Week 2

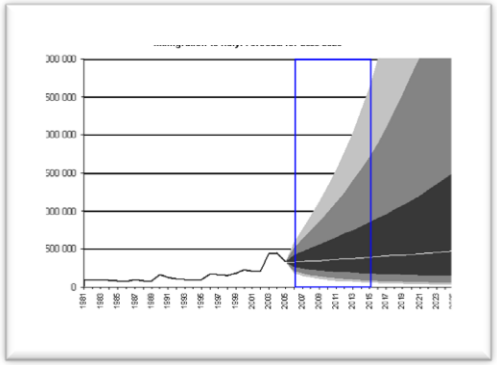
Selected Topics in Bayesian modelling

Jakub Bijak, Jon Forster & Jason Hilton

Elicitation of prior distributions

Expert judgement

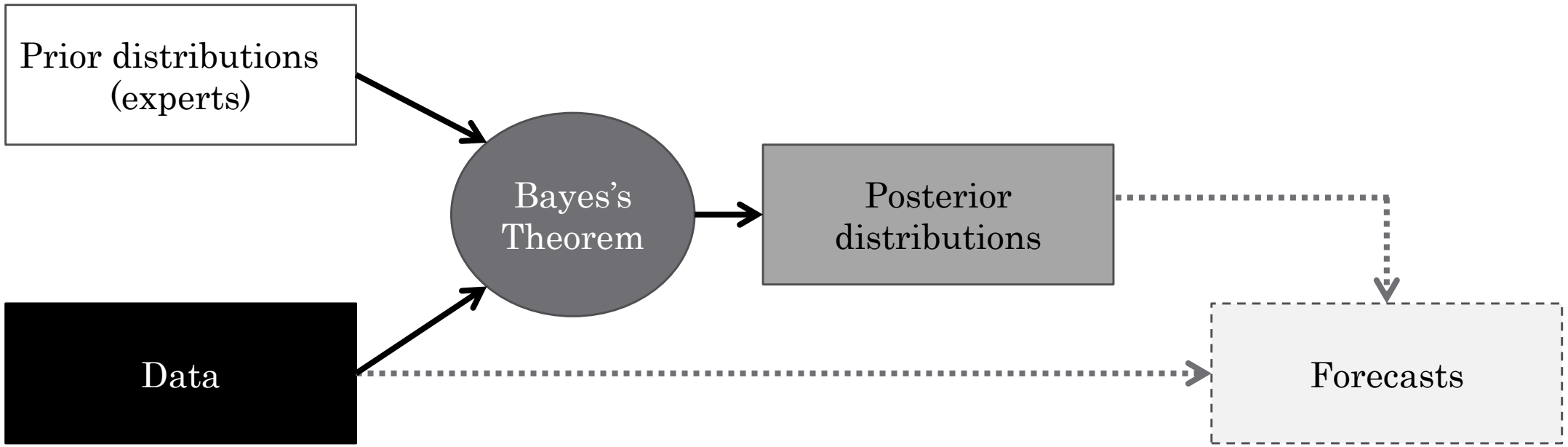
- Bayesian estimation and prediction can explicitly take advantage of knowledge of experts in the field
- Experts can provide **informative prior distributions**
- This is important especially for
  - Small samples and missing data
  - Making statements about the future



The key challenge is **how to elicit** expert judgement

Expert judgement

- Assumptions *a priori* and the data moderate each other through the Bayes theorem (recall week 1)



Elicitation – general remarks

- Questions concern expert's **intuition** on a given topic
- Questions should be clear, possibly in a plain language to ensure correct interpretation by the experts
- Words and statements which can be seen as emotional, normative, (dis)favouring, etc. should be avoided
- Too much information should be avoided: Armstrong (1985) **"lack of information is better than worthless information"**
- Appropriate formulation of questions, especially concerning numbers, avoiding the "anchoring" and "bias" of answers
- **Pre-testing** of the questionnaire – verification, whether the questions are correctly and intuitively understood

Delphi method

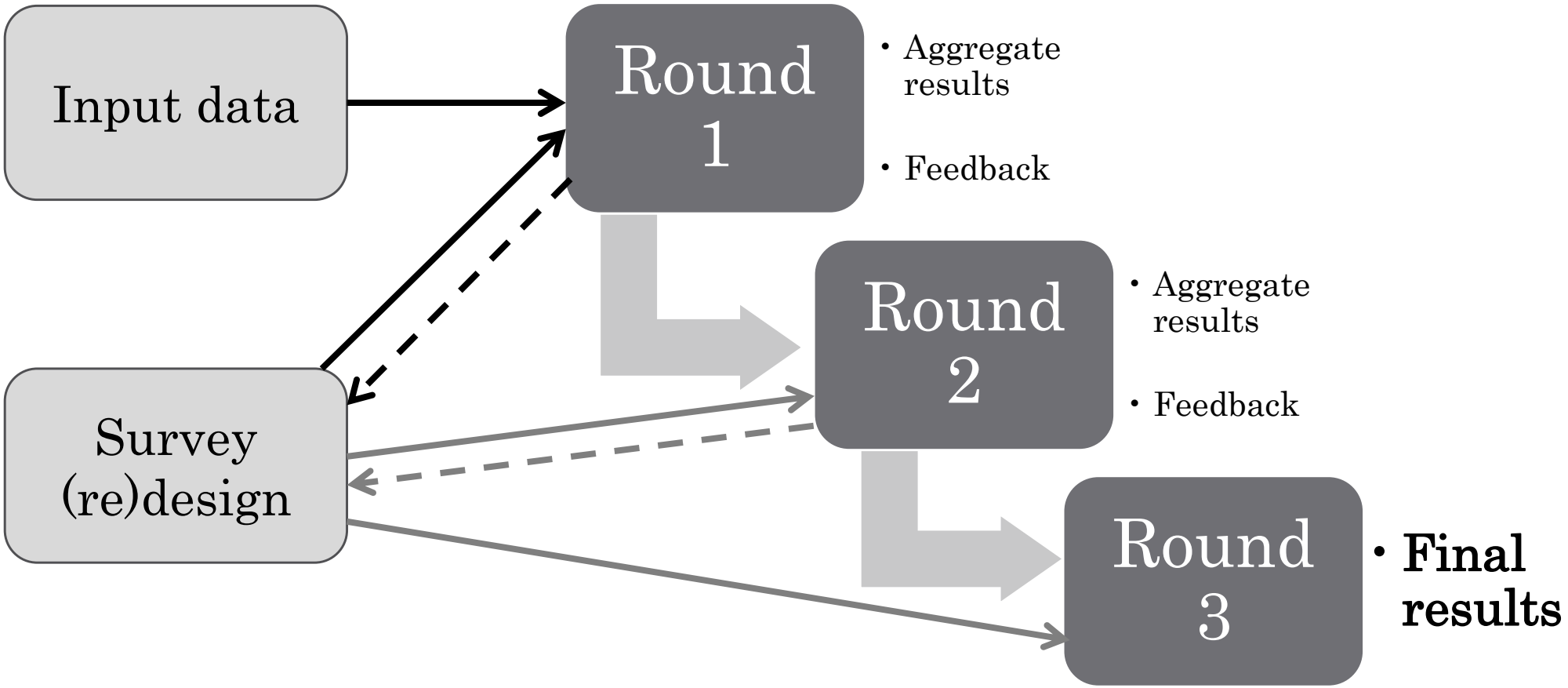
- A technique of obtaining information and opinion concerning the future by means of an **iterative survey** (Dalkey 1967)
- The aim is to facilitate an **informed consensus**

Henry Martin (The New Yorker Collection/The Cartoon Bank), via: https://methodsblog.wordpress.com/2015/05/12/the-delphi-technique/

Delphi method

- **Characteristics:**
  - Respondents are **experts** in a given field
  - Respondents remain **anonymous** and reply **independently**
  - Opinions are obtained **iteratively**
  - The answers can be statistically **aggregated**
  - Experts are informed about **aggregate results** of the preceding round and **anonymous justifications** for answers
  - Respondents with "extreme" answers can be asked to provide **rationale** for such views

Delphi method

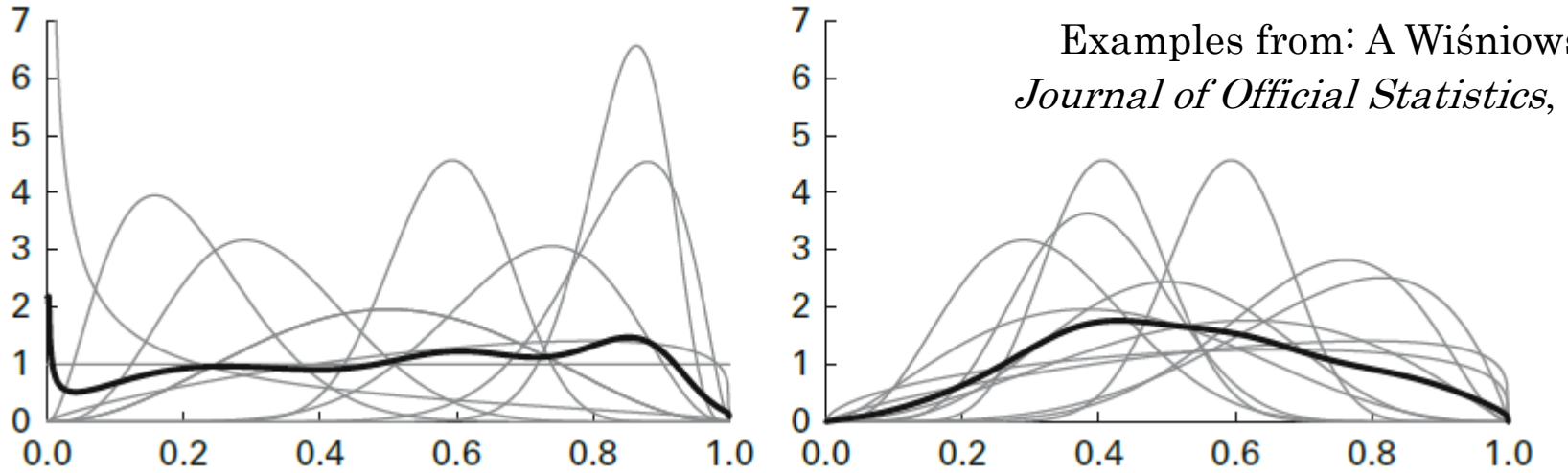


Delphi method – practicalities

- The experts should possess appropriate domain knowledge
- The joint knowledge of experts should encompass the whole problem domain → heterogeneous profile of respondents
- The group should consist of **5–20 experts**. More might cause:
  - Excess of redundant information
  - Increased risk of conflicting opinions
  - "Information noise"
- Number of experts is arbitrary – it depends on available resources, the problem, and the expected quality of answers

Elicitation – outcomes

- **NB:** Many experts – additional source of uncertainty
- Question of aggregation: e.g a probabilistic mixture



Examples from: A Wiśniowski et al. (2013) *Journal of Official Statistics*, 29(4), 583–607.

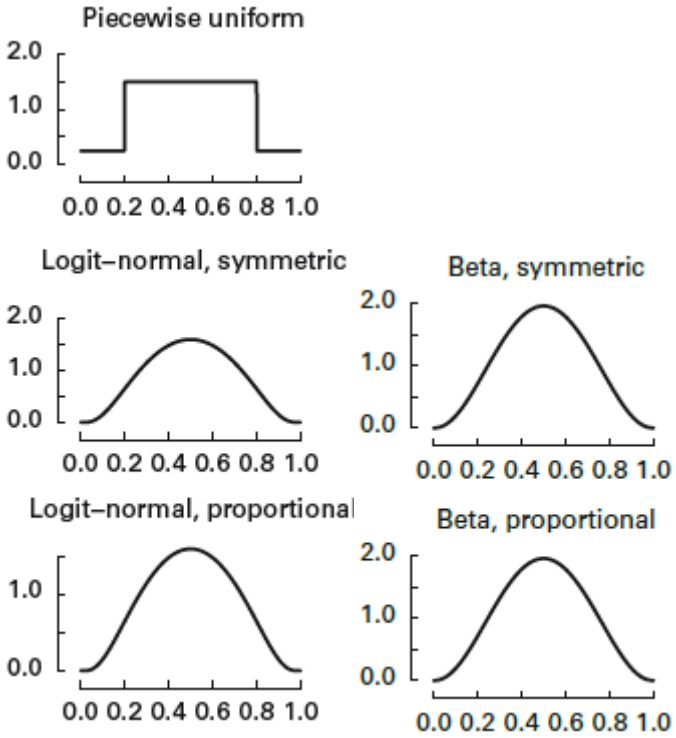
- **Open issue:** should the weights be equal, or for example related to past performance? (Cooke 1991)

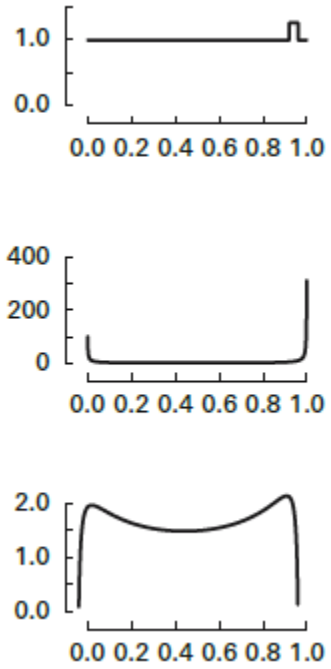
Elicitation – caveats

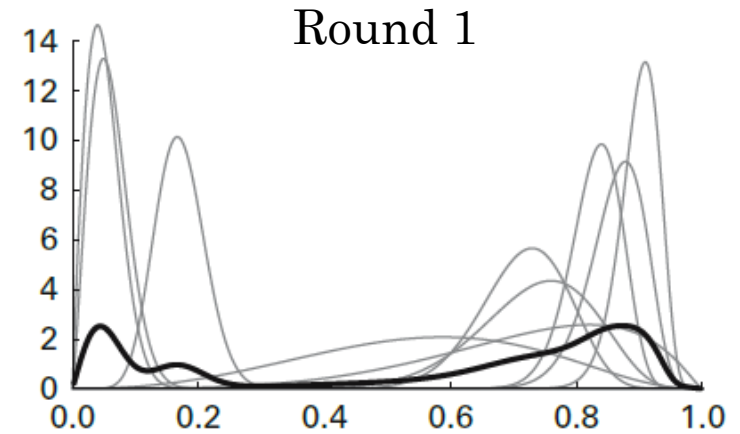
- Problems with questions about **probability** (direct) and **odds** (indirect – lack of normalisation)
- Interpretation of probability in frequentist terms
- Tendency to perceive the uniqueness of events, not as just one possibility – the role of context and analogies
- Problem of overestimating the accuracy of judgement :
  - Probabilities summing up to over one
  - Too narrow uncertainty intervals
  - Too light or too heavy distribution tails

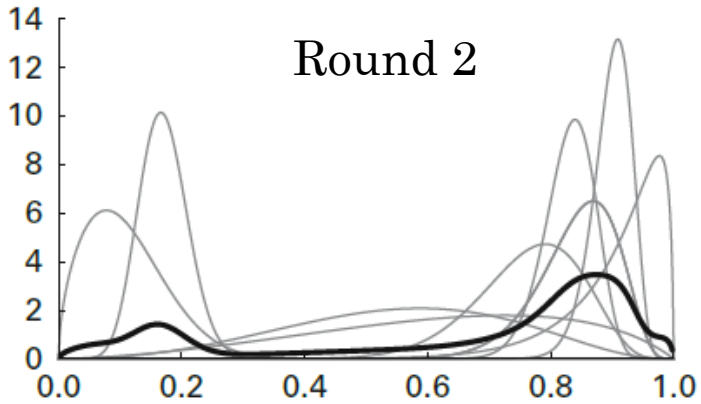
Caveats - illustrations

Specification matters – Interpretation matters – Convergence matters





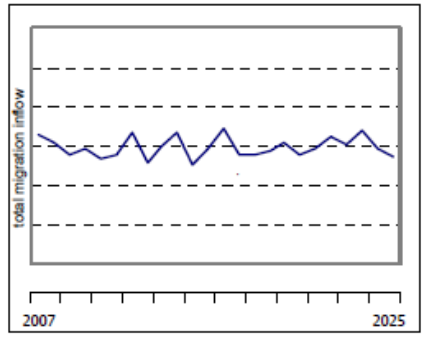


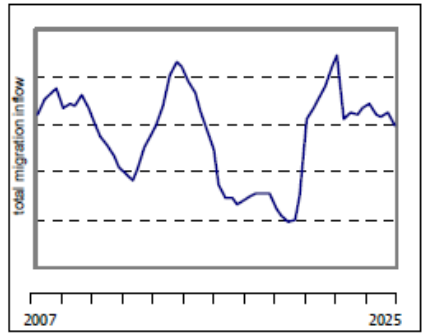


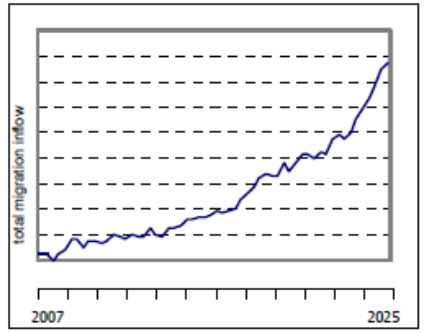
Examples from: A Wiśniowski et al. (2013)
Journal of Official Statistics, 29(4), 583–607.

Lessons learned

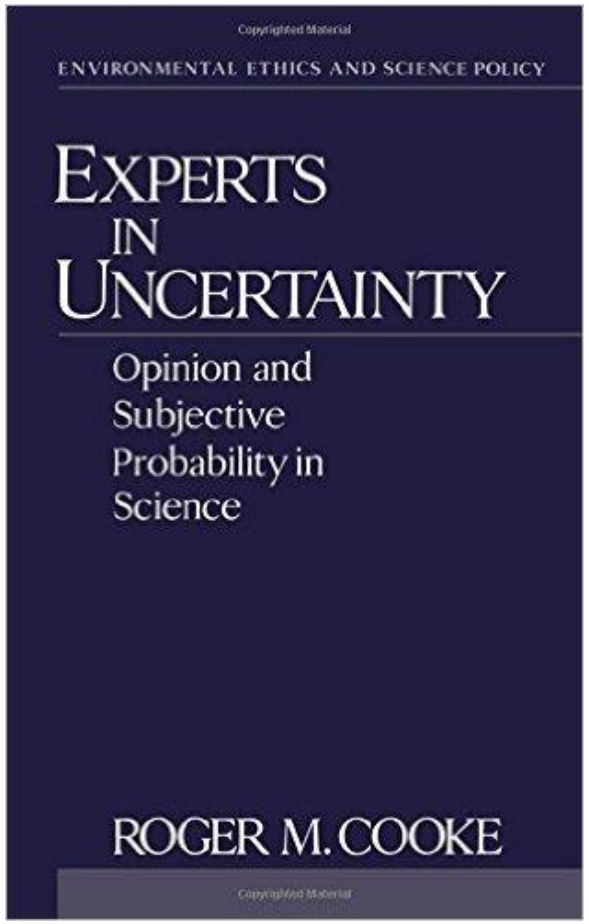
- **Implications:** Questions need to be clear, but still formal and unambiguous
- **Taboo words:** probability, random variable, stochastic process, distribution, expected value, variance, quantile, stationarity, (auto)regression, unit root, (co)integration …
- **Solution:** visualisations of various processes
- **Designing and testing** of the questionnaire: interdisciplinary work



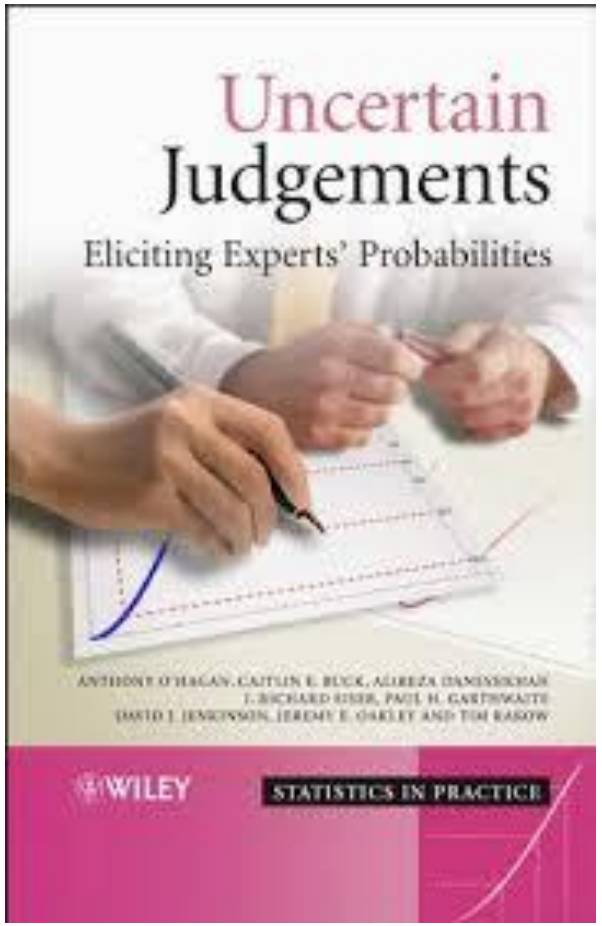




Selected reading



Roger M Cooke
(1991) OUP



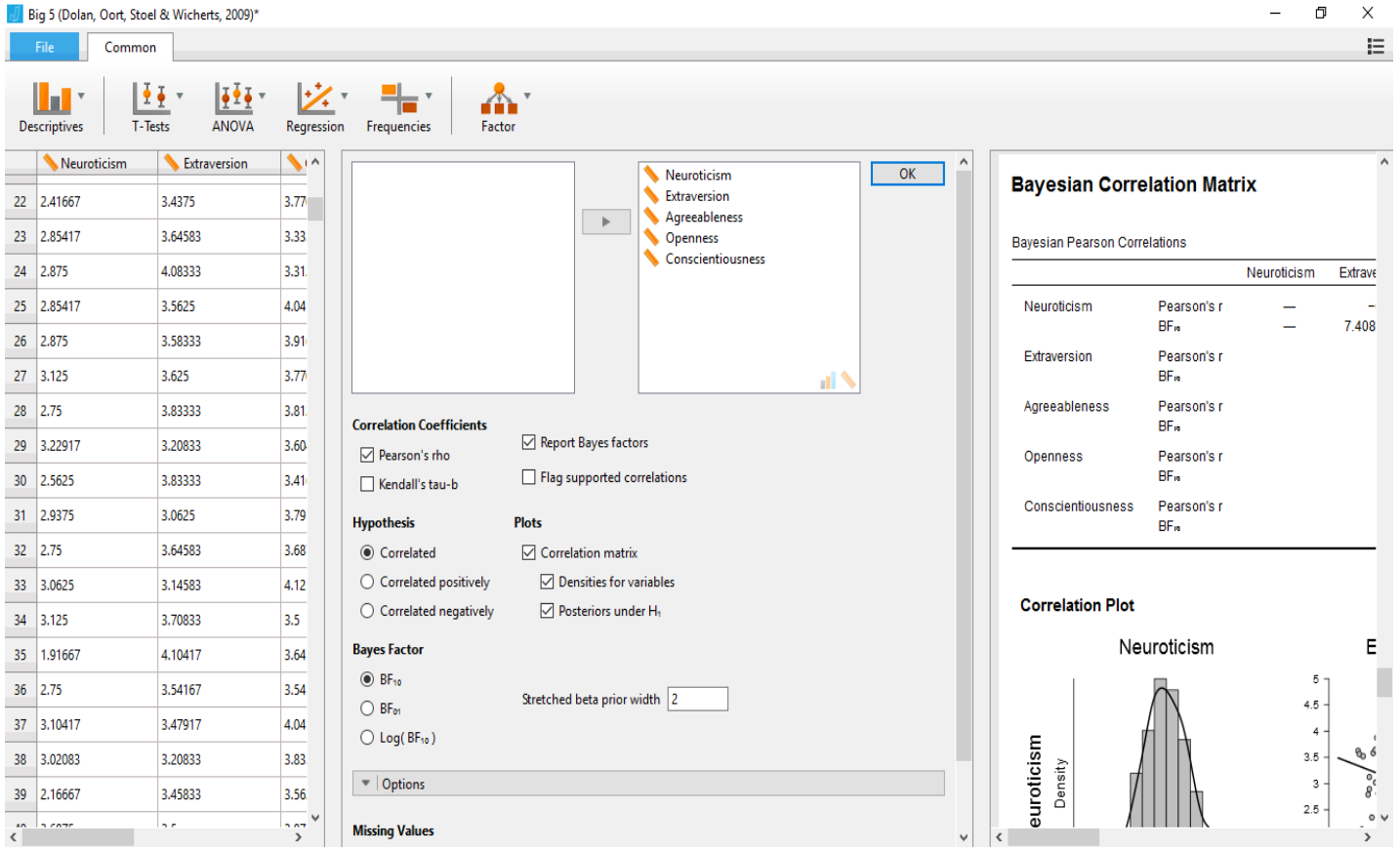
Anthony O'Hagan
et al. (2006), Wiley

Bijak J and Wiśniowski A (2010) Bayesian forecasting of immigration to selected European countries by using expert knowledge. *Journal of the Royal Statistical Society Series A*, 173(4), 775–796. **DOI:**10.1111/j.1467-985X.2009.00635.x

Wiśniowski A, Bijak J, Christiansen S, Forster JJ, Keilman N, Raymer J and Smith PWF (2013) Utilising expert opinion to improve the measurement of international migration in Europe. *Journal of Official Statistics*, 29(4), 583–607. **DOI:** 10.2478/jos-2013-0041.

A brief overview of Bayesian software

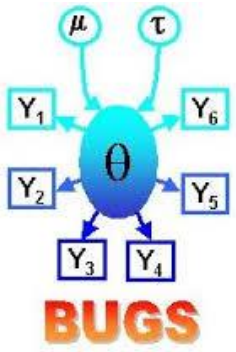
From the more user-friendly…

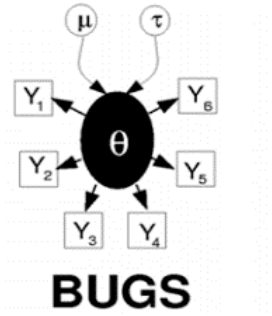


JASP – recall the workshop two weeks ago

https://jasp-stats.org/

…to more advanced





The BUGS project: High-level language

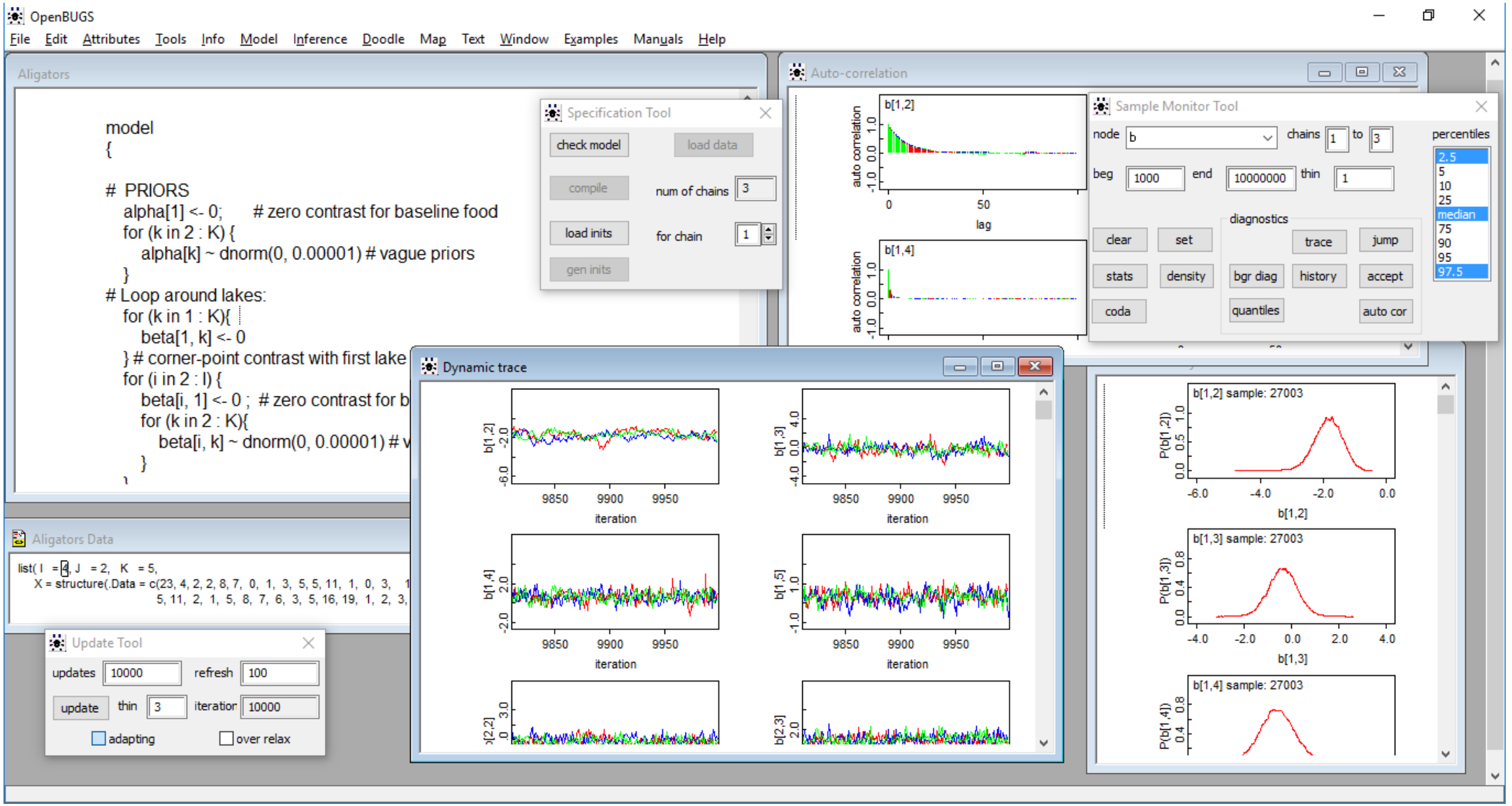
Bayesian inference Using Gibbs Sampling

- **BUGS** (now obsolete)
- **WinBUGS**
- **OpenBUGS** (open source, current focus)
- Also: JAGS (Just Another Gibbs Sampler)

http://www.mrc-bsu.cam.ac.uk/software/bugs/

http://www.openbugs.net

OpenBUGS


OpenBUGS
File Edit Attributes Tools Info Model Inference Doodle Map Text Window Examples Manuals Help
Aligators
model
{
# PRIORS
alpha[1] <- 0; # zero contrast for baseline food
for (k in 2 : K) {
alpha[k] ~ dnorm(0, 0.00001) # vague priors
}
# Loop around lakes:
for (k in 1 : K){
beta[1, k] <- 0
} # corner-point contrast with first lake
for (i in 2 : I) {
beta[i, 1] <- 0 ; # zero contrast for b
for (k in 2 : K){
beta[i, k] ~ dnorm(0, 0.00001) # v
}
Specification Tool
check model
load data
compile
num of chains 3
load inits
for chain 1
gen inits
Auto-correlation
b[1,2]
b[1,4]
auto correlation
lag
Sample Monitor Tool
node b
chains 1 to 3
beg 1000
end 10000000
thin 1
percentiles
2.5
5
10
25
median
75
90
95
97.5
diagnostics
clear
set
trace
jump
stats
density
bgr diag
history
accept
coda
quantiles
auto cor
Dynamic trace
b[1,2]
b[1,3]
b[1,4]
b[1,5]
b[2,3]
iteration
Aligators Data
list(I = 4, J = 2, K = 5,
X = structure(.Data = c(23, 4, 2, 2, 8, 7, 0, 1, 3, 5, 5, 11, 1, 0, 3,
5, 11, 2, 1, 5, 8, 7, 6, 3, 5, 16, 19, 1, 2, 3,
Update Tool
updates 10000
refresh 100
update
thin 3
iteration 10000
adapting
over relax
b[1,2] sample: 27003
b[1,3] sample: 27003
b[1,4] sample: 27003
P(b[1,2])
P(b[1,3])
P(b[1,4])

…to even more advanced

STAN

- Similar idea to BUGS, but different numerical algorithms used (such as Hamiltonian Monte Carlo)
- Computationally more efficient

http://mc-stan.org/

SAS

https://support.sas.com/rnd/app/stat/procedures/BayesianAnalysis.html

PROC MCMC : General purpose

Requires SAS/STAT 9.2 or higher

BCHOICE : discrete choice models

FMM : finite mixture models

GENMOD : GLMs

LIFEREG : parametric models for failure time data

PHREG : survival analysis with the Cox proportional hazards model

Interfaces

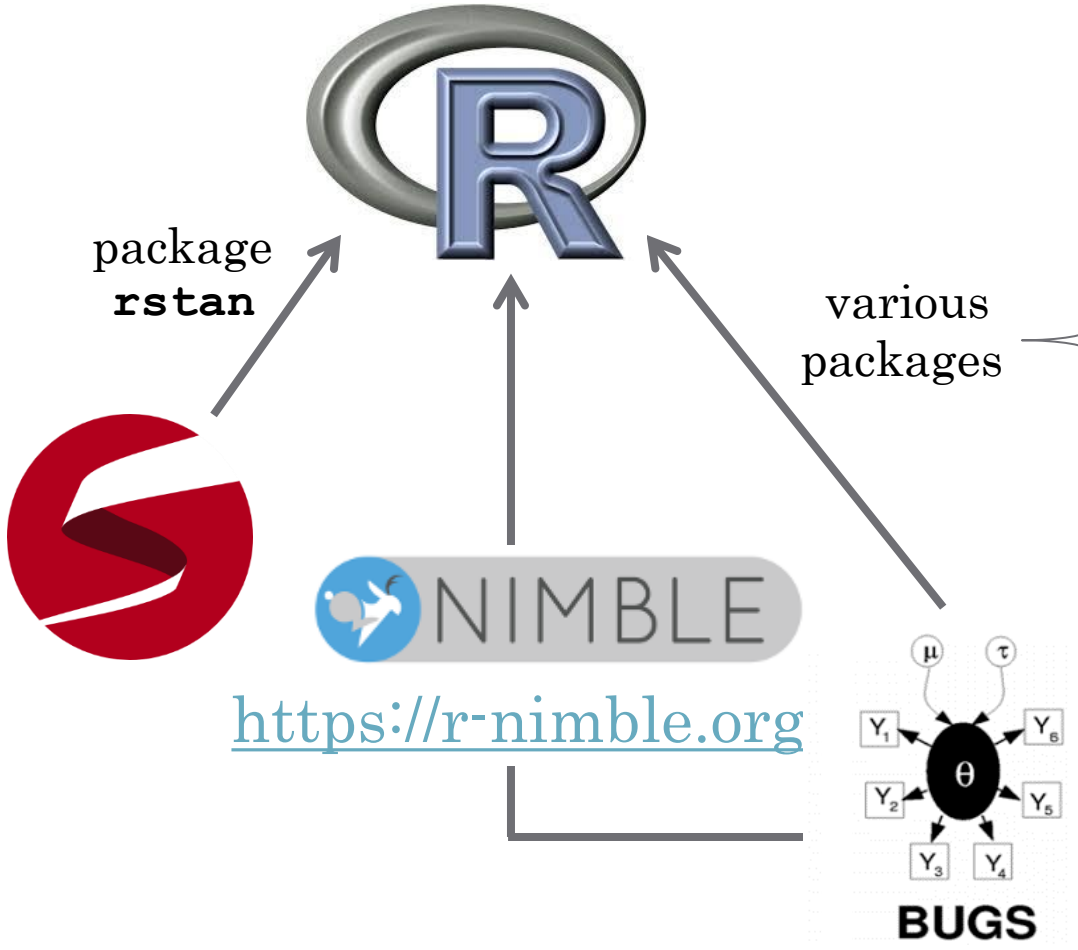


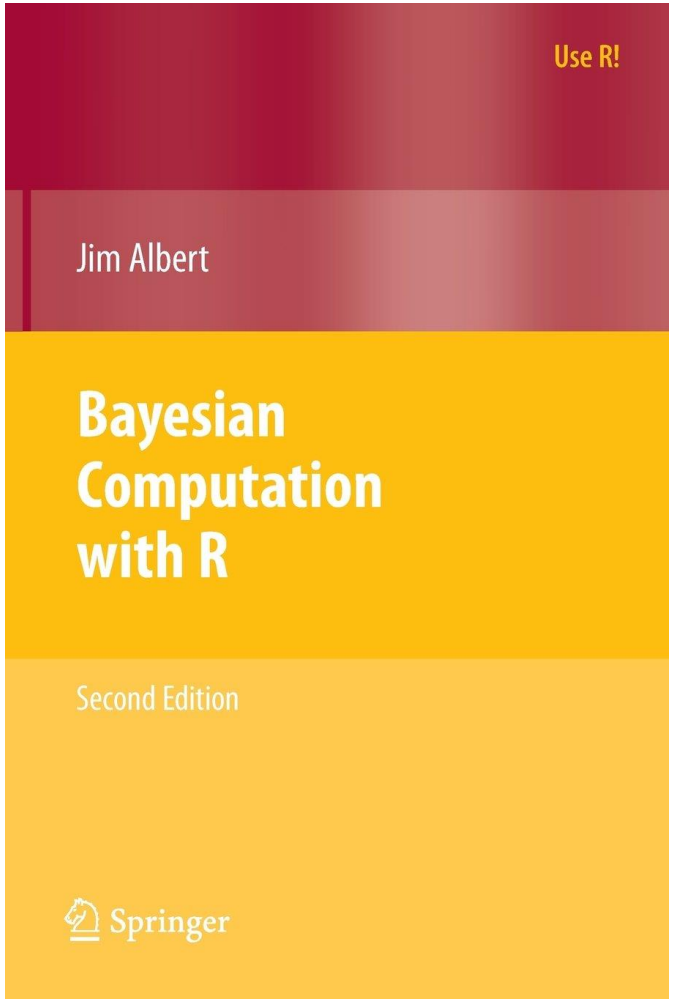
Table 9.1 Selected R packages for Bayesian analysis

Package	Description
Arm	Bayesian inference in linear, generalized linear, ordered logit and probit models
bayesm	Analysis of linear regression models, multinomial logit, multinomial probit, multivariate probit, multivariate mixture of Normal distributions normals (including clustering), density estimation using finite mixtures of Normal distributions, Dirichlet Process priors, hierarchical linear models, hierarchical multinomial logit, hierarchical negative binomial regression models, and linear instrumental variable models
DPpackage	R functions for Bayesian nonparametric and semi-parametric models. The package includes, among others, semi-parametric models for density estimation, censored data, binary regression models and generalized linear mixed models.
MCMCpack	Model-specific MCMC algorithms for inference in regression models (linear regression, logit, ordinal probit, probit, Poisson regression, etc.), measurement models (item response theory and factor models), changepoint models (binary and Poisson), and models for ecological inference. It contains a generic Metropolis sampler.
LearnBayes	R functions and sample data for the Albert (2007) book on R in the Bayesian computations
BMA	Bayesian averaging of linear, generalized linear and survival models
BAYSTAR	Bayesian analysis of threshold autoregressive models
MSBVAR	Estimation of Bayesian VAR and structural VAR models
BRugs	Provides R interface to the OpenBUGS environment (for Windows)
R2WinBUGS	Provides R interface to the WinBUGS environment (for Windows)
rbugs	Provides R interface to OpenBUGS (for Linux)
boa	Post-estimation tool for diagnostics, summarising and visualising the MCMC output, also useful for importing from the BUGS format
coda	The *Convergence Diagnosis and Output Analysis* (CODA) package – a set of functions to summarize, plot, and diagnose convergence from MCMC samples, including a possibility of import from the BUGS format.

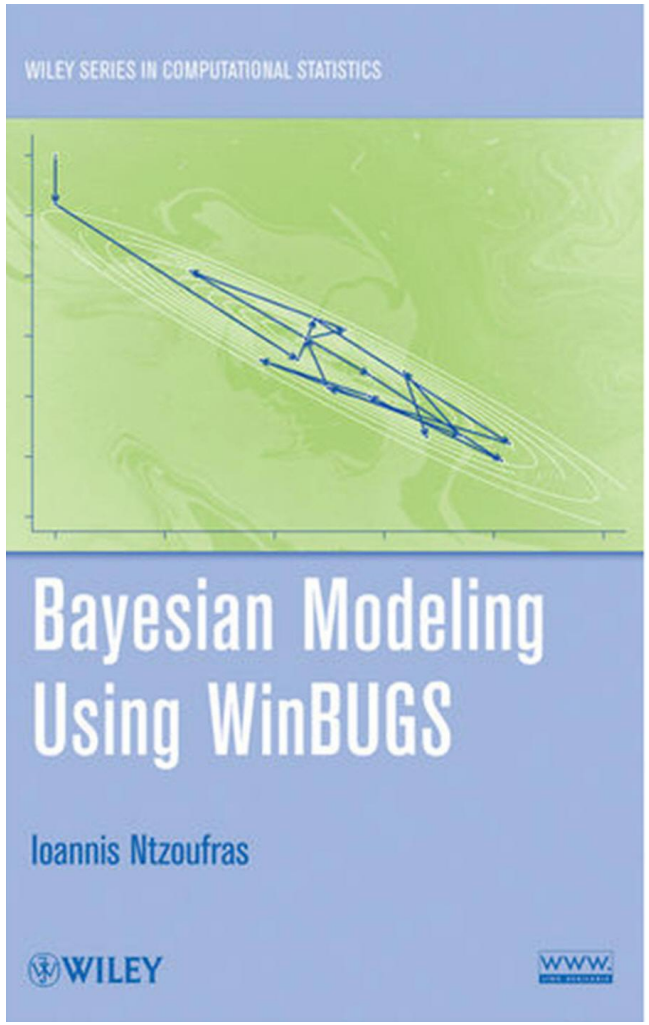
Source: Park et al. (2009); http://cran.r-project.org/web/views/Bayesian.html.

After: A Wiśniowski, Chapter 9 on “Bayesian Computing in Practice” in Bijak (2010), p. 176.

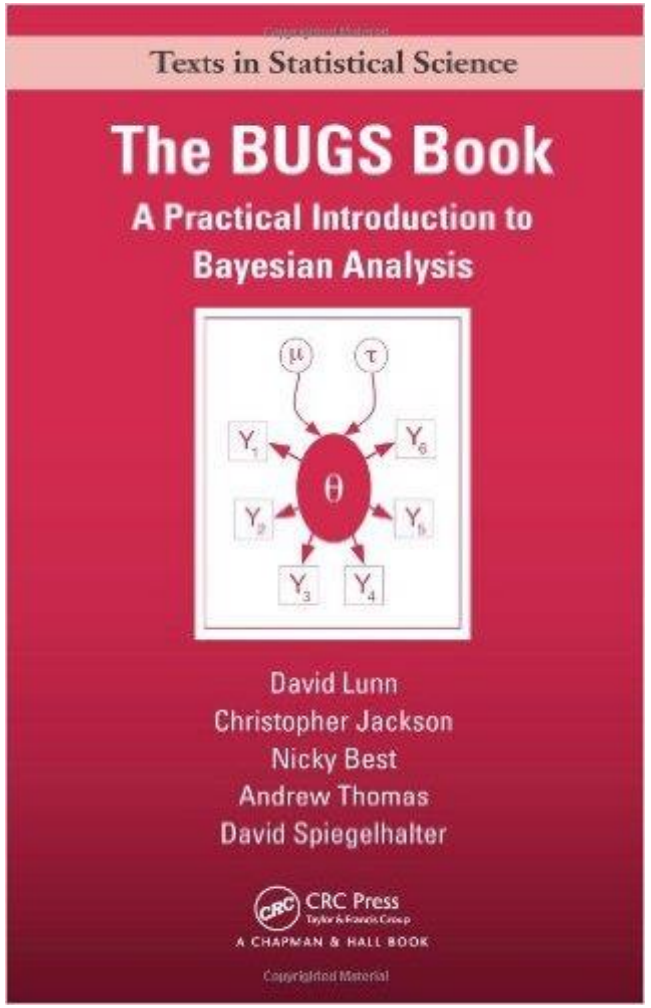
Books – examples



Jim Albert (2007)

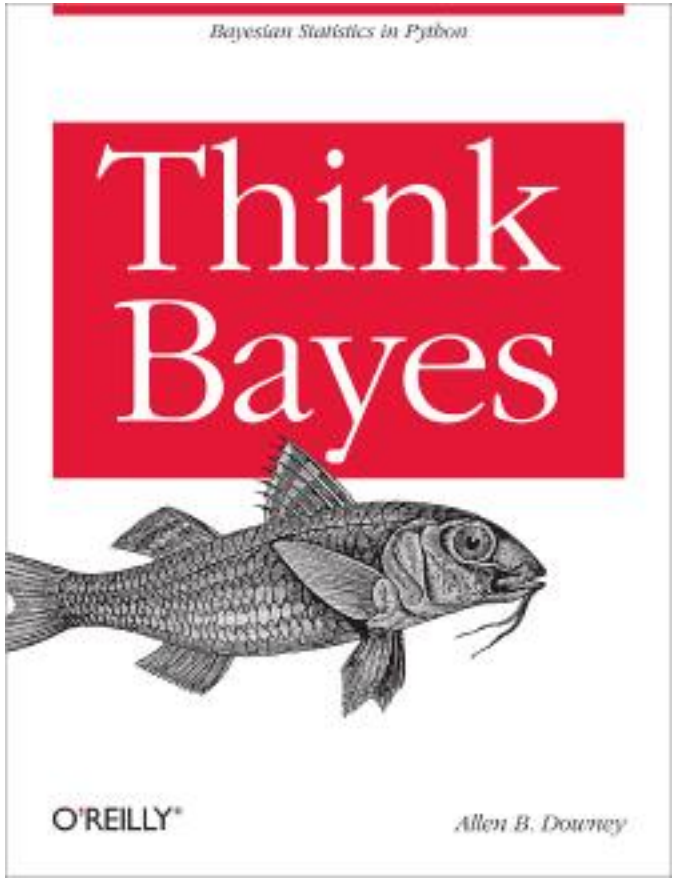


Ioannis Ntzoufras (2009)

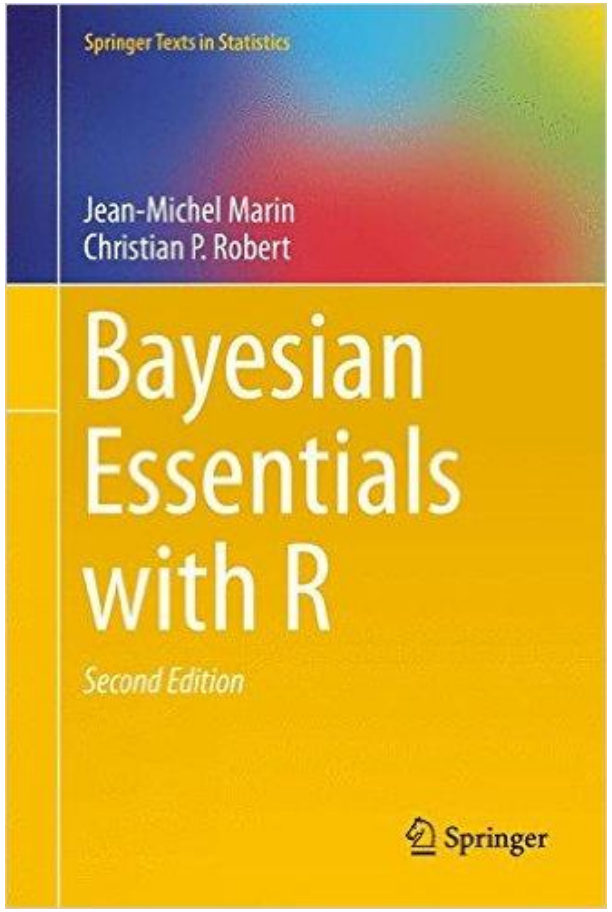


David Lunn, *et al.* (2012)

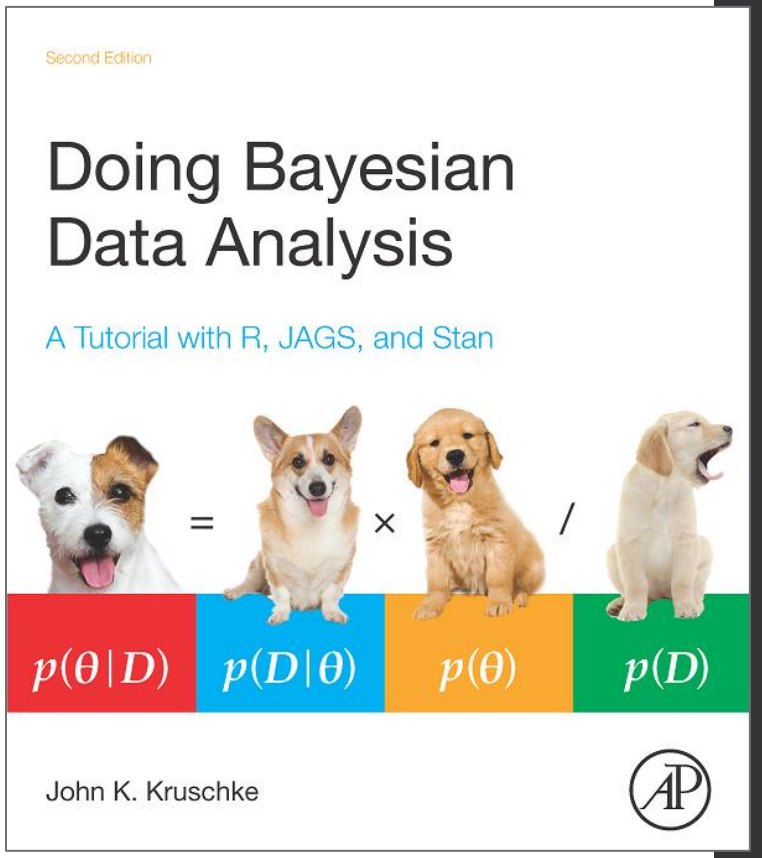
Books – examples



Allen B Downey (2012)
Free e-book (CC-BY-NC), via:
http://www.greenteapress.com/thinkbayes/thinkbayes.pdf



Jean-Michel Marin and Christian Robert (2014)



John K Kruschke (2014)

Thank you!

The workshop starts at 13:30
Room 65 / 2141 (Avenue)

Jakub Bijak, Jon Forster & Jason Hilton

(with credit to Arek Wiśniowski)