Bayes factors for research workers

Ly, A.

Citation for published version (APA):
In this dissertation we advocate the use of Bayes factors in empirical research to replace or complement standard null hypothesis tests based on p-values. These Bayes factors were specifically designed to quantify the evidence for or against the existence of an effect. This was done by comparing two models with the same distributional assumptions, where the alternative model is an extension of the null model by incorporating one extra parameter. Instead of returning a decision to “reject” or “not reject”, a Bayes factor $BF_{10}(d)$ returns a non-negative number that represents the evidence provided by the observed data $d$ for the model that includes the effect. The returned number can be seen as a refinement of the binary decision with $BF_{10}(d) = \infty$ and $BF_{10}(d) = 0$ corresponding to definite rejection and acceptance of the null, respectively. Moreover, the Bayes factor allows its users to forgo the binary decision and acknowledge uncertainty, so that the evidence can be updated continually in light of new data, directly and easily. For empirical scientists to be able to use these Bayes factors, we implemented them in Jeffreys’s Amazing Statistics Program, JASP, which is freely available and open-source.

(url: https://jasp-stats.org)
Bayes Factors for Research Workers

Alexander Ly
Bayes Factors for Research Workers

ACADEMISCH PROEFSCHRIFT
ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties ingestelde
commissie, in het openbaar te verdedigen in de Aula der Universiteit
op 19 januari 2018, te 13.00 uur
door Alexander Ly
geboren te Heerlen
Promotiecommissie

Promotor: Prof. dr. E. M. Wagenmakers
Copromotor: Dr. M. Marsman

Overige leden: Prof. dr. J. O. Berger
Prof. dr. M. D. Lee
Prof. dr. P. D. Grünwald
Prof. dr. H. L. J. van der Maas
Dr. L. J. Waldorp
Dr. R. P. P. P. Grasman

Faculteit: Faculteit der Maatschappij- en Gedragswetenschappen
Voor opa Minh Chung Ly
mijn vader Han Giang Ly
mijn moeder Phuong Hoa Tran
mijn zus Julia-Selina Ly
mijn liefste Gracia Edwards
my family
## Contents

1 Introduction  
  1.1 Bayesian model learning .......................... 1  
  1.2 Chapter outline .................................. 2  

I Bayes Factor Rationale  

2 Harold Jeffreys’s Default Bayes Factor Hypothesis Tests: Explanation, Extension, and Application in Psychology  
  2.1 Introduction ...................................... 10  
  2.2 Historical and philosophical background of the Bayes factor .......... 11  
  2.3 Jeffreys’s procedure for constructing a default Bayes factor ........ 16  
  2.4 Jeffreys’s Bayes factor for the test of the nullity of a normal mean:  
      The Bayesian t-test .................................. 19  
  2.5 Jeffreys’s Bayes factor for the test of the nullity of a correlation ... 25  
  2.6 Conclusion ........................................ 36  
  2.A The default Bayes factor hypothesis tests proposed by Jeffreys in  
      ToP .................................................. 38  
  2.B The hypergeometric function ................................ 38  
  2.C The stretched beta density ................................ 38  
  2.D Translation of Jeffreys’s notation in ToP ....................... 39  

3 An Evaluation of Alternative Methods for Testing Hypotheses,  
   from the Perspective of Harold Jeffreys  
  3.1 Introduction ...................................... 42  
  3.2 Rejoinder to Robert .............................. 42  
  3.3 Rejoinder to Chandramouli and Shiffrin ..................... 51  
  3.4 Conclusion ........................................ 65  

II Bayes Factors for Common Designs  

4 Bayesian Inference for Kendall’s Rank Correlation Coefficient  
  4.1 Introduction ...................................... 69  
  4.2 Methods ........................................... 72  
  4.3 Results ............................................ 75  
  4.4 Concluding comments ............................ 77
5 Informed Bayesian $t$-Tests
  5.1 Introduction ................................................. 81
  5.2 Jeffreys’s default Bayes factor ............................. 83
  5.3 One-sample and paired samples $t$-test ....................... 84
  5.4 Two-sample $t$-test .......................................... 89
  5.5 Example III: Reanalysis of 593 $t$-tests ..................... 95
  5.6 Quantifying evidence for $H_0$ ............................. 97
  5.7 Concluding comments ........................................ 99

6 A Limit-Consistent Bayes Factor for Testing the Equality of
Two Poisson Rates .................................................. 101
  6.1 Introduction .................................................... 101
  6.2 Jeffreys’s Bayes factor for the comparison of two Poisson rates ............................. 104
  6.3 A limit-consistent Bayes factor for the comparison of two Poisson rates .................... 107
  6.4 Discussion ..................................................... 111

III Scientific Learning with Bayes Factors .......................... 113

7 Four Requirements for an Acceptable Research Programme ............. 115
  7.1 The power fallacy .............................................. 116
  7.2 The fallacy of the transposed conditional ..................... 116
  7.3 Requirements of a research programme ....................... 117
  7.4 Concluding comments ........................................ 120

8 Bayesian Reanalyses from Summary Statistics: A Guide for
Academic Consumers ............................................... 123
  8.1 Introduction .................................................... 124
  8.2 The Festinger & Carlsmith (1959) cognitive dissonance study ............................. 125
  8.3 Bayesian reanalysis ............................................ 126
  8.4 Concluding comments ........................................ 128

9 Replication Bayes Factors from Evidence Updating ..................... 129
  9.1 Introduction .................................................... 129
  9.2 The Bayes factor .............................................. 131
  9.3 Bayesian updating in action .................................. 132
  9.4 The replication Bayes factor reconceptualised ................. 134
  9.5 Example 1: A $t$-test to assess whether superstition improves performance ..................... 137
  9.6 Example 2: A contingency table analysis to test whether more valuable stimuli are judged to be relatively rare ..................... 138
  9.7 Concluding comments ........................................ 140
  9.A Deriving the $t$-value across all data sets .................... 141
  9.B Replication Bayes factors as conditional Bayes factors ..................... 143
  9.C Replication paradox and solution ................................ 143
## IV Analytic Results

### 10 Analytic Posteriors for Pearson’s Correlation Coefficient

- 10.1 Introduction ................................................................. 149
- 10.2 Notation and result ...................................................... 150
- 10.3 Analytic posteriors for the case $\beta = 0$ .......................... 155
- 10.A A lemma distilled from the Bateman Project ....................... 157

### 11 Analytic Posteriors for the Binomial Rate Parameters, and the Odds Ratio

- 11.1 Introduction ................................................................. 159
- 11.2 Binomial distribution .................................................... 159
- 11.3 Products of generalised beta prime distributions and the odds ratio ...................................................... 166
- 11.4 Concluding remarks ..................................................... 168

## V Two Tutorials

### 12 A Tutorial on Bridge Sampling

- 12.1 Introduction ................................................................. 171
- 12.2 Four sampling methods to approximate the marginal likelihood ...................................................... 174
- 12.3 Case study: Bridge sampling for reinforcement learning models ...................................................... 193
- 12.4 Discussion ................................................................. 205
- 12.A The bridge sampling estimator as a general case of methods 1 – 3 ...................................................... 207
- 12.B Bridge sampling implementation: Avoiding numerical issues ...................................................... 207
- 12.C Correcting for the probit transformation ...................................................... 208
- 12.D Details on the application of bridge sampling to the individual-level EV model ...................................................... 210
- 12.E Details on the application of bridge sampling to the hierarchical EV model ...................................................... 211

### 13 A Tutorial on Fisher Information

- 13.1 Introduction ................................................................. 213
- 13.2 The role of Fisher information in frequentist statistics ...................................................... 217
- 13.3 The role of Fisher information in Bayesian statistics ...................................................... 222
- 13.4 The role of Fisher information in minimum description length ...................................................... 231
- 13.5 Concluding comments ..................................................... 242
- 13.A Generalisation to vector-valued parameters: The Fisher information matrix ...................................................... 244
- 13.B Frequentist statistics based on asymptotic normality ...................................................... 245
- 13.C Bayesian use of the Fisher-Rao metric: The Jeffreys’s prior ...................................................... 249
- 13.D MDL: Coding theoretical background ...................................................... 256
- 13.E Regularity conditions ..................................................... 260

## VI Conclusion

### 14 Discussion and Future Directions
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>References</td>
<td>271</td>
</tr>
<tr>
<td>Nederlandse Samenvatting</td>
<td>299</td>
</tr>
<tr>
<td>Acknowledgements — Dankwoord</td>
<td>305</td>
</tr>
<tr>
<td>Publications</td>
<td>307</td>
</tr>
</tbody>
</table>