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 17

Unknown Polynomial Order

The Unknown Polynomial Order package fits polynomials to two column Ascii data when both the order of the polynomial and the polynomial coefficients are unknown. The interface to this package is shown in Figure17.1. This interface differs from most others in one respect, there are no parameter ranges to enter, so use of the interface is particularly simple. To use this package, you must do the following:

Select the Polynomial Models package from the Package menu. When selected this menu will bring up the “Given” and “Unknown” polynomial model interface.

Check the “Unknown Order” box to select the Unknown Polynomial Order package. When this check box is activated the “Set Order” widget becomes inactive. This is illustrated in Fig. 17.1 where the “Unknown Order” has been checked, and, consequently, the“Set Order” widget has been grayed out.

Load one two column Ascii data sets. The data may be loaded using the Files menu. You can also load an arrayed Fid and then use a single cursor to mark the center of a peak and use the “Get Peak” button on the bottom right of the Fid viewer. Finally, the “Files/Load Ascii/Bayes Analyze” button can be used to load an Ascii data set from the amplitudes estimated by Bayes Analyze. When a data set is successfully loaded the data is plotted in the Ascii Data viewer. This package does not allow you to run with multiple data sets. If you attempt to do so, you will be prompted to remove all but a single file.

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.

Figure 17.1: Unknown Polynomial Order Package Interface

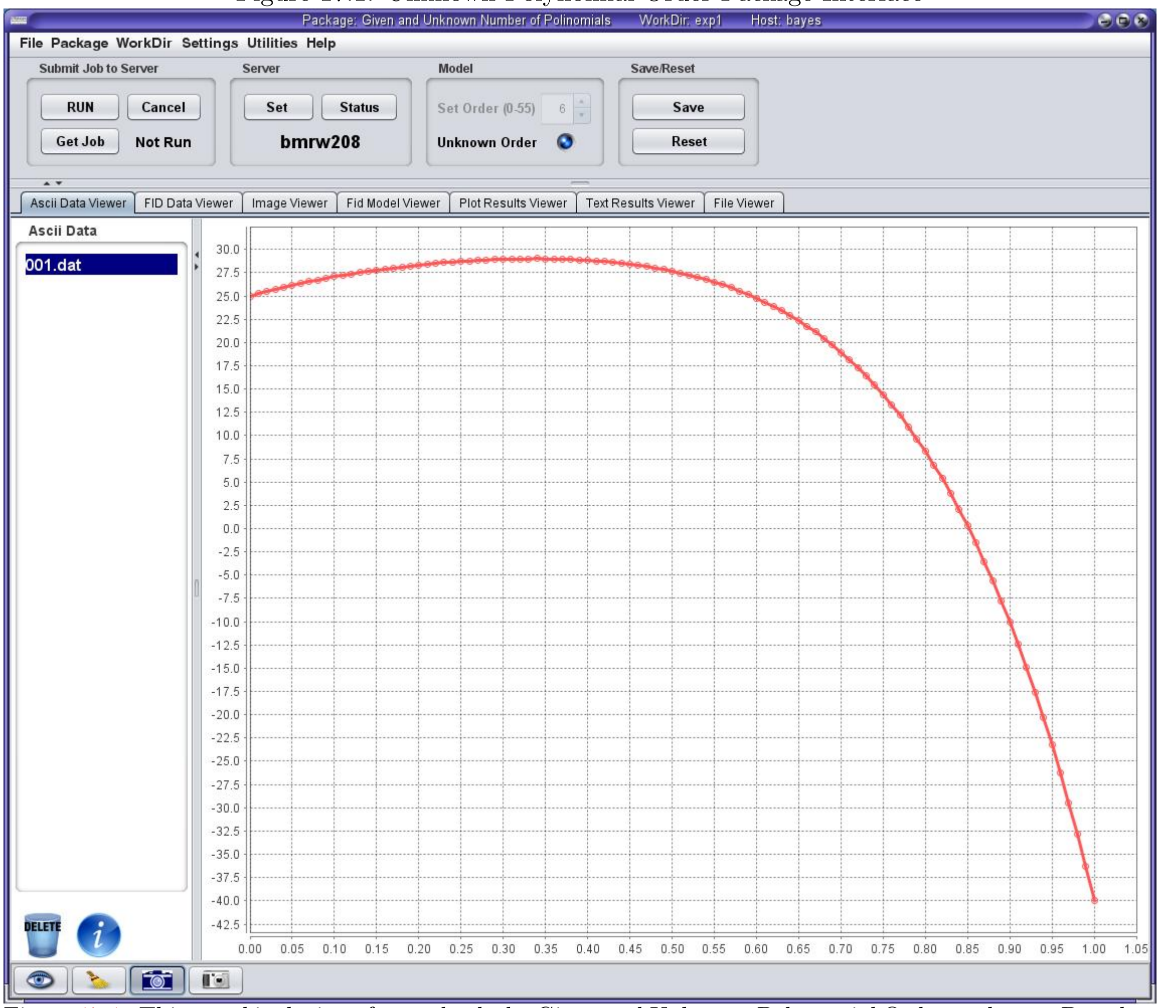

Figure 17.1: This panel is the interface to both the Given and Unknown Polynomial Order packages. Data has already been loaded. Note that in this example, the “Unknown Order” check box has been set. Consequently, the “Set Order” spinner has been deactivated. Because of the way this calculation is done very high orders are possible and numerically stable. However, the high orders, above 40, require very high signal-to-noise and even then roundoff errors degrade the accuracy to 4 or 5 decimal places.

17.1 Bayesian Calculations

The Unknown Polynomial Order model is just that, its a model in which a polynomial is fit to the data:

$$d_i = \sum_{j=0}^{m} A_j t_i^j + n_i \tag{17.1}$$

where A_j is the amplitude of the jth polynomial, m is the unknown order of the polynomial expansion, and n_i represents noise in the data. As in Chapter 16, we introduce a change of function and a change of variables and we refer the reader to that chapter for a discussion of the change of function and variables. The change of function is to orthonormal polynomials designated by $G_j(t_i)$, so the expansion given in Eq. (17.1) becomes:

$$d_i = \sum_{j=0}^{m} B_j G_j(t_i) + n_i \tag{17.2}$$

where B_j are the amplitudes in the orthonormal expansion. Note this change of function and variables is an identify, so

$$\sum_{j=0}^{m} A_j t_i^j = \sum_{j=0}^{m} B_j G_j(t_i). \tag{17.3}$$

The Bayesian calculation is implemented using Markov chain Monte Carlo with simulated annealing to draw samples from the joint posterior probability for the parameters, $P(mB_0B_1 \ldots B_m|DI)$. From these samples we then compute the marginal posterior probabilities for the amplitudes and the polynomial order. The joint posterior probability for the parameters is computed by application of Bayes' theorem

$$P(mB_0B_1 \ldots B_m|DI) \propto P(mB_0B_1 \ldots B_m|I)P(D|mB_0B_1 \ldots B_mI) \tag{17.4}$$

where $P(mB_0B_1 \ldots B_m|I)$ is the joint prior probability for the amplitudes and the polynomial order, and $P(D|mB_0B_1 \ldots B_mI)$ is the direct probability for the data given the parameters and the polynomial order. We factor the joint prior probability for the parameters, $P(mB_0B_1 \ldots B_m|I)$, into a series of independent prior probabilities:

$$P(mB_0B_1 \ldots B_m|I) = P(m|I) \prod_{j=0}^{m} P(B_j|I) \tag{17.5}$$

where $P(m|I)$ is the prior probability for the polynomial order and $P(B_j|I)$ is the prior probability for the jth amplitude. Substituting, Eq. (17.5) into Eq. (17.4) one obtains

$$P(mB_0B_1 \ldots B_m|DI) \propto P(m|I) \left[\prod_{j=0}^{m} P(B_j|I)\right] P(D|mB_0B_1 \ldots B_mI) \tag{17.6}$$

as the joint posterior probability for all of the parameters, including the polynomial order.

17.1.1 Assigning Priors

Before we assign the likelihood, we are going to assign the prior probability for the polynomial order, $P(m|I)$, and the prior probability for the amplitudes, the $P(B_j|I)$. The prior probability for the polynomial order was assigned as a discrete Gaussian with lower bound zero, an upper bound of 50, a mean value of 5, and a standard deviation of 10:

$$P(m|I) = \begin{cases} \frac{1}{C}\exp\left\{-\frac{(5-m)^2}{2\times 10^2}\right\} & m \in \{0,1,\ldots,50\} \\ 0 & \text{otherwise} \end{cases}, \tag{17.7}$$

with the normalization constant C set so that the sum over the total models is one:

$$C = \sum_{m=0}^{50} \exp\left\{-\frac{(5-m)^2}{2\times 10^2}\right\}. \tag{17.8}$$

Which expresses a belief that the polynomial order should be small, we think it very unlikely that the polynomial order would be as large as 50; but we think it reasonably possible for the order to be in the tens or twenties.

The prior probabilities for the amplitudes will be assigned exactly the same way they were when the did the Given Polynomial Order package, Chapter 16. That prior was given by Eq. 16.15 and we simply use that prior here:

$$P(B_j|I) = \left(2\pi\delta^2\right)^{-\frac{1}{2}} \exp\left\{-\frac{(T_j - B_j)^2}{2\delta^2}\right\} \tag{17.9}$$

where δ is the standard deviation of this prior probability and indicates how strongly we believe the expected amplitude is T_j. How we set δ is explained shortly. The expected amplitude, T_j, is given by

$$T_j \equiv \sum_{i=1}^{N} d_i G_j(t_i). \tag{17.10}$$

In Chapter 16 when this prior was used, we knew the order of the expansion polynomial and thus could determine the mean-square residual. We could use the mean-square residual to set δ to a value much wider than any amplitude supported by the data. So this prior probability just acted as a guide to the Markov chain Monte Carlo simulations. Here setting δ is harder because we don't know which model to use. However, we still want to set δ to a value that will guide the Markov chain Monte Carlo simulations but not make δ so large that the simulations never converge. In Chapter 16 we noted that because these amplitudes appear in the model in a linear fashion, we could solve the problem analytically, we don't have to use Markov chain Monte Carlo at all. The only reason for using a Markov chain is for consistency with the other packages in our Bayesian Analysis software. However, there is nothing to stop us from computing $P(m|DI)$ analytically and using that to set δ. Without going into the details of this calculation, the posterior probability for polynomial of order m is given by:

$$P(m|DI) = P(m|I)\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{N-m}{2}\right)\left[\overline{h_m^2}\right]^{-\frac{m}{2}}\left[\frac{\overline{d^2}-\overline{h_m^2}}{2}\right]^{-\frac{N-m}{2}} \tag{17.11}$$

where $P(m|I)$ is given by Eq. (17.7) and the prior probability for the amplitudes was assigned as a normalized unbounded Gaussian with mean zero and standard deviation γ; which was marginalized out of the problem using a series of approximations given in [2] The quantity, $\overline{d^2} - \overline{h_m^2}$, is the total-squared residual given a polynomial of order m. The sufficient statistic, $\overline{h_m^2}$, is the total-squared projection of the data onto the given polynomial model and is defined as

$$\overline{h_m^2} \equiv \sum_{k=0}^{m} T_k^2. \tag{17.12}$$

The expected standard deviation of the noise independent of the model order is given by:

$$\sqrt{\langle\sigma^2\rangle} = \sum_{m=0}^{Max} P(m|DI)\sqrt{\frac{\overline{d^2} - \overline{h_m^2}}{N}}. \tag{17.13}$$

Finally, δ was set to

$$\delta = 10\sqrt{\langle\sigma^2\rangle}. \tag{17.14}$$

While rather complicated, this calculation was used for two reasons: I needed an estimate of the standard deviation of the noise, which this gives by simple straightforward calculation; and I needed a way to determine where the maximum of the posterior probability for the polynomial order was. I needed this maximum to determine where to center the distribution of simulations that is printed out while this program is running. The problem is illustrated in Fig. 17.2. The maximum order of the polynomial is 50, but there is only room to print out 10 of these probabilities. So the output window must be shifted to cover the maximum posterior probability for the polynomial order. To do that, I needed to know where the maximum was. This calculation solved both of these problems at one time and it did so using Bayesian probability theory. See my book, [2], for more on this calculation and where each of these terms comes from.

17.1.2 Assigning The Joint Posterior Probability

Having assigned the prior probabilities, we can now proceed with assigning the joint posterior probability for the parameters, Eq. (17.6). First, however, we assign the direct probability for the data. The direct probability, $P(D|mB_0B_1\ldots B_mI)$, is a marginal probability and is computed from the joint probability for the data and the standard deviation of the noise

$$P(D|mB_0B_1\ldots B_mI) = \int P(\sigma D|mB_0B_1\ldots B_mI)d\sigma \tag{17.15}$$

which we factor as

$$P(D|mB_0B_1\ldots B_mI) = \int P(\sigma|I)P(D|\sigma mB_0B_1\ldots B_mI)d\sigma. \tag{17.16}$$

Assigning a Jeffreys' prior to $P(\sigma|I)$ and a Gaussian likelihood, one obtains

$$P(mB_0B_1\ldots B_m|DI) \propto P(m|I)\left[\prod_{j=0}^{M} P(B_j|I)\right]\int \frac{1}{\sigma}\left(2\pi\sigma^2\right)^{-\frac{N}{2}}\exp\left\{-\frac{Q}{2\sigma^2}\right\}d\sigma \tag{17.17}$$

Figure 17.2: The Distribution of Models On The Console Log

Phase	Annl Param	<Likelihood>	<StdDevLike>	3	4	5	6	7	8	9	10	11	12
Annealing	0.000	-2.8665E+01	-1.4089E+02	4	5	2	0	4	4	3	4	2	0
2	0.004	-3.1841E+01	-5.2177E+01	5	5	7	4	3	5	2	4	0	0
3	0.007	-3.1615E+01	9.4579E+00	6	5	14	4	5	2	3	1	1	0
4	0.012	-3.2178E+01	5.4344E+01	4	7	15	8	4	4	0	1	1	0
5	0.017	-3.4670E+01	1.0615E+02	2	6	14	13	7	3	0	1	1	0
6	0.025	-3.6671E+01	1.3447E+02	0	4	19	8	7	6	3	1	0	0
7	0.039	-3.6100E+01	1.5631E+02	0	0	21	13	7	2	3	2	0	0
8	0.065	-3.4743E+01	1.7613E+02	0	0	16	18	7	6	1	0	0	0
9	0.088	-3.3993E+01	1.9458E+02	0	0	11	22	12	3	0	0	0	0
10	0.109	-3.4516E+01	2.0296E+02	0	0	8	24	10	4	2	0	0	0
.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.
Phase	Frac	<Likelihood>	<StdDevLike>	3	4	5	6	7	8	9	10	11	12
Sampling	0.500	-3.1813E+01	2.6462E+02	0	0	0	48	0	0	0	0	0	0
12	0.550	-3.1991E+01	2.6438E+02	0	0	0	46	2	0	0	0	0	0
13	0.600	-3.1907E+01	2.6411E+02	0	0	0	47	1	0	0	0	0	0
14	0.650	-3.1989E+01	2.6459E+02	0	0	0	46	2	0	0	0	0	0
15	0.700	-3.1817E+01	2.6430E+02	0	0	0	48	0	0	0	0	0	0
16	0.750	-3.1904E+01	2.6432E+02	0	0	0	47	1	0	0	0	0	0
17	0.800	-3.1994E+01	2.6414E+02	0	0	0	46	2	0	0	0	0	0
18	0.850	-3.1992E+01	2.6429E+02	0	0	0	46	2	0	0	0	0	0
19	0.900	-3.1909E+01	2.6386E+02	0	0	0	47	1	0	0	0	0	0
20	0.950	-3.1903E+01	2.6447E+02	0	0	0	47	1	0	0	0	0	0
Phase	Frac	<Likelihood>	<StdDevLike>	3	4	5	6	7	8	9	10	11	12
Sampling	1.000	-3.1898E+01	2.6484E+02	0	0	0	47	1	0	0	0	0	0

Figure 17.2: While the unknown Polynomial Order package is running, it prints a listing that shows the distribution of the model indicator as a function of the annealing parameter. The 10 columns to the right are the number of simulations in model 0, 1, etc. As the annealing parameter increases these should cluster into one or two columns and the distribution of these simulations is the posterior probability for the polynomial order. In this example the data were a 6th order polynomial. Notice that as soon as the annealing parameter begins to increase the simulations quickly move to high order models and eventually they usually all end up in the 6th order polynomial model.

where we have left the prior probabilities in their symbolic form. Evaluating the integral over σ one obtains

$$P(mB_0B_1\ldots B_m|DI) \propto P(m|I)\left[\prod_{j=0}^{m} P(B_j|I)\right]\left[\frac{Q}{2}\right]^{-\frac{N}{2}} \tag{17.18}$$

as the posterior probability for the parameters including the polynomial order, where

$$\begin{aligned} Q &\equiv \sum_{i=1}^{N}\left(d_i - \sum_{j=0}^{m} B_j G_j(t_i)\right)^2 \\ &= N\overline{d^2} - 2\sum_{j=0}^{m} B_j T_j + \sum_{j=0}^{m} B_j^2. \end{aligned} \tag{17.19}$$

In evaluating the integral over σ there were a number of constants that were dropped. In model selection problems that is usually a bad thing to do and will, almost always, cause problem. Here we could to it because each polynomial model contains exactly the same constants and so they always cancel. Finally, substituting the prior probability for the polynomial order, Eq. (17.7), the prior probability for the amplitudes, Eq. (17.9) into Eq. (17.18), the joint posterior probability for the parameters is given by:

$$P(mB_0B_1\ldots B_m|DI) \propto \exp\left\{-\frac{(5-m)^2}{2\times 10^2}\right\}\left[\prod_{j=0}^{m}\left(2\pi\delta^2\right)^{-\frac{1}{2}}\exp\left\{-\frac{(T_j - B_j)^2}{2\delta^2}\right\}\right]\left[\frac{Q}{2}\right]^{-\frac{N}{2}}. \tag{17.20}$$

It is this joint probability density function that is targeted by the Markov chain Monte Carlo simulations.

If one were to compute the posterior probability for the polynomial order using Eq. (17.20) and compare it to that given by Eq. (17.11). You would find you get different results. That's because in computing these two sets of equations we used slightly different prior probabilities for the amplitudes. While these prior probabilities were not much different, they are nonetheless different and that difference would manifest itself as a slight difference in the final calculations. As far as which is right, they are both correct given the two sets of prior information. Regardless, they won't differ by much and almost certainly, given the discrete nature of the posterior probability, won't differ at all after you normalize the finial posterior probability.

17.2 Outputs From the Unknown Polynomial Order Package

The Text outputs from the Unknown Polynomial Order package consist of: "Bayes.prob.model," "BayesPolUnknown.mcmc.values," "Bayes.params," "Console.log" see Fig. 17.2, "Bayes.accepted" and a condensed output file "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.

The main new output from the Unknown Polynomial Order package is a plot of the posterior probability for the polynomial order, Fig. 17.3. The data used to generate this figure are the

Figure 17.3: The Posterior Probability For The Polynomial Order

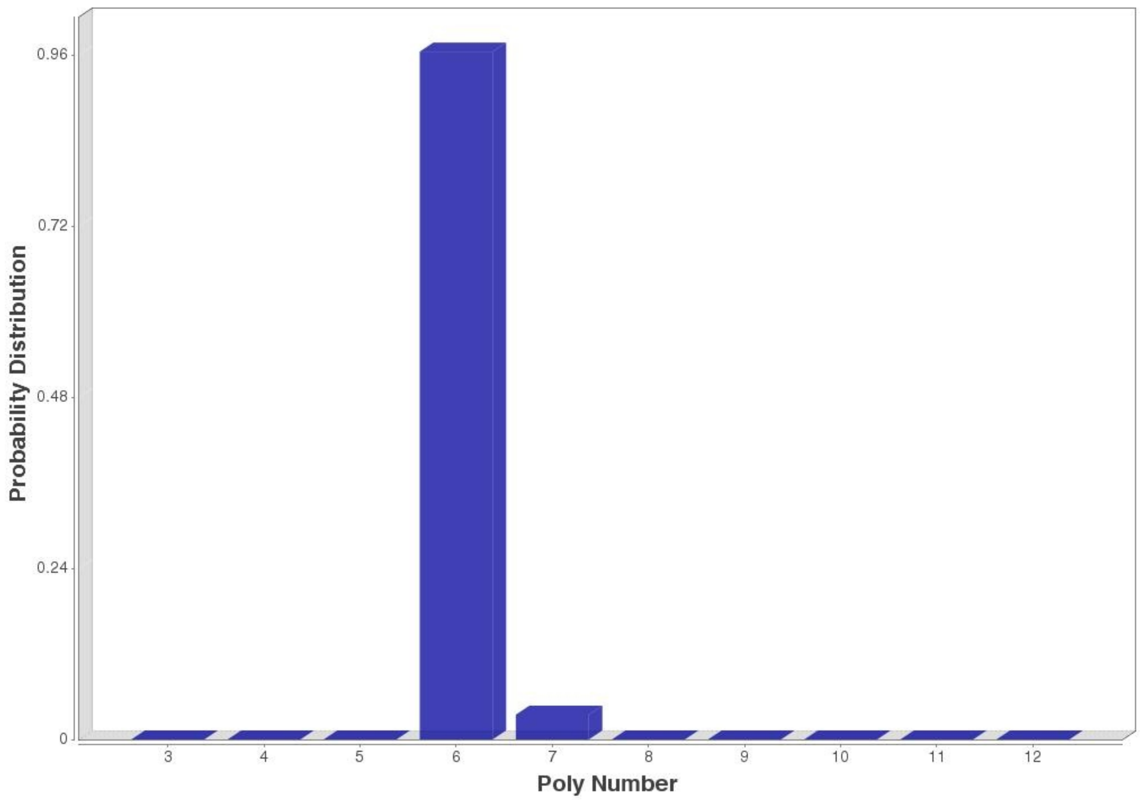

Figure 17.3: The main new output in the Unknown Polynomial Order Package is the posterior probability for the polynomial order, here called polynomial number. This output figure contains 10 probabilities centered around the peak in the posterior probability.

Polynomials.6th.order.dat distributed in the Bayes.test.data. This figure consists of a bar chart of the posterior probability. However, there are only 10 output probabilities, while the posterior probability contains a maximum of 51 probabilities (orders 0 through 50). So these probabilities are centered around the location of the maximum, see Section 17.1.1 for a discussion of how this maximum is located. Usually all of the probability is concentrated in one or two probabilities around the maximum with an abrupt lower bound and a more gentle drop off as you go to higher orders. Consequently, after locating the maximum of the posterior probability for the model order, the lowest output probability is set 3 orders below the maximum and the maximum output probability is 6 orders above the maximum probability order. All this is illustrated in Fig. 17.3, there the maximum is 6, the lowest output order is $6 - 3 = 3$, and the highest is $6 + 6 = 12$.

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.