Bayesian Analysis Users Guide
Release 4.00, Manual Version 1

G. Larry Bretthorst
Biomedical MR Laboratory
Washington University School Of Medicine,
Campus Box 8227
Room 2313, East Bldg.,
4525 Scott Ave.
St. Louis MO 63110
http://bayes.wustl.edu
Email: larry@bayes.wustl.edu

July 13, 2018
Contents

Manual Status 16

1 An Overview Of The Bayesian Analysis Software 19
  1.1 The Server Software ............................................. 19
  1.2 The Client Interface ............................................. 22
    1.2.1 The Global Pull Down Menus .............................. 24
    1.2.2 The Package Interface ..................................... 24
    1.2.3 The Viewers ................................................ 27

2 Installing the Software 29

3 the Client Interface 33
  3.1 The Global Pull Down Menus .................................. 35
    3.1.1 the Files menu ............................................ 35
    3.1.2 the Packages menu ....................................... 40
    3.1.3 the WorkDir menu ........................................ 45
    3.1.4 the Settings menu ....................................... 46
    3.1.5 the Utilities menu ....................................... 50
    3.1.6 the Help menu ........................................... 50
  3.2 The Submit Job To Server area ............................... 51
  3.3 The Server area ................................................ 52
  3.4 Interface Viewers .............................................. 52
    3.4.1 the Ascii Data Viewer .................................... 53
    3.4.2 the fid Data Viewer ...................................... 53
    3.4.3 Image Viewer ............................................... 59
      3.4.3.1 the Image List area ................................ 59
      3.4.3.2 the Set Image area ................................... 62
      3.4.3.3 the Image Viewing area .............................. 62
      3.4.3.4 the Grayscale area on the bottom .................. 63
      3.4.3.5 the Pixel Info area ................................ 63
      3.4.3.6 the Image Statistics area ........................... 64
    3.4.4 Prior Viewer ............................................... 65
    3.4.5 Fid Model Viewer ......................................... 68
      3.4.5.1 The fid Model Format ................................ 70
### 3.4.5.2 The Fid Model Reports

- Page 71

### 3.4.6 Plot Results Viewer

- Page 71

### 3.4.7 Text Results Viewer

- Page 74

### 3.4.8 Files Viewer

- Page 80

### 3.5 Common Interface Plots

- Page 80

#### 3.5.1 Data, Model And Residual Plot

- Page 81

#### 3.5.2 Posterior Probability For A Parameter

- Page 82

#### 3.5.3 Maximum Entropy Histograms

- Page 83

#### 3.5.4 Markov Monte Carlo Samples

- Page 83

#### 3.5.5 Probability Vs Parameter Samples plot

- Page 86

#### 3.5.6 Expected Log Likelihood Plot

- Page 88

#### 3.5.7 Scatter Plots

- Page 88

#### 3.5.8 Logarithm of the Posterior Probability Plot

- Page 91

#### 3.5.9 Fortran/C Code Viewer

- Page 93

##### 3.5.9.1 Fortran/C Model Viewer Popup Editor

- Page 93

---

### 4 An Introduction to Bayesian Probability Theory

#### 4.1 The Rules of Probability Theory

- Page 99

#### 4.2 Assigning Probabilities

- Page 102

#### 4.3 Example: Parameter Estimation

- Page 109

##### 4.3.1 Define The Problem

- Page 110

###### 4.3.1.1 The Discrete Fourier Transform

- Page 110

###### 4.3.1.2 Aliases

- Page 113

##### 4.3.2 State The Model—Single-Frequency Estimation

- Page 114

##### 4.3.3 Apply Probability Theory

- Page 115

##### 4.3.4 Assign The Probabilities

- Page 118

##### 4.3.5 Evaluate The Sums and Integrals

- Page 120

##### 4.3.6 How Probability Generalizes The Discrete Fourier Transform

- Page 123

##### 4.3.7 Aliasing

- Page 126

##### 4.3.8 Parameter Estimates

- Page 132

#### 4.4 Summary and Conclusions

- Page 136

---

### 5 Given Exponential Model

#### 5.1 The Bayesian Calculation

- Page 139

#### 5.2 Outputs From The Given Exponential Package

- Page 141

---

### 6 Unknown Number of Exponentials

#### 6.1 The Bayesian Calculations

- Page 145

#### 6.2 Outputs From The Unknown Number of Exponentials Package

- Page 148

---

### 7 Inversion Recovery

#### 7.1 The Bayesian Calculation

- Page 153

#### 7.2 Outputs From The Inversion Recovery Package

- Page 154
8 Bayes Analyze
8.1 Bayes Model .................................................. 159
8.2 The Bayes Analyze Model Equation .................................. 161
8.3 The Bayesian Calculations ........................................ 167
8.4 Levenberg-Marquardt And Newton-Raphson ............................ 171
8.5 Outputs From The Bayes Analyze Package ....................... 176
  8.5.1 The “bayes.params.mnnn” Files .................................. 177
      8.5.1.1 The Bayes Analyze File Header ................................ 178
      8.5.1.2 The Global Parameters ........................................ 182
      8.5.1.3 The Model Components ........................................ 184
  8.5.2 The “bayes.model.mnnn” Files .................................... 185
  8.5.3 The “bayes.output.mnnn” File .................................... 186
  8.5.4 The “bayes.probabilities.mnnn” File ................................. 190
  8.5.5 The “bayes.log.mnnn” File ....................................... 193
  8.5.6 The “bayes.status.mnnn” and “bayes.accepted.mnnn” Files ......... 196
  8.5.7 The “bayes.model.mnnn” File .................................... 197
  8.5.8 The “bayes.summary1.mnnn” File .................................. 198
  8.5.9 The “bayes.summary2.mnnn” File .................................. 199
  8.5.10 The “bayes.summary3.mnnn” File .................................. 200
  8.6 Bayes Analyze Error Messages ..................................... 200

9 Big Peak/Little Peak .............................................. 207
9.1 The Bayesian Calculation .......................................... 209
9.2 Outputs From The Big Peak/Little Peak Package .................. 216

10 Metabolic Analysis ............................................... 219
10.1 The Metabolic Model .............................................. 223
10.2 The Bayesian Calculation .......................................... 225
10.3 The Metabolite Models ............................................. 228
    10.3.1 The IPGD_D2O Metabolite ...................................... 228
    10.3.2 The Glutamate.2.0 Metabolite .................................. 232
    10.3.3 The Glutamate.3.0 Metabolite .................................. 235
10.4 The Example Metabolite ............................................ 236
10.5 Outputs From The Bayes Metabolite Package ..................... 238

11 Find Resonances .................................................. 239
11.1 The Bayesian Calculations ........................................ 241
11.2 Outputs From The Bayes Find Resonances Package ................ 246

12 Diffusion Tensor Analysis ......................................... 247
12.1 The Bayesian Calculation ........................................ 249
12.2 Using The Package ................................................ 254

13 Big Magnetization Transfer ........................................ 259
13.1 The Bayesian Calculation ........................................ 259
13.2 Outputs From The Big Magnetization Transfer Package .......... 262
# Table of Contents

14 Magnetization Transfer .............................................. 265
   14.1 The Bayesian Calculation ................................... 267
   14.2 Using The Package ........................................... 271

15 Magnetization Transfer Kinetics ................................. 275
   15.1 The Bayesian Calculation ................................... 277
   15.2 Using The Package ........................................... 281

16 Given Polynomial Order ............................................ 285
   16.1 The Bayesian Calculation ................................... 287
   16.1.1 Gram-Schmidt ............................................. 287
   16.1.2 The Bayesian Calculation ................................ 288
   16.2 Outputs From the Given Polynomial Order Package ....... 290

17 Unknown Polynomial Order ......................................... 293
   17.1 Bayesian Calculations ....................................... 295
   17.1.1 Assigning Priors .......................................... 296
   17.1.2 Assigning The Joint Posterior Probability ............. 297
   17.2 Outputs From the Unknown Polynomial Order Package ... 299

18 Errors In Variables ................................................ 303
   18.1 The Bayesian Calculation ................................... 305
   18.2 Outputs From The Errors In Variables Package ........... 308

19 Behrens-Fisher ........................................................ 311
   19.1 Bayesian Calculation ....................................... 311
   19.1.1 The Four Model Selection Probabilities ................. 314
   19.1.1.1 The Means And Variances Are The Same ............... 315
   19.1.1.2 The Mean Are The Same And The Variances Differ ... 317
   19.1.1.3 The Means Differ And The Variances Are The Same ... 318
   19.1.1.4 The Means And Variances Differ ...................... 319
   19.1.2 The Derived Probabilities ................................ 320
   19.1.3 Parameter Estimation ..................................... 321
   19.2 Outputs From Behrens-Fisher Package ..................... 322

20 Enter Ascii Model ................................................... 329
   20.1 The Bayesian Calculation ................................... 331
   20.1.1 The Bayesian Calculations Using Eq. (20.1) ............ 331
   20.1.2 The Bayesian Calculations Using Eq. (20.2) ............ 332
   20.2 Outputs From The Enter Ascii Model Package ............. 335

21 Enter Ascii Model Selection ....................................... 337
   21.1 The Bayesian Calculations ................................... 339
   21.1.1 The Direct Probability With No Amplitude Marginalization ... 340
   21.1.2 The Direct Probability With Amplitude Marginalization ... 342
   21.1.2.1 Marginalizing the Amplitudes ....................... 343
   21.1.2.2 Marginalizing The Noise Standard Deviation ......... 348
List of Figures

1.1 The Start Up Window .................................................. 23
1.2 Example Package Exponential Interface ............................. 25
2.1 Installation Kit For The Bayesian Analysis Software ............ 31
3.1 The Start Up Window .................................................. 34
3.2 The Files Menu ...................................................... 35
3.3 The Files/Load Image Submenu ....................................... 37
3.4 The Packages Menu .................................................... 41
3.5 The Working Directory Menu ......................................... 46
3.6 The Working Directory Information Popup ......................... 47
3.7 The Settings Pull Down Menu ......................................... 47
3.8 The MCMC Parameters Popup ......................................... 48
3.9 The Edit Server Popup .................................................. 49
3.10 The Submit Job Widgets ............................................... 51
3.11 The Server Widgets Group ............................................ 52
3.12 The Ascii Data Viewer ................................................. 54
3.13 The Fid Data Viewer .................................................. 55
3.14 Fid Data Display Type ................................................ 56
3.15 Fid Data Options Menu ............................................... 58
3.16 The Image Viewer ..................................................... 60
3.17 The Image Viewer Right Mouse Popup Menu ....................... 61
3.18 The Prior Probability Viewer ......................................... 66
3.19 The Fid Model Viewer .................................................. 69
3.20 The Plot Results Viewer ............................................... 72
3.21 Plot Information Popup ............................................... 73
3.22 The Text Results Viewer ............................................... 75
3.23 The Bayes Condensed File ............................................ 78
3.24 Data, Model, And Resid Plot ......................................... 81
3.25 The Parameter Posterior Probabilities ............................... 82
3.26 The Maximum Entropy Histograms .................................. 84
3.27 The Parameter Samples Plot .......................................... 85
3.28 Posterior Probability Vs Parameter Value ........................... 86
3.29 Posterior Probability Vs Parameter Value, A Skewed Example ... 87
3.30 The Expected Value Of The Logarithm Of The Likelihood ........ 89
3.31 The Scatter Plots ................................................. 90
3.32 The Logarithm Of The Posterior Probability By Repeat Plot .......... 92
3.33 The Fortran/C Model Viewer .................................. 94
3.34 The Fortran/C Code Editor .................................. 95
4.1 Frequency Estimation Using The DFT .......................... 112
4.2 Aliases ......................................................... 113
4.3 Nonuniformly Nonsimultaneously Sampled Sinusoid ...................... 127
4.4 Alias Spacing .................................................. 128
4.5 Which Is The Critical Time ..................................... 130
4.6 Example, Frequency Estimation .................................. 131
4.7 Estimating The Sinusoids Parameters .................................. 133
5.1 The Given And Unknown Number Of Exponential Package Interface .... 138
6.1 The Unknown Exponential Interface ................................ 144
6.2 The Distribution Of Models ..................................... 149
6.3 The Posterior Probability For Exponential Model ......................... 150
7.1 The Inversion Recovery Interface .................................. 152
8.1 Bayes Analyze Interface ......................................... 156
8.2 Bayes Analyze Fid Model Viewer .................................. 160
8.3 The Bayes Analyze File Header .................................. 179
8.4 The bayes.noise File .............................................. 180
8.5 Bayes Analyze Global Parameters .................................. 183
8.6 The Third Section Of The Parameter File ............................. 184
8.7 Example Of An Initial Model In The Output File ......................... 187
8.8 Base 10 Logarithm Of The Odds .................................. 187
8.9 A Small Sample Of The Output Report ................................ 188
8.10 Bayes Analyze Uncorrelated Output ................................ 189
8.11 The bayes.probabilities.nnnn File ................................. 191
8.12 The bayes.log.nnnn File ........................................... 193
8.13 The bayes.status.nnnn File ....................................... 196
8.14 The bayes.model.nnnn File ....................................... 197
8.15 The bayes.model.nnnn File Uncorrelated Resonances ...................... 198
8.16 Bayes Analyze Summary Header .................................. 198
8.17 The Summary2 (Best Summary) .................................... 199
8.18 The Summary3 Report .............................................. 201
9.1 The Big Peak/Little Peak Interface .................................. 208
9.2 The Time Dependent Parameters .................................... 218
10.1 The Bayes Metabolite Interface .................................... 220
10.2 The Bayes Metabolite Viewer ..................................... 222
10.3 Bayes Metabolite Parameters And Probabilities List ...................... 227
10.4 The IPGD_D20 Metabolite ........................................ 229
10.5 Bayes Metabolite IPGD,D20 Spectrum .................................................. 230
10.6 Bayes Metabolite, The Fraction of Glucose ........................................... 231
10.7 Glutamate Example Spectrum .............................................................. 233
10.8 Estimating The $F_{c0}$, $y$ and $F_{a0}$ Parameters ...................................... 236
10.9 Bayes Metabolite, The Ethyl Ether Example .......................................... 237

11.1 The Find Resonances Interface With The Ethyl Ether Spectrum ................. 240

12.1 The Diffusion Tensor Package Interface ............................................... 248
12.2 Diffusion Tensor Parameter Estimates .................................................. 256
12.3 Diffusion Tensor Posterior Probability For The Model ............................ 257

13.1 The Big Magnetization Package Interface ............................................. 260
13.2 Big Magnetization Transfer Example Fid .............................................. 263
13.3 Big Magnetization Transfer Expansion ................................................ 263
13.4 Big Magnetization Transfer Peak Pick .................................................. 264

14.1 The Magnetization Transfer Package Interface ..................................... 266
14.2 Magnetization Transfer Package Peak Picking ...................................... 272
14.3 Magnetization Transfer Example Data .................................................. 273
14.4 Magnetization Transfer Example Spectrum ......................................... 274

15.1 Magnetization Transfer Kinetics Package Interface ................................ 276
15.2 Magnetization Transfer Kinetics Package Arrhenius Plot ......................... 282
15.3 Magnetization Transfer Kinetics Water Viscosity Table ........................... 283

16.1 Given Polynomial Order Package Interface ......................................... 286
16.2 Given Polynomial Order Scatter Plot .................................................. 291

17.1 Unknown Polynomial Order Package Interface ..................................... 294
17.2 The Distribution of Models On The Console Log ................................... 298
17.3 The Posterior Probability For The Polynomial Order ............................... 300

18.1 The Errors In Variables Package Interface ......................................... 304
18.2 The McMC Values File Produced By The Errors In Variables Package ........ 310

19.1 The Behrens-Fisher Interface .............................................................. 312
19.2 Behrens-Fisher Hypotheses Tested ....................................................... 313
19.3 Behrens-Fisher Console Log ................................................................. 323
19.4 Behrens-Fisher Status Listing .............................................................. 324
19.5 Behrens-Fisher McMC Values File, The Preamble .................................. 325
19.6 Behrens-Fisher McMC Values File, The Middle ..................................... 326
19.7 Behrens-Fisher McMC Values File, The End .......................................... 327

20.1 Enter Ascii Model Package Interface .................................................... 330

21.1 The Enter Ascii Model Selection Package Interface ................................ 338
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>26.1 Absorption Model Images</td>
<td>396</td>
</tr>
<tr>
<td>26.2 The Interface To The Image Phasing Package</td>
<td>397</td>
</tr>
<tr>
<td>26.3 Linear Phasing Package The Console Log</td>
<td>403</td>
</tr>
<tr>
<td>27.1 Nonlinear Phasing Example</td>
<td>406</td>
</tr>
<tr>
<td>27.2 The Interface To The Nonlinear Phasing Package</td>
<td>410</td>
</tr>
<tr>
<td>28.1 The Interface To The Analyze Image Pixels Package</td>
<td>412</td>
</tr>
<tr>
<td>29.1 The Interface To The Image Model Selection Package</td>
<td>416</td>
</tr>
<tr>
<td>29.2 Single Exponential Example Image</td>
<td>419</td>
</tr>
<tr>
<td>29.3 Single Exponential Example Data</td>
<td>420</td>
</tr>
<tr>
<td>29.4 Posterior Probability For The ExpOneNoConst Model</td>
<td>421</td>
</tr>
<tr>
<td>A.1 Ascii Data File Format</td>
<td>424</td>
</tr>
<tr>
<td>D.1 The McMC Values Report Header</td>
<td>450</td>
</tr>
<tr>
<td>D.2 McMC Values Report, The Middle</td>
<td>451</td>
</tr>
<tr>
<td>D.3 The McMC Values Report, The End</td>
<td>452</td>
</tr>
<tr>
<td>E.1 Writing Models A Fortran Example</td>
<td>456</td>
</tr>
<tr>
<td>E.2 Writing Models A C Example</td>
<td>457</td>
</tr>
<tr>
<td>E.3 Writing Models, The Parameter File</td>
<td>459</td>
</tr>
<tr>
<td>E.4 Writing Models Fortran Declarations</td>
<td>463</td>
</tr>
<tr>
<td>E.5 Writing Models Fortran Example</td>
<td>466</td>
</tr>
<tr>
<td>E.6 Writing Models The Parameter File</td>
<td>467</td>
</tr>
<tr>
<td>G.1 Example FDF File Header</td>
<td>473</td>
</tr>
<tr>
<td>H.1 The Posterior Probability For The Number of Outliers</td>
<td>476</td>
</tr>
<tr>
<td>H.2 The Data, Model and Residual Plot With Outliers</td>
<td>478</td>
</tr>
</tbody>
</table>
## List of Tables

8.1 Multiplet Relative Amplitudes ............................................. 165  
8.2 Bayes Analyze Models ...................................................... 181  
8.3 Bayes Analyze Short Descriptions ....................................... 195
Chapter 18

Errors In Variables

The “Errors in Variables” package fits polynomials to data when you have errors in both the abscissa \( X \) and the data value \( Y \). The interface to this package is shown in Fig. 18.1. This interface is used to configure the Errors In Variables Package by setting both the polynomial order and by indicating what errors are known or given. Additionally, depending on the settings of the “Given Errors In” widget, the interface will load two, three or four column Ascii data. To use this package, you must do the following:

Select the Errors In Variables Package from the Package menu.

Using “Given Errors In” pull down menu select the type of Errors In Variables problem you wish to solve. Your choices are:

1. “Not Given” solves the Errors In Variables problem when the errors in both \( X \) and \( Y \) are not given. This option requires two column Ascii data, see Section 18.2 for a description of these files.
2. “\( X \) and \( Y \)” solves the Errors In Variables problem when the errors in \( X \) and \( Y \) are given. This option requires four column Ascii data.
3. “\( X \) Only” solves the Errors In Variables problem when the errors in \( X \) are given. This option requires three column Ascii data.
4. “\( Y \) Only” solves the Errors In Variables problem when the errors in \( Y \) are given. This option requires three column Ascii data.

Load one data set appropriate to the selected problem. Only “Not Given” uses standard two column Ascii data. Consequently, when this option is selected both Bayes Analyze and a Peak Pick can server as input. All of the other options require input of either three or four column Ascii files which can only be loaded using the “Files/Load Ascii/Files” menu. When you have successfully loaded a data set it will be plotted in the Ascii Data Viewer. However, the plot is a simple XY plot and no attempt is made by the interface to show the error bars in either \( X \) or \( Y \).

Using the “Order” pull down menu, set the order of the polynomial to fit to the data.
Figure 18.1: The Errors In Variables Package Interface

To use the Errors In Variables package:

1. Indicate which errors are known using the “Given Errors In:” combo box.
2. Load an Ascii file having 2, 3 or four columns as indicated.
3. Select the polynomial order you wish to analyze.
4. Select the server to run the analysis.
5. Run the analysis using the “Run” button.
6. Use “Get Job” to get the results from the server.

Figure 18.1: The Errors In Variables Package solves the problem of fitting polynomials when there are errors in both $X$ and $Y$ where $X$ is an abscissa and $Y$ is some function of $X$. Here this function is assumed to be a polynomial of a given order.
Select the server that is to process the analysis.

Check the status of the selected server to determine if the server is busy, change to another server if the selected server is busy.

Run the analysis on the selected server by activating the Run button.

Get the results of the analysis by activating the Get Job button. If the analysis is running, this button will return the Accepted report containing the status of the current run. Otherwise, it will fetch and display the results from the current analysis.

18.1 The Bayesian Calculation

The problem we are considering is one in which the data consists of measured pairs of numbers, \((\hat{x}_i, \hat{y}_i)\), where both the measured abscissa data value, \(\hat{x}_i\), and the measured ordinate, \(\hat{y}_i\), contain errors, i.e., noise, and we wish to fit a polynomial to these measured data pairs. If we designate the “true” but unknown values of \((x_i, y_i)\) as \((x_i, y_i)\) then the polynomial one would normally fit is given by

\[
y = F(x) = \sum_{j=0}^{m} B_j H_j(x)
\]

where the polynomial, \(F(x)\), is evaluated at the true value of \(x\) giving a true value of \(y\), \(B_j\) is the amplitude of the polynomial \(H_j(x)\). These polynomials are the same polynomials discussed in Chapter 16 and are the Gram-Schmidt polynomials generated from \(x^j\). Unfortunately, neither the \(y\) values nor the \(x\) values are known. An abscissa data value, \(\hat{x}_i\), is related to the unknown true abscissa, \(x_i\), by

\[
\hat{x}_i = x_i + e_i
\]

where \(e_i\) is the error in the measured abscissa. Similarly, a measured ordinate value, \(\hat{y}_i\), is related to the unknown true ordinate, \(y_i\), by

\[
\hat{y}_i = y_i + n_i
\]

where \(n_i\) is the error in the measured ordinate.

In the following calculation we will assume that the errors in the abscissa and ordinate are both unknown. We do this for the simple reason that it is the harder, i.e., more interesting calculation. However, in some cases one or both of these errors are actually known and the computer program that implements this package implements four different cases: 1) errors in both abscissa and ordinate known, 2) errors in the abscissa know but the ordinate error is unknown, etc.

If we Taylor expand the polynomial, \(F(x)\), around \(\hat{x}_i\), then

\[
F(x_i) \approx F(\hat{x}_i) + F'(\hat{x}_i)(\hat{x}_i - x_i)
\]

where

\[
F'(\hat{x}_i) = \frac{dF(\hat{x}_i)}{d\hat{x}_i}
\]

This first order Taylor expansion will gives a reasonable approximation to \(F(x)\) if the errors in the unknown abscissa are not large, i.e., if the polynomial is approximately linear over the likely error in
the abscissa. We make the Taylor expansion because, now the unknown true values of $x_i$ appear in
the model in a linear fashion and consequently we will be able to remove them by marginalization.

The probabilities we want to compute are the marginal posterior probabilities for the amplitudes
of the polynomials and standard deviations of the noise prior probabilities. To calculate these, the
numerical simulations must target the joint posterior probability for the amplitudes, $\{B_0, \ldots, B_m\}$,
and the noise standard deviations, $\sigma_x$ and $\sigma_y$ given the abscissa and ordinate data. This joint
posterior probability is represented symbolically by $P(B_0 \cdots B_m \sigma_x \sigma_y | \hat{x} \hat{y} I)$, where $\hat{x}$ and $\hat{y}$ represent
the abscissa and ordinate data.

The joint posterior probability for the amplitudes and the noise standard deviations targeted by
the Markov chain Monte Carlo simulation, $P(B_0 \cdots B_m \sigma_x \sigma_y | \hat{x} \hat{y} I)$, is a marginal probability:

$$P(B_0 \cdots B_m \sigma_x \sigma_y | \hat{x} \hat{y} I) = \int P(B_0 \cdots B_m \sigma_x \sigma_y | x) | \hat{x} \hat{y} I) d\{x\} \tag{18.6}$$

where the integrals are over all of the unknown abscissa values. To compute this posterior probability,
one factors the right-hand side of this equation using Bayes’ theorem and the product rule to obtain

$$P(B_0 \cdots B_m \sigma_x \sigma_y | x) | \hat{x} \hat{y} I) \propto P(\sigma_x | I)P(\sigma_y | I) \prod_{j=0}^{m} P(B_j | I) \times \prod_{i=1}^{N} P(x_i | I) \times \prod_{i=1}^{N} P(\hat{x}_i | x_i \sigma_x I) \times \prod_{i=1}^{N} P(\hat{y}_i | x_i B_0 \cdots B_m \sigma_y I). \tag{18.7}$$

The prior probability for the standard deviation of the noise for the abscissa data, $P(\sigma_x | I)$, will be
assigned as a bound Jeffreys’ prior

$$P(\sigma_x | I) = \begin{cases} \frac{1}{R_x \sigma_x} & \text{if } \sigma_{xL} \leq \sigma_x \leq \sigma_{xH} \\ 0 & \text{otherwise} \end{cases} \tag{18.8}$$

where $\sigma_{xL}$ and $\sigma_{xH}$ are a lower and upper bound on $\sigma_x$. The normalization constant $R_x$ is given by

$$R_x = \int_{\sigma_{xL}}^{\sigma_{xH}} \frac{d\sigma_x}{\sigma_x} = \log(\sigma_{xH}/\sigma_{xL}). \tag{18.9}$$

The bounds, $\sigma_{xL}$ and $\sigma_{xH}$, are set rather pragmatically within the program that implements this
package. The program computes the mean-square deviation, $\langle \hat{x}^2 - \bar{x}^2 \rangle$, using the $\hat{x}$ data, and then
sets the lower and upper bounds

$$\sigma_{xL} = \frac{\langle \hat{x}^2 - \bar{x}^2 \rangle}{10} \quad \text{and} \quad \sigma_{xH} = 10 \langle \hat{x}^2 - \bar{x}^2 \rangle, \tag{18.10}$$

which restricts $\sigma_x$ to a two order of magnitude variation. Similarly, $P(\sigma_y | I)$, will be assigned

$$P(\sigma_y | I) = \begin{cases} \frac{1}{R_y \sigma_y} & \text{if } \sigma_{yL} \leq \sigma_y \leq \sigma_{yH} \\ 0 & \text{otherwise} \end{cases} \tag{18.11}$$
where the normalization constant, \( R_y \), and the lower and upper bounds are computed analogously to corresponding abscissa values. The prior probability for the amplitudes, \( P(B_j|I) \), is assigned a bounded zero mean Gaussian. The bounds are set at plus and minus ten times the largest projection of the model onto the data:

\[
P(B_j|I) = \begin{cases} 
(2\pi \sigma_B^2)^{-1/2} \exp \left\{ -\frac{B_j^2}{2\sigma_B^2} \right\} & \text{if } B_L \leq B_j \leq B_H \\
0 & \text{otherwise}
\end{cases}
\]  

(18.12)

with \( B_L = -B_H \) and \( B_H = \max(\sum_{i=1}^{N} \hat{y}_i H_j(\hat{x}_i)) \). The standard deviation of this prior is \( \sigma_B = B_H/3 \). So the interval, \( B_H - B_L \), represents a 6 standard deviation interval and the prior serves as little more than a way to keep the amplitudes from wandering into an unphysical region of the parameter space.

In this problem the unknown true values, the \( x_i \), are parameters and one must assign a prior probability to all such parameters. The prior probability for the \( x_i \), the \( P(x_i|I) \), are assigned as unbounded Gaussians having mean equal to \( \hat{x}_i \) and standard deviation, \( \sigma_p \):

\[
P(x_i|I) = (2\pi \sigma_p^2)^{-1/2} \exp \left\{ -\frac{(\hat{x}_i - x_i)^2}{2\sigma_p^2} \right\}
\]  

(18.13)

where \( \hat{x}_i \) is the sampling point we thought we were measuring, and we set \( \sigma_p \) equal to the average \( x \) interval. We did this because it greatly simplified the formulas that must be programmed and so makes for a faster program, without changing the results to within the error bars.

The likelihood for a measured \( \hat{x}_i \) data value, \( P(\hat{x}_i|x_i\sigma_x I) \), was assigned a Gaussian

\[
P(\hat{x}_i|x_i\sigma_x I) = (2\pi \sigma_x^2)^{-1/2} \exp \left\{ -\frac{(\hat{x}_i - x_i)^2}{2\sigma_x^2} \right\}
\]  

(18.14)

and the likelihood for a measured \( \hat{y}_i \) data value, \( P(\hat{y}_i|\{A\}\sigma_y I) \), was assigned a Gaussian likelihood of the form:

\[
P(\hat{y}_i|B_0 \cdots B_m\sigma_y I) = (2\pi \sigma_y^2)^{-1/2} \exp \left\{ -\frac{(\hat{y}_i - F(\hat{x}_i) - F'(\hat{x}_i)[\hat{x}_i - x_i])^2}{2\sigma_y^2} \right\}.
\]  

(18.15)

If we now accumulate all of the priors, Eqs. (18.8,18.11,18.13), and likelihoods, Eqs.(18.14,18.15),
and substitute them into Eq. (18.6) one obtains

\[ P(B_0 \cdots B_m \sigma_x \sigma_y | \hat{x} \hat{y}) \propto \int_{-\infty}^{+\infty} dx_1 \cdots dx_N \frac{1}{R_x \sigma_x} \frac{1}{R_y \sigma_y} \]

\[ \times (2\pi \sigma_B^2)^{-\frac{m+1}{2}} \exp \left\{ -\sum_{j=0}^{m} \frac{B_j^2}{2\sigma_B^2} \right\} \]

\[ \times (2\pi \sigma_p^2)^{-\frac{N}{2}} \exp \left\{ -\sum_{i=1}^{N} \frac{\hat{x}_i - x_i}{2\sigma_p^2} \right\} \]

\[ \times (2\pi \sigma_x^2)^{-\frac{N}{2}} \exp \left\{ -\sum_{i=1}^{N} \frac{\hat{x}_i - x_i}{2\sigma_x^2} \right\} \]

\[ \times (2\pi \sigma_y^2)^{-\frac{N}{2}} \exp \left\{ -\sum_{i=1}^{N} \frac{\hat{y}_i - F(\hat{x}_i) - F'(\hat{x}_i)[\hat{x}_i - x_i]}{2\sigma_y^2} \right\} \]

(18.16)

as the posterior probability for the amplitudes and noise standard deviations. Evaluating the \( N \) integrals over the \( x_i \), one obtains

\[ P(B_0 \cdots B_m \sigma_x \sigma_y | \hat{x} \hat{y}) \propto \frac{1}{\sigma_x \sigma_y} \exp \left\{ -\sum_{j=0}^{m} \frac{B_j^2}{2\sigma_B^2} \right\} \prod_{i=1}^{N} \left( \frac{\sigma_x \sigma_y}{\sigma_i} \right) \exp \left\{ -\frac{Q_i}{2\sigma_i^2} \right\} \]

(18.17)

where we have dropped some constants that cancel when this probability density function is normalized. The function, \( Q_i \), is given by

\[ Q_i \equiv \left[ \hat{y}_i - F(\hat{x}_i) \right]^2 \frac{\sigma_y^2}{\sigma_i^2} + \left[ \hat{x}_i - x_i \right]^2 \frac{\left( \hat{y}_i - F(\hat{x}_i) \right)}{\sigma_y^2} + \left( \hat{y}_i - F(\hat{x}_i) \right)^2 \frac{2\sigma_y^2}{\sigma_i^2} \]

(18.18)

where

\[ \sigma_i \equiv \sqrt{\sigma_x^2 \sigma_y^2 + \sigma_p^2 \sigma_x^2 + F'(\hat{x}_i) \sigma_p^2 \sigma_x^2} \]

(18.19)

In the special case that the Errors In Variables package implements, \( \hat{x}_i = x_i \), \( Q_i \) simplifies and one obtains

\[ P(B_0 \cdots B_m \sigma_x \sigma_y | \hat{x} \hat{y}) \propto \frac{1}{\sigma_x \sigma_y} \exp \left\{ -\sum_{j=0}^{m} \frac{B_j^2}{2\sigma_B^2} \right\} \prod_{i=1}^{N} \left( \frac{\sigma_x \sigma_y}{\sigma_i} \right) \exp \left\{ -\frac{\left[ \hat{y}_i - F(\hat{x}_i) \right]^2 \left( \sigma_y^2 + \sigma_x^2 \right)}{2\sigma_i^2} \right\} \]

(18.20)

Computationally, this special case is simpler to calculate, so the program runs faster without given up the ability to estimate both \( \sigma_x \) and \( \sigma_y \) and it is this probability density function that is targeted by the Markov chain Monte Carlo simulation.

### 18.2 Outputs From The Errors In Variables Package

These output files can be viewed using the Text Viewer or they can be viewed using File Viewer by navigating to the current working directory and then selecting the files. The format of the mcmc.values report is discussed in Appendix D and the other reports are discussed in Chapter 3. Additionally, the “Plot Results Viewer” can be used to view the output probability density functions. In addition to the standard data, model and residual plots there are probability density functions for the decay rate constants, decay times, the amplitudes for each data set for each exponential and finally there are probability density functions for the standard deviation of the noise in each data set.

The only thing the least bit unusual about this package is the Ascii data that is required. In most Ascii packages the data are two columns, an abscissa and an ordinate. However, here there are four different file formats:

1. When the errors in both the abscissa and ordinate are unknown, two column Ascii data is required. Column one is the abscissa and column two is the ordinate. These data may be generated and or loaded using the files menu.

2. When the errors are known in the abscissa but not in the ordinate, three column Ascii data is required. Column one is the abscissa, column two is the ordinate, and column three is the error in the abscissa. This input can only be loaded using the “Files/Load Ascii/File” menu.

3. When the errors are known in the ordinate but not in the abscissa, three column Ascii data is required. Column one is the abscissa, column two is the ordinate, and column three is the error in the ordinate. This input can only be loaded using the “Files/Load Ascii/File” menu.

4. When the errors are known in both the abscissa and the ordinate, four column Ascii data is required. Column one is the abscissa, column two is the ordinate, column three is the error in the abscissa and column four is the error in the ordinate. This input can only be loaded using the “Files/Load Ascii/File” menu.

There are four test data sets located in the “Bayes.test.data/ErrorsInVariables” directory that may be used for testing. These four data sets are named: ErrInVar_given_x.dat, ErrInVar_given_xy.dat, ErrInVar_given_y.dat, and ErrInVar_not_given.dat respectively. An example of the McMC values report generated by the Errors In Variables package is shown in Fig. 18.2. This report was generated using the “ErrInVar_given_xy.dat” data. The top section of the report contains some configuration information followed by information about the priors. This is followed by some averages over the various probabilities including the posterior probability for the model. This is followed by the parameters that maximized the joint marginal posterior probability for the parameters. Finally, the mean and standard deviation estimates of the amplitudes of the polynomials are given. For more on the general layout of the McMC value file, see Appendix D.

In addition to the McMC values file there are the standard Data, Model and Residual plots which can be viewed using the Plot Results Viewer. There are posterior probability density functions for the uncertainty in the abscissa data values, this plot is named “Sigma X” and there is a posterior probability density for the uncertainty in the ordinate, named “Sigma Y.” Additionally, there is one output probability density function for each amplitude in the polynomial being analyzed. So if you are analyzing a 6th order polynomial, there are seven output probability density functions. Finally, there are the output plots that come at the end of the plot list. These include the expected logarithm of the likelihood as a function of the annealing parameter, the scatter plots and the logarithm of the posterior probability for each simulation as a function of repeat number.
The Parameter File Listing for the Errors in Variables Package

! BayesErrInVarsGiven Package
! Created 10-Feb-2012 10:11:27 by larry
!
   Output Dir = BayesOtherAnalysis
Number Of Abscissa = 1
Number Of Columns = 1
Number Of Sets = 1
   File Name = BayesOtherAnalysis/ErrInVar_given_xy.dat
McMC Simulations = 48
McMC Repeats = 21
Minimum Annealing Steps = 21
Histogram Type = Binned
Outlier Detection = Disabled
Number Of Priors = 0
Package Parameters = 2
Total Mcmc Samples = 1008
   Kill Count = 4

McMC Values Report for the Errors In Variables package

<table>
<thead>
<tr>
<th>Param Desc</th>
<th>Low</th>
<th>Mean</th>
<th>High</th>
<th>Sigma</th>
<th>Norm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coef 0.1</td>
<td>-1.504E+03</td>
<td>0.000E+00</td>
<td>1.504E+03</td>
<td>4.513E+02</td>
<td>-3.626E+00</td>
</tr>
<tr>
<td>Coef 1.1</td>
<td>-1.504E+03</td>
<td>0.000E+00</td>
<td>1.504E+03</td>
<td>4.513E+02</td>
<td>-3.626E+00</td>
</tr>
<tr>
<td>X 1.1</td>
<td>-5.445E-02</td>
<td>-4.785E-03</td>
<td>4.488E-02</td>
<td>9.933E-03</td>
<td>-3.222E+00</td>
</tr>
<tr>
<td>X 101.1</td>
<td>9.488E-01</td>
<td>9.985E-01</td>
<td>1.048E+00</td>
<td>9.933E-03</td>
<td>-3.222E+00</td>
</tr>
</tbody>
</table>

Avg. Sd.

The Average Log Posterior Probability: 194.0262 122.93507
The Average Log Prior Params: -9.0257 0.00079
The Average Log Likelihood: 4.06109543E+02 1.37923E+00
The Log Posterior Probability For The Model: 3.74864004E+02

The parameters that maximized the posterior probability are:

<table>
<thead>
<tr>
<th>Parameter Description</th>
<th>Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Std Dev in X</td>
<td>4.65331682E-04</td>
</tr>
<tr>
<td>Std Dev in Y</td>
<td>1.10193510E+00</td>
</tr>
<tr>
<td>Coef 0.1</td>
<td>1.01151226E+01</td>
</tr>
<tr>
<td>Coef 1.1</td>
<td>9.68326780E+00</td>
</tr>
</tbody>
</table>

The expected parameter values (mean value of the probability distributions):

<table>
<thead>
<tr>
<th>Parameter Description</th>
<th>Mean Value</th>
<th>Std. Dev.</th>
<th>Peak Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sigma X</td>
<td>1.09472E-03</td>
<td>3.34846E-04</td>
<td>4.65332E-04</td>
</tr>
<tr>
<td>Sigma Y</td>
<td>1.09118E+00</td>
<td>7.44062E-02</td>
<td>1.10194E+00</td>
</tr>
<tr>
<td>Coef 1.0</td>
<td>1.01672E+01</td>
<td>2.01540E-01</td>
<td>1.0151E+01</td>
</tr>
<tr>
<td>Coef 1.1</td>
<td>9.59826E+00</td>
<td>3.49706E-01</td>
<td>9.68327E+00</td>
</tr>
</tbody>
</table>

Figure 18.2: This is the Errors In Variables mcmc.values file. It is the primary printed output from the Errors In Variables package. This report was generated using test data found in Bayes.test.data, the ErrorsInVariables subdirectory, the data file was ErrInVar_given_xy.dat. The top section of the report contains some configuration information followed by information about the priors. This is followed by some averages over the various probabilities including the posterior probability for the model. This is followed by the parameters that maximized the joint marginal posterior probability. Finally, the mean and standard deviation estimates of the amplitudes of the polynomials are given.
Bibliography


479


[45] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller (1953), “Equation of State Calculations by Fast Computing Machines,” Journal of Chemical Physics. The previous link is to the American Institute of Physics and if you do not have access to Science Sitations you many not be able to retrieve this paper.


