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 20

Enter Ascii Model

The Enter Ascii Model Package allows you to enter a model of your own and then use Bayesian probability theory to analyze that model.[1] To use this package you do not have to have either Fortran or C installed on your server. However, If you do not have either Fortran or C installed, the only models you will be able to use are the system models. Consequently, installing both Fortran and C is strongly recommended. The interface to this package is shown in Fig. 20.1 To use this package, you must do the following:

Select the "Enter Ascii Model" package from the Package menu.

Load a Fortran or C model using the "System" or "User" buttons in the "Load And Build Model" widget group.

Load one or more Ascii data sets using the Files menu. When a data set is successfully loaded the data is plotted in the Ascii Data viewer. The format of the Ascii data that must be loaded is dependent on the model. Usually the data are two column Ascii, however, in general this package takes multicolumn Ascii data with a multicolumn abscissa. See Appendix A for a detailed description of the Ascii data files used by the Bayesian Analysis software.

Build the model using the "Build" button.

Check the Analysis Options/Find Outliers box if you suspect outliers are present in the data.

Review the prior probabilities for the loaded model using the Prior Viewer.

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.

[1] I would like to build a system library of predefined models. If you have models that you think would be of general use, I would like to hear from you. To have one of your models included, I would need the source code, the parameter file, a brief description of the model equations and data requirements.

Figure 20.1: Enter Ascii Model Package Interface

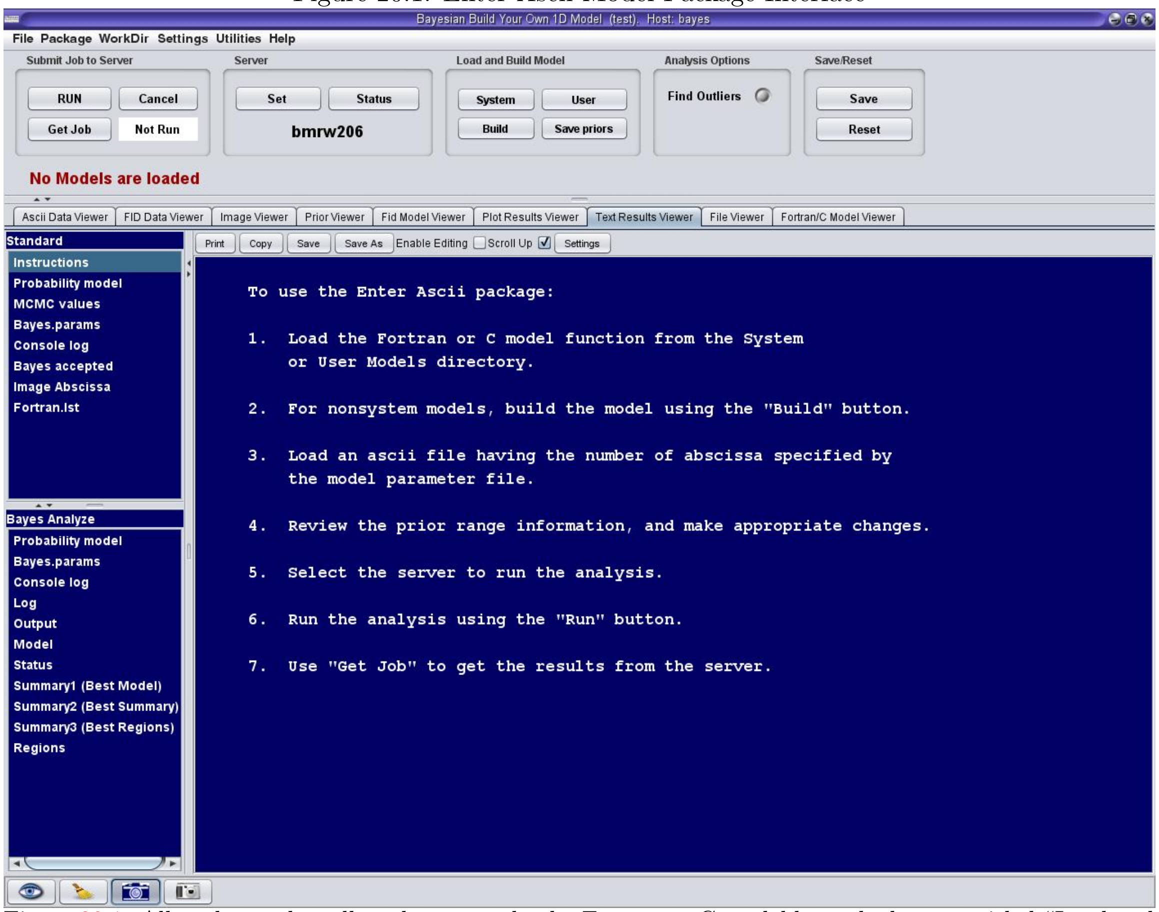

Figure 20.1: All packages that allow the user to load a Fortran or C model have the buttons titled "Load and Build Model." These buttons allow you to load a model from either the system directory or from you user directory. They allow you to compile a model and save the current prior settings. Additionally, using the "Fortran/C Model Viewer" you can edit, modify and create models, see Appendix E for more on creating Fortran and C models.

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.

20.1 The Bayesian Calculation

The calculation done by Enter Ascii Model Package is a parameter estimation calculation. However, there are two distinct functional forms for the model that are used: one using marginalization over the amplitudes, and one that does not. The model function that does not use marginalization is given by:

$$d_j(t_i) = U_j(t_i, r_1, r_2, \ldots) + n_j(t_i) \tag{20.1}$$

where $d_j(t_i)$ represents a data item in the jth data set at abscissa value t_i and t_i may be vector valued. $U_j(t_i, r_1, r_2, \ldots)$ is the model function. r_j are the various parameters appearing in the model including any amplitudes that may be present, and $n_j(t_i)$ represents noise in the jth data set at abscissa t_i. Because, this model does not marginalize out the amplitudes, it is possible to restrict the amplitudes ranges using the prior probabilities.

The other model used by this package assumes the amplitudes are to be marginalized from the joint posterior probability for the parameters. The model equation that uses marginalization is similar

$$d_j(t_i) = \sum_{\ell=1}^{m} A_{jk} G_{j\ell}(t_i, r_1, r_2, \ldots) + n_j(t_i) \tag{20.2}$$

where the amplitudes are labeled A_{jk} meaning the kth amplitude in the jth data set, the sum is over all of the amplitudes in the model, $G_{j\ell}(t_i, r_1, r_2, \ldots)$ is the ℓth model function in the jth data set evaluated at abscissa t_i and this model equation implicitly assumes that each data set contains same number of amplitudes.

20.1.1 The Bayesian Calculations Using Eq. (20.1)

To compute the marginal posterior probability for each parameter using Eq. (20.1), a Markov chain Monte Carlo simulation is run targeting the joint posterior probability for all of the parameters. This joint posterior probability is represented symbolically by $P(r_1 r_2 \ldots | DI)$. The joint posterior probability for the parameters is factored using Bayes' theorem to obtain

$$P(r_1 r_2 \ldots | DI) \propto P(r_1 r_2 \ldots | I) P(D | r_1 r_2 \ldots I) \tag{20.3}$$

where D stands for all of the data in all of the data sets, $P(r_1 r_2 \ldots \sigma_1 \ldots | I)$, is factored into independent prior probabilities for each parameter:

$$P(r_1 r_2 \ldots | DI) \propto \left[\prod_{l=1}^{m} P(r_l | I) \right] P(D | r_1 r_2 \ldots I) \tag{20.4}$$

where m is the total number of parameters in the model, The priors, $P(r_j | I)$, are specified in the input parameter file that describes the model. These prior are either the defaults, if you loaded the model from the system directory, or they are the priors set using the interface. Because we don't know the functional form of these priors, we are going to leave them in symbolic form. Factoring the direct probability for the data into an independent direct probability for each data set, one obtains

$$P(r_1 r_2 \ldots |DI) \propto \left[\prod_{l=1}^{m} P(r_l|I)\right] \prod_{j=1}^{n} P(D_j|r_1 r_2 \ldots I) \tag{20.5}$$

as the joint posterior probability for the parameters. The direct probability for the data is a marginal likelihood, because the standard deviation of the noise prior probability is not present. Introducing a standard deviation of the noise prior probability, σ_j, for each data set, and using the rules of probability theory to remove these parameters, one obtains:

$$P(r_1 r_2 \ldots |DI) \propto \left[\prod_{l=1}^{m} P(r_l|I)\right] \prod_{j=1}^{n} \left[\int P(\sigma_j|I) P(D_j|\sigma_j r_1 r_2 \ldots I) d\sigma_j\right]. \tag{20.6}$$

We have reached the point in this calculation where one has no other choice than to assign probabilities to represent each of these probabilities and then to perform the indicated integrals. Assign a Jeffreys' prior to the prior probability for the noise standard deviation:

$$P(\sigma_j|I) \propto \frac{1}{\sigma_j}, \tag{20.7}$$

and assigning the direct probability for the data using a Gaussian of standard deviation σ_j one obtains

$$P(r_1 r_2 \ldots |DI) \propto \left[\prod_{l=1}^{m} P(r_l|I)\right] \prod_{j=1}^{n} \left[\int \sigma_j^{-(N_j+1)} \exp\left\{-\frac{Q_j(t_i, r_1, r_2, \ldots)}{2\sigma_j^2}\right\} d\sigma_j\right]. \tag{20.8}$$

as the joint posterior probability for the parameters, where $Q_j(t_i, r_1, r_2, \ldots)$ is given by:

$$Q_j(t_i, r_1, r_2, \ldots) \equiv \sum_{i=1}^{N_j} \left[d_j(t_i) - U_j(t_i, r_1, r_2, \ldots)\right]^2, \tag{20.9}$$

and is the total squared residual and is essentially χ^2. Evaluating the integral over the standard deviation of the noise, one obtains

$$P(r_1 r_2 \ldots |DI) \propto \left[\prod_{l=1}^{m} P(r_l|I)\right] \prod_{j=1}^{n} \left[\frac{Q_j(t_i, r_1, r_2, \ldots)}{2}\right]^{-\frac{N_j}{2}} \tag{20.10}$$

as the joint posterior probability for the parameters, where we have dropped a number of constants that make no difference in this parameter estimation problem.

20.1.2 The Bayesian Calculations Using Eq. (20.2)

To compute the marginal posterior probability for each parameter using Eq. (20.2), a Markov chain Monte Carlo simulation is run targeting the joint posterior probability for all of the nonlinear parameters. In this context, nonlinear means all of the parameters appearing in the model in a nonlinear fashion, i.e., all of the parameters except the amplitudes. This joint posterior probability is represented symbolically by $P(r_1r_2\ldots|DI)$. The joint posterior probability for the nonlinear parameters is factored using Bayes' theorem to obtain

$$P(r_1r_2\ldots|DI) \propto P(r_1r_2\ldots|I)P(D|r_1r_2\ldots I) \tag{20.11}$$

where D stands for all of the data in all of the data sets, the prior probability for all of the nonlinear parameters is represented by, $P(r_1r_2\ldots\sigma_1\ldots|I)$, and we will factor it into independent prior probabilities for each parameter. Consequently, the joint posterior probability for all of the nonlinear parameters is given by

$$P(r_1r_2\ldots|DI) \propto \left[\prod_{j=1}^{m} P(r_j|I)\right] P(D|r_1r_2\ldots I) \tag{20.12}$$

where m is the total number of nonlinear parameters in the model, The priors, $P(r_j|I)$, are specified in the input parameter file that describes the model. These prior are either the defaults, if you loaded the model from the system directory, or they are the priors set using the interface. Because we don't know the functional form of these priors, we are going to leave them in symbolic form. Factoring the direct probability for the data into an independent direct probability for each data set, one obtains

$$P(r_1r_2\ldots|DI) \propto \left[\prod_{l=1}^{m} P(r_l|I)\right] \prod_{j=1}^{n} P(D_j|r_1r_2\ldots I) \tag{20.13}$$

as the joint posterior probability for the nonlinear parameters.

The direct probability for the data is a marginal likelihood, because neither the standard deviation of the noise prior probability nor the amplitudes are present. To proceed with this calculation, these parameters must be reintroduced into the joint posterior probability for the nonlinear parameters. Representing the standard deviation of the noise prior probability for each data set as σ_j and $\{A\}_j$ as all of the amplitudes in the jth data set, one obtains

$$P(r_1r_2\ldots|DI) \propto \left[\prod_{l=1}^{m} P(r_l|I)\right] \prod_{j=1}^{n} \int P(D_j\sigma_j\{A\}_j|r_1r_2\ldots I)d\sigma_j d\{A\}_j \tag{20.14}$$

as the joint posterior probability for the parameters. Factoring the right-hand side of this equation, one obtains

$$P(r_1r_2\ldots|DI) \propto \left[\prod_{l=1}^{m} P(r_l|I)\right] \prod_{j=1}^{n} \int P(\sigma_j|I)P(\{A\}_j|I)P(D_j|\sigma_j\{A\}_j r_1r_2\ldots I)d\sigma_j d\{A\}_j \tag{20.15}$$

where $P(\sigma_j|I)$ is the prior probability for the standard deviation of the noise prior probability in the jth data set. Similarly, $P(\{A\}_j|I)$ is the joint prior probability for the amplitudes in the jth data set. If we assume the amplitudes are logically independent, then the joint prior probability for the amplitudes, $P(\{A\}_j|I)$, can be factored into a product of prior probabilities for each amplitude:

$$P(r_1r_2\ldots|DI) \propto \left[\prod_{l=1}^{m} P(r_l|I)\right] \prod_{j=1}^{n} \int P(\sigma_j|I) \left[\prod_{k=1}^{\nu} P(A_{jk}|I)\right] P(D_j|\sigma_j\{A\}_j r_1r_2\ldots I)d\sigma_j d\{A\}_j \tag{20.16}$$

where $P(A_{jk}|I)$ is the prior probability for the kth amplitude in the jth data set, and ν is the number of data sets. We will assign a zero-mean Gaussian prior probability for each amplitudes. This Gaussian prior probability is given by

$$P(A_{jk}|I) \propto \left(\frac{2\pi\sigma_j^2}{\gamma^2 g_{jkk}}\right)^{-\frac{1}{2}} \exp\left\{-\frac{A_{jk}^2\gamma^2 g_{jkk}}{2\sigma_j^2}\right\} \tag{20.17}$$

where

$$g_{jk\ell} \equiv \sum_{i=1}^{N_j} G_{jk}(t_i)G_{j\ell}(t_i) \tag{20.18}$$

and γ is used to control the width of this prior probability. The reason for this particular functional form is that it allows one to evaluate the integrals over the amplitudes in a concise functional form that aids in doing the numerical calculations. Substituting the prior probability for the amplitudes, Eq. (20.17), into the joint posterior probability for the parameters, Eq. 20.16,

$$\begin{aligned} P(r_1r_2\ldots|DI) \quad &\propto \quad \left[\prod_{l=1}^{m} P(r_l|I)\right] \\ &\times \quad \prod_{j=1}^{n}\Bigg[\int \frac{1}{\sigma_j}\left(\frac{2\pi\sigma_j^2}{\gamma^2 g_{j11}\cdots g_{j\nu\nu}}\right)^{-\frac{\nu}{2}} \\ &\times \quad \exp\left\{-\sum_{k=1}^{\nu}\frac{A_{jk}^2\gamma^2 g_{jkk}}{2\sigma_j^2}\right\} \\ &\times \quad P(D_j|\sigma_j\{A\}_j r_1 r_1 \ldots I)d\sigma_j d\{A\}_j\Bigg] \end{aligned} \tag{20.19}$$

and assigning a Gaussian for the direct probability for the data, $P(D_j|\sigma_j\{A\}_j r_1 r_2 \ldots I)$, one obtains:

$$\begin{aligned} P(r_1r_2\ldots|DI) \quad &\propto \quad \left[\prod_{l=1}^{m} P(r_l|I)\right] \\ &\times \quad \prod_{j=1}^{n}\Bigg[\int \frac{1}{\sigma_j}\left(\frac{2\pi\sigma_j^2}{\gamma^2 g_{j11}\cdots g_{j\nu\nu}}\right)^{-\frac{\nu}{2}} \\ &\times \quad \exp\left\{-\sum_{k=1}^{\nu}\frac{A_{jk}^2\gamma^2 g_{jkk}}{2\sigma_j^2}\right\} \\ &\times \quad \left(2\pi\sigma_j^2\right)^{-\frac{N_j}{2}} \exp\left\{-\sum_{i=1}^{N_j}\frac{\left(d_{ji}-\sum_{k=1}^{\mu} A_{jk}G_{jk}(t_i,r_1\cdots)\right)^2}{2\sigma_j}\right\} d\sigma_j d\{A\}_j\Bigg]. \end{aligned} \tag{20.20}$$

After evaluating the integrals over the amplitudes, one obtains

$$P(r_1r_2\ldots|DI) \propto \left[\prod_{l=1}^{m} P(r_l|I)\right] \prod_{j=1}^{n}\left[\frac{\gamma^2}{g_{j11}\cdots g_{j\nu\nu}}\right] |g_{jkl}|^{-\frac{1}{2}} \left[\frac{Q_j(r_1r_2\ldots)}{2}\right]^{-\frac{N_j}{2}} \tag{20.21}$$

with

$$Q_j(r_1 r_2 \ldots) \equiv \sum_{i=1}^{N_j} \left[d_{ji} - \sum_{\ell=1}^{\nu} \hat{A}_{j\ell} G_{j\ell}(t_i r_1 r_2 \ldots) \right]^2, \tag{20.22}$$

$|g_{jkl}|$ is the magnitude of the determinate of the g_{jkl} matrix defined in Eq. (20.18) and the amplitudes $\hat{a}_{j\ell}$ are given by the solution to

$$\sum_{k=1}^{\nu} g_{jk\ell} \hat{A}_{j\ell} = T_{j\ell} \tag{20.23}$$

with the right-hand side of this equation given by:

$$T_{j\ell} = \sum_{i=1}^{N_j} d_j(t_i) G_{j\ell}(t_i). \tag{20.24}$$

See [2], and [11] for more on how the integrals over the amplitudes are evaluated. Equation 20.21 is the joint posterior probability for the nonlinear parameters that is targeted by the Markov chain Monte Carlo simulations. These simulations only vary the nonlinear parameters, the amplitudes simply do not appear in the posterior probability. However, the amplitudes are output from the simulation. The output amplitudes are given by Eq. (20.23). Because these amplitudes are estimated for each value of the nonlinear parameters, there is as many samples from the distributions of the amplitudes as there is for each of the nonlinear parameters. Consequently, the model that use marginalization do output density functions for the amplitudes.

20.2 Outputs Form The Enter Ascii Model Package

The Text outputs files from the Enter Ascii Model packages consist of: "Bayes.prob.model," "BayesModelAscii.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 each parameter in the currently loaded Fortran/C model. These output probability density functions are named

`ModelFileName.ParamName`

where `ModelFileName` is the name of the currently loaded model. For example, if you have a model named `MyFunnyExp` model, and it has a decay rate named `FunnyRate` the output file containing the posterior probability for `FunnyRate` would be named:

`MyFunnyExp.FunnyRate.`

This naming convention also applies to derived parameters. So, if in addition to generating samples for `FunnyRate`, you also generated samples from a derived inverse decay rate, which was called `FunnyDecayTime` then there would also be an output file named

```
MyFunnyExp.FunnyDecayTime
```

containing the posterior probability for the decay time. For more on writing Ascii models in either Fortran or C, see Appendix E.

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.