# 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

October 21, 2016

# Contents

# List of Figures

# List of Tables

# 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

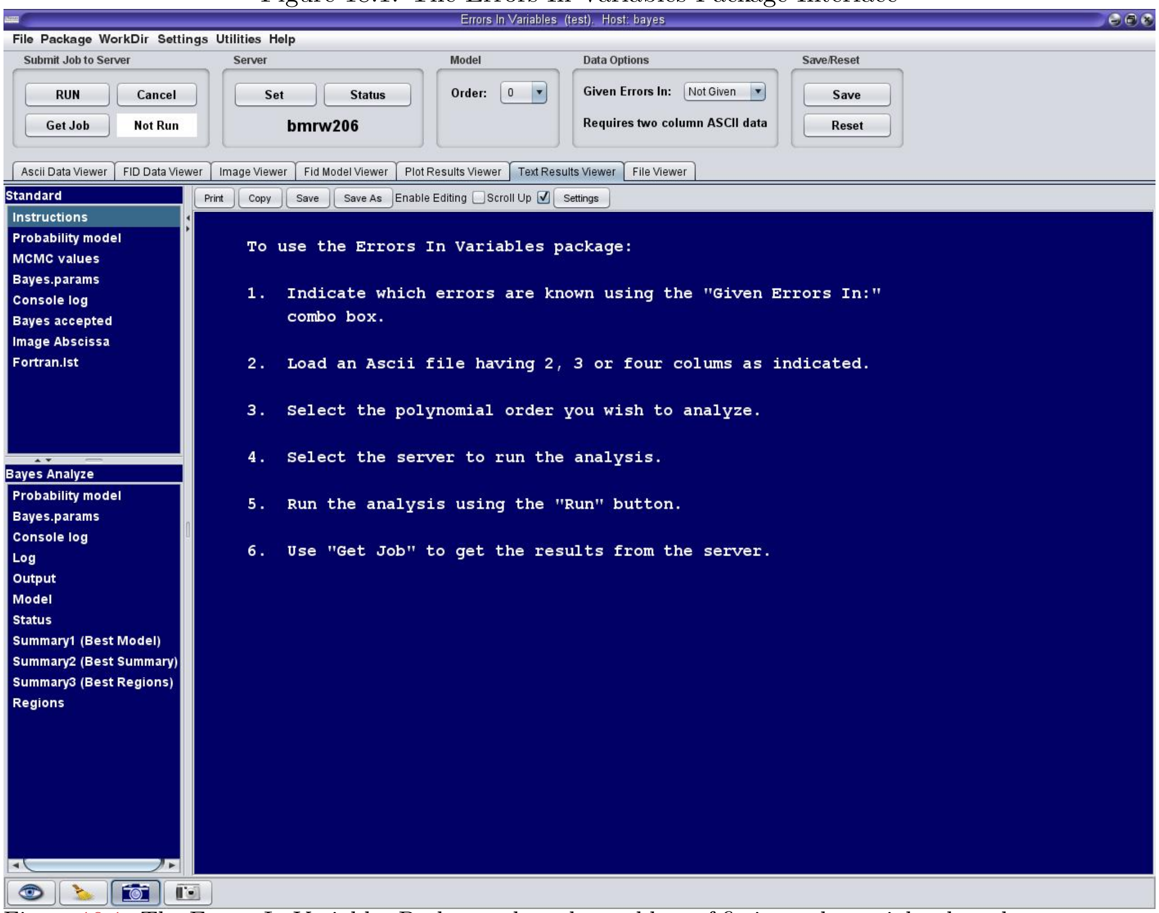


Figure 18.1: The Errors In Variables Package solves the problem of fitting polynomials when there are errors in both $X$ and $Y$ where $X$ is an abscissa and $Y$ is some function of $X$. Here this function is assumed to be a polynomial of a given order.

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

**Get** the 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 $(\hat{x}_i, \hat{y}_i)$ as $(x_i, y_i)$ then the polynomial one would normally fit is given by

$$y = F(x) = \sum_{j=0}^{m} B_j H_j(x) \tag{18.1}$$

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

$$\hat{x}_i = x_i + e_i \tag{18.2}$$

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

$$\hat{y}_i = y_i + n_i \tag{18.3}$$

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

$$F(x_i) \approx F(\hat{x}_i) + F'(\hat{x}_i)(\hat{x}_i - x_i) \tag{18.4}$$

where

$$F'(\hat{x}_i) = \frac{dF(\hat{x}_i)}{d\hat{x}_i}. \tag{18.5}$$

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

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

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

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

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, $\langle \bar{x^2} - \bar{x}^2 \rangle$, using the $\hat{x}$ data, and then sets the lower and upper bounds

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

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

$$P(\sigma_y|I) = \begin{cases} \dfrac{1}{R_y \sigma_y} & \text{if } \sigma_{yL} \leq \sigma_y \leq \sigma_{yH} \\ 0 & \text{otherwise} \end{cases} \tag{18.11}$$

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:

$$P(B_j|I) = \begin{cases} (2\pi\sigma_B^2)^{-1/2} \exp\left\{-\dfrac{B_j^2}{2\sigma_B^2}\right\} & \text{if } B_L \le B_j \le B_H \\ 0 & \text{otherwise} \end{cases} \tag{18.12}$$

with $B_L = -B_H$ and $B_H = \max(\text{abs}[\sum_{i=1}^{N} \hat{y}_i H_j(\hat{x}_i)])$. 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 $\tilde{x}_i$ and standard deviation, $\sigma_p$:

$$P(x_i|I) = (2\pi\sigma_p^2)^{-1/2} \exp\left\{-\frac{(\tilde{x}_i - x_i)^2}{2\sigma_p^2}\right\} \tag{18.13}$$

where $\tilde{x}_i$ 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

$$P(\hat{x}_i|x_i\sigma_x I) = (2\pi\sigma_x^2)^{-1/2} \exp\left\{-\frac{(\hat{x}_i - x_i)^2}{2\sigma_x^2}\right\} \tag{18.14}$$

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:

$$P(\hat{y}_i|B_0 \cdots B_m\sigma_y I) = (2\pi\sigma_y^2)^{-1/2} \exp\left\{-\frac{\left(\hat{y}_i - F(\hat{x}_i) - F'(\hat{x}_i)[\hat{x}_i - x_i]\right)^2}{2\sigma_y^2}\right\}. \tag{18.15}$$

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

$$\begin{aligned} P(B_0 \cdots B_m \sigma_x \sigma_y | \hat{x}\hat{y} I) &\propto \int_{-\infty}^{+\infty} dx_1 \cdots dx_N \frac{1}{R_x \sigma_x} \frac{1}{R_y \sigma_y} \\ &\times (2\pi\sigma_B^2)^{-\frac{m+1}{2}} \exp\left\{ -\sum_{j=0}^{m} \frac{B_j^2}{2\sigma_B^2} \right\} \\ &\times (2\pi\sigma_p^2)^{-\frac{N}{2}} \exp\left\{ -\sum_{i=1}^{N} \frac{(\tilde{x}_i - x_i)^2}{2\sigma_p^2} \right\} \\ &\times (2\pi\sigma_x^2)^{-\frac{N}{2}} \exp\left\{ -\sum_{i=1}^{N} \frac{(\hat{x}_i - x_i)^2}{2\sigma_x^2} \right\} \\ &\times (2\pi\sigma_y^2)^{-\frac{N}{2}} \exp\left\{ -\sum_{i=1}^{N} \frac{\left(\hat{y}_i - F(\hat{x}_i) - F'(\hat{x}_i)[\hat{x}_i - x_i]\right)^2}{2\sigma_y^2} \right\} \end{aligned} \tag{18.16}$$

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

$$P(B_0 \cdots B_m \sigma_x \sigma_y | \hat{x}\hat{y}) \propto \frac{1}{\sigma_x \sigma_y} \exp\left\{ -\sum_{j=0}^{m} \frac{B_j^2}{2\sigma_B^2} \right\} \prod_{i=1}^{N} \left( \frac{\sigma_x \sigma_y}{\sigma_i} \right) \exp\left\{ -\frac{Q_i}{2\sigma_i^2} \right\} \tag{18.17}$$

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

$$Q_i \equiv [\hat{y}_i - F(\hat{x}_i)]^2 [\sigma_p^2 + \sigma_x^2] - 2F'(\hat{x}_i)\sigma_x^2 [\hat{x}_i - \tilde{x}_i][\hat{y}_i - F(\hat{x}_i)] + [F'(\hat{x}_i)^2 \sigma_x^2 + \sigma_y^2][\hat{x}_i - \tilde{x}_i]^2 \tag{18.18}$$

where

$$\sigma_i \equiv \sqrt{\sigma_x^2 \sigma_y^2 + \sigma_p^2 \sigma_y^2 + F'(\hat{x}_i)\sigma_p^2 \sigma_x^2}. \tag{18.19}$$

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

$$P(B_0 \cdots B_m \sigma_x \sigma_y | \hat{x}\hat{y}) \propto \frac{1}{\sigma_x \sigma_y} \exp\left\{ -\sum_{j=0}^{m} \frac{B_j^2}{2\sigma_B^2} \right\} \prod_{i=1}^{N} \left( \frac{\sigma_x \sigma_y}{\sigma_i} \right) \exp\left\{ -\frac{[\hat{y}_i - F(\hat{x}_i)]^2 [\sigma_p^2 + \sigma_x^2]}{2\sigma_i^2} \right\}. \tag{18.20}$$

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

The Text outputs files from the Errors In Variables packages consist of: "Bayes.prob.model," "BayesErrInVarsGiven.mcmc.values," "Bayes.params," "Console.log," "Bayes.accepted" and a "Bayes.Condensed.File

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.

Figure 18.2: The McMC Values File Produced By The Errors In Variables Package

```
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

                            Priors Used In This Run
Param Desc            Low         Mean         High         Sigma         Norm
Coef 0.1           -1.504E+03    0.000E+00    1.504E+03    4.513E+02   -3.626E+00
Coef 1.1           -1.504E+03    0.000E+00    1.504E+03    4.513E+02   -3.626E+00
X 1.1              -5.445E-02   -4.785E-03    4.488E-02    9.933E-03   -3.222E+00
X 101.1             9.488E-01    9.985E-01    1.048E+00    9.933E-03   -3.222E+00

                                                       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:
Parameter Description                Parameter
Std Dev in X                       4.65331682E-04
Std Dev in Y                       1.10193510E+00
Coef 0.1                           1.01151226E+01
Coef 1.1                           9.68326780E+00

The expected parameter values (mean value of the probability distributions):
Parameter Description                Mean Value     Std. Dev.     Peak Value
Sigma X                             1.09472E-03    5.34846E-04    4.65332E-04
Sigma Y                             1.09118E+00    7.44062E-02    1.10194E+00
Coef 1.0                            1.01672E+01    2.01540E-01    1.01151E+01
Coef 1.1                            9.59826E+00    3.49706E-01    9.68327E+00
```

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


[1] Rev. Thomas Bayes (1763), "An Essay Toward Solving a Problem in the Doctrine of Chances," *Philos. Trans. R. Soc. London*, **53,** pp. 370-418; reprinted in *Biometrika*, **45**, pp. 293-315 (1958), and *Facsimiles of Two Papers by Bayes,* with commentary by W. Edwards Deming, New York, Hafner, 1963.

[2] G. Larry Bretthorst (1988), "Bayesian Spectrum Analysis and Parameter Estimation," in *Lecture Notes in Statistics,* **48**, J. Berger, S. Fienberg, J. Gani, K. Krickenberg, and B. Singer (eds), Springer-Verlag, New York, New York.

[3] G. Larry Bretthorst (1990), "An Introduction to Parameter Estimation Using Bayesian Probability Theory," in *Maximum Entropy and Bayesian Methods,* Dartmouth College 1989, P. Fougère *ed.,* pp. 53-79, Kluwer Academic Publishers, Dordrecht the Netherlands.

[4] G. Larry Bretthorst (1990), "Bayesian Analysis I. Parameter Estimation Using Quadrature NMR Models" *J. Magn. Reson.*, **88**, pp. 533-551.

[5] G. Larry Bretthorst (1990), "Bayesian Analysis II. Signal Detection And Model Selection" *J. Magn. Reson.*, **88**, pp. 552-570.

[6] G. Larry Bretthorst (1990), "Bayesian Analysis III. Examples Relevant to NMR" *J. Magn. Reson.*, **88**, pp. 571-595.

[7] G. Larry Bretthorst (1991), "Bayesian Analysis. IV. Noise and Computing Time Considerations," *J. Magn. Reson.*, **93**, pp. 369-394.

[8] G. Larry Bretthorst (1992), "Bayesian Analysis. V. Amplitude Estimation for Multiple Well-Separated Sinusoids," *J. Magn. Reson.*, **98**, pp. 501-523.

[9] G. Larry Bretthorst (1992), "Estimating The Ratio Of Two Amplitudes In Nuclear Magnetic Resonance Data," in *Maximum Entropy and Bayesian Methods,* C. R. Smith *et al.* (eds.), pp. 67-77, Kluwer Academic Publishers, the Netherlands.

[10] G. Larry Bretthorst (1993), "On The Difference In Means," in Physics & Probability Essays in honor of Edwin T. Jaynes, W. T. Grandy and P. W. Milonni (eds.), pp. 177-194, Cambridge University Press, England.

[11] G. Larry Bretthorst (1996), "An Introduction To Model Selection Using Bayesian Probability Theory," in *Maximum Entropy and Bayesian Methods*, G. R. Heidbreder, ed., pp. 1-42, Kluwer Academic Publishers, Printed in the Netherlands.

[12] G. Larry Bretthorst (1999), "The Near-Irrelevance of Sampling Frequency Distributions," in *Maximum Entropy and Bayesian Methods,* W. von der Linden *et al.* (eds.), pp. 21-46, Kluwer Academic Publishers, the Netherlands.

[13] G. Larry Bretthorst (2001), "Nonuniform Sampling: Bandwidth and Aliasing," in *Maximum Entropy and Bayesian Methods in Science and Engineering,* Joshua Rychert, Gary Erickson and C. Ray Smith *eds.*, pp. 1-28, American Institute of Physics, USA.

[14] G. Larry Bretthorst, Christopher D. Kroenke, and Jeffrey J. Neil (2004), "Characterizing Water Diffusion In Fixed Baboon Brain," in Bayesian Inference And Maximum Entropy Methods In Science And Engineering, Rainer Fischer, Roland Preuss and Udo von Toussaint *eds.*, AIP conference Proceedings, **735**, pp. 3-15.

[15] G. Larry Bretthorst, William C. Hutton, Joel R. Garbow, and Joseph J.H. Ackerman (2005), "Exponential parameter estimation (in NMR) using Bayesian probability theory," *Concepts in Magnetic Resonance*, 27A, Issue 2, pp. 55-63.

[16] G. Larry Bretthorst, William C. Hutton, Joel R. Garbow, and Joseph J. H. Ackerman (2005), "Exponential model selection (in NMR) using Bayesian probability theory," *Concepts in Magnetic Resonance*, 27A, Issue 2, pp. 64-72.

[17] G. Larry Bretthorst, William C. Hutton, Joel R. Garbow, and Joseph J.H. Ackerman (2005), "How accurately can parameters from exponential models be estimated? A Bayesian view," *Concepts in Magnetic Resonance*, 27A, Issue 2, pp. 73-83.

[18] G. Larry Bretthorst, W. C. Hutton, J. R. Garbow, and Joseph J. H. Ackerman (2008), "High Dynamic Range MRS Time-Domain Signal Analysis," *Magn. Reson. in Med.*, **62**, pp. 1026-1035.

[19] V. Chandramouli, K. Ekberg, W. C. Schumann, S. C. Kalhan, J. Wahren, and B. R. Landau (1997), "Quantifying gluconeogenesis during fasting," *American Journal of Physiology*, **273**, pp. H1209-H1215.

[20] R. T. Cox (1961), "The Algebra of Probable Inference," Johns Hopkins Univ. Press, Baltimore.

[21] André d'Avignon, G. Larry Bretthorst, Marlyn Emerson Holtzer, and Alfred Holtzer (1998), "Site-Specific Thermodynamics and Kinetics of a Coiled-Coil Transition by Spin Inversion Transfer NMR," *Biophysical Journal*, **74**, pp. 3190-3197.

[22] André d'Avignon, G. Larry Bretthorst, Marlyn Emerson Holtzer, and Alfred Holtzer (1999), "Thermodynamics and Kinetics of a Folded-Folded Transition at Valine-9 of a GCN4-Like Leucine Zipper," *Biophysical Journal*, **76**, pp. 2752-2759.

[23] David Freedman, and Persi Diaconis (1981), "On the histogram as a density estimator: $L_2$ theory," *Zeitschrift f¨r Wahrscheinlichkeitstheorie und verwandte Gebiete*, **57**, 4, pp. 453-476.

[24] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter (1996), "Markov Chain Monte Carlo in Practice," Chapman & Hall, London.

[25] Paul M. Goggans, and Ying Chi (2004), "Using Thermodynamic Integration to Calculate the Posterior Probability in Bayesian Model Selection Problems," in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 23rd International Workshop*, **707**, pp. 59-66.

[26] Marlyn Emerson Holtzer, G. Larry Bretthorst, D. André d'Avignon, Ruth Hogue Angelette, Lisa Mints, and Alfred Holtzer (2001), "Temperature Dependence of the Folding and Unfolding Kinetics of the GCN4 Leucine Lipper via 13C alpha-NMR," *Biophysical Journal*, **80**, pp. 939-951.

[27] E. T. Jaynes (1968), "Prior Probabilities," *IEEE Transactions on Systems Science and Cybernetics*, SSC-4, pp. 227-241; reprinted in [30].

[28] E. T. Jaynes (1978), "Where Do We Stand On Maximum Entropy?" in *The Maximum Entropy Formalism*, R. D. Levine and M. Tribus *Eds.*, pp. 15-118, Cambridge: MIT Press, Reprinted in [30].

[29] E. T. Jaynes (1980), "Marginalization and Prior Probabilities," in *Bayesian Analysis in Econometrics and Statistics*, A. Zellner *ed.,* North-Holland Publishing Company, Amsterdam; reprinted in [30].

[30] E. T. Jaynes (1983), "Papers on Probability, Statistics and Statistical Physics," a reprint collection, D. Reidel, Dordrecht the Netherlands; second edition Kluwer Academic Publishers, Dordrecht the Netherlands, 1989.

[31] E. T. Jaynes (1957), "How Does the Brain do Plausible Reasoning?" unpublished Stanford University Microwave Laboratory Report No. 421; reprinted in *Maximum-Entropy and Bayesian Methods in Science and Engineering* **1,** pp. 1-24, G. J. Erickson and C. R. Smith *Eds.,* 1988.

[32] E. T. Jaynes (2003), "Probability Theory—The Logic of Science," edited by G. Larry Bretthorst, Cambridge University Press, Cambridge UK.

[33] Sir Harold Jeffreys (1939), "Theory of Probability," Oxford Univ. Press, London; Later editions, 1948, 1961.

[34] John G. Jones, Michael A. Solomon, Suzanne M. Cole, A. Dean Sherry, and Craig R. Malloy (2001) "An integrated $^{2}$H and $^{13}$C NMR study of gluconeogenesis and TCA cycle flux in humans," *American Journal of Physiology, Endocrinology, and Metabolism,* **281,** pp. H848-H856.

[35] John Kotyk, N. G. Hoffman, W. C. Hutton, G. Larry Bretthorst, and J. J. H. Ackerman (1992), "Comparison of Fourier and Bayesian Analysis of NMR Signals. I. Well-Separated Resonances (The Single-Frequency Case)," *J. Magn. Reson.*, **98**, pp. 483–500.

[36] Pierre Simon Laplace (1814), "A Philosophical Essay on Probabilities," John Wiley & Sons, London, Chapman & Hall, Limited 1902. Translated from the 6th edition by F. W. Truscott and F. L. Emory.

[37] N. Lartillot, and H. Philippe (2006), "Computing Bayes Factors Using Thermodynamic Integration," *Systematic Biology*, **55** (2), pp. 195-207.

[38] D. Le Bihan, and E. Breton (1985), "Imagerie de diffusion in-vivo par rsonance," Comptes rendus de l'Acadmie des Sciences (Paris), **301** (15), pp. 1109-1112.

[39] N. R. Lomb (1976), "Least-Squares Frequency Analysis of Unevenly Spaced Data," *Astrophysical and Space Science*, **39**, pp. 447-462.

[40] T. J. Loredo (1990), "From Laplace To SN 1987A: Bayesian Inference In Astrophysics," in *Maximum Entropy and Bayesian Methods,* P. F. Fougere (ed), Kluwer Academic Publishers, Dordrecht, The Netherlands.

[41] Craig R. Malloy, A. Dean Sherry, and Mark Jeffrey (1988), "Evaluation of Carbon Flux and Substrate Selection through Alternate Pathways Involving the Citric Acid Cycle of the Heart by 13C NMR Spectroscopy," *Journal of Biological Chemistry*, **263** (15), pp. 6964-6971.

[42] Craig R. Malloy, Dean Sherry, and Mark Jeffrey (1990), "Analysis of tricarboxylic acid cycle of the heart using $^{13}$C isotope isomers," *American Journal of Physiology*, **259**, pp. H987-H995.

[43] Lawrence R. Mead and Nikos Papanicolaou, "Maximum entropy in the problem of moments," *J. Math. Phys.* **25**, 2404–2417 (1984).

[44] K. Merboldt, Wolfgang Hanicke, and Jens Frahm (1969), "Self-diffusion NMR imaging using stimulated echoes," Journal of Magnetic Resonance, **64** (3), pp. 479-486.

[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 Americain Institute of Physics and if you do not have access to Science Sitations you many not be able to retrieve this paper.

[46] Radford M. Neal (1993), "Probabilistic Inference Using Markov Chain Monte Carlo Methods," technical report CRG-TR-93-1, Dept. of Computer Science, University of Toronto.

[47] Jeffrey J. Neil, and G. Larry Bretthorst (1993), "On the Use of Bayesian Probability Theory for Analysis of Exponential Decay Data: An Example Taken from Intravoxel Incoherent Motion Experiments," *Magn. Reson. in Med.*, **29**, pp. 642–647.

[48] H. Nyquist (1924), "Certain Factors Affecting Telegraph Speed," *Bell System Technical Journal*, **3**, pp. 324-346.

[49] H. Nyquist (1928), "Certain Topics in Telegraph Transmission Theory," *Transactions AIEE*, **3**, pp. 617-644.

[50] William H. Press, Saul A. Teukolsky, William T. Vetterling and Brian P. Flannery (1992), "Numerical Recipes The Art of Scientific Computing Second Edition," Cambridge University Press, Cambridge UK.

[51] Emanuel Parzen (1962), "On Estimation of a Probability Density Function and Mode," *Annals of Mathematical Statistics* **33**, 1065–1076

[52] Karl Pearson (1895), "Contributions to the Mathematical Theory of Evolution. II. Skew Variation in Homogeneous Material," *Phil. Trans. R. Soc. A* **186**, 343–326.

[53] Murray Rosenblatt, "Remarks on Some Nonparametric Estimates of a Density Function," *Annals of Mathematical Statistics* **27**, 832–837 (1956).

[54] Jeffery D. Scargle (1981), "Studies in Astronomical Time Series Analysis I. Random Process In The Time Domain," *Astrophysical Journal Supplement Series*, **45**, pp. 1-71.

[55] Jeffery D. Scargle (1982), "Studies in Astronomical Time Series Analysis II. Statistical Aspects of Spectral Analysis of Unevenly Sampled Data," *Astrophysical Journal*, **263**, pp. 835-853.

[56] Jeffery D. Scargle (1989), "Studies in Astronomical Time Series Analysis. III. Fourier Transforms, Autocorrelation Functions, and Cross-correlation Functions of Unevenly Spaced Data," *Astrophysical Journal*, **343**, pp. 874-887.

[57] Arthur Schuster (1905), "The Periodogram and its Optical Analogy," *Proceedings of the Royal Society of London*, **77**, p. 136-140.

[58] Claude E. Shannon (1948), "A Mathematical Theory of Communication," *Bell Syst. Tech. J.*, **27**, pp. 379-423.

[59] John E. Shore, and Rodney W. Johnson (1981), "Properties of cross-entropy minimization," *IEEE Trans. on Information Theory*, **IT-27,** No. 4, pp. 472-482.

[60] John E. Shore and Rodney W. Johnson (1980), "Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy," *IEEE Trans. on Information Theory*, **IT-26** (1), pp. 26-37.

[61] Devinderjit Sivia, and John Skilling (2006), "Data Analysis: A Bayesian Tutorial," Oxford University Press, USA.

[62] Edward O. Stejskal and Tanner, J. E. (1965), "Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient." Journal of Chemical Physics, **42** (1), pp. 288-292.

[63] D. G. Taylor and Bushell, M. C. (1985), "The spatial mapping of translational diffusion coefficients by the NMR imaging technique," Physics in Medicine and Biology, **30** (4), pp. 345-349.

[64] Myron Tribus (1969), "Rational Descriptions, Decisions and Designs," Pergamon Press, Oxford.

[65] P. M. Woodward (1953), "Probability and Information Theory, with Applications to Radar," McGraw-Hill, N. Y. Second edition (1987); R. E. Krieger Pub. Co., Malabar, Florida.

[66] Arnold Zellner (1971), "An Introduction to Bayesian Inference in Econometrics," John Wiley and Sons, New York.