the model expressed as a Directed Acyclic Graph

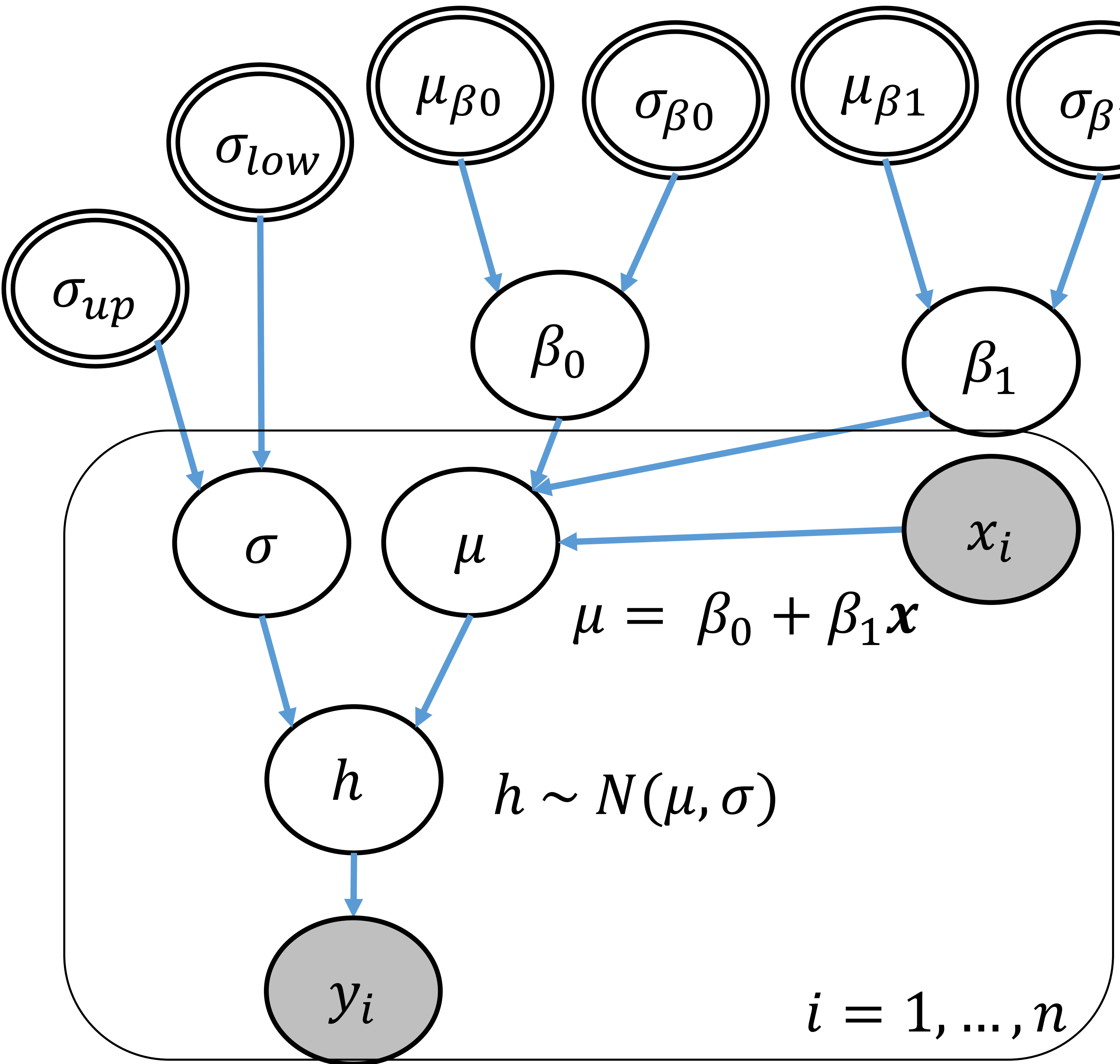
$$y \sim N(\beta_0 + \beta_1 \boldsymbol{x}, \sigma)$$

$$\beta_0 \sim N(\mu_{\beta 0}, \sigma_{\beta 0})$$

$$\beta_1 \sim N(\mu_{\beta 1}, \sigma_{\beta 1})$$

$$\sigma \sim U(\sigma_{low}, \sigma_{up})$$

Add data in the model by indexes and plates
Data is gray nodes



$$i = 1, \dots, n$$

$$y_i \sim N(\beta_0 + \beta_1 \boldsymbol{x_i}, \sigma)$$

$$\beta_0 \sim N(\mu_{\beta 0}, \sigma_{\beta 0})$$

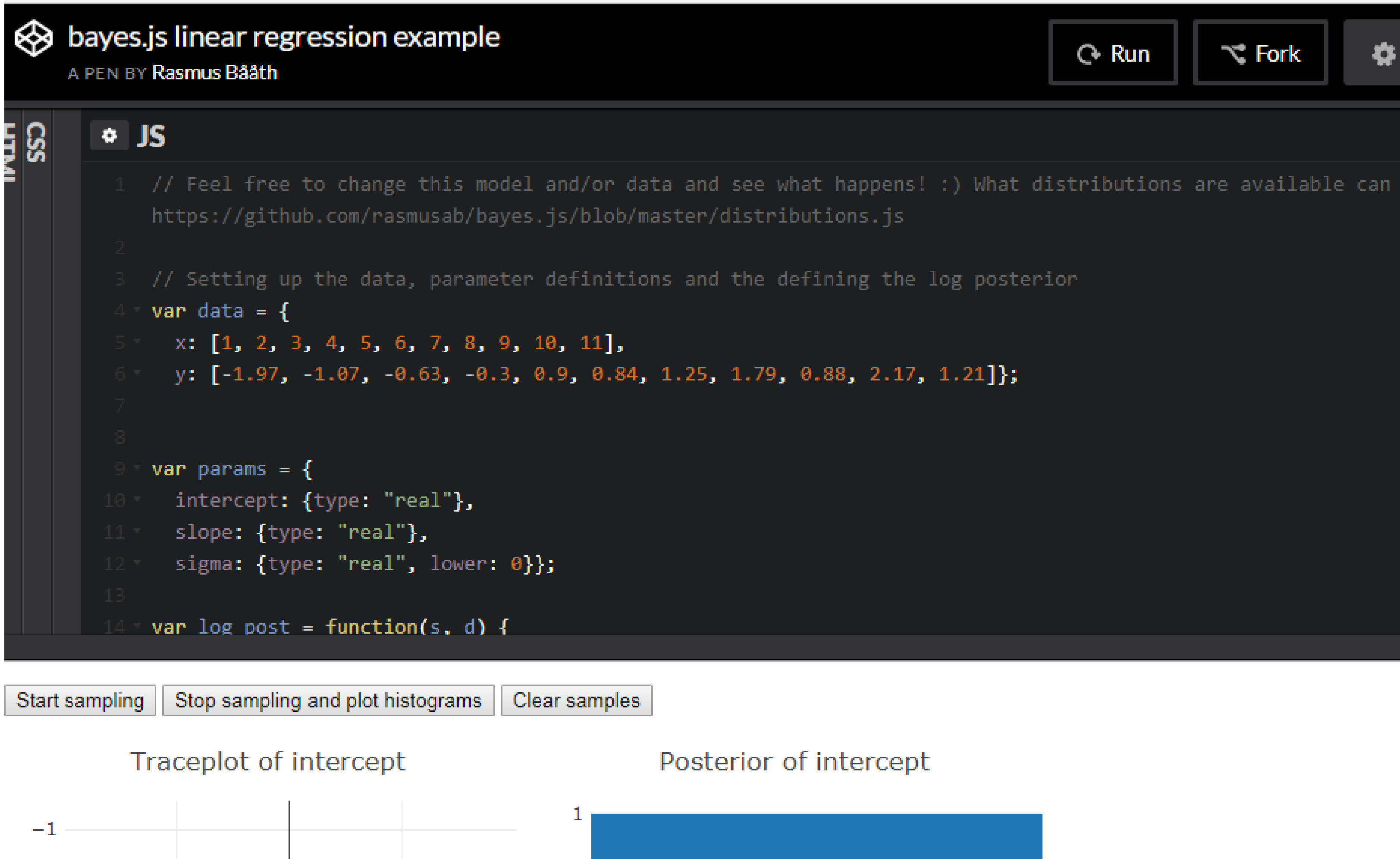
$$\beta_1 \sim N(\mu_{\beta 1}, \sigma_{\beta 1})$$

$$\sigma \sim U(\sigma_{low}, \sigma_{up})$$

MCMC sampling

Try Rasmus nice java code
for MCMC sampling

- Write yourself
- BUGS, JAGS, Stan
- ShinyStan



MCMC sampling

- Stan – uses Hamiltonian MC sampling
- A lot of nice features
- Usually faster than Gibbs sampling

Bayesian Hierarchical Models

- Common statistical models GLMMs
- MCMC – e.g glmmADMB
- Laplace Approximations - INLA

glmmADMB

The `glmmADMB` package, built on the open-source AD Model Builder platform, is an R package for fitting generalized linear mixed models (GLMMs).

Its capabilities include:

- a wide range of families (respo
binomial (type 1 and 2), Beta,
the standard exponential famil
- a wide range of link functions:
- Zero-inflation (currently only a
- Single or multiple random effe
- Markov chain Monte Carlo (M(

In order to use glmmADMB effective
requires familiarity with (i) generalize
regression) and (ii) 'modern' mixed r
by manipulating sums of squares).

Journal of the Royal Statistical Society
SERIES B Statistical Methodology

J. R. Statist. Soc. B (2009)
71, *Part* 2, *pp.* 319–392

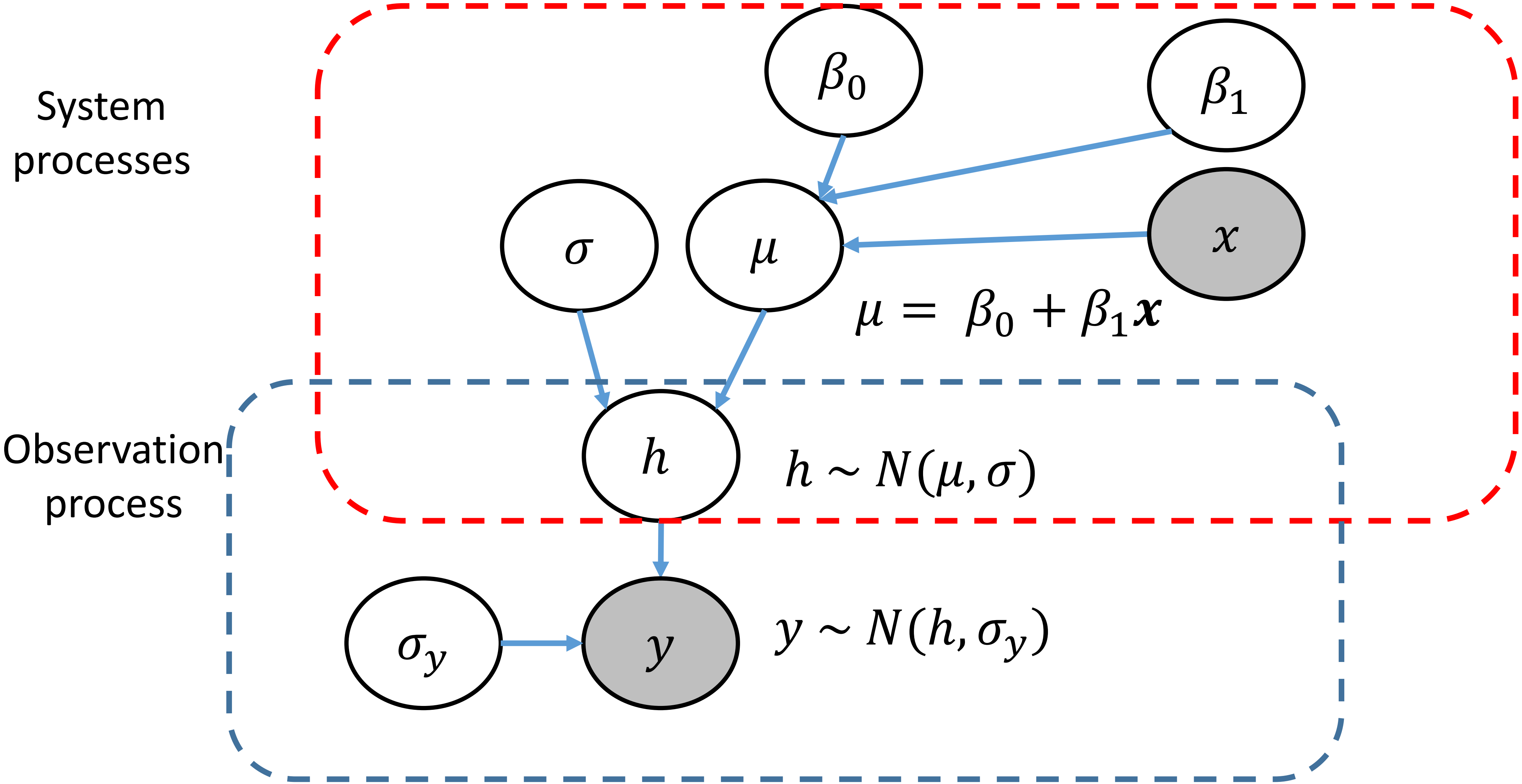
Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations

Håvard Rue and Sara Martino
Norwegian University for Science and Technology, Trondheim, Norway

and Nicolas Chopin
Centre de Recherche en Economie et Statistique and Ecole Nationale de la Statistique et de l'Administration Economique, Paris, France

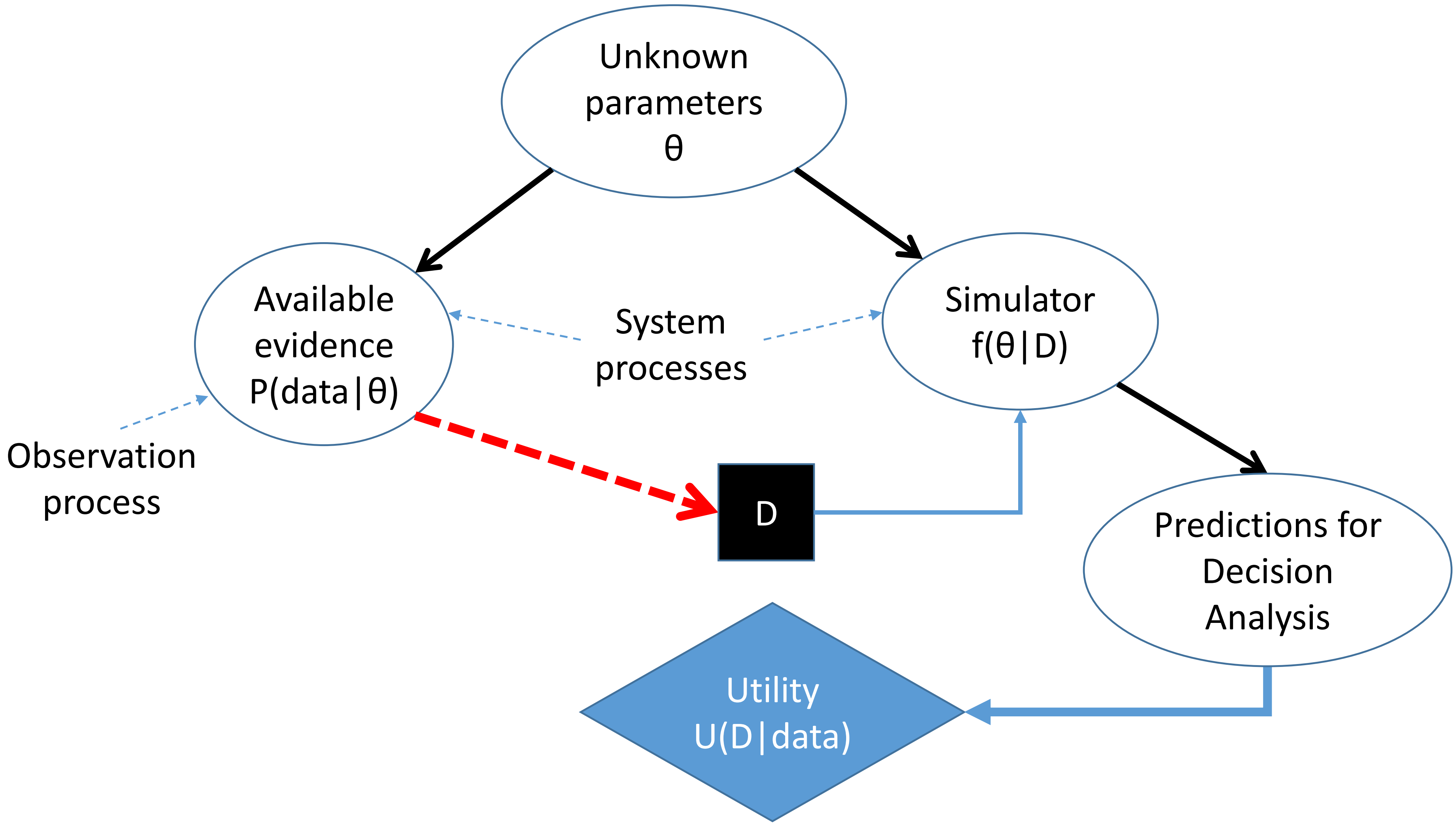
[*Read before* The Royal Statistical Society *at a meeting organized by the* Research Section *on Wednesday*, *October 15th, 2008*, Professor I. L. Dryden *in the Chair*]

Summary. Structured additive regression models are perhaps the most commonly used class of models in statistical applications. It includes, among others, (generalized) linear models, (generalized) additive models, smoothing spline models, state space models, semiparametric regres-

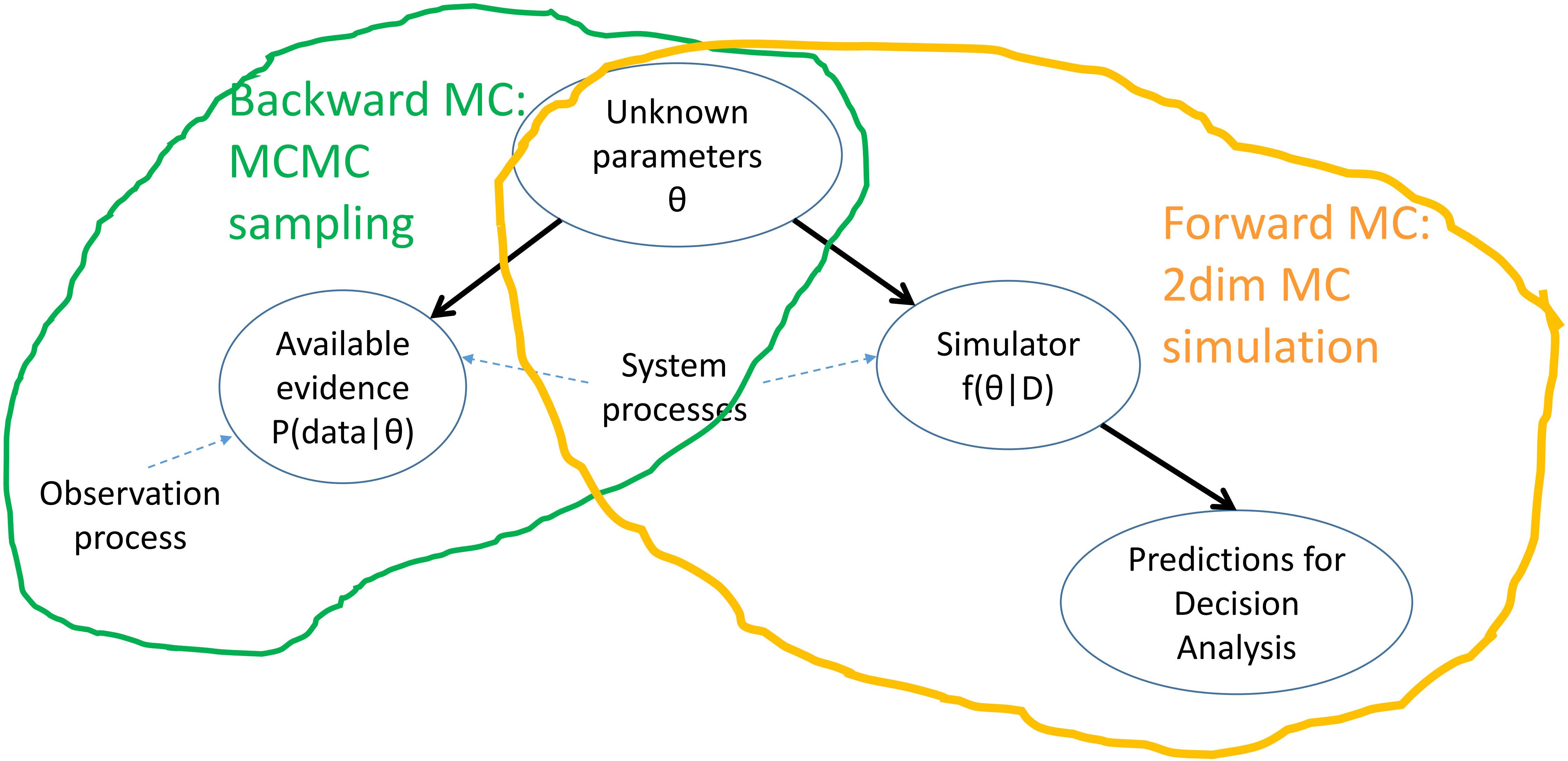


σ_y is observation error. If this is known, we can use the model for meta-anlysis

Bayesian Evidence Synthesis

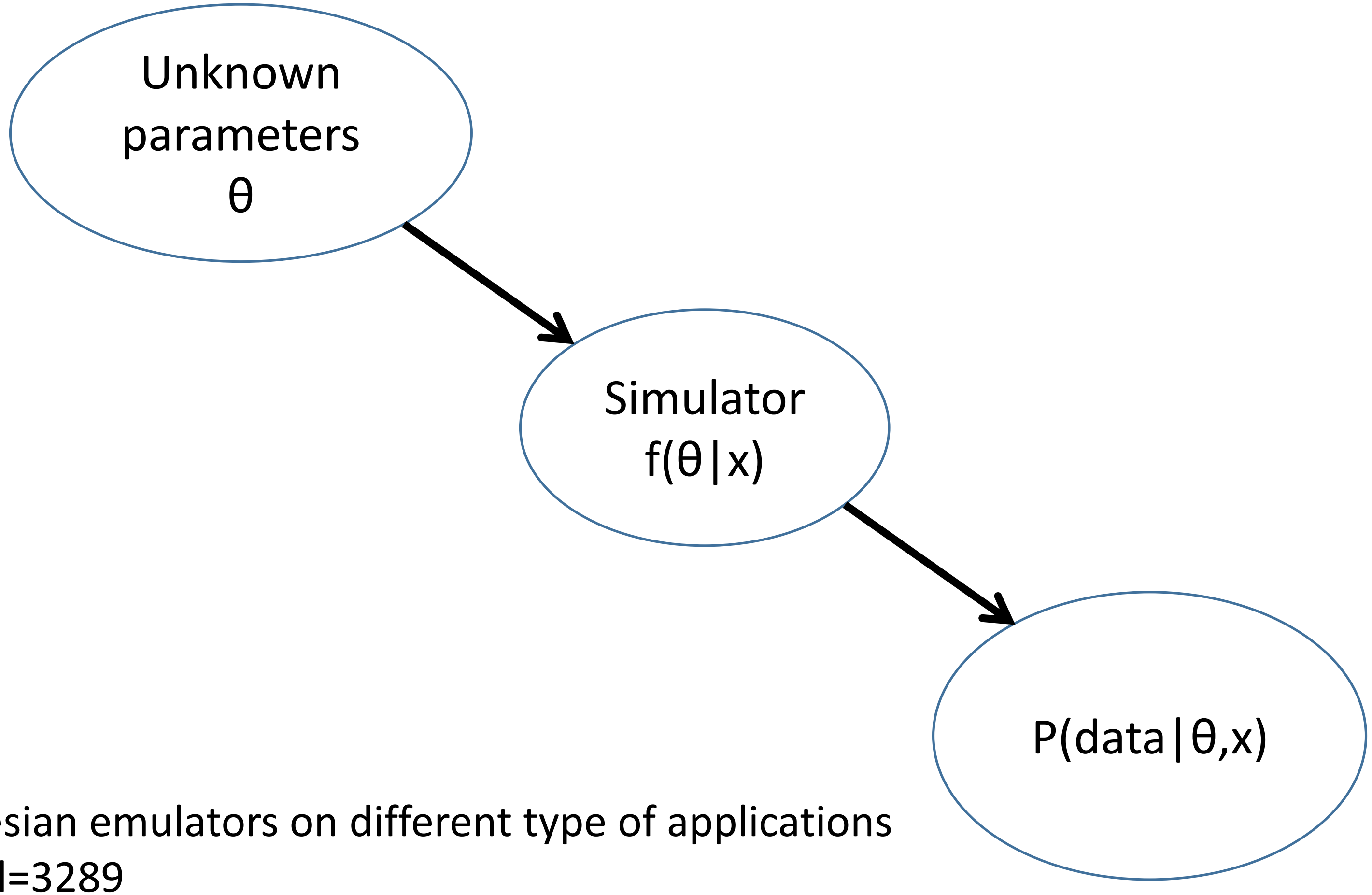


Bayesian Evidence Synthesis



Bayesian calibration of simulation models

- Inverse modelling
- Approximate Bayesian Computation
- Emulators
- Gaussian Processes

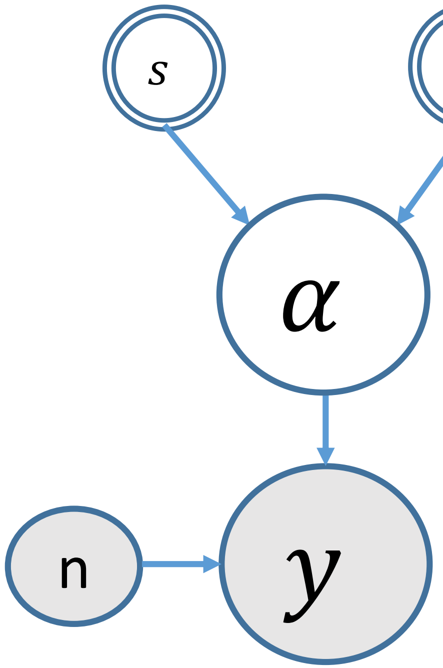


For examples: Ian Vernon, Durham University, Bayesian emulators on different type of applications
https://www.dur.ac.uk/research/directory/staff/?id=3289

Before computers – analytical solutions

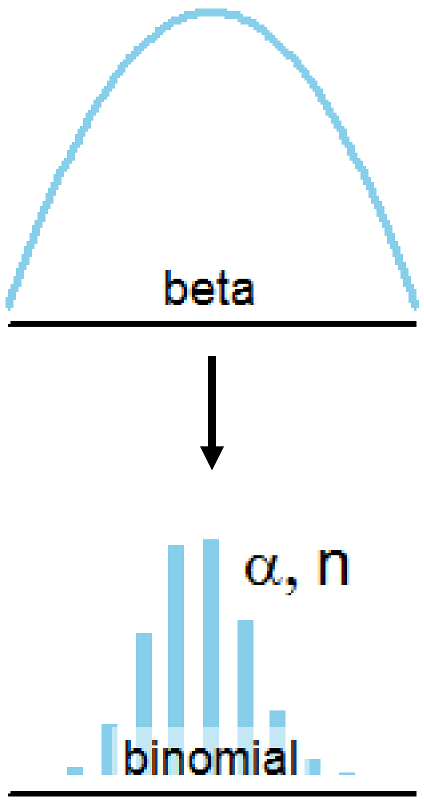
- Conjugate distributions
- Prior and posterior is of the same distribution
- Still useful
- Example Beta-Binomial model

Beta-Binomial model



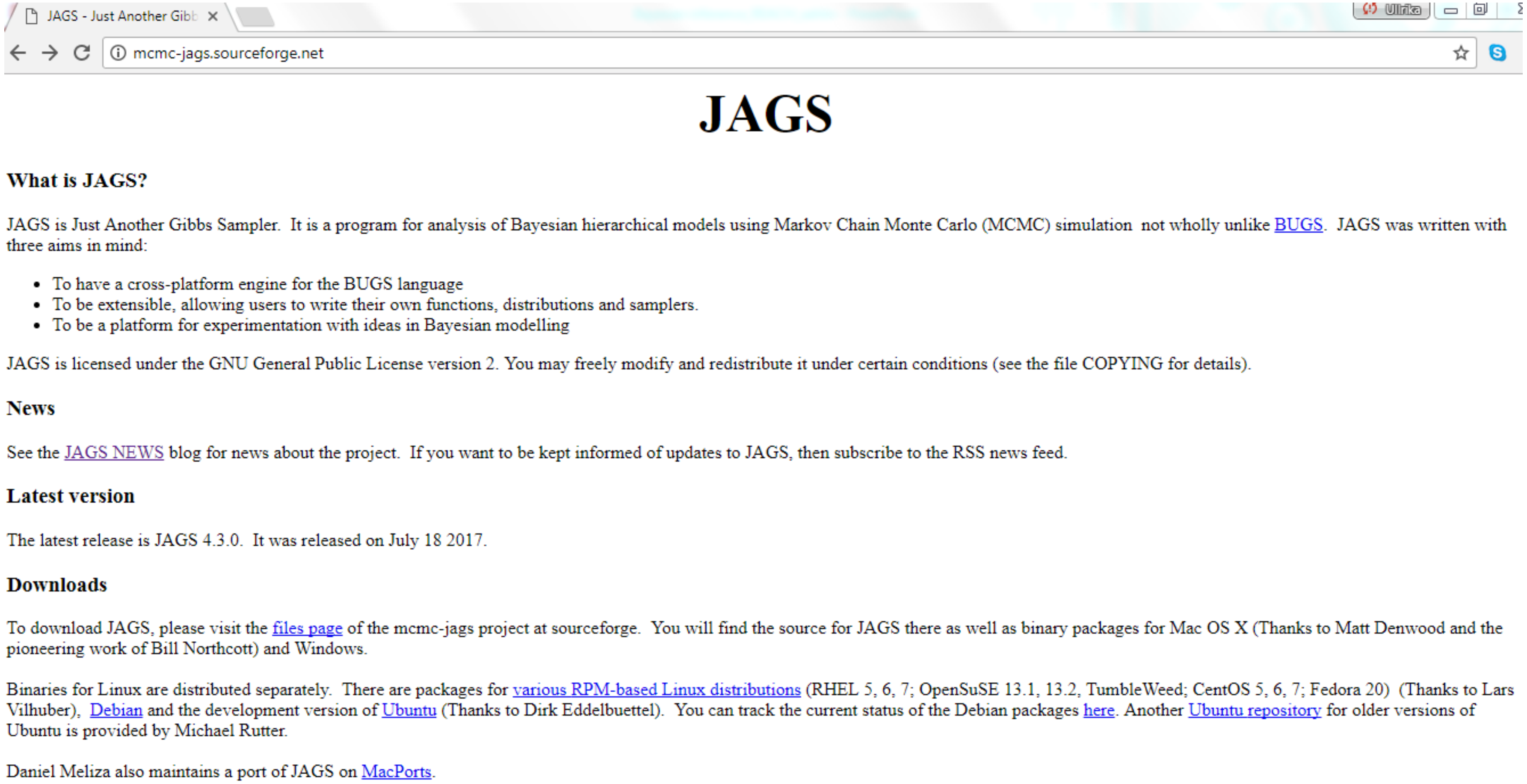
$\alpha \sim Beta(s \cdot t, s \cdot (1-t))$

$y \sim Bin(\alpha, n)$



$$\alpha | y, n \sim Beta(s \cdot t + y, s \cdot (1-t) + (n-y))$$

MCMC sampling with JAGS



JAGS - Just Another Gibb

mcmc-jags.sourceforge.net

JAGS

What is JAGS?

JAGS is Just Another Gibbs Sampler. It is a program for analysis of Bayesian hierarchical models using Markov Chain Monte Carlo (MCMC) simulation not wholly unlike BUGS. JAGS was written with three aims in mind:

- To have a cross-platform engine for the BUGS language
- To be extensible, allowing users to write their own functions, distributions and samplers.
- To be a platform for experimentation with ideas in Bayesian modelling

JAGS is licensed under the GNU General Public License version 2. You may freely modify and redistribute it under certain conditions (see the file COPYING for details).

News

See the JAGS NEWS blog for news about the project. If you want to be kept informed of updates to JAGS, then subscribe to the RSS news feed.

Latest version

The latest release is JAGS 4.3.0. It was released on July 18 2017.

Downloads

To download JAGS, please visit the files page of the mcmc-jags project at sourceforge. You will find the source for JAGS there as well as binary packages for Mac OS X (Thanks to Matt Denwood and the pioneering work of Bill Northcott) and Windows.

Binaries for Linux are distributed separately. There are packages for various RPM-based Linux distributions (RHEL 5, 6, 7; OpenSuSE 13.1, 13.2, TumbleWeed; CentOS 5, 6, 7; Fedora 20) (Thanks to Lars Vilhuber), Debian and the development version of Ubuntu (Thanks to Dirk Eddelbuettel). You can track the current status of the Debian packages here. Another Ubuntu repository for older versions of Ubuntu is provided by Michael Rutter.

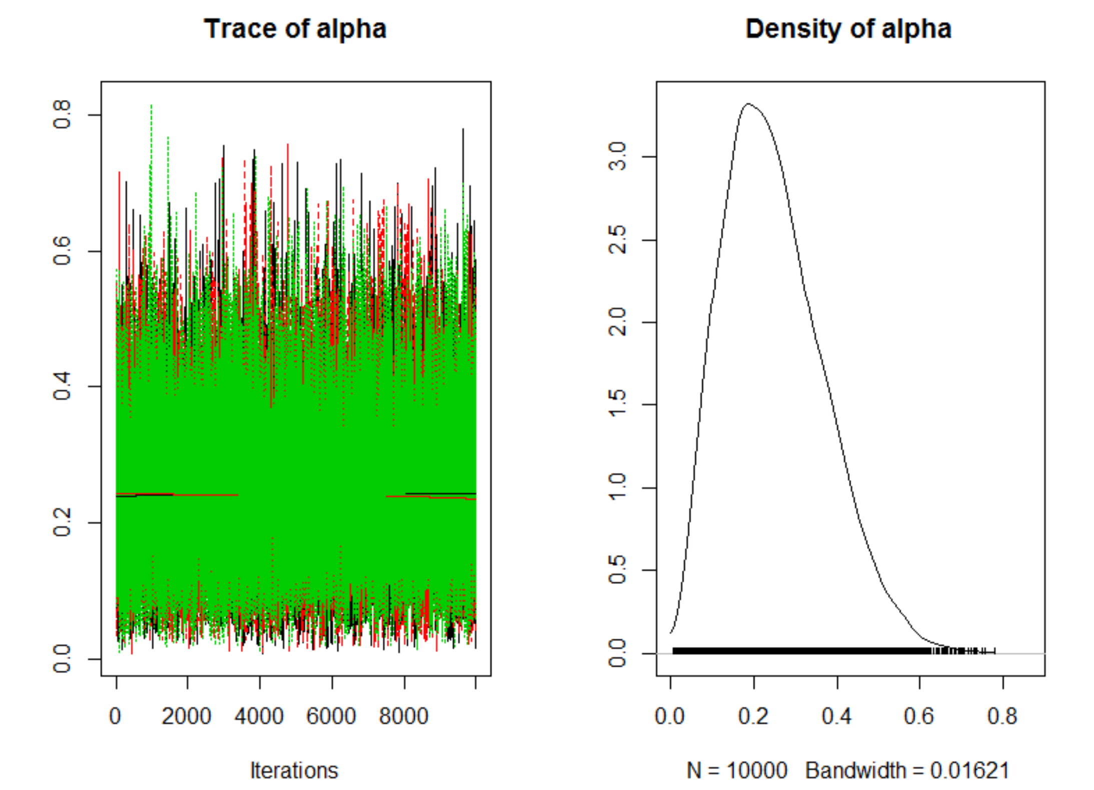
Daniel Meliza also maintains a port of JAGS on MacPorts.

MCMC sampling of the Beta-Binomial model

betabinomial.R

Source on Save

```
library(rjags)

ms = "
model {
  alpha ~ dbeta(s*t, s*(1-t))
  y ~ dbinom(alpha,n)
}"

data_to_model = list(s=2, t = 0.5,
                     n = 10, y = 2)

attach(data_to_model)
m = jags.model(textConnection(ms), data=data_to_model,
               n.adapt = 10^6, n.chains=3)
sam = coda.samples(m, c('alpha'), n.iter=10000, thin=1)
plot(sam)
summary(sam)

# conjugate distribution
# posterior mean
(s*t+y) / ((s*t+y) + (s*(1-t)+(n-y)))

```



MCMC sampling of the Beta-Binomial model

- The mean of the posterior sample is 0.249814
- The theoretical mean using the Beta-Binomial model is 0.25

```
> summary(sam)

Iterations = 1:10000
Thinning interval = 1
Number of chains = 3
Sample size per chain = 10000

1. Empirical mean and standard deviation for each variable,
   plus standard error of the mean:

          Mean             SD       Naive SE Time-series SE
      0.249814       0.120205       0.000694       0.000694

2. Quantiles for each variable:

   2.5%     25%     50%     75%   97.5%
0.05948 0.15991 0.23541 0.32580 0.51695

> (s*t+y) / ((s*t+y) + (s*(1-t)+(n-y)))
[1] 0.25
```

A principle for learning

$$P(\theta) + data \rightarrow P(\theta|data)$$

$$P(data|\theta)$$

$$P(\theta|data) \cdot P(data) = P(data|\theta) \cdot P(\theta)$$

Bayes Rule:

$$P(\theta|data) = \frac{P(data|\theta) \cdot P(\theta)}{P(data)}$$

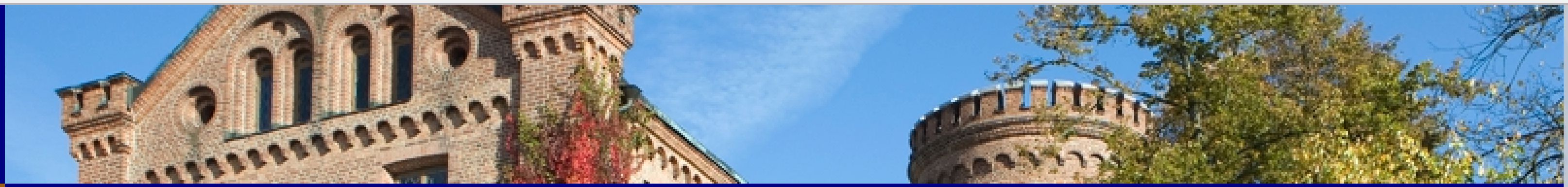
Bayesian inference can be used for any data analysis

Purpose

- Hypothesis testing
- Estimation
- Assessment
- Quantification of uncertainty
- Decision analysis

So

- Gives you what you want
- Curse of too high degree of freedom
- Got to understand what you are doing (see also the Folk Theorem)
- Integrate data and expert knowledge
- Sometimes time consuming
- Fun



Bayes@Lund

Welcome!

This is the Bayesian statistics mailing list at Lund University. The purpose of this mailing list is to connect all those at Lund University interested in Bayesian statistics. To send a message to the mailing list just send an e-mail to bayes@listserver.lu.se . Topics suitable for messages to this mailing list includes information regarding upcoming seminars, talks, courses and workshops related to Bayesian statistics.

An archive of the Bayes@Lund mailing list is available here:
https://groups.google.com/forum/#!forum/bayes_lund

To subscribe (and to unsubscribe) to the mailing list use the form below, after submitting the form you should shortly receive an e-mail where you have to confirm your subscription.

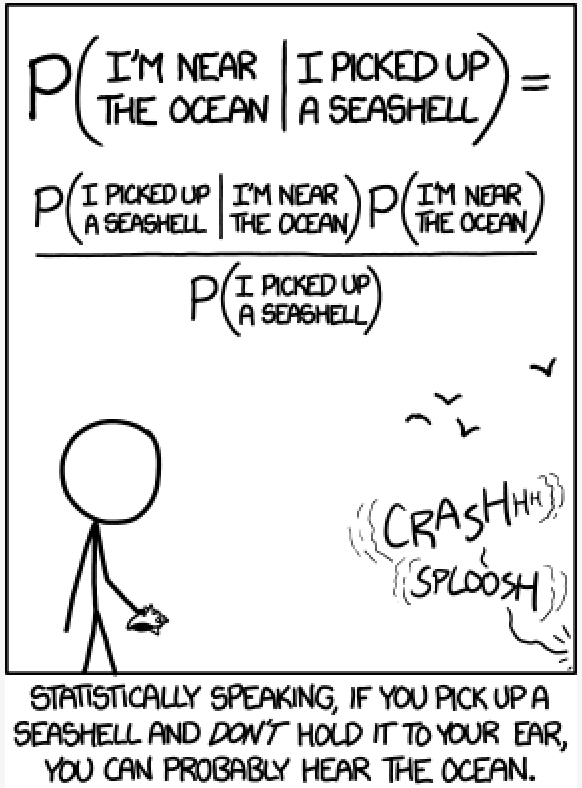
Subscription form

- (•) Subscribe
- ( ) Unsubscribe

Your e-mail:

[   ] Submit

The administrator of this mailing list is Rasmus Bååth, mail him at rasmus.baath@lucs.lu.se if something is not working properly for you.



Even if the the mailing list is mainly for information regarding upcoming talks and seminars there might also be the occasional Bayesian comic strip...

http://sumsar.net/

+

=

Bayes @Lund

Session 1

13.35 – 13.55 An introduction to Bayesian and hierarchical modelling – *Johan Lindström, Mathematical Statistics*

13.55 – 14.05 An example of a Bayesian model in BUGS and R – *Yf Jiang, Biology*

14.05 – 14.25 A generalized approach to modeling and estimating indirect effects in ecology – *Yann Clough, Centre of Environmental and Climate Research*

14.25 – 14.45 Resting time for migrating birds, and x-ray time variability of galaxies: using Baysian Cramér-Rao bounds – *Dragi Anevski, Mathematical Statistics*

14.45 Fika break

Session 2

15.15 – 15.35 Bayesian First Aid: Replacing null hypothesis tests by Bayesian estimation. – *Rasmus Bååth, Cognitive Science*

15.35 – 15.55 Bayesian approach, non-observed variables, and collecting long term data – *Krzysztof Podgorski, Statistics*

15.55 – 16.15 Reasons to be Bayesian – *Ullrika Sahlin, Centre of Environmental and Climate Research*

16.15 – 16.35 Incorporating uncertainty when evaluating subsidy effects on farmland bird biodiversity - a Bayesian wannabe analysis – *Martin Stjernman, Biology*

16.35 General discussion moderated by Ullrika and Rasmus. The final discussion will be on the slow adoption of Bayesian methods at Lund University. Topics of discussion will be: Why are Bayesian methods not well represented in courses at Lund University? How to deal with aversions to the application of Bayesian methods? How to strengthen the role of Bayesian methods at Lund University?

17.00 End of the day

The post-conference discussion will be held at Bishops Arms (S:t Petri Kyrkogata 7).

BAYES @LUND

2014

PROGRAM

Pre-conference hands on session by Ullrika Sahlin and Paul Caplat

Bayes from a frequentist point of view, Krzysztof Podgorski, Departement of Statistics

Teaching Bayesian data analysis in psychology, Geoffrey R. Patching, Department of Psychology

Lindley's paradox, Bengt Ringnér, Mathematical Statistics

Pollen based spatial reconstruction of past land cover, Behnaz Pirzamanbin, Mathematical Statistics

Keynote speaker: Mattias Villani, Linköping University – Bayesian model inference – why, what and how?

Tiny data, approximate Bayesian Computation and the socks of Karl Broman, Rasmus Bååth, Cognitive Science

Data-Cloning ABC for (approximate) maximum likelihood estimation, Umberto Picchini, Mathematical Statistics

Joint cell population identification through Bayesian hierarchical modeling, Kerstin Johnsson, Centre for Mathematical Sciences

Distributing a collapsed sampler for topic models, Måns Magnusson, Lindköping University

Estimation of local moose population using Bayesian hierarchical modelling, Jonas Wallin, Matematiska vetenskaper, Chalmers

Performance of Bayesian prediction of treatment differences using a two-factor linear mixed-effects model, Johannes Forkman, SLU

Bayesian estimation of optimal portfolio, Stepan Mazur, Department of Statistics

How Bayesian belief networks can help save the world, Ullrika Sahlin, Centre for Environmental and Climate Research

PROGRAM

Pre-conference tutorial by Eric-Jan Wagenmakers

An Introduction to Bayesian computation and evidence synthesis using STAN, Robert Grant, Faculty of Health, Social Care and Education, University of London.

Keynote: Bayesian Benefits for the Pragmatic Researcher, Eric-Jan Wagenmakers, Department of Psychology, University of Amsterdam.

Bayesian Meta Analysis and Bias Modeling: A Case Study with Relative Clause Processing in Mandarin Chinese, Shravan Vasishth and Lena Jaeger, Department of Linguistics, University of Potsdam.

A Bayesian reflection on the meaning of evidence, Ullrika Sahlin, Centre for Environmental and Climate Research, Lund University.

The Bootstrap is a Bayesian procedure, but that doesn't mean it's any good, Rasmus Bååth, Lund University Cognitive Science.

Bayesian methods in epidemiological research – why so seldom used? Jonas Björk, Division of occupational and environmental medicine, Lund University.

Regularized supervised topic models for high-dimensional multi-class regression, Måns Magnusson, Department of Computer and Information Science, Linköping University.

Modeling the growth of Swedish Scots pines, Henrike Häbel, Department of Mathematical Sciences, Chalmers University of Technology.

PROGRAM

Pre-conference tutorial by Rasmus Bååth

Keynote Darren Wilkinson: Hierarchical modelling of genetic interaction in budding yeast

Stefan Wiens, Making the most of your ANOVAs: From NHST to Bayesian analyses

Martin Stjernman, Joint species modelling -- beautiful in theory, tricky in practice

Shravan Vasishth, Finite mixture modeling: a case study involving retrieval processes in sentence comprehension

Keynote Richard McElreath: Understanding Bayesian statistics without frequentist language

Judith Butepage: Learning to make decisions under uncertainty

Mark Andrews: Teaching Bayesian Data Analysis to Social Scientists

Thomas Hamelrick: Potentials of mean force for protein structure prediction: from hack to math

Junpeng Lao: Statistical Inferences of Eye movement data using Bayesian smoothing

Richard Torkar: Convincing researchers to transition to Bayesian statistics - the case of software engineering

Bertil Wegmann: Experiences from teaching Bayesian inference to students familiar with frequentist statistics

Erik Lindström: Multilevel Monte Carlo methods for inference in multivariate diffusions

Ullrika Sahlin: Using expert's knowledge in Bayesian analysis

BAYES
@LUND

2018 !!!