



Generalised Bayesian Inference under Prior-Data Conflict

Gero Walter

Department of Statistics
Ludwig-Maximilians-Universität München (LMU)

October 25th, 2013





Overview

1. Bayesian inference & prior-data conflict
2. Generalised Bayesian inference with sets of priors
(joint work with Thomas Augustin and Frank Coolen)
3. Common-cause failure modeling
(joint work with Matthias Troffaes and Dana Kelly)



Overview

1. Bayesian inference & prior-data conflict
 2. Generalised Bayesian inference with sets of priors
(joint work with Thomas Augustin and Frank Coolen)
 3. Common-cause failure modeling
(joint work with Matthias Troffaes and Dana Kelly)



Overview

Contents



Overview

1. Introduction

~~1.1 Declinaciones~~

Contents





Overview

1. Introduction

1.1 Preliminaries

1.3. Dirichlet-Multinomial Model for Common-Cause Failure	20
1.3.1. An Example for the Need to Model Common-Cause Failure	20
1.3.2. The Basic Parameter Model	21
1.3.3. The Alpha-Factor Model	22
1.3.4. Dirichlet Prior for Alpha-Factors	23
1.3.5. Usual Handling of Epistemic Information for Alpha-Factors	24
1.3.6. Cautious Epistemic Information for Alpha-Factors	26
1.3.6.1. Fixed Learning Parameter	27
1.3.6.2. Interval for Learning Parameter	28
1.3.6.3. Conclusion	30

1.3.4. The Normal-Normal Model	24
1.3.5. The Dirichlet-Multinomial Model	26
1.3. Dirichlet-Multinomial Model for Common-Cause Failure	28
1.3.1. An Example for the Need to Model Common-Cause Failure	28
1.3.2. The Basic Parameter Model	28
1.3.3. The Alpha-Factor Model	29
1.3.4. Dirichlet Prior for Alpha-Factors	31
1.3.5. Usual Handling of Epistemic Information for Alpha-Factors	34
1.3.6. Cautious Epistemic Information for Alpha-Factors	35
1.3.6.1. Fixed Learning Parameter	37
1.3.6.2. Interval for Learning Parameter	38
1.3.6.3. Conclusion	38



Overview

vi

Contents

2. Imprecise Probability as Foundation of Generalised Bayesian Inference	31
2.1. Imprecise or Interval Probability	32
2.1.1. General Concept and Basic Interpretation	32
2.1.2. Main Formulations	33
2.1.2.1. Lower Previsions	34
2.1.2.2. Upper Previsions and Strong Supremum	35
2.1.2.3. Sets of Exchangeable Gamblers	36
2.1.2.4. Sets of Probability Distributions	36
2.1.2.5. Conditioning and the Generalised Bayes' Rule	37
2.1.3. Generalised Bayesian Inference / Procedure	38
2.1.3.1. Relation to Bayesian Sensitivity Analysis	38
2.1.3.2. Critique	40
2.1.4. A Brief Glance on Related Concepts	42
2.1.4.1. Belief Functions	42
2.1.4.2. Fuzzy Measures	43
2.1.4.3. Possibilistic Approaches	43
2.2. Motives for the Use of Imprecise Probability Methods	46
2.2.1. The Fundamental Motive	46
2.2.2. Risk and Ambiguity	46
2.2.3. Motives from a Bayesian Perspective	46
2.2.3.1. Foundational Motives	46
2.2.3.2. Weakly Informative Priors	48
2.2.3.3. Prior-Data Conflict	48
2.2.4. Critique and Discussion of Some Alternatives	48
2.2.4.1. Other Gains to Imprecise Probability Models	48
2.2.4.2. Hierarchical Models	49
3. Generalised Bayesian Inference with Sets of Conjugate Priors in Exponential Families	55
3.1. Model Overview and Discussion	56
3.1.1. The General Framework	56
3.1.2. The Beta-Binomial Model	57
3.1.3. The HDM and other Prior Non-Informative Models	58
3.1.4. Substantial Prior Information and Sensitivity to Prior-Data Conflict	58
3.2. Alternative Models Using Sets of Priors	59
3.2.1. Some Alternative Model Frameworks	59
3.2.1.1. Neighborhood Models	59
3.2.1.2. The Density Ratio Class	60
3.2.2. Some Approaches Based on Conjugate Priors	70
3.2.2.1. The Model by Cooley [1991b; 1994]	70
3.2.2.2. The Model by Mackie [1991]	72
3.2.2.3. Some of the Models Studied by Peerschke and Walley [1999]	72
3.2.2.4. Some Other Approaches Using Sets of Priors	73
3.2.2.4.1. The Model by Hindenlang et al. [2013]	73

vii

Contents

Contents	viii
1. Introduction	1
1.1. Imprecise Probability and its Application in Reliability Engineering	1
1.2. Imprecise Probability and its Application in Bayesian Inference	1
1.3. Imprecise Probability and its Application in Common-Cause Failure Modelling	1
1.4. Summary	1
1.5. Acknowledgements	1
1.6. References	1
1.7. Author biography	1
2. Imprecise Probability as Foundation of Generalised Bayesian Inference	1
2.1. Imprecise or Interval Probability	1
2.1.1. General Concept and Basic Interpretation	1
2.1.2. Main Formulations	1
2.1.2.1. Lower Previsions	1
2.1.2.2. Upper Previsions and Strong Supremum	1
2.1.2.3. Sets of Exchangeable Gamblers	1
2.1.2.4. Sets of Probability Distributions	1
2.1.2.5. Conditioning and the Generalised Bayes' Rule	1
2.1.3. Generalised Bayesian Inference / Procedure	1
2.1.3.1. Relation to Bayesian Sensitivity Analysis	1
2.1.3.2. Critique	1
2.1.4. A Brief Glance on Related Concepts	1
2.1.4.1. Belief Functions	1
2.1.4.2. Fuzzy Measures	1
2.1.4.3. Possibilistic Approaches	1
2.2. Motives for the Use of Imprecise Probability Methods	1
2.2.1. The Fundamental Motive	1
2.2.2. Risk and Ambiguity	1
2.2.3. Motives from a Bayesian Perspective	1
2.2.3.1. Foundational Motives	1
2.2.3.2. Weakly Informative Priors	1
2.2.3.3. Prior-Data Conflict	1
2.2.4. Critique and Discussion of Some Alternatives	1
2.2.4.1. Other Gains to Imprecise Probability Models	1
2.2.4.2. Hierarchical Models	1
2.2.5. Summary	1
2.2.6. Acknowledgements	1
2.2.7. References	1
2.2.8. Author biography	1
2.3. Generalised Bayesian Inference with Sets of Conjugate Priors in Exponential Families	1
2.3.1. Model Overview and Discussion	1
2.3.1.1. The General Framework	1
2.3.1.2. The Beta-Binomial Model	1
2.3.1.3. The HDM and other Prior Non-Informative Models	1
2.3.1.4. Substantial Prior Information and Sensitivity to Prior-Data Conflict	1
2.3.2. Alternative Models Using Sets of Priors	1
2.3.2.1. Some Alternative Model Frameworks	1
2.3.2.1.1. Neighborhood Models	1
2.3.2.1.2. The Density Ratio Class	1
2.3.2.2. Some Approaches Based on Conjugate Priors	1
2.3.2.2.1. The Model by Cooley [1991b; 1994]	1
2.3.2.2.2. The Model by Mackie [1991]	1
2.3.2.2.3. Some of the Models Studied by Peerschke and Walley [1999]	1
2.3.2.2.4. Some Other Approaches Using Sets of Priors	1
2.3.2.2.4.1. The Model by Hindenlang et al. [2013]	1
2.3.3. Summary	1
2.3.4. Acknowledgements	1
2.3.5. References	1
2.3.6. Author biography	1
2.4. Summary	1
2.5. Acknowledgements	1
2.6. References	1
2.7. Author biography	1

Contents	xii
1. Introduction	1
1.1. Imprecise Probability and its Application in Reliability Engineering	1
1.2. Imprecise Probability and its Application in Bayesian Inference	1
1.3. Imprecise Probability and its Application in Common-Cause Failure Modelling	1
1.4. Summary	1
1.5. Acknowledgements	1
1.6. References	1
1.7. Author biography	1
2. Imprecise Probability as Foundation of Generalised Bayesian Inference	1
2.1. Imprecise or Interval Probability	1
2.1.1. General Concept and Basic Interpretation	1
2.1.2. Main Formulations	1
2.1.2.1. Lower Previsions	1
2.1.2.2. Upper Previsions and Strong Supremum	1
2.1.2.3. Sets of Exchangeable Gamblers	1
2.1.2.4. Sets of Probability Distributions	1
2.1.2.5. Conditioning and the Generalised Bayes' Rule	1
2.1.3. Generalised Bayesian Inference / Procedure	1
2.1.3.1. Relation to Bayesian Sensitivity Analysis	1
2.1.3.2. Critique	1
2.1.4. A Brief Glance on Related Concepts	1
2.1.4.1. Belief Functions	1
2.1.4.2. Fuzzy Measures	1
2.1.4.3. Possibilistic Approaches	1
2.2. Motives for the Use of Imprecise Probability Methods	1
2.2.1. The Fundamental Motive	1
2.2.2. Risk and Ambiguity	1
2.2.3. Motives from a Bayesian Perspective	1
2.2.3.1. Foundational Motives	1
2.2.3.2. Weakly Informative Priors	1
2.2.3.3. Prior-Data Conflict	1
2.2.4. Critique and Discussion of Some Alternatives	1
2.2.4.1. Other Gains to Imprecise Probability Models	1
2.2.4.2. Hierarchical Models	1
2.2.5. Summary	1
2.2.6. Acknowledgements	1
2.2.7. References	1
2.2.8. Author biography	1
2.2.9. Summary	1
2.2.10. Acknowledgements	1
2.2.11. References	1
2.2.12. Author biography	1
2.2.13. Summary	1
2.3. Generalised Bayesian Inference with Sets of Conjugate Priors in Exponential Families	1
2.3.1. Model Overview and Discussion	1
2.3.1.1. The General Framework	1
2.3.1.2. The Beta-Binomial Model	1
2.3.1.3. The HDM and other Prior Non-Informative Models	1
2.3.1.4. Substantial Prior Information and Sensitivity to Prior-Data Conflict	1
2.3.2. Alternative Models Using Sets of Priors	1
2.3.2.1. Some Alternative Model Frameworks	1
2.3.2.1.1. Neighborhood Models	1
2.3.2.1.2. The Density Ratio Class	1
2.3.2.2. Some Approaches Based on Conjugate Priors	1
2.3.2.2.1. The Model by Cooley [1991b; 1994]	1
2.3.2.2.2. The Model by Mackie [1991]	1
2.3.2.2.3. Some of the Models Studied by Peerschke and Walley [1999]	1
2.3.2.2.4. Some Other Approaches Using Sets of Priors	1
2.3.2.2.4.1. The Model by Hindenlang et al. [2013]	1
2.3.3. Summary	1
2.3.4. Acknowledgements	1
2.3.5. References	1
2.3.6. Author biography	1
2.3.7. Summary	1
2.3.8. Acknowledgements	1
2.3.9. References	1
2.3.10. Author biography	1
2.3.11. Summary	1



Overview

vi

Contents

2. Imprecise Probability as Foundation of Generalised Bayesian Inference	31
2.1. Imprecise or Interval Probability	32
2.1.1. General Concept and Basic Interpretation	32
2.1.2. Main Formulations	33
2.1.2.1. Lower Previsions	34
2.1.2.2. Sets of Acceptable Betting Sets	35
2.1.2.3. Sets of Derivable Gambling Sets	36
2.1.2.4. Sets of Probability Distributions	37
2.1.2.5. Conditioning and the Generalised Bayes' Rule	38

2.2.3.3. Prior-Data Conflict 50

2.1.4.5. Examples for Instrumental Applications	49
2.2. Matrices for the Use of Imprecise Probability Methods	49
2.2.1. The Fundamental Matrix	49
2.2.2. Risk and Ambiguity	49
2.2.3. Motives from a Bayesian Perspective	49
2.2.3.1. Foundational Motives	49
2.2.3.2. Weakly Informative Priors	50
2.2.4. Prior-Data Conflict	50
2.2.4.1. Critique and Discussion of Some Alternatives	50
2.2.4.2. Objectives to Imprecise Probability Models	50
2.2.4.3. Hierarchical Models	52

3. Generalised Bayesian Inference with Sets of Conjugate Priors in Exponential Families	55
3.1. Model Overview and Discussion	56
3.1.1. The General Framework	56
3.1.2. Some Applications	56
3.1.3. The HDM and other Prior Non-Invariance Models	58
3.1.4. Substantial Prior Information and Sensitivity to Prior-Data Conflict	64
3.2. Alternative Models Using Sets of Priors	68
3.2.1. Some Alternative Model Frameworks	68
3.2.1.1. Neighborhood Models	68
3.2.1.2. The Density Ratio Class	69
3.2.2. Some Approaches Based on Conjugate Priors	70
3.2.2.1. The Model by Cooley [1991b; 1994]	70
3.2.2.2. The Model by Mackie [1992]	72
3.2.2.3. Some of the Models Studied by Peerschke and Walley [1999]	72
3.2.2.4. Some Other Approaches Using Sets of Priors	73
3.2.2.4.1. The Model by Hindenlang et al. [2013]	73



Overview

vii Contents

2. Imprecise Probability as Foundation of Generalised Bayesian Inference	31
2.1. Imprecise or Interval Probability	31
2.1.1. General Concept and Basic Interpretation	31

2.2. Motives for the Use of Imprecise Probability Methods 46

 2.2.3. Conditioning and the Generalised Bayes Rule 31

Contents	31
2.2.2.2. The Dumbbell Problem [108]	31
2.2.2.3. Imprecise Bayesian Models & Extended Bayesian Inference	31
2.2.2.3.1. Imprecise Bayesian Models	31
2.2.2.3.2. Extended Bayesian Inference	31
2.2.2.3.3. Imprecise Bayesian Models and Extended Bayesian Inference	31
2.2.2.3.4. Imprecise Bayesian Models and Extended Bayesian Inference	31
2.2.2.3.5. Imprecise Bayesian Models and Extended Bayesian Inference	31
2.2.2.3.6. Imprecise Bayesian Models and Extended Bayesian Inference	31
2.2.2.3.7. Imprecise Bayesian Models and Extended Bayesian Inference	31
2.2.2.3.8. Imprecise Bayesian Models and Extended Bayesian Inference	31
2.2.2.3.9. Imprecise Bayesian Models and Extended Bayesian Inference	31
2.2.2.3.10. Imprecise Bayesian Models and Extended Bayesian Inference	31
2.2.2.3.11. Imprecise Bayesian Models and Extended Bayesian Inference	31
2.2.2.3.12. Imprecise Bayesian Models and Extended Bayesian Inference	31

Contents	31
2.2.2.4. Dumbbell Rule	31
2.2.2.4.1. Dumbbell Rule	31
2.2.2.4.2. Dumbbell Rule	31
2.2.2.4.3. Dumbbell Rule	31
2.2.2.4.4. Dumbbell Rule	31
2.2.2.4.5. Dumbbell Rule	31
2.2.2.4.6. Dumbbell Rule	31
2.2.2.4.7. Dumbbell Rule	31
2.2.2.4.8. Dumbbell Rule	31
2.2.2.4.9. Dumbbell Rule	31
2.2.2.4.10. Dumbbell Rule	31
2.2.2.4.11. Dumbbell Rule	31
2.2.2.4.12. Dumbbell Rule	31

2.2.3.3. Prior-Data Conflict 50

2.2.3.3. Prior-Data Conflict	50
2.2.3.3.1. Motives for the Use of Imprecise Probability Methods 50	50
2.2.3.3.1.1. Foundational Motives 50	50
2.2.3.3.1.2. Risk and Ambiguity 50	50
2.2.3.3.1.3. Motives from a Bayesian Perspective 50	50
2.2.3.3.1.3.1. Foundational Motives 50	50
2.2.3.3.1.3.2. Weakly Informative Priors 50	50
2.2.3.3.1.3.3. Prior-Data Conflict 50	50
2.2.3.3.1.3.3.1. Critique and Discussion of Some Alternatives 50	50
2.2.3.3.1.3.3.1.1. Other Gains to Imprecise Probability Models 50	50
2.2.3.3.1.3.3.1.2. Hierarchical Models 50	50

3. Generalised Bayesian Inference with Sets of Conjugate Priors in Exponential Families	55
3.1. Model Overview and Discussion 55	55
3.1.1. The General Framework 55	55
3.1.2. The General Structure 55	55
3.1.3. The HDM and other Prior Non-Informative Models 55	55
3.1.4. Substantial Prior Information and Sensitivity to Prior-Data Conflict 55	55
3.2. Alternative Models Using Sets of Priors 55	55
3.2.1. Some Alternative Model Frameworks 55	55
3.2.1.1. Neighbourhood Models 55	55
3.2.1.2. The Density Ratio Class 55	55
3.2.2. Some Approaches Based on Conjugate Priors 55	55
3.2.2.1. The Model by Cooley [1991b; 1994] 55	55
3.2.2.2. The Model by Mackie [1991] 55	55
3.2.2.3. Some of the Models Studied by Peerschke and Walley [1993] 55	55
3.2.2.4. Some Other Approaches Using Sets of Priors 55	55
3.2.2.4.1. The Model by Hindenlang et al. [2013] 55	55



Overview

2. Imprecise Probability as Foundation of Generalised Bayesian Inference	31
2.1. Imprecise or Interval Probability	32
2.1.1. General Concept and Basic Interpretation	32
2.2. Motives for the Use of Imprecise Probability Methods	46
2.2.1. The Empirical Motive	46
2.2.1.1. Conditioning and the Generalised Bayes Rule	46
2.2.1.2. Bayesian Inference and Bayesian Decision Theory	47
2.2.1.3. Prior-Data Conflict	48
2.2.1.3.1. Motives for the Use of Imprecise Probability Methods	48
2.2.1.3.1.1. Foundational Motives	48
2.2.1.3.1.2. Motives from a Bayesian Perspective	48
2.2.1.3.1.3. Weakly Informative Priors	49
2.2.1.3.2. Prior-Data Conflict	49
2.2.1.3.2.1. Critique and Discussion of Some Alternatives	49
2.2.1.3.2.2. Objectives to Imprecise Probability Models	50
2.2.1.3.2.3. Hierarchical Models	52
2.2.2. Motives from a Bayesian Perspective	52
2.2.3. Motives from a Bayesian Perspective	53
2.2.3.1. Foundational Motives	53
2.2.3.2. Motives from a Bayesian Perspective	53
2.2.3.3. Weakly Informative Priors	54
2.2.3.4. Prior-Data Conflict	54
2.2.3.4.1. Critique and Discussion of Some Alternatives	54
2.2.3.4.2. Objectives to Imprecise Probability Models	55
2.2.3.4.3. Hierarchical Models	57
2.2.4. Substantive Prior Information and Sensitivity to Prior-Data Conflict	57
2.3. Generalised Bayesian Inference with Sets of Conjugate Priors in Exponential Families	58
2.3.1. Model Overview and Discussion	58
2.3.1.1. The General Framework	58
2.3.1.2. The Beta-Binomial Model	59
2.3.1.3. The HDM and other Prior Non-Informative Models	61
2.3.1.4. Substantial Prior Information and Sensitivity to Prior-Data Conflict	64
2.3.1.5. Alternative Models Using Sets of Priors	65
2.3.1.5.1. Some Alternative Model Frameworks	65
2.3.1.5.1.1. Neighborhood Models	65
2.3.1.5.1.2. The Density Ratio Class	65
2.3.1.5.2. Some Approaches Based on Conjugate Priors	70
2.3.1.5.2.1. The Model by Cooley [1991b; 1994]	70
2.3.1.5.2.2. The Model by Mackie [1991]	72
2.3.1.5.2.3. Some of the Models Studied by Peherstorfer and Walley [1999]	72
2.3.1.5.3. Some Other Approaches Using Sets of Priors	73
2.3.1.5.3.1. The Model by Hindenlang et al. [2013]	73



Overview

2. Imprecise Probability as Foundation of Generalised Bayesian Inference	31
2.1. Imprecise or Interval Probability	32
2.1.1. General Concept and Basic Interpretation	32
2.1.1.1. General Concept and Basic Interpretation	32
2.1.1.2. Imprecise Probability	32
2.1.1.3. The General Framework	32
2.1.1.4. Properties and Criteria	32
2.1.1.5. The IDM and other Prior Near-Ignorance Models	32
2.1.1.6. Substantial Prior Information and Sensitivity to Prior-Data Conflict	32
2.1.2. Motives for the Use of Imprecise Probability Methods	46
2.1.2.1. The Foundational Motives	46
2.1.2.1.1. Motives for the Use of Imprecise Probability Methods	46
2.1.2.1.2. The Fundamental Motive	46
2.1.2.1.3. Risk and Ambiguity	46
2.1.2.1.4. Motives from a Bayesian Perspective	46
2.1.2.1.5. Foundational Motives	46
2.1.2.1.6. Weakly Informative Priors	46
2.1.2.1.7. Prior-Data Conflict	46
2.1.2.1.8. Critique and Discussion of Some Alternatives	46
2.1.2.1.9. Objectives to Imprecise Probability Models	46
2.1.2.1.10. Hierarchical Models	46
2.1.2.2. Generalised Bayesian Inference with Sets of Conjugate Priors in Exponential Families	55
2.1.2.2.1. Model Overview and Discussion	55
2.1.2.2.2. The General Framework	55
2.1.2.2.3. Properties and Criteria	55
2.1.2.2.4. The IDM and other Prior Near-Ignorance Models	55
2.1.2.2.5. Substantial Prior Information and Sensitivity to Prior-Data Conflict	55
3. Generalised Bayesian Inference with Sets of Conjugate Priors in Exponential Families	55
3.1. Model Overview and Discussion	56
3.1.1. The General Framework	56
3.1.2. Properties and Criteria	59
3.1.3. The IDM and other Prior Near-Ignorance Models	61
3.1.4. Substantial Prior Information and Sensitivity to Prior-Data Conflict	64



Overview

C ontents	vii
J.3.3.1. The Model by Whitcomb [181]	75
J.3. Imprecise and Prior-Data Conflict in Generalized Bayesian Inference	75
J.3.1. Introduction	75
J.3.2. Traditional Bayesian Inference and LUCk-models	82
J.3.3. Imprecise Priors for Inference in LUCk-models	84
J.3.3.1. LUCk-models	84
J.3.3.2. LUCk and Prior-Data Conflict	88
J.3.3.3. Imprecise Priors for Inference in LUCk-models	90
J.3.3.4. Generalization of the Generalized LUCk-model	92
J.3.3.5. Concluding Remarks	94
J.4. The Jack Package	95
J.5. On Prior-Data Conflict in Predictive Bernoulli Inferences	101
J.5.1. Introduction	101
J.5.2. Imprecise Beta-Binomial Models	105
J.5.2.1. The Framework	105
J.5.2.2. Wake's pdfs-BBBM	108
J.5.2.3. The Beta-Binomial Shrink Prior	110
J.5.2.4. Interim Beta-Binomial	114
J.5.3. Weighted Inference	114
J.5.3.1. The Basic Model	114
J.5.3.2. The Extended Model	127
J.5.3.3. Weighted Inference Model Properties	138
J.5.4. Insights and Challenges	139
J.6. Concluding Remarks	123
J.6.1. Summary	123
J.6.2. Discussion	124
J.6.3. Outlook	128
A Appendix	135
A.1. Bayesian Linear Regression: Differing of Conjugate Models and Their [In] Semantics Due to Prior-Data Conflict	135
A.1.1. Introduction	135
A.1.2. Prior-data Conflict in the I.I.D. Case	135
A.1.2.1. Samples from a scaled Normal distribution	135
A.1.2.2. Samples from a Multinomial distribution	138
A.1.3. The Standard Approach for Bayesian Linear Regression (SCR)	140
A.1.3.1. Update of β σ^2	140
A.1.3.2. Update of σ^2 β	142
A.1.3.3. Update of β σ^2	143
A.1.4. An Alternative Approach for Conjugate Priors in Bayesian Linear Regression (CCP)	144
A.1.4.1. Update of β σ^2	144

	Page
1. <i>Introduction</i>	1
2. <i>Methodology</i>	2
3. <i>Results</i>	3
4. <i>Discussion</i>	4
5. <i>Conclusion</i>	5
6. <i>References</i>	6
7. <i>Appendix A</i>	7
8. <i>Appendix B</i>	8
9. <i>Appendix C</i>	9
10. <i>Appendix D</i>	10
11. <i>Appendix E</i>	11
12. <i>Appendix F</i>	12
13. <i>Appendix G</i>	13
14. <i>Appendix H</i>	14
15. <i>Appendix I</i>	15
16. <i>Appendix J</i>	16
17. <i>Appendix K</i>	17
18. <i>Appendix L</i>	18
19. <i>Appendix M</i>	19
20. <i>Appendix N</i>	20
21. <i>Appendix O</i>	21
22. <i>Appendix P</i>	22
23. <i>Appendix Q</i>	23
24. <i>Appendix R</i>	24
25. <i>Appendix S</i>	25
26. <i>Appendix T</i>	26
27. <i>Appendix U</i>	27
28. <i>Appendix V</i>	28
29. <i>Appendix W</i>	29
30. <i>Appendix X</i>	30
31. <i>Appendix Y</i>	31
32. <i>Appendix Z</i>	32
33. <i>Appendix AA</i>	33
34. <i>Appendix BB</i>	34
35. <i>Appendix CC</i>	35
36. <i>Appendix DD</i>	36
37. <i>Appendix EE</i>	37
38. <i>Appendix FF</i>	38
39. <i>Appendix GG</i>	39
40. <i>Appendix HH</i>	40
41. <i>Appendix II</i>	41
42. <i>Appendix JJ</i>	42
43. <i>Appendix KK</i>	43
44. <i>Appendix LL</i>	44
45. <i>Appendix MM</i>	45
46. <i>Appendix NN</i>	46
47. <i>Appendix OO</i>	47
48. <i>Appendix PP</i>	48
49. <i>Appendix RR</i>	49
50. <i>Appendix TT</i>	50
51. <i>Appendix YY</i>	51
52. <i>Appendix ZZ</i>	52
53. <i>Appendix AAAA</i>	53
54. <i>Appendix BBBB</i>	54
55. <i>Appendix CCCC</i>	55
56. <i>Appendix DDDD</i>	56
57. <i>Appendix EEEE</i>	57
58. <i>Appendix FFFF</i>	58
59. <i>Appendix GGGG</i>	59
60. <i>Appendix HHHH</i>	60
61. <i>Appendix IIII</i>	61
62. <i>Appendix JJJJ</i>	62
63. <i>Appendix KKKK</i>	63
64. <i>Appendix LLLL</i>	64
65. <i>Appendix MLLL</i>	65
66. <i>Appendix NLLL</i>	66
67. <i>Appendix OLLL</i>	67
68. <i>Appendix PLLL</i>	68
69. <i>Appendix QLLL</i>	69
70. <i>Appendix RLLL</i>	70
71. <i>Appendix SLLL</i>	71
72. <i>Appendix TLLL</i>	72
73. <i>Appendix ULLL</i>	73
74. <i>Appendix VLLL</i>	74
75. <i>Appendix WLLL</i>	75
76. <i>Appendix XLLL</i>	76
77. <i>Appendix YLLL</i>	77
78. <i>Appendix ZLLL</i>	78
79. <i>Appendix AAAAA</i>	79
80. <i>Appendix BBBBB</i>	80
81. <i>Appendix CCCCC</i>	81
82. <i>Appendix DDDDD</i>	82
83. <i>Appendix EEEEE</i>	83
84. <i>Appendix FFFFF</i>	84
85. <i>Appendix GGGGG</i>	85
86. <i>Appendix HHHHH</i>	86
87. <i>Appendix IIIIII</i>	87
88. <i>Appendix JJJJJ</i>	88
89. <i>Appendix KKKKK</i>	89
90. <i>Appendix LLLLL</i>	90
91. <i>Appendix MLLLL</i>	91
92. <i>Appendix NLLLL</i>	92
93. <i>Appendix OLLLL</i>	93
94. <i>Appendix PLLLL</i>	94
95. <i>Appendix QLLLL</i>	95
96. <i>Appendix RLLLL</i>	96
97. <i>Appendix SLLLL</i>	97
98. <i>Appendix TLLLL</i>	98
99. <i>Appendix ULLLL</i>	99
100. <i>Appendix VLLLL</i>	100
101. <i>Appendix WLLLL</i>	101
102. <i>Appendix XLLLL</i>	102
103. <i>Appendix YLLLL</i>	103
104. <i>Appendix ZLLLL</i>	104
105. <i>Appendix AAAAAA</i>	105
106. <i>Appendix BBBBBB</i>	106
107. <i>Appendix CCCCCC</i>	107
108. <i>Appendix DDDDDD</i>	108
109. <i>Appendix EEEEEE</i>	109
110. <i>Appendix FFFFFF</i>	110
111. <i>Appendix GGGGGG</i>	111
112. <i>Appendix HHHHHH</i>	112
113. <i>Appendix IIIIII</i>	113
114. <i>Appendix JJJJJ</i>	114
115. <i>Appendix KKKKK</i>	115
116. <i>Appendix LLLLLL</i>	116
117. <i>Appendix MLLLLL</i>	117
118. <i>Appendix NLLLLL</i>	118
119. <i>Appendix OLLLLL</i>	119
120. <i>Appendix PLLLLL</i>	120
121. <i>Appendix QLLLLL</i>	121
122. <i>Appendix RLLLLL</i>	122
123. <i>Appendix SLLLLL</i>	123
124. <i>Appendix TLLLLL</i>	124
125. <i>Appendix ULLLLL</i>	125
126. <i>Appendix VLLLLL</i>	126
127. <i>Appendix WLLLLL</i>	127
128. <i>Appendix XLLLLL</i>	128
129. <i>Appendix YLLLLL</i>	129
130. <i>Appendix ZLLLLL</i>	130
131. <i>Appendix AAAAAA</i>	131
132. <i>Appendix BBBBBB</i>	132
133. <i>Appendix CCCCCC</i>	133
134. <i>Appendix DDDDDD</i>	134
135. <i>Appendix EEEEEE</i>	135
136. <i>Appendix FFFFFF</i>	136
137. <i>Appendix GGGGGG</i>	137
138. <i>Appendix HHHHHH</i>	138
139. <i>Appendix IIIIII</i>	139
140. <i>Appendix JJJJJ</i>	140
141. <i>Appendix KKKKK</i>	141
142. <i>Appendix LLLLLL</i>	142
143. <i>Appendix MLLLLL</i>	143
144. <i>Appendix NLLLLL</i>	144
145. <i>Appendix OLLLLL</i>	145
146. <i>Appendix PLLLLL</i>	146
147. <i>Appendix QLLLLL</i>	147
148. <i>Appendix RLLLLL</i>	148
149. <i>Appendix SLLLLL</i>	149
150. <i>Appendix TLLLLL</i>	150
151. <i>Appendix ULLLLL</i>	151
152. <i>Appendix VLLLLL</i>	152
153. <i>Appendix WLLLLL</i>	153
154. <i>Appendix XLLLLL</i>	154
155. <i>Appendix YLLLLL</i>	155
156. <i>Appendix ZLLLLL</i>	156
157. <i>Appendix AAAAAA</i>	157
158. <i>Appendix BBBBBB</i>	158
159. <i>Appendix CCCCCC</i>	159
160. <i>Appendix DDDDDD</i>	160
161. <i>Appendix EEEEEE</i>	161
162. <i>Appendix FFFFFF</i>	162
163. <i>Appendix GGGGGG</i>	163
164. <i>Appendix HHHHHH</i>	164
165. <i>Appendix IIIIII</i>	165
166. <i>Appendix JJJJJ</i>	166
167. <i>Appendix KKKKK</i>	167
168. <i>Appendix LLLLLL</i>	168
169. <i>Appendix MLLLLL</i>	169
170. <i>Appendix NLLLLL</i>	170
171. <i>Appendix OLLLLL</i>	171
172. <i>Appendix PLLLLL</i>	172
173. <i>Appendix QLLLLL</i>	173
174. <i>Appendix RLLLLL</i>	174
175. <i>Appendix SLLLLL</i>	175
176. <i>Appendix TLLLLL</i>	176
177. <i>Appendix ULLLLL</i>	177
178. <i>Appendix VLLLLL</i>	178
179. <i>Appendix WLLLLL</i>	179
180. <i>Appendix XLLLLL</i>	180
181. <i>Appendix YLLLLL</i>	181
182. <i>Appendix ZLLLLL</i>	182
183. <i>Appendix AAAAAA</i>	183
184. <i>Appendix BBBBBB</i>	184
185. <i>Appendix CCCCCC</i>	185
186. <i>Appendix DDDDDD</i>	186
187. <i>Appendix EEEEEE</i>	187
188. <i>Appendix FFFFFF</i>	188
189. <i>Appendix GGGGGG</i>	189
190. <i>Appendix HHHHHH</i>	190
191. <i>Appendix IIIIII</i>	191
192. <i>Appendix JJJJJ</i>	192
193. <i>Appendix KKKKK</i>	193
194. <i>Appendix LLLLLL</i>	194
195. <i>Appendix MLLLLL</i>	195
196. <i>Appendix NLLLLL</i>	196
197. <i>Appendix OLLLLL</i>	197
198. <i>Appendix PLLLLL</i>	198
199. <i>Appendix QLLLLL</i>	199
200. <i>Appendix RLLLLL</i>	200
201. <i>Appendix SLLLLL</i>	201
202. <i>Appendix TLLLLL</i>	202
203. <i>Appendix ULLLLL</i>	203
204. <i>Appendix VLLLLL</i>	204
205. <i>Appendix WLLLLL</i>	205
206. <i>Appendix XLLLLL</i>	206
207. <i>Appendix YLLLLL</i>	207
208. <i>Appendix ZLLLLL</i>	208
209. <i>Appendix AAAAAA</i>	209
210. <i>Appendix BBBBBB</i>	210
211. <i>Appendix CCCCCC</i>	211
212. <i>Appendix DDDDDD</i>	212
213. <i>Appendix EEEEEE</i>	213
214. <i>Appendix FFFFFF</i>	214
215. <i>Appendix GGGGGG</i>	215
216. <i>Appendix HHHHHH</i>	216
217. <i>Appendix IIIIII</i>	217
218. <i>Appendix JJJJJ</i>	218
219. <i>Appendix KKKKK</i>	219
220. <i>Appendix LLLLLL</i>	220
221. <i>Appendix MLLLLL</i>	221
222. <i>Appendix NLLLLL</i>	222
223. <i>Appendix OLLLLL</i>	223
224. <i>Appendix PLLLLL</i>	224
225. <i>Appendix QLLLLL</i>	225
226. <i>Appendix RLLLLL</i>	226
227. <i>Appendix SLLLLL</i>	227
228. <i>Appendix TLLLLL</i>	228
229. <i>Appendix ULLLLL</i>	229
230. <i>Appendix VLLLLL</i>	230
231. <i>Appendix WLLLLL</i>	231
232. <i>Appendix XLLLLL</i>	232
233. <i>Appendix YLLLLL</i>	233
234. <i>Appendix ZLLLLL</i>	234
235. <i>Appendix AAAAAA</i>	235
236. <i>Appendix BBBBBB</i>	236
237. <i>Appendix CCCCCC</i>	237
238. <i>Appendix DDDDDD</i>	238
239. <i>Appendix EEEEEE</i>	239
240. <i>Appendix FFFFFF</i>	240
241. <i>Appendix GGGGGG</i>	241
242. <i>Appendix HHHHHH</i>	242
243. <i>Appendix IIIIII</i>	243
244. <i>Appendix JJJJJ</i>	244
245. <i>Appendix KKKKK</i>	245
246. <i>Appendix LLLLLL</i>	246
247. <i>Appendix MLLLLL</i>	247
248. <i>Appendix NLLLLL</i>	248
249. <i>Appendix OLLLLL</i>	249
250. <i>Appendix PLLLLL</i>	250
251. <i>Appendix QLLLLL</i>	251
252. <i>Appendix RLLLLL</i>	252
253. <i>Appendix SLLLLL</i>	253
254. <i>Appendix TLLLLL</i>	254
255. <i>Appendix ULLLLL</i>	255
256. <i>Appendix VLLLLL</i>	256
257. <i>Appendix WLLLLL</i>	257
258. <i>Appendix XLLLLL</i>	258
259. <i>Appendix YLLLLL</i>	259
260. <i>Appendix ZLLLLL</i>	260
261. <i>Appendix AAAAAA</i>	261
262. <i>Appendix BBBBBB</i>	262
263. <i>Appendix CCCCCC</i>	263
264. <i>Appendix DDDDDD</i>	264
265. <i>Appendix EEEEEE</i>	265
266. <i>Appendix FFFFFF</i>	266
267. <i>Appendix GGGGGG</i>	267
268. <i>Appendix HHHHHH</i>	268
269. <i>Appendix IIIIII</i>	269
270. <i>Appendix JJJJJ</i>	270
271. <i>Appendix KKKKK</i>	271
272. <i>Appendix LLLLLL</i>	272
273. <i>Appendix MLLLLL</i>	273
274. <i>Appendix NLLLLL</i>	274
275. <i>Appendix OLLLLL</i>	275
276. <i>Appendix PLLLLL</i>	276
277. <i>Appendix QLLLLL</i>	277
278. <i>Appendix RLLLLL</i>	278
279. <i>Appendix SLLLLL</i>	279
280. <i>Appendix TLLLLL</i>	280
281. <i>Appendix ULLLLL</i>	281
282. <i>Appendix VLLLLL</i>	282
283. <i>Appendix WLLLLL</i>	283
284. <i>Appendix XLLLLL</i>	284
285. <i>Appendix YLLLLL</i>	285
286. <i>Appendix ZLLLLL</i>	286
287. <i>Appendix AAAAAA</i>	287
288. <i>Appendix BBBBBB</i>	288
289. <i>Appendix CCCCCC</i>	289
290. <i>Appendix DDDDDD</i>	290
291. <i>Appendix EEEEEE</i>	291
292. <i>Appendix FFFFFF</i>	292
293. <i>Appendix GGGGGG</i>	293
294. <i>Appendix HHHHHH</i>	294
295. <i>Appendix IIIIII</i>	295
296. <i>Appendix JJJJJ</i>	296
297. <i>Appendix KKKKK</i>	297
298. <i>Appendix LLLLLL</i>	298
299. <i>Appendix MLLLLL</i>	299
300. <i>Appendix NLLLLL</i>	300
301. <i>Appendix OLLLLL</i>	301
302. <i>Appendix PLLLLL</i>	302
303. <i>Appendix QLLLLL</i>	303
304. <i>Appendix RLLLLL</i>	304
305. <i>Appendix SLLLLL</i>	305
306. <i>Appendix TLLLLL</i>	306
307. <i>Appendix ULLLLL</i>	307
308. <i>Appendix VLLLLL</i>	308
309. <i>Appendix WLLLLL</i>	309
310. <i>Appendix XLLLLL</i>	310
311. <i>Appendix YLLLLL</i>	311
312. <i>Appendix ZLLLLL</i>	312
313. <i>Appendix AAAAAA</i>	313
314. <i>Appendix BBBBBB</i>	314
315. <i>Appendix CCCCCC</i>	315
316. <i>Appendix DDDDDD</i>	316
317. <i>Appendix EEEEEE</i>	317
318. <i>Appendix FFFFFF</i>	318
319. <i>Appendix GGGGGG</i>	319
320. <i>Appendix HHHHHH</i>	320
321. <i>Appendix IIIIII</i>	321
322. <i>Appendix JJJJJ</i>	322
323. <i>Appendix KKKKK</i>	323
324. <i>Appendix LLLLLL</i>	324
325. <i>Appendix MLLLLL</i>	325
326. <i>Appendix NLLLLL</i>	326
327. <i>Appendix OLLLLL</i>	327
328. <i>Appendix PLLLLL</i>	328
329. <i>Appendix QLLLLL</i>	329
330. <i>Appendix RLLLLL</i>	330
331. <i>Appendix SLLLLL</i>	331
332. <i>Appendix TLLLLL</i>	332
333. <i>Appendix ULLLLL</i>	333
334. <i>Appendix VLLLLL</i>	334
335. <i>Appendix WLLLLL</i>	335
336. <i>Appendix XLLLLL</i>	336
337. <i>Appendix YLLLLL</i>	337
338. <i>Appendix ZLLLLL</i>	338
339. <i>Appendix AAAAAA</i>	339
340. <i>Appendix BBBBBB</i>	340
341. <i>Appendix CCCCCC</i>	341
342. <i>Appendix DDDDDD</i>	342
343. <i>Appendix EEEEEE</i>	343
344. <i>Appendix FFFFFF</i>	344
345. <i>Appendix GGGGGG</i>	345
346. <i>Appendix HHHHHH</i>	346
347. <i>Appendix IIIIII</i>	347
348. <i>Appendix JJJJJ</i>	348
349. <i>Appendix KKKKK</i>	349
350. <i>Appendix LLLLLL</i>	350
351. <i>Appendix MLLLLL</i>	351
352. <i>Appendix NLLLLL</i>	352
353. <i>Appendix OLLLLL</i>	353
354. <i>Appendix PLLLLL</i>	354
355. <i>Appendix QLLLLL</i>	355
356. <i>Appendix RLLLLL</i>	356
357. <i>Appendix SLLLLL</i>	357
358. <i>Appendix TLLLLL</i>	358
359. <i>Appendix ULLLLL</i>	359
360. <i>Appendix VLLLLL</i>	360
361. <i>Appendix WLLLLL</i>	361
362. <i>Appendix XLLLLL</i>	362
363. <i>Appendix YLLLLL</i>	363
364. <i>Appendix ZLLLLL</i>	364
365. <i>Appendix AAAAAA</i>	365
366. <i>Appendix BBBBBB</i>	366
367. <i>Appendix CCCCCC</i>	367
368. <i>Appendix DDDDDD</i>	368
369. <i>Appendix EEEEEE</i>	369
370. <i>Appendix FFFFFF</i>	370
371. <i>Appendix GGGGGG</i>	371
372. <i>Appendix HHHHHH</i>	372
373. <i>Appendix IIIIII</i>	373
374. <i>Appendix JJJJJ</i>	374
375. <i>Appendix KKKKK</i>	375
376. <i>Appendix LLLLLL</i>	376
377. <i>Appendix MLLLLL</i>	377
378. <i>Appendix NLLLLL</i>	378
379. <i>Appendix OLLLLL</i>	379
380. <i>Appendix PLLLLL</i>	380
381. <i>Appendix QLLLLL</i>	381
382. <i>Appendix RLLLLL</i>	382
383. <i>Appendix SLLLLL</i>	383
384. <i>Appendix TLLLLL</i>	384
385. <i>Appendix ULLLLL</i>	385
386. <i>Appendix VLLLLL</i>	386
387. <i>Appendix WLLLLL</i>	387
388. <i>Appendix XLLLLL</i>	388
389. <i>Appendix YLLLLL</i>	389
390. <i>Appendix ZLLLLL</i>	390
391. <i>Appendix AAAAAA</i>	391
392. <i>Appendix BBBBBB</i>	392
393. <i>Appendix CCCCCC</i>	393
394. <i>Appendix DDDDDD</i>	394
395. <i>Appendix EEEEEE</i>	395
396. <i>Appendix FFFFFF</i>	396
397. <i>Appendix GGGGGG</i>	397
398. <i>Appendix HHHHHH</i>	398
399. <i>Appendix IIIIII</i>	399
400. <i>Appendix JJJJJ</i>	400
401. <i>Appendix KKKKK</i>	401
402. <i>Appendix LLLLLL</i>	402
403. <i>Appendix MLLLLL</i>	403
404. <i>Appendix NLLLL</i>	



Overview

Contents	vii
1.3.3.4. The Model by Whitcomb (1985)	75
1.3.4. Impression and Prior-Data Conflict in Generalised Bayesian Inference	75
1.3.1. Introduction	75
1.3.2. Traditional Bayesian Inference and LUCkS-models	81
1.3.3. Impression Priors for Inference in LUCkS-models	84
1.3.3.1. Impression Priors in LUCkS-models	84
1.3.3.2. If LUCkS-models are Prone to Impression Priors Conflict	88
1.3.4. Improved Impression Priors for Inference in LUCkS-models	98
1.3.5. Illustration of the Generalised LUCkS-model	92
1.3.5.1. Computing Results	93
1.4. The Jack Function	93
1.5. Impression Data Conflict in Predictive Bernoulli Inference	101
1.5.1. Introduction	101
1.5.2. Impression Beta-Binomial Models	103

3.3. Imprecision and Prior-Data Conflict in Generalised Bayesian Inference	79
3.3.1. Introduction	79
3.3.2. Traditional Bayesian Inference and LUCK-models	82
3.3.3. Imprecise Priors for Inference in LUCK-models	84
3.3.3.1. iLUCK-models	84
3.3.3.2. iLUCK-models and Prior-Data Conflict	88
3.3.4. Improved Imprecise Priors for Inference in LUCK-models	90
3.3.5. Illustration of the Generalised iLUCK-model	92

A.1.3. C points of β	141
A.1.4. An Alternative Approach for Conjugate Priors in Bayesian Linear Regression (CCCP)	144
3.1.4.1. Unests of β σ^2	148



Overview

Contents	vii
1.3.3.4. The Model by Whitehead [1995]	75
1.3. Impression and Prior-Data Conflict in Generalised Bayesian Inference	79
1.3.1. Introduction	79
1.3.2. Traditional Bayesian Inference and LUCkS-models	81
1.3.3. Impression Priors for inference in LUCkS-models	84
1.3.3.1. Impression Priors for inference in LUCkS-models	84
1.3.3.2. If LUCkS-models and Prior-Data Conflict	88
1.3.4. Improved Impression Priors for inference in LUCkS-models	90
1.3.5. Illustrating of the Generalised LUCkS-model	92
1.3.6. Summary Remarks	93
1.4. The Jack Package	93
1.5. Impression Priors in Predictive Bayesian Inference	181
1.5.1. Introduction	181
1.5.2. Impression Rate-Binomial Models	181

3.3. Imprecision and Prior-Data Conflict in Generalised Bayesian Inference	79
3.3.1. Introduction	79
3.3.2. Traditional Bayesian Inference and LUCK-models	82
3.3.3. Imprecise Priors for Inference in LUCK-models	84
3.3.3.1. iLUCK-models	84
3.3.3.2. iLUCK-models and Prior-Data Conflict	88
3.3.4. Improved Imprecise Priors for Inference in LUCK-models	90
3.3.5. Illustration of the Generalised iLUCK-model	92

Journal of Statistical Theory and Applications, Vol. 13, No. 4, pp. 113–124, December 2014
ISSN: 1561-7366 print/1561-7366 online © 2014 by Marcel Dekker, Inc.

3.4. The <code>luck</code> Package	98
3.5. On Rcpp, R + C, Shiny, Prediction, Probability, Inf	102



Overview

viii	Contents
	1-8
A.4.4.1. Update of σ^2	1-8
A.4.4.2. Update of β_1	1-13
A.4.5. Diagonal Covariance Matrix	1-17
A.5. A Bayesian Set Shape for Strong Prior-Dominant Modeling	1-17
A.5.1. A Novel Parameterization of Canonical Conjugate Priors	1-17
A.5.2. Informative Reference for Block-Signed Probability Sets	1-18
A.5.3. The Box Shape	1-18
A.5.3.1. Basic Definition	1-18
A.5.3.2. Finding the Touchpoints for the Boxic Set	1-21
A.5.3.3. Strong Prior-Data Agreement Property	1-24
A.5.3.4. General Update with $\sigma > \frac{1}{2}$	1-25
A.5.4. Discussion and Outlook	1-26
Bibliography	1-71



Overview

viii	Contents
A.4.4.2. Updates of θ^0	130
A.4.4.3. Update of β	131
A.4.5. Discussion and Outlook	134
A.2. A Parameter Set Shape for Strong Prior-Data Agreement of Modeling	137
A.2.1. A Novel Parameterisation of Canonical Conjugate Priors	137
A.2.2. The Shape of the Parameter Set Shaped by Item-Specific Parameters	142
A.2.3. The Bootstrap	153
A.2.3.1. Basic Definition	153
A.2.3.2. The Bootstrap Distribution	153
A.2.3.3. Finding the Touchpoints for the Basic Set	153
A.2.3.4. Strong Prior-Data Agreement Property	154
A.2.3.5. General Update with $s > 2$	155
A.2.4. Discussion and Outlook	159
Bibliography	171

A.2. A Parameter Set Shape for Strong Prior-Data Agreement Modelling	157
A.2.1. A Novel Parametrisation of Canonical Conjugate Priors	157
A.2.2. If σ_0^2 Is Large, If σ_0^2 Is Small, If σ_0^2 Is Intermediate	158



Bayesian Inference & Prior-Data Conflict

The Bayesian approach to statistical inference

prior $p(\vartheta)$ + likelihood $f(\mathbf{x} | \vartheta)$ \rightarrow posterior $p(\vartheta | \mathbf{x})$

All inferences are based on the posterior (e.g., point estimate, ...)



Bayesian Inference & Prior-Data Conflict

The Bayesian approach to statistical inference

prior $p(\vartheta)$ + likelihood $f(\mathbf{x} | \vartheta)$ \rightarrow posterior $p(\vartheta | \mathbf{x})$

All inferences are based on the posterior (e.g., point estimate, ...)

Assigning a certain prior distribution on ϑ

= defining a conglomerate of probability statements (on ϑ)



Bayesian Inference & Prior-Data Conflict

The Bayesian approach to statistical inference

prior $p(\vartheta)$ + likelihood $f(\mathbf{x} | \vartheta)$ \rightarrow posterior $p(\vartheta | \mathbf{x})$

All inferences are based on the posterior (e.g., point estimate, ...)

Assigning a certain prior distribution on ϑ

= defining a conglomerate of probability statements (on ϑ)

Prior-Data Conflict

- ▶ *informative prior beliefs* and *trusted data* (sampling model correct, no outliers, etc.) are in conflict
- ▶ “[...] the prior [places] its mass primarily on distributions in the sampling model for which the observed data is surprising” (Evans & Moshonov, 2006)
- ▶ there are not enough data to overrule the prior



Prior-Data Conflict: Basic Example

- ▶ Bernoulli observations: 0/1 observations (team wins no/yes)



Prior-Data Conflict: Basic Example

- ▶ Bernoulli observations: 0/1 observations (team wins no/yes)
- ▶ given: a set of observations and strong prior information



Prior-Data Conflict: Basic Example

- ▶ Bernoulli observations: 0/1 observations (team wins no/yes)
- ▶ given: a set of observations and strong prior information
- ▶ we are, e.g., interested in (predictive) probability P that team wins in the next match



Prior-Data Conflict: Basic Example

- ▶ Bernoulli observations: 0/1 observations (team wins no/yes)
- ▶ given: a set of observations and strong prior information
- ▶ we are, e.g., interested in (predictive) probability P that team wins in the next match

Beta-Binomial Model

data :	$s \theta$	\sim	$\text{Binom}(n, \theta)$
conjugate prior:	$\theta n^{(0)}, y^{(0)}$	\sim	$\text{Beta}(n^{(0)}, y^{(0)})$
posterior:	$\theta n^{(n)}, y^{(n)}$	\sim	$\text{Beta}(n^{(n)}, y^{(n)})$

where s = number of wins in the n matches observed



Beta-Binomial Model (BBM)

Beta-Binomial Model

data :	$s \theta$	\sim	$\text{Binom}(n, \theta)$
conjugate prior:	$\theta n^{(0)}, y^{(0)}$	\sim	$\text{Beta}(n^{(0)}, y^{(0)})$
posterior:	$\theta n^{(n)}, y^{(n)}$	\sim	$\text{Beta}(n^{(n)}, y^{(n)})$



Beta-Binomial Model (BBM)

Beta-Binomial Model

data :	$s \theta$	\sim	$\text{Binom}(n, \theta)$
conjugate prior:	$\theta n^{(0)}, y^{(0)}$	\sim	$\text{Beta}(n^{(0)}, y^{(0)})$
posterior:	$\theta n^{(n)}, y^{(n)}$	\sim	$\text{Beta}(n^{(n)}, y^{(n)})$

$$P = E[\theta | n^{(n)}, y^{(n)}]$$



Beta-Binomial Model (BBM)

Beta-Binomial Model

data :	$s \theta$	\sim	$\text{Binom}(n, \theta)$
conjugate prior:	$\theta n^{(0)}, y^{(0)}$	\sim	$\text{Beta}(n^{(0)}, y^{(0)})$
posterior:	$\theta n^{(n)}, y^{(n)}$	\sim	$\text{Beta}(n^{(n)}, y^{(n)})$

$$P = E[\theta | n^{(n)}, y^{(n)}] = y^{(n)} = \frac{n^{(0)}}{n^{(0)} + n} \cdot y^{(0)} + \frac{n}{n^{(0)} + n} \cdot \frac{s}{n}$$



Beta-Binomial Model (BBM)

Beta-Binomial Model

data :	$s \theta$	\sim	$\text{Binom}(n, \theta)$
conjugate prior:	$\theta n^{(0)}, y^{(0)}$	\sim	$\text{Beta}(n^{(0)}, y^{(0)})$
posterior:	$\theta n^{(n)}, y^{(n)}$	\sim	$\text{Beta}(n^{(n)}, y^{(n)})$

$$P = E[\theta | n^{(n)}, y^{(n)}] = y^{(n)} = \frac{n^{(0)}}{n^{(0)} + n} \cdot y^{(0)} + \frac{n}{n^{(0)} + n} \cdot \frac{s}{n}$$

$$n^{(n)} = n^{(0)} + n$$



Beta-Binomial Model (BBM)

Beta-Binomial Model

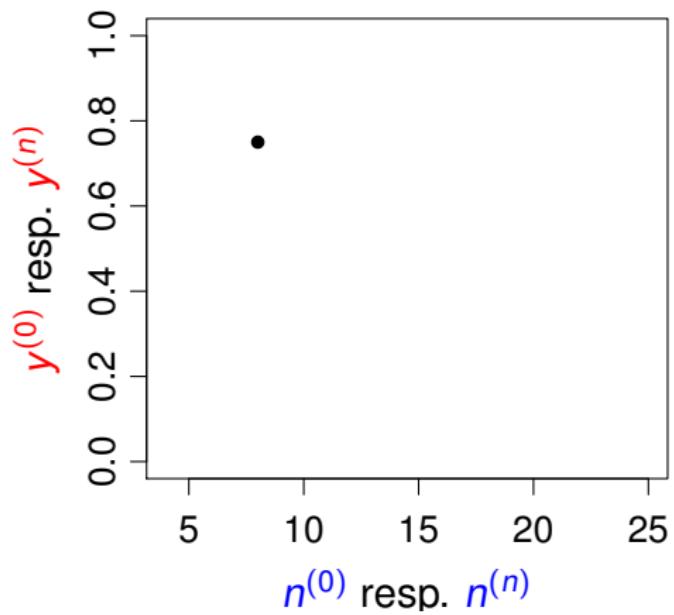
data :	$s \theta$	\sim	$\text{Binom}(n, \theta)$
conjugate prior:	$\theta n^{(0)}, y^{(0)}$	\sim	$\text{Beta}(n^{(0)}, y^{(0)})$
posterior:	$\theta n^{(n)}, y^{(n)}$	\sim	$\text{Beta}(n^{(n)}, y^{(n)})$

$$P = E[\theta | n^{(n)}, y^{(n)}] = y^{(n)} = \frac{n^{(0)}}{n^{(0)} + n} \cdot y^{(0)} + \frac{n}{n^{(0)} + n} \cdot \frac{s}{n}$$

$$n^{(n)} = n^{(0)} + n \quad \text{Var}(\theta | n^{(n)}, y^{(n)}) = \frac{y^{(n)}(1 - y^{(n)})}{n^{(n)} + 1}$$



Beta-Binomial Model (BBM)

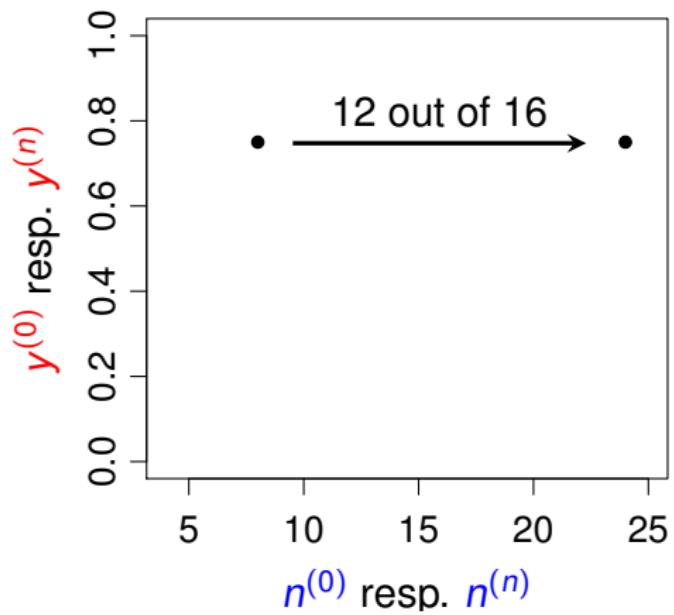


no conflict:

prior $n^{(0)} = 8$, $y^{(0)} = 0.75$
data $s/n = 12/16 = 0.75$



Beta-Binomial Model (BBM)



no conflict:

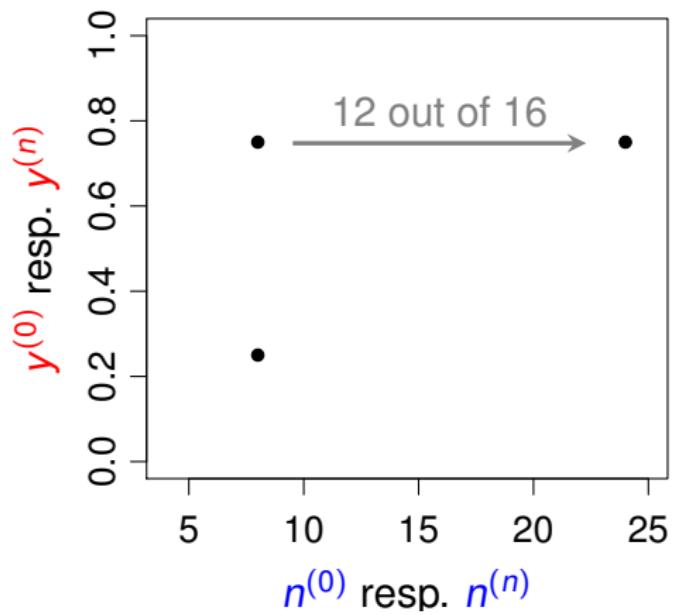
prior $n^{(0)} = 8$, $y^{(0)} = 0.75$

data s/n = $12/16 = 0.75$

$n^{(n)} = 24$, $y^{(n)} = 0.75$



Beta-Binomial Model (BBM)



no conflict:

prior $n^{(0)} = 8$, $y^{(0)} = 0.75$
data $s/n = 12/16 = 0.75$

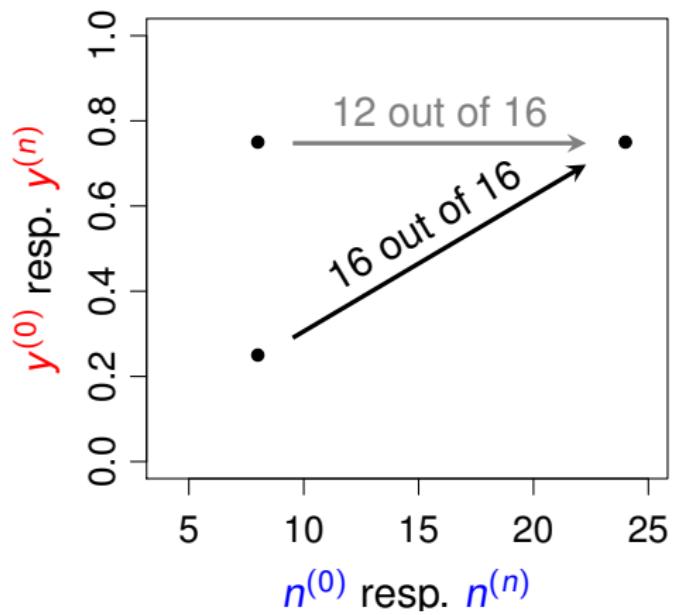
$n^{(n)} = 24$, $y^{(n)} = 0.75$

prior-data conflict:

prior $n^{(0)} = 8$, $y^{(0)} = 0.25$
data $s/n = 16/16 = 1$



Beta-Binomial Model (BBM)



no conflict:

prior $n^{(0)} = 8, y^{(0)} = 0.75$
data $s/n = 12/16 = 0.75$

$n^{(n)} = 24, y^{(n)} = 0.75$

prior-data conflict:

prior $n^{(0)} = 8, y^{(0)} = 0.25$
data $s/n = 16/16 = 1$



Canonical Conjugate Priors

Weighted average structure is underneath **all common conjugate priors** for exponential family sampling distributions!

$(x_1, \dots, x_n) = \mathbf{x} \stackrel{iid}{\sim}$ canonical exponential family

$$p(\mathbf{x} | \vartheta) \propto \exp \left\{ \langle \psi, \tau(\mathbf{x}) \rangle - n b(\psi) \right\} \quad [\psi \text{ transformation of } \vartheta]$$

(includes Binomial, Multinomial, Normal, Poisson, Exponential, ...)



Canonical Conjugate Priors

Weighted average structure is underneath **all common conjugate priors** for exponential family sampling distributions!

$(x_1, \dots, x_n) = \mathbf{x} \stackrel{iid}{\sim}$ canonical exponential family

$$p(\mathbf{x} | \vartheta) \propto \exp \left\{ \langle \psi, \tau(\mathbf{x}) \rangle - n b(\psi) \right\} \quad [\psi \text{ transformation of } \vartheta]$$

(includes Binomial, Multinomial, Normal, Poisson, Exponential, ...)

► conjugate prior: $p(\psi | n^{(0)}, y^{(0)}) \propto \exp \left\{ n^{(0)} [\langle \psi, y^{(0)} \rangle - b(\psi)] \right\}$



Canonical Conjugate Priors

Weighted average structure is underneath **all common conjugate priors** for exponential family sampling distributions!

$(x_1, \dots, x_n) = \mathbf{x} \stackrel{iid}{\sim}$ canonical exponential family

$$p(\mathbf{x} | \vartheta) \propto \exp \left\{ \langle \psi, \tau(\mathbf{x}) \rangle - n b(\psi) \right\} \quad [\psi \text{ transformation of } \vartheta]$$

(includes Binomial, Multinomial, Normal, Poisson, Exponential, ...)

- conjugate prior: $p(\psi | n^{(0)}, y^{(0)}) \propto \exp \left\{ n^{(0)} [\langle \psi, y^{(0)} \rangle - b(\psi)] \right\}$
- (conjugate) posterior: $p(\psi | n^{(0)}, y^{(0)}, \mathbf{x}) \propto \exp \left\{ n^{(n)} [\langle \psi, y^{(n)} \rangle - b(\psi)] \right\}$

where $y^{(n)} = \frac{n^{(0)}}{n^{(0)} + n} \cdot y^{(0)} + \frac{n}{n^{(0)} + n} \cdot \frac{\tau(\mathbf{x})}{n}$ and $n^{(n)} = n^{(0)} + n$



Canonical Conjugate Priors

► (conjugate) posterior: $p(\psi | n^{(n)}, y^{(n)}) \propto \exp\{n^{(n)}[\langle \psi, y^{(n)} \rangle - b(\psi)]\}$

where $y^{(n)} = \frac{n^{(0)}}{n^{(0)} + n} \cdot y^{(0)} + \frac{n}{n^{(0)} + n} \cdot \frac{\tau(\mathbf{x})}{n}$ and $n^{(n)} = n^{(0)} + n$



Canonical Conjugate Priors

► (conjugate) posterior: $p(\psi | n^{(n)}, y^{(n)}) \propto \exp\{n^{(n)}[\langle \psi, y^{(n)} \rangle - b(\psi)]\}$

where $y^{(n)} = \frac{n^{(0)}}{n^{(0)} + n} \cdot y^{(0)} + \frac{n}{n^{(0)} + n} \cdot \frac{\tau(\mathbf{x})}{n}$ and $n^{(n)} = n^{(0)} + n$

Interpretation

- $n^{(0)}$ = determines spread and learning speed
- $y^{(0)}$ = prior expectation of $\tau(\mathbf{x})/n$



Canonical Conjugate Priors

► (conjugate) posterior: $p(\psi | n^{(n)}, y^{(n)}) \propto \exp\{n^{(n)}[\langle \psi, y^{(n)} \rangle - b(\psi)]\}$

where $y^{(n)} = \frac{n^{(0)}}{n^{(0)} + n} \cdot y^{(0)} + \frac{n}{n^{(0)} + n} \cdot \frac{\tau(\mathbf{x})}{n}$ and $n^{(n)} = n^{(0)} + n$

Interpretation

- $n^{(0)}$ = determines spread and learning speed
- $y^{(0)}$ = prior expectation of $\tau(\mathbf{x})/n$

Example: Scaled Normal Data

Data :	$\mathbf{x} \mu \sim N(\mu, 1)$
conjugate prior:	$\mu n^{(0)}, y^{(0)} \sim N(y^{(0)}, 1/n^{(0)})$
posterior:	$\mu n^{(n)}, y^{(n)} \sim N(y^{(n)}, 1/n^{(n)}) \quad (\frac{\tau(\mathbf{x})}{n} = \bar{\mathbf{x}})$



Why Generalise Bayesian Inference?

Bayesian theory lacks the ability to specify the degree of uncertainty in probability statements encoded in a (prior, posterior) distribution.

Foundational arguments regarding over-precision of the classical Bayesian framework

Comments on hierarchical modelling



Why Generalise Bayesian Inference?

Bayesian theory lacks the ability to specify the degree of uncertainty in probability statements encoded in a (prior, posterior) distribution.

Variance or stretch of a distribution for describing uncertainty?

Foundational arguments regarding over-precision of the classical Bayesian framework

Comments on hierarchical modelling



Why Generalise Bayesian Inference?

Bayesian theory lacks the ability to specify the degree of uncertainty in probability statements encoded in a (prior, posterior) distribution.

Variance or stretch of a distribution for describing uncertainty?

- Does not work in the case of prior-data conflict:
In conjugate updating, the posterior variance does not depend on the degree of prior-data conflict in most cases.
- How to express the precision of a probability statement?

Foundational arguments regarding over-precision of the classical Bayesian framework

Comments on hierarchical modelling



Imprecision

Add **imprecision** as new model dimension:
Sets of priors model uncertainty in probability statements
and allow to better model partial information on ϑ



Imprecision

Add **imprecision** as new model dimension:
Sets of priors model uncertainty in probability statements
and allow to better model partial information on ϑ

Interpretation

smaller sets = more precise probability statements

Lottery A

Number of winning tickets:
exactly known as 5 out of 100

$$\rightarrow P(\text{win}) = 5/100$$

Lottery B

Number of winning tickets:
not exactly known, supposedly
between 1 and 7 out of 100

$$\rightarrow P(\text{win}) = [1/100, 7/100]$$



Bayesian Inference with Sets of Priors

Standard Bayesian inference procedure

prior + likelihood \rightarrow posterior

using Bayes' Rule

All inferences are based on the posterior



Bayesian Inference with Sets of Priors

Standard Bayesian inference procedure

prior + likelihood → posterior

using Bayes' Rule

All inferences are based on the posterior

Generalised Bayesian inference procedure

set of priors + likelihood → set of posteriors

Coherence (consistency of inferences) ensured by using

Generalised Bayes' Rule (GBR, Walley 1991)

= element-wise application of Bayes' Rule

All inferences are based on the set of posteriors



Bayesian Inference with Sets of Priors

Standard Bayesian inference procedure

prior + likelihood → posterior

using Bayes' Rule

All inferences are based on the posterior

Generalised Bayesian inference procedure

set of priors + likelihood → set of posteriors

Coherence (consistency of inferences) ensured by using

Generalised Bayes' Rule (GBR, Walley 1991)

= element-wise application of Bayes' Rule

All inferences are based on the set of posteriors

Let hyperparameters ($n^{(0)}$, $y^{(0)}$) vary in a set → set of priors $\mathcal{M}^{(0)}$



Bayesian Inference with Sets of Priors

Set of posteriors $\mathcal{M}^{(n)}$ via $= \{(n^{(n)}, y^{(n)}): (n^{(0)}, y^{(0)}) \in \quad\}$

single posterior $p(n^{(n)}, y^{(n)}) \rightarrow$ set of posteriors $\mathcal{M}^{(n)}$ (via)

$E[\psi | n^{(n)}, y^{(n)}] \rightarrow [\underline{E}[\psi | \quad], \bar{E}[\psi | \quad]]$

$P(\psi \in A | n^{(n)}, y^{(n)}) \rightarrow [\underline{P}[\psi \in A | \quad], \bar{P}[\psi \in A | \quad]]$

HPD interval \rightarrow union of HPD intervals

Lower/upper bounds by min/max over set of posteriors



Taking the Convex Hull as the Set of Priors

Convex Set of Priors

$$\mathcal{M}^{(0)} = \text{conv}\left(\left\{p(\psi | n^{(0)}, \mathbf{y}^{(0)}) : (n^{(0)}, \mathbf{y}^{(0)}) \in \quad \right\}\right)$$

$\mathcal{M}^{(0)}$ = finite convex mixtures of canonical conjugate priors defined by



Taking the Convex Hull as the Set of Priors

Convex Set of Priors

$$\mathcal{M}^{(0)} = \text{conv}\left(\left\{p(\psi | n^{(0)}, \mathbf{y}^{(0)}): (n^{(0)}, \mathbf{y}^{(0)}) \in \quad \right\}\right)$$

$\mathcal{M}^{(0)}$ = finite convex mixtures of canonical conjugate priors defined by

Updating & mixture commute → set of posteriors can be written as...

Convex Set of Posteriors

$$\mathcal{M}^{(n)} = \text{conv}\left(\left\{p(\psi | n^{(n)}, \mathbf{y}^{(n)}): (n^{(n)}, \mathbf{y}^{(n)}) \in \quad \right\}\right)$$

$\mathcal{M}^{(n)}$ = finite convex mixtures of canonical conjugate posteriors defined by set of updated hyperparameters



Taking the Convex Hull as the Set of Priors

Convex Set of Priors

$$\mathcal{M}^{(0)} = \text{conv}\left(\left\{p(\psi | n^{(0)}, \mathbf{y}^{(0)}): (n^{(0)}, \mathbf{y}^{(0)}) \in \quad \right\}\right)$$

$\mathcal{M}^{(0)}$ = finite convex mixtures of canonical conjugate priors defined by

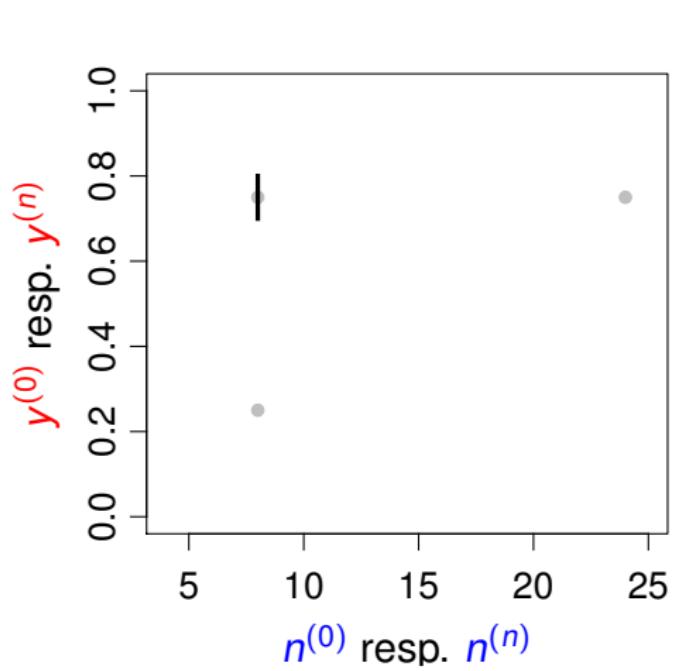
Updating & mixture commute → set of posteriors can be written as...

Convex Set of Posteriors

$$\mathcal{M}^{(n)} = \text{conv}\left(\left\{p(\psi | n^{(n)}, \mathbf{y}^{(n)}): (n^{(n)}, \mathbf{y}^{(n)}) \in \quad \right\}\right)$$

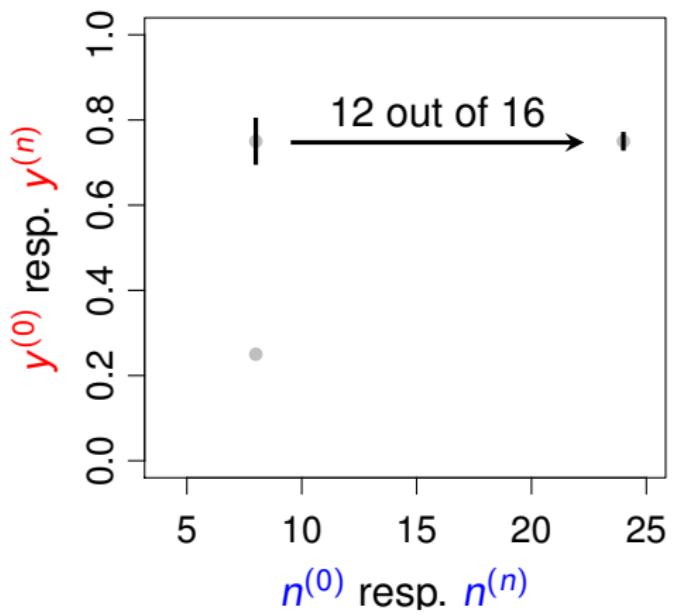
$\mathcal{M}^{(n)}$ = finite convex mixtures of canonical conjugate posteriors defined by set of updated hyperparameters

Convex sets make the procedure very general (mixture distributions), but are useful only for inferences that are *linear* in the posteriors (expectations: yes, variances: no)

Imprecise BBM with $n^{(0)}$ fixed: IDM (Walley 1996)
Quaghebeur & de Cooman (2005)

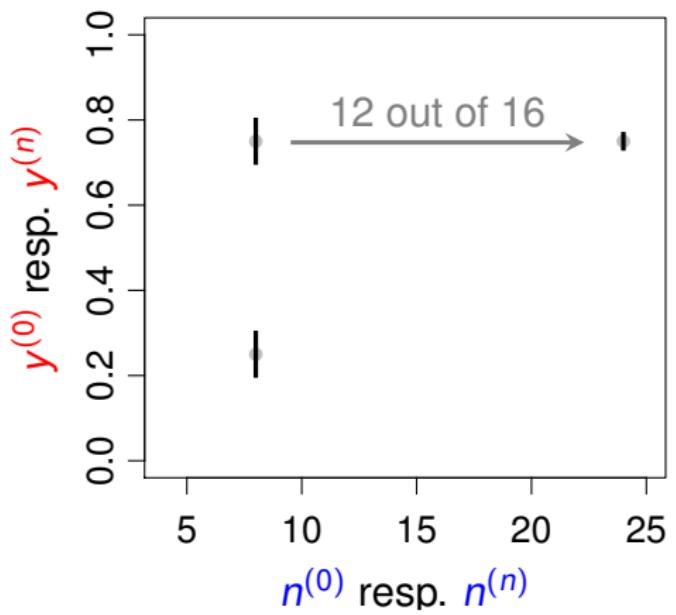
no conflict:

prior $n^{(0)} = 8$, $y^{(0)} \in [0.7, 0.8]$
data $s/n = 12/16 = 0.75$

Imprecise BBM with $n^{(0)}$ fixed: IDM (Walley 1996)
Quaghebeur & de Cooman (2005)

no conflict:

prior $n^{(0)} = 8$, $y^{(0)} \in [0.7, 0.8]$ data $s/n = 12/16 = 0.75$ $n^{(n)} = 24$, $y^{(n)} \in [0.73, 0.77]$

Imprecise BBM with $n^{(0)}$ fixed: IDM (Walley 1996)
Quaghebeur & de Cooman (2005)

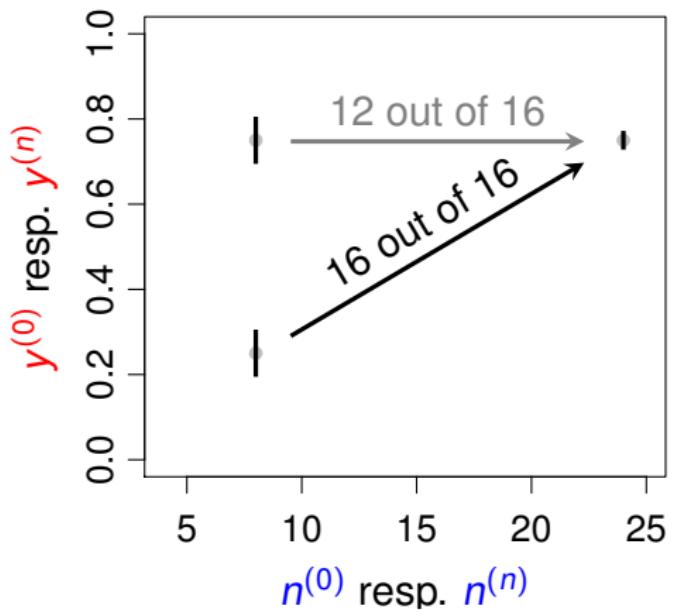
no conflict:

prior $n^{(0)} = 8$, $y^{(0)} \in [0.7, 0.8]$
data $s/n = 12/16 = 0.75$

$n^{(n)} = 24$, $y^{(n)} \in [0.73, 0.77]$

prior-data conflict:

prior $n^{(0)} = 8$, $y^{(0)} \in [0.2, 0.3]$
data $s/n = 16/16 = 1$

Imprecise BBM with $n^{(0)}$ fixed: IDM (Walley 1996)
Quaghebeur & de Cooman (2005)

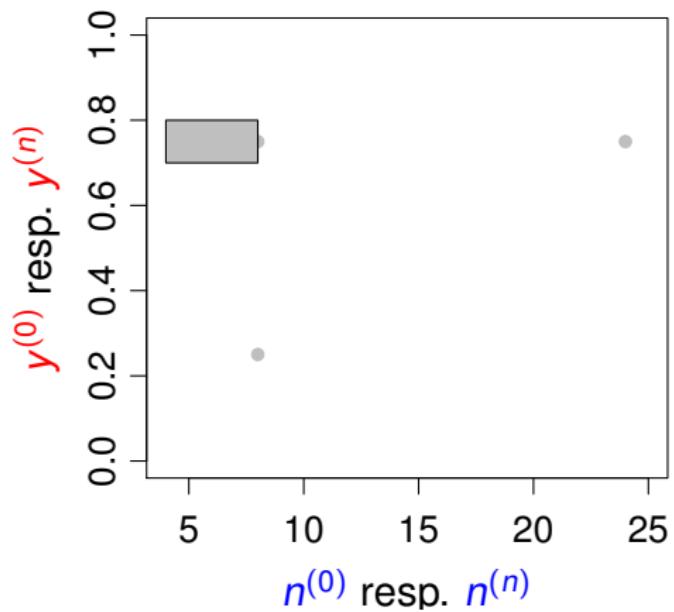
no conflict:

prior $n^{(0)} = 8$, $y^{(0)} \in [0.7, 0.8]$
data $s/n = 12/16 = 0.75$

$n^{(n)} = 24$, $y^{(n)} \in [0.73, 0.77]$

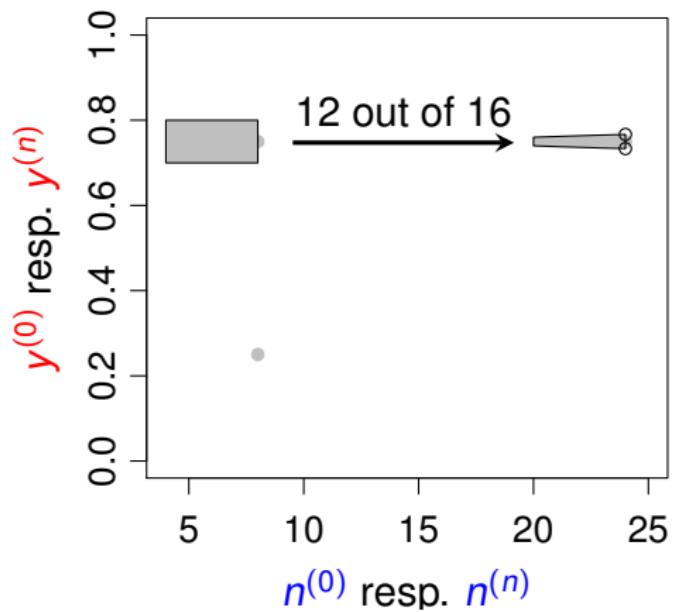
prior-data conflict:

prior $n^{(0)} = 8$, $y^{(0)} \in [0.2, 0.3]$
data $s/n = 16/16 = 1$

Imprecise BBM with $[\underline{n}^{(0)}, \bar{n}^{(0)}]$: Walley (1991, §5.4.3)
Walter & Augustin (2009)

no conflict:

prior $n^{(0)} \in [4, 8]$, $y^{(0)} \in [0.7, 0.8]$
data s/n = $12/16 = 0.75$

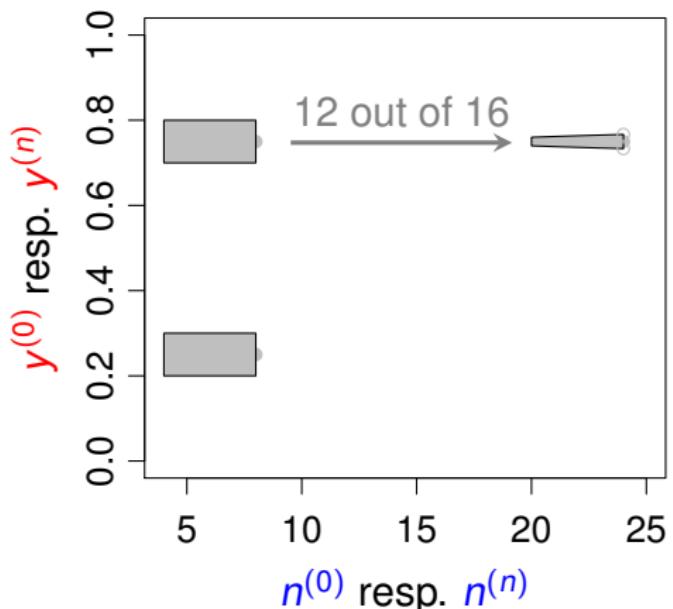
Imprecise BBM with $[\underline{n}^{(0)}, \bar{n}^{(0)}]$: Walley (1991, §5.4.3)
Walter & Augustin (2009)

no conflict:

prior $n^{(0)} \in [4, 8]$, $y^{(0)} \in [0.7, 0.8]$
data s/n = $12/16 = 0.75$

$$y^{(n)} \in [0.73, 0.77]$$



Imprecise BBM with $[n^{(0)}, \bar{n}^{(0)}]$: Walley (1991, §5.4.3)
Walter & Augustin (2009)

no conflict:

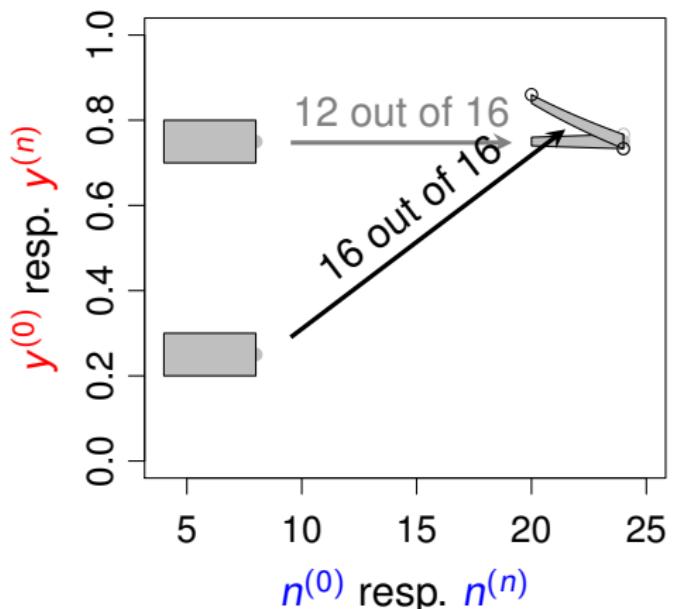
prior $n^{(0)} \in [4, 8]$, $y^{(0)} \in [0.7, 0.8]$
data s/n = $12/16 = 0.75$

$$y^{(n)} \in [0.73, 0.77]$$

prior-data conflict:

prior $n^{(0)} \in [4, 8]$, $y^{(0)} \in [0.2, 0.3]$
data s/n = $16/16 = 1$



Imprecise BBM with $[n^{(0)}, \bar{n}^{(0)}]$: Walley (1991, §5.4.3)
Walter & Augustin (2009)

no conflict:

prior $n^{(0)} \in [4, 8]$, $y^{(0)} \in [0.7, 0.8]$
data $s/n = 12/16 = 0.75$

$$y^{(n)} \in [0.73, 0.77]$$

prior-data conflict:

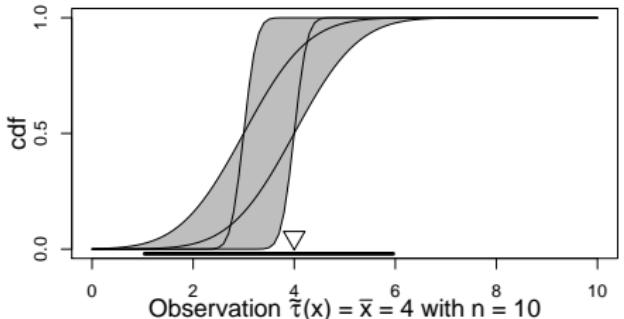
prior $n^{(0)} \in [4, 8]$, $y^{(0)} \in [0.2, 0.3]$
data $s/n = 16/16 = 1$

$$y^{(n)} \in [0.73, 0.86]$$

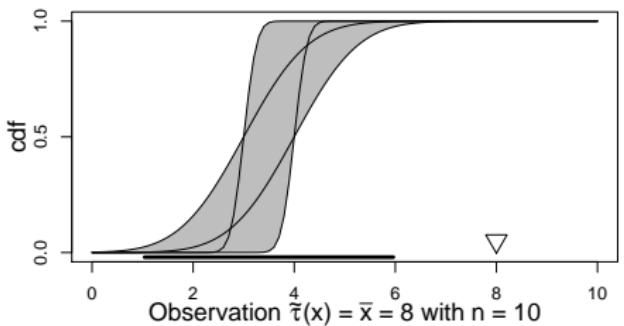


Scaled Normal Data Example ($\mathbf{x} \stackrel{iid}{\sim} N(\mu, 1)$)

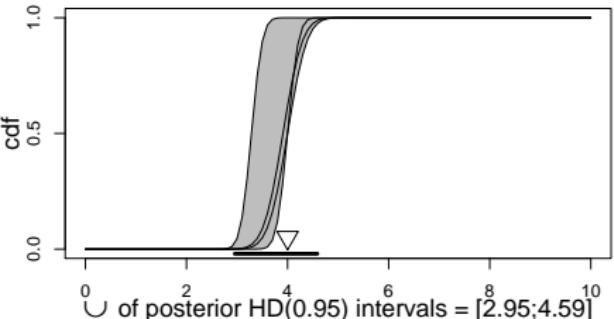
Set of priors: $y^{(0)} \in [3;4]$ and $n^{(0)} \in [1;25]$



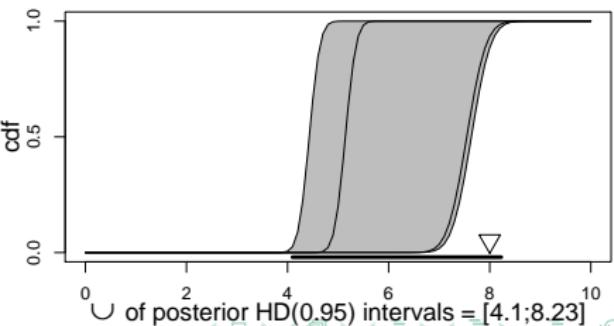
Set of priors: $y^{(0)} \in [3;4]$ and $n^{(0)} \in [1;25]$



Set of posteriors: $y^{(1)} \in [3.29;4]$ and $n^{(1)} \in [11;35]$



Set of posteriors: $y^{(1)} \in [4.43;7.64]$ and $n^{(1)} \in [11;35]$





General Model Properties

Favourable inference properties (cf. other models based on sets of priors)

- ▶ $n \rightarrow \infty$



General Model Properties

Favourable inference properties (cf. other models based on sets of priors)

- $n \rightarrow \infty \rightarrow y^{(n)}$ stretch in $\rightarrow 0$



General Model Properties

Favourable inference properties (cf. other models based on sets of priors)

- $n \rightarrow \infty \rightarrow y^{(n)}$ stretch in $\rightarrow 0 \rightarrow$ precise inferences



General Model Properties

Favourable inference properties (cf. other models based on sets of priors)

- $n \rightarrow \infty \rightarrow y^{(n)}$ stretch in $\rightarrow 0 \rightarrow$ precise inferences
- larger $n^{(0)} \rightarrow$ larger \rightarrow more vague inferences



General Model Properties

Favourable inference properties (cf. other models based on sets of priors)

- ▶ $n \rightarrow \infty \rightarrow y^{(n)}$ stretch in $\rightarrow 0 \rightarrow$ precise inferences
- ▶ larger $n^{(0)} \rightarrow$ larger \rightarrow more vague inferences
- ▶ larger range of $y^{(0)}$ in \rightarrow larger range of $y^{(n)}$ in
 \rightarrow more vague inferences



General Model Properties

Favourable inference properties (cf. other models based on sets of priors)

- $n \rightarrow \infty \rightarrow y^{(n)}$ stretch in $\rightarrow 0 \rightarrow$ precise inferences
- larger $n^{(0)} \rightarrow$ larger \rightarrow more vague inferences
- larger range of $y^{(0)}$ in \rightarrow larger range of $y^{(n)}$ in
 \rightarrow more vague inferences

Model very easy to handle:

- Hyperparameter set defines set of priors $\mathcal{M}^{(0)}$



General Model Properties

Favourable inference properties (cf. other models based on sets of priors)

- ▶ $n \rightarrow \infty \rightarrow y^{(n)}$ stretch in $\rightarrow 0 \rightarrow$ precise inferences
- ▶ larger $n^{(0)} \rightarrow$ larger \rightarrow more vague inferences
- ▶ larger range of $y^{(0)}$ in \rightarrow larger range of $y^{(n)}$ in
 \rightarrow more vague inferences

Model very easy to handle:

- ▶ Hyperparameter set defines set of priors $\mathcal{M}^{(0)}$
- ▶ Hyperparameter set defines set of posteriors $\mathcal{M}^{(n)}$



General Model Properties

Favourable inference properties (cf. other models based on sets of priors)

- ▶ $n \rightarrow \infty \rightarrow y^{(n)}$ stretch in $\rightarrow 0 \rightarrow$ precise inferences
- ▶ larger $n^{(0)} \rightarrow$ larger \rightarrow more vague inferences
- ▶ larger range of $y^{(0)}$ in \rightarrow larger range of $y^{(n)}$ in
 \rightarrow more vague inferences

Model very easy to handle:

- ▶ Hyperparameter set defines set of priors $\mathcal{M}^{(0)}$
- ▶ Hyperparameter set defines set of posteriors $\mathcal{M}^{(n)}$
- ▶ \rightarrow is easy: $n^{(n)} = n^{(0)} + n$, $y^{(n)} = \frac{n^{(0)}}{n^{(0)}+n}y^{(0)} + \frac{n}{n^{(0)}+n}\tau(\mathbf{x})$



General Model Properties

Favourable inference properties (cf. other models based on sets of priors)

- ▶ $n \rightarrow \infty \rightarrow y^{(n)}$ stretch in $\rightarrow 0 \rightarrow$ precise inferences
- ▶ larger $n^{(0)} \rightarrow$ larger \rightarrow more vague inferences
- ▶ larger range of $y^{(0)}$ in \rightarrow larger range of $y^{(n)}$ in \rightarrow more vague inferences

Model very easy to handle:

- ▶ Hyperparameter set defines set of priors $\mathcal{M}^{(0)}$
- ▶ Hyperparameter set defines set of posteriors $\mathcal{M}^{(n)}$
- ▶ \rightarrow is easy: $n^{(n)} = n^{(0)} + n$, $y^{(n)} = \frac{n^{(0)}}{n^{(0)}+n}y^{(0)} + \frac{n}{n^{(0)}+n}\tau(\mathbf{x})$
- ▶ For quantities linear in posteriors, bounds are attained at “pure” posteriors $p(\psi | n^{(n)}, y^{(n)}) \rightarrow$ optimise over (not over $\mathcal{M}^{(n)}$)



General Model Properties

Favourable inference properties (cf. other models based on sets of priors)

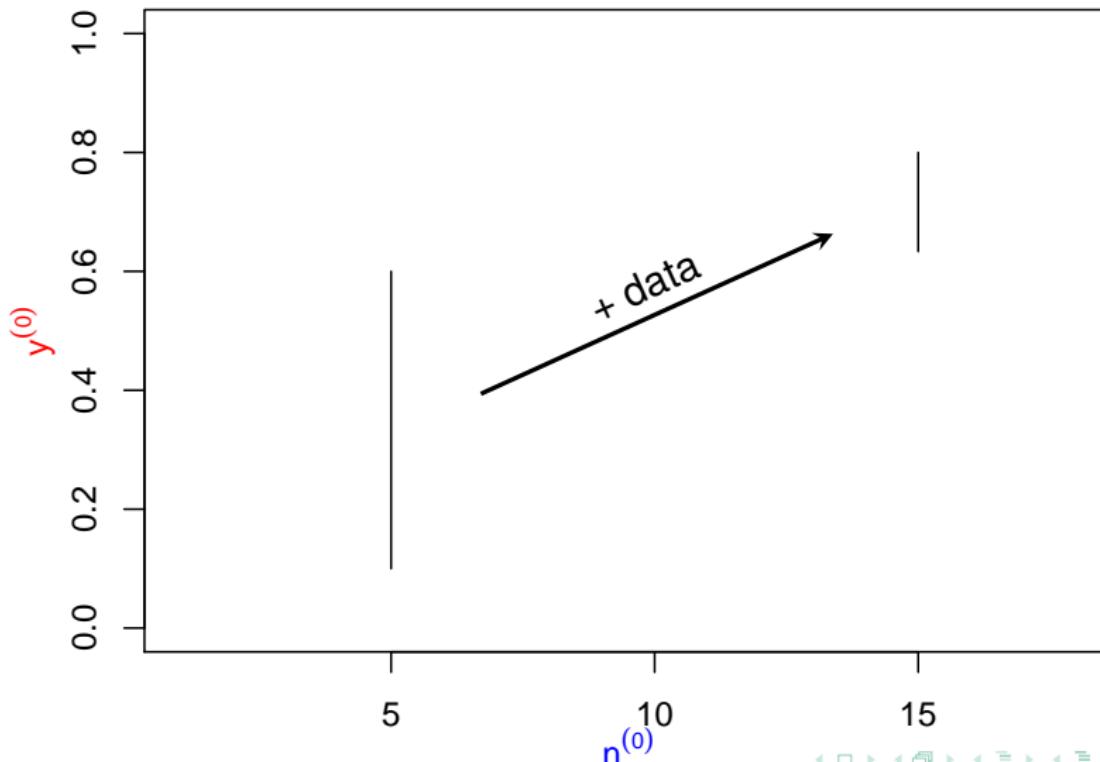
- $n \rightarrow \infty \rightarrow y^{(n)}$ stretch in $\rightarrow 0 \rightarrow$ precise inferences
- larger $n^{(0)} \rightarrow$ larger \rightarrow more vague inferences
- larger range of $y^{(0)}$ in \rightarrow larger range of $y^{(n)}$ in
 \rightarrow more vague inferences

Model very easy to handle:

- Hyperparameter set defines set of priors $\mathcal{M}^{(0)}$
- Hyperparameter set defines set of posteriors $\mathcal{M}^{(n)}$
- \rightarrow is easy: $n^{(n)} = n^{(0)} + n$, $y^{(n)} = \frac{n^{(0)}}{n^{(0)}+n}y^{(0)} + \frac{n}{n^{(0)}+n}\frac{\tau(\mathbf{x})}{n}$
- For quantities linear in posteriors, bounds are attained at “pure” posteriors $p(\psi | n^{(n)}, y^{(n)}) \rightarrow$ optimise over (not over $\mathcal{M}^{(n)}$)
- Often, optimising over $(n^{(n)}, y^{(n)}) \in$ is also easy:
closed form solution for $y^{(n)}$ = posterior ‘guess’ for $\frac{\tau(\mathbf{x})}{n}$ (think: $\bar{\mathbf{x}}$)
given has ‘nice’ shape

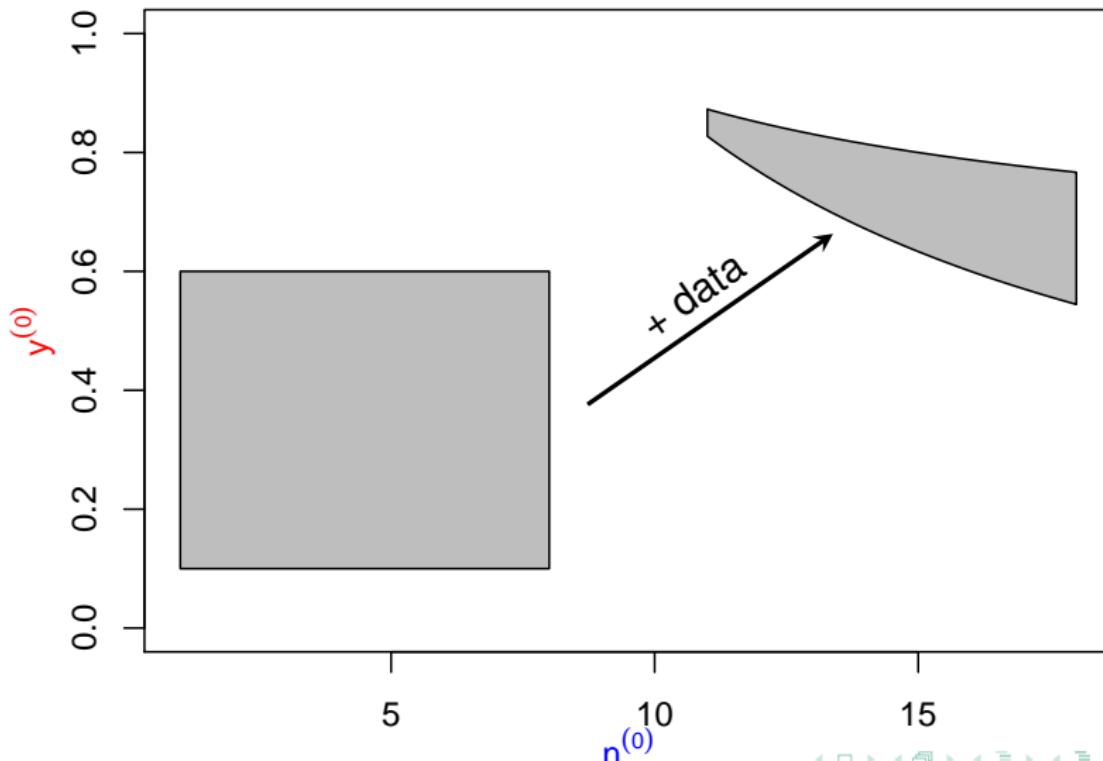


Parameter Set Shapes



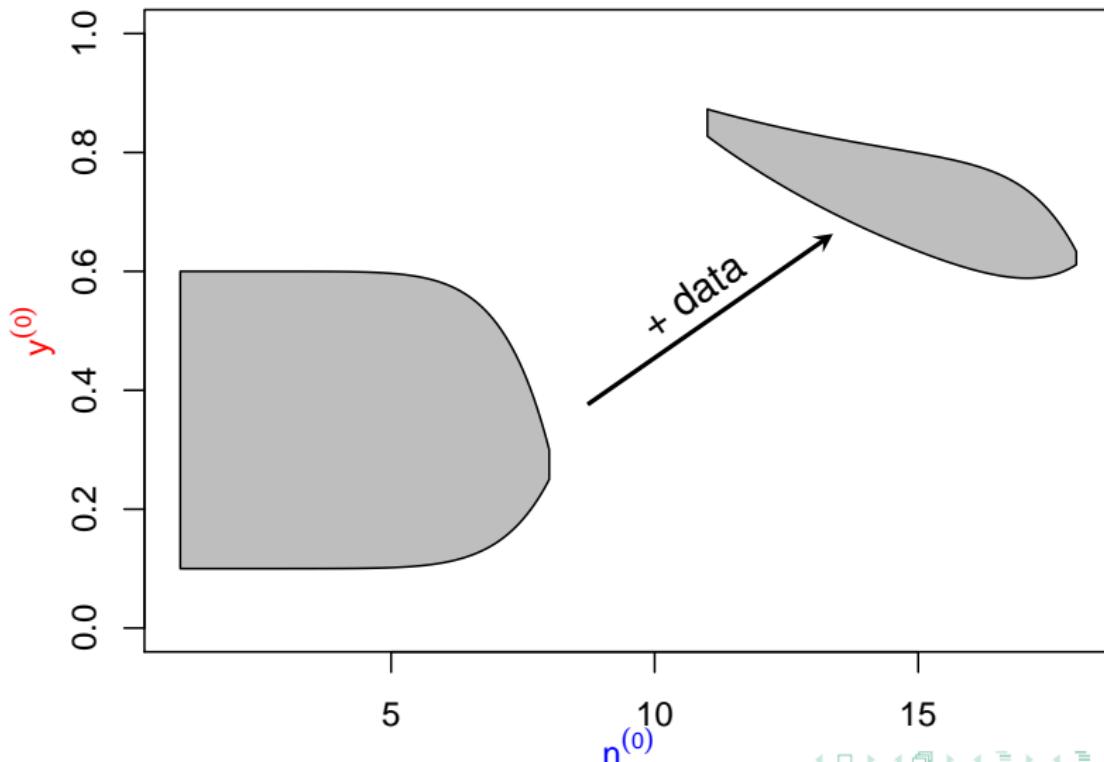


Parameter Set Shapes





Parameter Set Shapes





Parameter Set Shapes

- ▶ Set shape is crucial modeling choice:
trade-off between model complexity and model behaviour



Parameter Set Shapes

- ▶ Set shape is crucial modeling choice:
trade-off between model complexity and model behaviour
- ▶ $= n^{(0)} \times [\underline{y}^{(0)}, \bar{y}^{(0)}]$ (Walley 1996; Quaghebeur & de Cooman 2005):
 $= n^{(n)} \times [\underline{y}^{(n)}, \bar{y}^{(n)}]$ → optimise over $[\underline{y}^{(n)}, \bar{y}^{(n)}]$ only,
but no prior-data conflict sensitivity



Parameter Set Shapes

- ▶ Set shape is crucial modeling choice:
trade-off between model complexity and model behaviour
- ▶ $= n^{(0)} \times [\underline{y}^{(0)}, \bar{y}^{(0)}]$ (Walley 1996; Quaghebeur & de Cooman 2005):
 $= n^{(n)} \times [\underline{y}^{(n)}, \bar{y}^{(n)}]$ → optimise over $[\underline{y}^{(n)}, \bar{y}^{(n)}]$ only,
but no prior-data conflict sensitivity
- ▶ $= [\underline{n}^{(0)}, \bar{n}^{(0)}] \times [\underline{y}^{(0)}, \bar{y}^{(0)}]$ (Walley 1991; Walter & Augustin 2009):
have non-trivial forms (banana / spotlight), but prior-data
conflict sensitivity and closed form for min / max $y^{(n)}$ over .
For other inferences, **R** package `luck` implements optimisation
over via box-constraint optimisation over

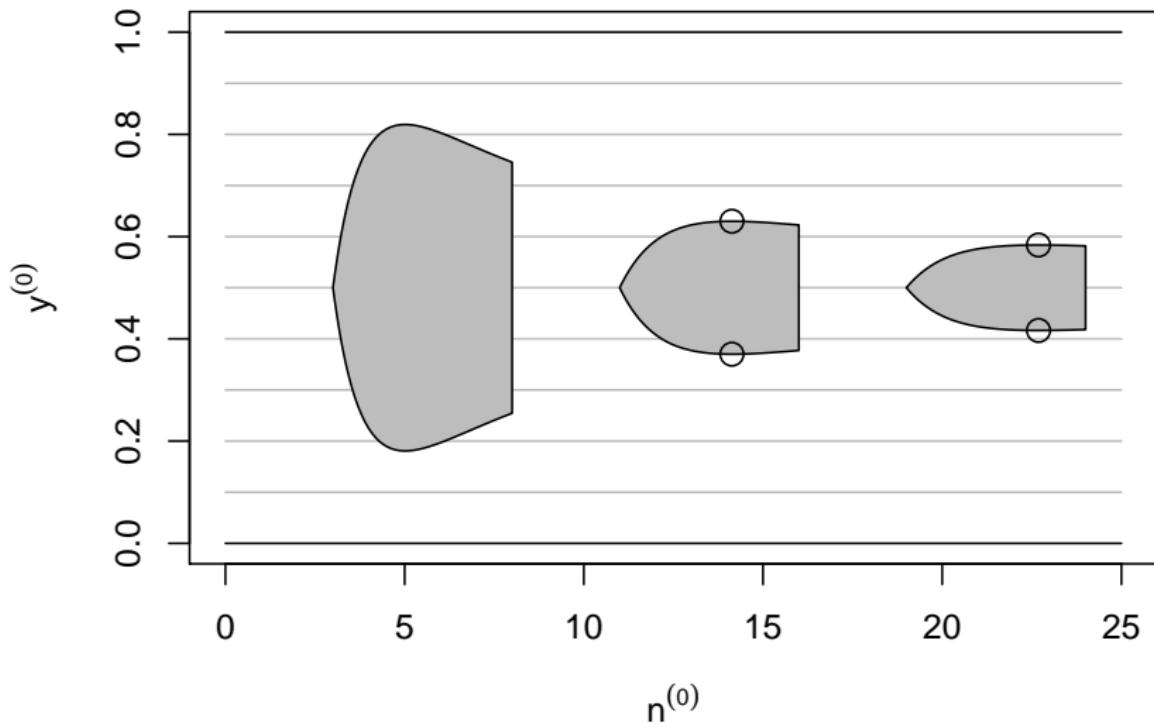


Parameter Set Shapes

- ▶ Set shape is crucial modeling choice:
trade-off between model complexity and model behaviour
- ▶ $= n^{(0)} \times [\underline{y}^{(0)}, \bar{y}^{(0)}]$ (Walley 1996; Quaghebeur & de Cooman 2005):
 $= n^{(n)} \times [\underline{y}^{(n)}, \bar{y}^{(n)}]$ → optimise over $[\underline{y}^{(n)}, \bar{y}^{(n)}]$ only,
but no prior-data conflict sensitivity
- ▶ $= [\underline{n}^{(0)}, \bar{n}^{(0)}] \times [\underline{y}^{(0)}, \bar{y}^{(0)}]$ (Walley 1991; Walter & Augustin 2009):
have non-trivial forms (banana / spotlight), but prior-data
conflict sensitivity and closed form for min / max $y^{(n)}$ over .
For other inferences, **R** package `luck` implements optimisation
over via box-constraint optimisation over
- ▶ Other set shapes are possible, but may be more difficult to handle

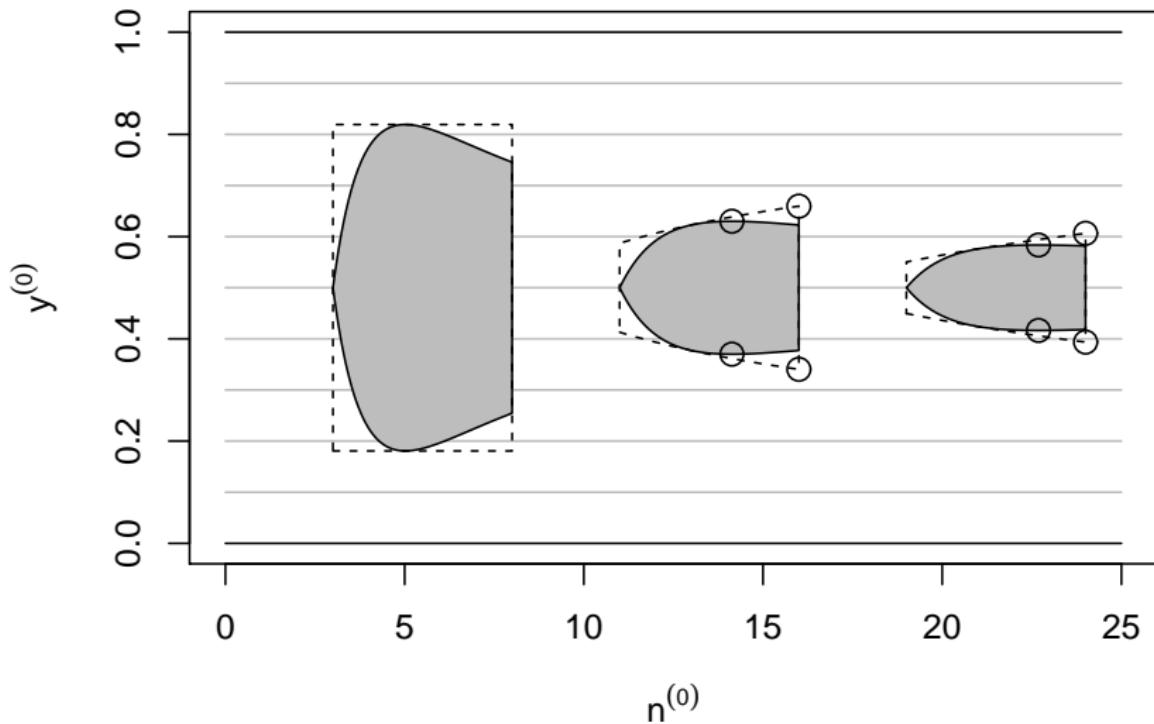


Parameter Set Shape for Strong Prior-Data Agreement





Parameter Set Shape for Strong Prior-Data Agreement





Common-Cause Failures



Source: Wikimedia Commons, http://commons.wikimedia.org/wiki/File:Fukushima_I_by_Digital_Globe.jpg ↗ ↘ ↙ ↛



Common-Cause Failures

common-cause failure

*simultaneous failure of several redundant components
due to a common or shared root cause (Høyland & Rausand, 1994)*



Source: Wikimedia Commons, http://commons.wikimedia.org/wiki/File:Fukushima_I_by_Digital_Globe.jpg



Common-Cause Failures

common-cause failure

simultaneous failure of several redundant components due to a common or shared root cause (Høyland & Rausand, 1994)



Source: <http://www.diakont.com/solutions/nuclear-energy-plant-systems/diesel-generator-control-systems/>

Source: Wikimedia Commons, http://commons.wikimedia.org/wiki/File:Fukushima_I_by_Digital_Globe.jpg



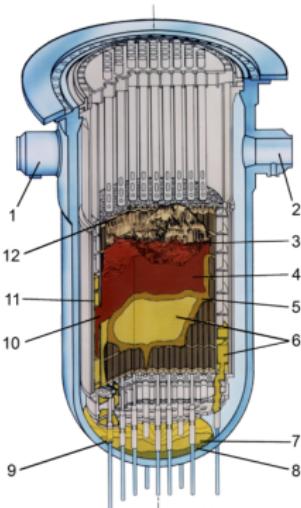
Common-Cause Failure Modelling



Above: CDC, <http://phil.cdc.gov/phil/> ID 1194

Right: Wikimedia Commons,

http://commons.wikimedia.org/wiki/File:Graphic_TMI-2_Core_End-State_Configuration.png





Alpha-Factor Model: Definition

Alpha-Factor Model

Multinomial distribution $M(\mathbf{n} | \boldsymbol{\alpha})$ for common-cause failures in a k -component system

$$p(\mathbf{n} | \boldsymbol{\alpha}) = \prod_{j=1}^k \alpha_j^{n_j}$$

where

- ▶ **alpha-factor** α_j := probability of j of the k components failing due to a common cause given that failure occurs
- ▶ **failure count** n_j := corresponding number of failures observed
- ▶ \mathbf{n} denotes (n_1, \dots, n_k) and $\boldsymbol{\alpha}$ denotes $(\alpha_1, \dots, \alpha_k)$

(the model actually serves to estimate failure *rates*, but the above is all what matters in this talk)



Alpha-Factor Model: Parameter Estimation

maximum likelihood estimator:

$$\hat{\alpha}_j = \frac{n_j}{n}, \quad \text{where } \sum_{j=1}^n n_j = n$$



Alpha-Factor Model: Parameter Estimation

maximum likelihood estimator:

$$\hat{\alpha}_j = \frac{n_j}{n}, \quad \text{where } \sum_{j=1}^n n_j = n$$

The Problem

- ▶ typically, for $j \geq 2$, the n_j are very low with zero being quite common for larger j
- ▶ zero counts = flat likelihoods → $\hat{\alpha}_j = ?$



Alpha-Factor Model: Parameter Estimation

maximum likelihood estimator:

$$\hat{\alpha}_j = \frac{n_j}{n}, \quad \text{where } \sum_{j=1}^n n_j = n$$

The Problem

- ▶ typically, for $j \geq 2$, the n_j are very low with zero being quite common for larger j
- ▶ zero counts = flat likelihoods → $\hat{\alpha}_j = ?$

→ need to rely on epistemic information: Bayesian inference



Bayesian Inference: Dirichlet Prior

Dirichlet-Multinomial Model

$$p(\alpha | n^{(0)}, \mathbf{y}^{(0)}) \propto \prod_{j=1}^k \alpha_j^{n^{(0)} y_j^{(0)} - 1}$$

where $(n^{(0)}, \mathbf{y}^{(0)})$
are *hyperparameters*

$$n^{(0)} > 0$$

$$\mathbf{y}^{(0)} \in \Delta = \left\{ (\mathbf{y}_1^{(0)}, \dots, \mathbf{y}_k^{(0)}): y_1^{(0)} \geq 0, \dots, y_k^{(0)} \geq 0, \sum_{j=1}^k y_j^{(0)} = 1 \right\}$$

$$p(\alpha | n^{(0)}, \mathbf{y}^{(0)}, \mathbf{n}) = p(\alpha | n^{(n)}, \mathbf{y}^{(n)}) = \prod_{j=1}^k \alpha_j^{n^{(n)} y_j^{(n)} - 1}$$



Bayesian Inference: Dirichlet Prior

Dirichlet-Multinomial Model

$$p(\alpha | n^{(0)}, \mathbf{y}^{(0)}) \propto \prod_{j=1}^k \alpha_j^{n^{(0)} y_j^{(0)} - 1}$$

where $(n^{(0)}, \mathbf{y}^{(0)})$
are *hyperparameters*

$$n^{(0)} > 0$$

$$\mathbf{y}^{(0)} \in \Delta = \left\{ (\mathbf{y}_1^{(0)}, \dots, \mathbf{y}_k^{(0)}): y_1^{(0)} \geq 0, \dots, y_k^{(0)} \geq 0, \sum_{j=1}^k y_j^{(0)} = 1 \right\}$$

$$p(\alpha | n^{(0)}, \mathbf{y}^{(0)}, \mathbf{n}) = p(\alpha | n^{(n)}, \mathbf{y}^{(n)}) = \prod_{j=1}^k \alpha_j^{n^{(n)} y_j^{(n)} - 1}$$

Interpretation

- ▶ $\mathbf{y}^{(0)}$ = prior expectation of α , i.e., a prior guess for $\frac{n_j}{n}$, $j = 1, \dots, n$
- ▶ $n^{(0)}$ = determines spread and learning speed



Bayesian Inference: Example

Example (Kelly & Atwood, 2011)

Consider a system with four redundant components ($k = 4$).

The analyst specifies the following prior expectation $\mu_{\text{spec},j}$ for each α_j :

$$\mu_{\text{spec},1} = 0.950 \quad \mu_{\text{spec},2} = 0.030 \quad \mu_{\text{spec},3} = 0.015 \quad \mu_{\text{spec},4} = 0.005$$

We have 36 observations, in which 35 showed one component failing, and 1 showed two components failing:

$$n_1 = 35$$

$$n_2 = 1$$

$$n_3 = 0$$

$$n_4 = 0$$



Bayesian Inference: Example

Example (Kelly & Atwood, 2011)

Consider a system with four redundant components ($k = 4$).

The analyst specifies the following prior expectation $\mu_{\text{spec},j}$ for each α_j :

$$\mu_{\text{spec},1} = 0.950 \quad \mu_{\text{spec},2} = 0.030 \quad \mu_{\text{spec},3} = 0.015 \quad \mu_{\text{spec},4} = 0.005$$

We have 36 observations, in which 35 showed one component failing, and 1 showed two components failing:

$$n_1 = 35$$

$$n_2 = 1$$

$$n_3 = 0$$

$$n_4 = 0$$

we focus on $E[\alpha_j | n^{(n)}, \mathbf{y}^{(n)}] = \mathbf{y}_j^{(n)} = \frac{n^{(0)}}{n^{(0)} + n} \cdot \mathbf{y}_j^{(0)} + \frac{n}{n^{(0)} + n} \cdot \frac{n_j}{n}$

(in a decision context, this expectation would typically end up in expressions for expected utility)



Example: Non-Informative Priors

large variation in posterior under different non-informative priors

- ▶ with constrained maximum entropy prior
(Atwood, 1996; Kelly & Atwood, 2011):

$$\begin{array}{ll} E[\alpha_1 | n^{(n)}, \mathbf{y}^{(n)}] = 0.967 & E[\alpha_2 | n^{(n)}, \mathbf{y}^{(n)}] = 0.028 \\ E[\alpha_3 | n^{(n)}, \mathbf{y}^{(n)}] = 0.003 & E[\alpha_4 | n^{(n)}, \mathbf{y}^{(n)}] = 0.001 \end{array}$$



Example: Non-Informative Priors

large variation in posterior under different non-informative priors

- ▶ with constrained maximum entropy prior
(Atwood, 1996; Kelly & Atwood, 2011):

$$\begin{array}{ll} \mathbb{E}[\alpha_1 | n^{(n)}, \mathbf{y}^{(n)}] = 0.967 & \mathbb{E}[\alpha_2 | n^{(n)}, \mathbf{y}^{(n)}] = 0.028 \\ \mathbb{E}[\alpha_3 | n^{(n)}, \mathbf{y}^{(n)}] = 0.003 & \mathbb{E}[\alpha_4 | n^{(n)}, \mathbf{y}^{(n)}] = 0.001 \end{array}$$

- ▶ with uniform prior ($\mathbf{y}_j^{(0)} = 0.25$ and $n^{(0)} = 4$):

$$\begin{array}{ll} \mathbb{E}[\alpha_1 | n^{(n)}, \mathbf{y}^{(n)}] = 0.9 & \mathbb{E}[\alpha_2 | n^{(n)}, \mathbf{y}^{(n)}] = 0.05 \\ \mathbb{E}[\alpha_3 | n^{(n)}, \mathbf{y}^{(n)}] = 0.025 & \mathbb{E}[\alpha_4 | n^{(n)}, \mathbf{y}^{(n)}] = 0.025 \end{array}$$



Example: Non-Informative Priors

large variation in posterior under different non-informative priors

- with constrained maximum entropy prior
(Atwood, 1996; Kelly & Atwood, 2011):

$$\begin{array}{ll} E[\alpha_1 | n^{(n)}, \mathbf{y}^{(n)}] = 0.967 & E[\alpha_2 | n^{(n)}, \mathbf{y}^{(n)}] = 0.028 \\ E[\alpha_3 | n^{(n)}, \mathbf{y}^{(n)}] = 0.003 & E[\alpha_4 | n^{(n)}, \mathbf{y}^{(n)}] = 0.001 \end{array}$$

- with uniform prior ($y_j^{(0)} = 0.25$ and $n^{(0)} = 4$):

$$\begin{array}{ll} E[\alpha_1 | n^{(n)}, \mathbf{y}^{(n)}] = 0.9 & E[\alpha_2 | n^{(n)}, \mathbf{y}^{(n)}] = 0.05 \\ E[\alpha_3 | n^{(n)}, \mathbf{y}^{(n)}] = 0.025 & E[\alpha_4 | n^{(n)}, \mathbf{y}^{(n)}] = 0.025 \end{array}$$

- with Jeffrey's prior ($y_j^{(0)} = 0.25$ and $n^{(0)} = 2$):

$$\begin{array}{ll} E[\alpha_1 | n^{(n)}, \mathbf{y}^{(n)}] = 0.9342 & E[\alpha_2 | n^{(n)}, \mathbf{y}^{(n)}] = 0.0395 \\ E[\alpha_3 | n^{(n)}, \mathbf{y}^{(n)}] = 0.0132 & E[\alpha_4 | n^{(n)}, \mathbf{y}^{(n)}] = 0.0132 \end{array}$$



Imprecise Dirichlet Model: Definition

Troffaes, Walter & Kelly (2013): model vague prior info more cautiously

Imprecise Dirichlet Model (IDM) for Common-Cause Failure

use a set of hyperparameters (Walley 1991, 1996):

$$= \left\{ (\underline{n}^{(0)}, \bar{n}^{(0)}, \underline{\mathbf{y}}^{(0)}, \bar{\mathbf{y}}^{(0)}): \underline{n}^{(0)} \in [\underline{n}^{(0)}, \bar{n}^{(0)}], \underline{\mathbf{y}}^{(0)} \in \Delta, \underline{\mathbf{y}}_j^{(0)} \in [\underline{y}_j^{(0)}, \bar{y}_j^{(0)}] \right\}$$



Imprecise Dirichlet Model: Definition

Troffaes, Walter & Kelly (2013): model vague prior info more cautiously

Imprecise Dirichlet Model (IDM) for Common-Cause Failure

use a set of hyperparameters (Walley 1991, 1996):

$$= \left\{ (\underline{n}^{(0)}, \underline{\mathbf{y}}^{(0)}): \underline{n}^{(0)} \in [\underline{n}^{(0)}, \bar{n}^{(0)}], \underline{\mathbf{y}}^{(0)} \in \Delta, \underline{y}_j^{(0)} \in [\underline{y}_j^{(0)}, \bar{y}_j^{(0)}] \right\}$$

Interpretation

- ▶ we are doing a sensitivity analysis (á la robust Bayes) over $(\underline{n}^{(0)}, \underline{\mathbf{y}}^{(0)}) \in$
- ▶ we take a set of priors based on as model for prior information



Imprecise Dirichlet Model: Definition

Troffaes, Walter & Kelly (2013): model vague prior info more cautiously

Imprecise Dirichlet Model (IDM) for Common-Cause Failure

use a set of hyperparameters (Walley 1991, 1996):

$$= \left\{ (\underline{n}^{(0)}, \underline{\mathbf{y}}^{(0)}): \underline{n}^{(0)} \in [\underline{n}^{(0)}, \bar{n}^{(0)}], \underline{\mathbf{y}}^{(0)} \in \Delta, \underline{y}_j^{(0)} \in [\underline{y}_j^{(0)}, \bar{y}_j^{(0)}] \right\}$$

Interpretation

- ▶ we are doing a sensitivity analysis (á la robust Bayes) over $(\underline{n}^{(0)}, \underline{\mathbf{y}}^{(0)}) \in$
- ▶ we take a set of priors based on as model for prior information

Analyst has to specify ('elicit')

bounds $[\underline{n}^{(0)}, \bar{n}^{(0)}]$ and bounds $[\underline{y}_j^{(0)}, \bar{y}_j^{(0)}]$ for each $j \in \{1, \dots, k\}$



Imprecise Dirichlet Model: Elicitation

- $[\underline{y}_j^{(0)}, \bar{y}_j^{(0)}]$: Cautious interpretation of prior specifications $\mu_{\text{spec},j}$:

$$[\underline{y}_1^{(0)}, \bar{y}_1^{(0)}] = [0.950, 1]$$

$$[\underline{y}_2^{(0)}, \bar{y}_2^{(0)}] = [0, 0.030]$$

$$[\underline{y}_3^{(0)}, \bar{y}_3^{(0)}] = [0, 0.015]$$

$$[\underline{y}_4^{(0)}, \bar{y}_4^{(0)}] = [0, 0.005]$$



Imprecise Dirichlet Model: Elicitation

- ▶ $[\underline{y}_j^{(0)}, \bar{y}_j^{(0)}]$: Cautious interpretation of prior specifications $\mu_{\text{spec},j}$:

$$[\underline{y}_1^{(0)}, \bar{y}_1^{(0)}] = [0.950, 1]$$

$$[\underline{y}_2^{(0)}, \bar{y}_2^{(0)}] = [0, 0.030]$$

$$[\underline{y}_3^{(0)}, \bar{y}_3^{(0)}] = [0, 0.015]$$

$$[\underline{y}_4^{(0)}, \bar{y}_4^{(0)}] = [0, 0.005]$$

- ▶ $[\underline{n}^{(0)}, \bar{n}^{(0)}]$: Good (1965):

reason about posterior expectations for hypothetical data



Imprecise Dirichlet Model: Elicitation

- ▶ $[\underline{y}_j^{(0)}, \bar{y}_j^{(0)}]$: Cautious interpretation of prior specifications $\mu_{\text{spec},j}$:

$$[\underline{y}_1^{(0)}, \bar{y}_1^{(0)}] = [0.950, 1]$$

$$[\underline{y}_2^{(0)}, \bar{y}_2^{(0)}] = [0, 0.030]$$

$$[\underline{y}_3^{(0)}, \bar{y}_3^{(0)}] = [0, 0.015]$$

$$[\underline{y}_4^{(0)}, \bar{y}_4^{(0)}] = [0, 0.005]$$

- ▶ $[\underline{n}^{(0)}, \bar{n}^{(0)}]$: Good (1965):

reason about posterior expectations for hypothetical data

$\bar{n}^{(0)}$ = number of one-component failures required
to reduce the upper probabilities of multi-component failure by half

$\underline{n}^{(0)}$ = number of multi-component failures required
to reduce the lower probability of one-component failure by half



Imprecise Dirichlet Model: Elicitation

$\bar{n}^{(0)}$ = number of one-component failures required
to reduce the upper probabilities of multi-component failure by half

$n^{(0)}$ = number of multi-component failures required
to reduce the lower probability of one-component failure by half



Imprecise Dirichlet Model: Elicitation

$\bar{n}^{(0)}$ = number of one-component failures required
to reduce the upper probabilities of multi-component failure by half

$\underline{n}^{(0)}$ = number of multi-component failures required
to reduce the lower probability of one-component failure by half

Reasonable values in example:

- ▶ $\bar{n}^{(0)} = 10$: after observing 10 one-component failures
 - halve upper probabilities of multi-component failures
- ▶ $\underline{n}^{(0)} = 1$: immediate multi-component failure
 - keen to reduce lower probability for one-component failure



Imprecise Dirichlet Model: Elicitation

$\bar{n}^{(0)}$ = number of one-component failures required
to reduce the upper probabilities of multi-component failure by half

$\underline{n}^{(0)}$ = number of multi-component failures required
to reduce the lower probability of one-component failure by half

Reasonable values in example:

- ▶ $\bar{n}^{(0)} = 10$: after observing 10 one-component failures
 - halve upper probabilities of multi-component failures
- ▶ $\underline{n}^{(0)} = 1$: immediate multi-component failure
 - keen to reduce lower probability for one-component failure

Difference between $\underline{n}^{(0)}$ and $\bar{n}^{(0)}$ reflects a level of caution:
The rate at which we reduce upper probabilities
is less than the rate at which we reduce lower probabilities



Imprecise Dirichlet Model: Inference

With $[\underline{n}^{(0)}, \bar{n}^{(0)}] = [1, 10]$, we get...

prior bounds + data \rightarrow posterior bounds

j	$\underline{y}_j^{(0)}$	$\bar{y}_j^{(0)}$	n_j	$E[\alpha_j \cdot]$	$\bar{E}[\alpha_j \cdot]$
1	0.950	1	35	0.967	0.978
2	0	0.030	1	0.0270	0.0283
3	0	0.015	0	0	0.00326
4	0	0.005	0	0	0.00109



Imprecise Dirichlet Model: Inference

With $[\underline{n}^{(0)}, \bar{n}^{(0)}] = [1, 10]$, we get...

prior bounds + data → posterior bounds

j	$\underline{y}_j^{(0)}$	$\bar{y}_j^{(0)}$	n_j	$E[\alpha_j \cdot]$	$\bar{E}[\alpha_j \cdot]$
1	0.950	1	35	0.967	0.978
2	0	0.030	1	0.0270	0.0283
3	0	0.015	0	0	0.00326
4	0	0.005	0	0	0.00109

- ▶ **Bounds**, rather than precise values, are desirable due to inferences being strongly sensitive to the prior particularly when faced with zero counts



Imprecise Dirichlet Model: Inference

With $[\underline{n}^{(0)}, \bar{n}^{(0)}] = [1, 10]$, we get...

prior bounds + data → posterior bounds

j	$\underline{y}_j^{(0)}$	$\bar{y}_j^{(0)}$	n_j	$E[\alpha_j \cdot]$	$\bar{E}[\alpha_j \cdot]$
1	0.950	1	35	0.967	0.978
2	0	0.030	1	0.0270	0.0283
3	0	0.015	0	0	0.00326
4	0	0.005	0	0	0.00109

- ▶ **Bounds**, rather than precise values, are desirable due to inferences being strongly sensitive to the prior particularly when faced with zero counts
- ▶ Simple ways to elicit the parameters of the model by **reasoning on hypothetical data**



Conclusion

- ▶ Conjugate priors are a convenient tool for Bayesian inference but there are some pitfalls
 - ▶ Hyperparameters $n^{(0)}$, $y^{(0)}$ are easy to interpret and elicit
 - ▶ Averaging property makes calculations simple, but leads to inadequate model behaviour in case of prior-data conflict



Conclusion

- ▶ Conjugate priors are a convenient tool for Bayesian inference but there are some pitfalls
 - ▶ Hyperparameters $n^{(0)}, y^{(0)}$ are easy to interpret and elicit
 - ▶ Averaging property makes calculations simple, but leads to inadequate model behaviour in case of prior-data conflict
- ▶ Sets of conjugate priors maintain advantages & mitigate issues
 - ▶ Hyperparameter set shape is important
 - ▶ Reasonable choice: $\text{rectangular } = [\underline{n}^{(0)}, \bar{n}^{(0)}] \times [\underline{y}^{(0)}, \bar{y}^{(0)}]$
(Walter & Augustin 2009: *generalised iLUCK-models*, luck)
 - ▶ Bounds for hyperparameters are easy to interpret and elicit
 - ▶ Additional imprecision in case of prior-data conflict leads to **cautious inferences if, and only if, caution is needed**
 - ▶ Shape for more precision in case of strong prior-data agreement is in development (joint work with Frank Coolen and Mik Bickis)



Works the Thesis is Based on

- Augustin, Thomas, Gero Walter, and Frank Coolen (2013). "Statistical Inference". In: *Introduction to Imprecise Probabilities*. Ed. by Frank Coolen et al. In preparation. Wiley.
- Troffaes, Matthias, Gero Walter, and Dana Kelly (2013). "A Robust Bayesian Approach to Modelling Epistemic Uncertainty in Common-Cause Failure Models". In: *Reliability Engineering & System Safety*. doi: [10.1016/j.ress.2013.05.022](https://doi.org/10.1016/j.ress.2013.05.022).
- Walter, Gero (2012). A Technical Note on the Dirichlet-Multinomial Model — The Dirichlet Distribution as the Canonically Constructed Conjugate Prior. Tech. rep. 131. Department of Statistics, LMU Munich. URL: <http://epub.ub.uni-muenchen.de/14068/>.
- Walter, Gero and Thomas Augustin (2009). "Imprecision and Prior-data Conflict in Generalized Bayesian Inference". In: *Journal of Statistical Theory and Practice* 3, pp. 255–271.
- Walter, Gero and Thomas Augustin (2010). "Bayesian linear regression — different conjugate models and their (in)sensitivity to prior-data conflict". In: *Statistical Modelling and Regression Structures*. Ed. by Thomas Kneib and Gerhard Tutz. Festschrift in Honour of Ludwig Fahrmeir. Springer, pp. 59–78. Extended version published as Tech. Rep. 69, Department of Statistics, LMU Munich. URL: <http://epub.ub.uni-muenchen.de/11050/1/tr069.pdf>.
- Walter, Gero, Thomas Augustin, and Frank P.A. Coolen (2011). "On Prior-Data Conflict in Predictive Bernoulli Inferences". In: *ISIPTA'11: Proceedings of the Seventh International Symposium on Imprecise Probabilities: Theories and Applications*. Ed. by Frank Coolen et al. SIPTA, pp. 391–400. URL: <http://www.sipta.org/isipta11/proceedings/046.html>.
- Walter, Gero and Norbert Krautbacher (2013). *luck: R package for Generalized iLUCK-models*. URL: <http://luck.r-forge.r-project.org/>.



Further References

- Atwood, C. (1996). "Constrained noninformative priors in risk assessment". In: *Reliability Engineering and System Safety* 53, pp. 37–46.

Benavoli, A. and M. Zaffalon (2012). "A model of prior ignorance for inferences in the one-parameter exponential family". In: *Journal of Statistical Planning and Inference* 142, pp. 1960–1979. doi: 10.1016/j.jspi.2012.01.023.

Berger, J. et al. (1994). "An overview of robust Bayesian analysis". In: *TEST* 3, pp. 5–124.

Evans, M. and H. Moshonov (2006). "Checking for Prior-Data Conflict". In: *Bayesian Analysis* 1, pp. 893–914.

Good, I. (1965). *The estimation of probabilities*. Cambridge (MA): MIT Press.

Høyland, Arnljot and Marvin Rausand (1994). *System reliability theory: models and statistical methods*. A Wiley interscience publication. New York, NY: Wiley. ISBN: 0-471-59397-4.

Kelly, Dana and Corwin Atwood (2011). "Finding a minimally informative Dirichlet prior distribution using least squares". In: *Reliability Engineering and System Safety* 96.3, pp. 398–402. ISSN: 0951-8320. doi: 10.1016/j.ress.2010.11.008.

Quaeghebeur, E. and G. de Cooman (2005). "Imprecise probability models for inference in exponential families". In: *ISIPTA '05. Proceedings of the Fourth International Symposium on Imprecise Probabilities and Their Applications*. Ed. by F. Cozman, R. Nau, and T. Seidenfeld. Manno: SIPTA, pp. 287–296.

Walley, Peter (1991). *Statistical Reasoning with Imprecise Probabilities*. London: Chapman and Hall. ISBN: 0-412-28660-2.

Walley, Peter (1996). "Inferences from multinomial data: Learning about a bag of marbles". In: *Journal of the Royal Statistical Society, Series B* 58.1, pp. 3–34.



Arguments for Imprecise Bayesian Inference

- ▶ The formal arguments for the Bayesian approach as a coherent way of inferential reasoning hold only if one can make “*arbitrarily fine discriminations in judgement about unknowns and utilities*” (Berger et al. 1994, p. 303). (Your subjective prior must exactly express your preferences, anticipating *anything* that may happen.)
- ▶ Walley’s (1991) framework for coherent inference using imprecise probabilities instead allows for incomplete and imprecise prior specifications.
 - ▶ more realistic description of uncertainties that are often hidden by spuriously precise models (through ‘arbitrary’ modeling decisions)
 - ▶ ‘overprecision’ is often compensated by ‘taking models not too seriously’ (“all models are wrong, but some are useful”)
 - ▶ why not use sensible, reliable models in the first place?



Hierarchical Modelling



- ▶ Usually: space of parametric priors indexed by parameter ϕ .
- ▶ Very useful if ϕ has a clear interpretation (e.g., as a global mean), such that a prior on ϕ can be meaningfully elicited.
- ▶ Noninformative priors on ϕ can be problematic (incoherence, improper posteriors).

Example (e.g., Walley 1991, p. 232)

$x_i \stackrel{iid}{\sim} N(\mu_i, 1)$, $\mu_i \stackrel{iid}{\sim} N(\mu, \sigma^2)$, $i = 1, \dots, n$.

For $p(\mu, \sigma^2) \propto \sigma^{-1}$, the posterior is improper.

- ▶ High-variance priors on ϕ do not express ignorance about ϕ , but a strong prior belief that $|\phi|$ is large (Walley 1991, p. 233).



Updating and Mixture Commute



Let

$$p_m(\vartheta | n_1^{(0)}, y_1^{(0)}, n_2^{(0)}, y_2^{(0)}, \kappa) := \kappa p(\vartheta | n_1^{(0)}, y_1^{(0)}) + (1 - \kappa) p(\vartheta | n_2^{(0)}, y_2^{(0)}),$$

with marginals

$$f_1(\mathbf{x}) = \int_{\Theta} f(\mathbf{x} | \vartheta) p(\vartheta | n_1^{(0)}, y_1^{(0)}) d\vartheta,$$

$$f_2(\mathbf{x}) = \int_{\Theta} f(\mathbf{x} | \vartheta) p(\vartheta | n_2^{(0)}, y_2^{(0)}) d\vartheta,$$

$$f_m(\mathbf{x}) = \int_{\Theta} f(\mathbf{x} | \vartheta) p_m(\vartheta | n_1^{(0)}, y_1^{(0)}, n_2^{(0)}, y_2^{(0)}, \kappa) d\vartheta$$

$$= \kappa \int_{\Theta} f(\mathbf{x} | \vartheta) p(\vartheta | n_1^{(0)}, y_1^{(0)}) d\vartheta + (1 - \kappa) \int_{\Theta} f(\mathbf{x} | \vartheta) p(\vartheta | n_2^{(0)}, y_2^{(0)}) d\vartheta$$

$$= \kappa f_1(\mathbf{x}) + (1 - \kappa) f_2(\mathbf{x}).$$



Updating and Mixture Commute

$$p_m(\vartheta \mid n_1^{(0)}, y_1^{(0)}, n_2^{(0)}, y_2^{(0)}, \kappa, \mathbf{x})$$

$$\begin{aligned} &= \frac{f(\mathbf{x} \mid \vartheta)}{f_m(\mathbf{x})} \left(\kappa p(\vartheta \mid n_1^{(0)}, y_1^{(0)}) + (1 - \kappa) p(\vartheta \mid n_2^{(0)}, y_2^{(0)}) \right) \\ &= \kappa \frac{f(\mathbf{x} \mid \vartheta) p(\vartheta \mid n_1^{(0)}, y_1^{(0)})}{f_m(\mathbf{x})} + (1 - \kappa) \frac{f(\mathbf{x} \mid \vartheta) p(\vartheta \mid n_2^{(0)}, y_2^{(0)})}{f_m(\mathbf{x})} \\ &= \kappa \frac{f_1(\mathbf{x}) p(\vartheta \mid n_1^{(n)}, y_1^{(n)})}{f_m(\mathbf{x})} + (1 - \kappa) \frac{f_2(\mathbf{x}) p(\vartheta \mid n_2^{(n)}, y_2^{(n)})}{f_m(\mathbf{x})} \\ &= p_m(\vartheta \mid n_1^{(n)}, y_1^{(n)}, n_2^{(n)}, y_2^{(n)}, \kappa^*), \end{aligned}$$

where $\kappa^* = \kappa \frac{f_1(\mathbf{x})}{f_m(\mathbf{x})} = \frac{\kappa f_1(\mathbf{x})}{\kappa f_1(\mathbf{x}) + (1 - \kappa) f_2(\mathbf{x})}.$



Updating and Mixture Commute



- ▶ updated mixture distribution is a mixture of the updated components with mixture parameter κ^* instead of κ .
- ▶ convex hull of prior components
= set of prior mixture distributions with $\kappa \in [0, 1]$.
- ▶ for any $\kappa \in [0, 1]$, the corresponding $\kappa^* \in [0, 1]$.
- ▶ in fact, $\{\kappa^* \mid \kappa \in [0, 1]\} = [0, 1]$.
- ▶ set of updated mixture distributions with $\kappa \in [0, 1]$
= convex hull of updated components ($\kappa^* \in [0, 1]$).
- ▶ arbitrary number of components by complete induction.



Other Models Based on Sets of Priors

◀ 3

◀ 10

► Neighbourhood models

- set of distributions 'close to' a central distribution P_0
- common in robust Bayesian approaches
- example: ε -contamination class: $\{P : P = (1 - \varepsilon) P_0 + \varepsilon Q, Q \in Q\}$
- not necessarily closed under Bayesian updating

► Density ratio class / interval of measures

- set of distributions by bounds for the density function $p(\vartheta)$:

$$\mathcal{M}_{l,u} = \left\{ p(\vartheta) : \exists c \in \mathbb{R}_{>0} : l(\vartheta) \leq cp(\vartheta) \leq u(\vartheta) \right\}$$

- posterior set is bounded by updated $l(\vartheta)$ and $u(\vartheta)$
- $u(\vartheta)/l(\vartheta)$ is constant under updating
 - ➡ size of the set does not decrease with n
 - ➡ too vague posterior inferences



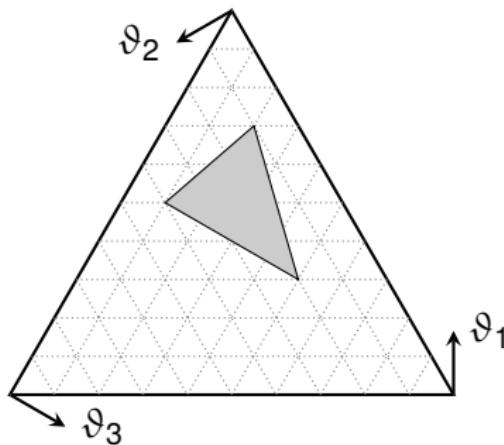
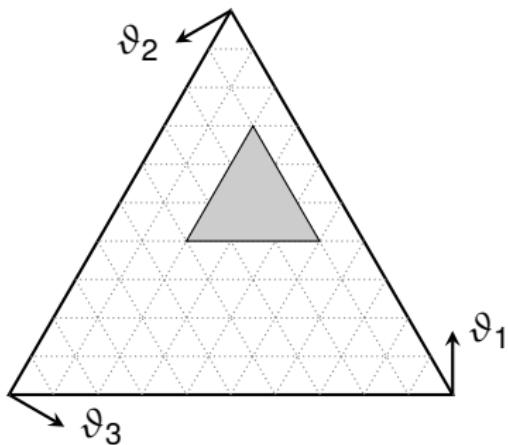
Other Models Based on Sets of Priors

◀ 3

◀ 10

► Discrete models

- discretize the parameter space as $\Theta = \{\theta_j\}_{j \in \{1, \dots, m\}}$
- set of distributions by bounds for $p(\vartheta_j)$ (or for expectations)
- posterior bounds determined by linear programming algorithm
- more flexibility at the cost of computational complexity
- no clear measure for weight of prior information as compared to n





R package luck

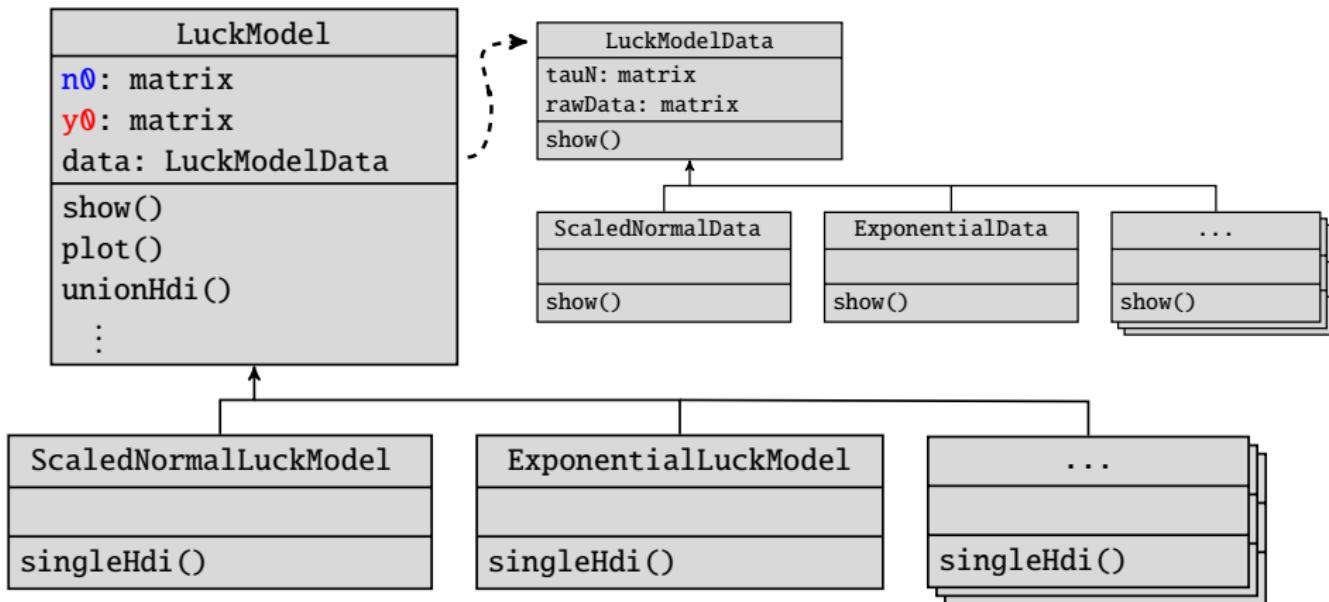


- ▶ S4 implementation of the general canonical prior parameter structure with rectangular sets $= [\underline{n}^{(0)}, \bar{n}^{(0)}] \times [\underline{y}^{(0)}, \bar{y}^{(0)}]$
- ▶ lean subclasses for concrete sample distributions (currently implemented: scaled normal, exponential)
- ▶ currently on R-Forge, to be submitted to CRAN

```
install.packages("luck", repos="http://R-Forge.R-project.org")
```

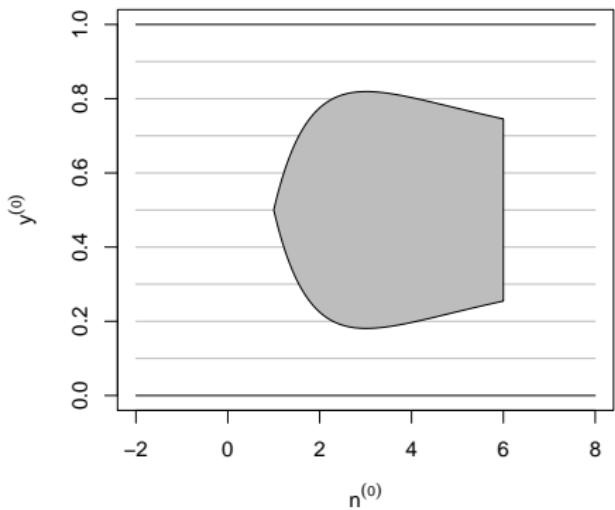
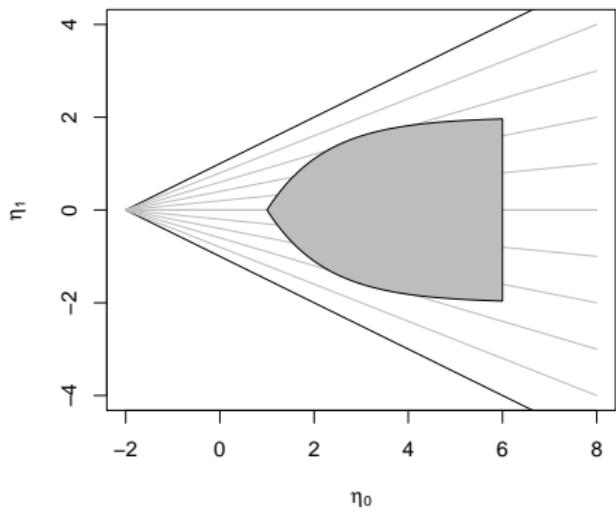


R package luck





Strong Prior-Data Agreement Modelling ⏪





'Noninformative' Priors

- ▶ noninformative priors can be useful as a technical device when they hardly influence the posterior
- ▶ they are often improper: problematic for testing and model selection
- ▶ they give a (precise) probability for any subset A of the parameter space, which seems incompatible with the notion of ignorance
- ▶ they express *indifference* instead of *ignorance*
- ▶ a set of priors with $P(A) = [0, 1]$ is noninformative
- ▶ model framework allows for near-noninformative sets of priors
 - ▶ IDM (Walley 1996): range of $y_j^{(0)} = (0, 1) \forall j$
 - ▶ Benavoli & Zaffalon (2012):
range of $y^{(0)} = (-\infty, +\infty)$, while $\bar{n}^{(0)}$ decreases with $y^{(0)}$
(to avoid $n^{(0)}|y^{(0)}| = \infty$, i.e. vacuous posterior inferences)