Bayesian Data-Analysis Toolbox
Release 4.23, Manual Version 3

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: gbretthorst@wustl.edu

September 18, 2018

Chapter 16

Given Polynomial Order

The Given polynomial Order package fits polynomials to two column Ascii data when the order of the polynomial is known. The interface to the Given Polynomial Order package is shown in Figure 16.1. 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.

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.

Set the Polynomial order using the "Set Order" selection widget. For the Given Polynomial Order, the order can be from 1 to 55.

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 16.1: Given Polynomial Order Package Interface

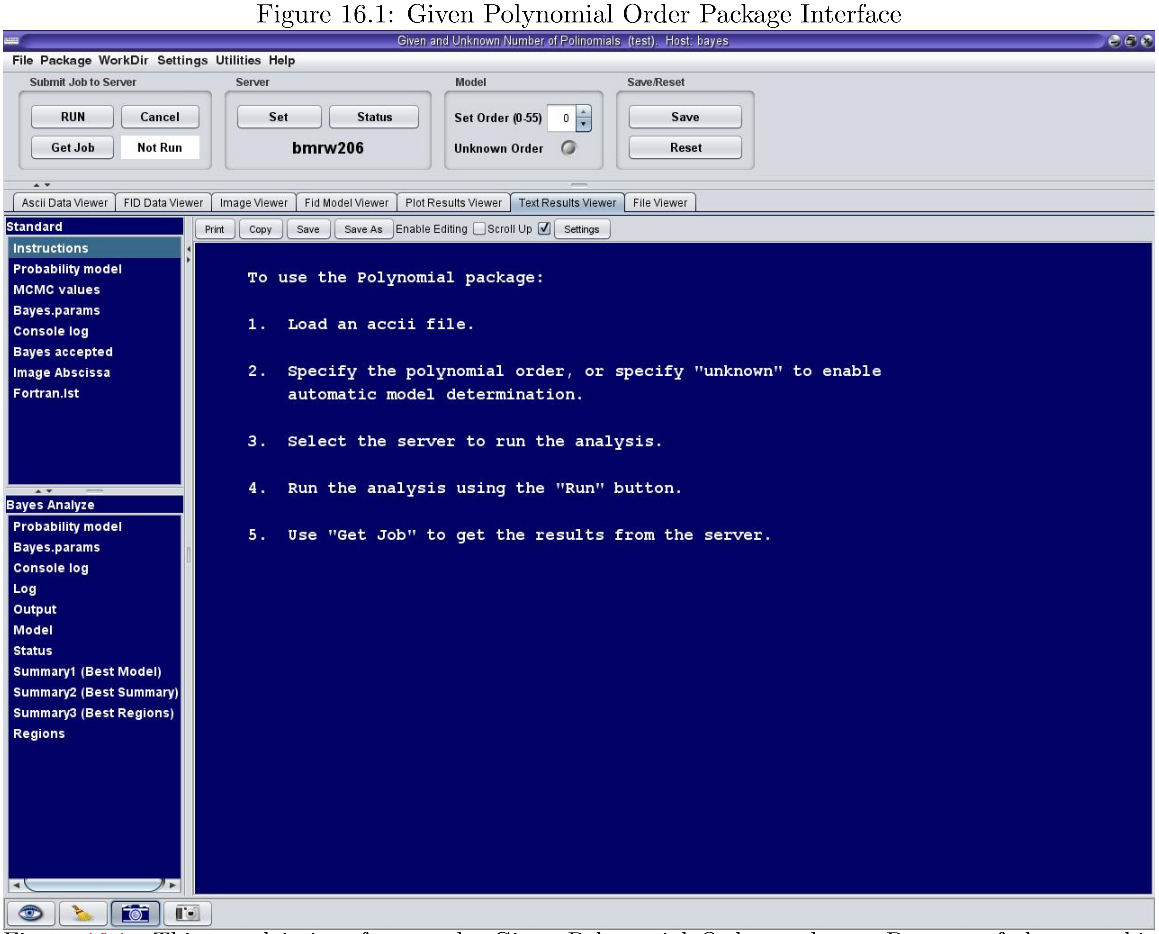

Figure 16.1: This panel is interface to the Given Polynomial Order package. Because of the way this calculation is done very high orders are possible and numerically stable. However, the high orders, above 40, require very high signal-to-noise and even then roundoff errors degrade the accuracy to 4 or 5 decimal places.

16.1 The Bayesian Calculation

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

$$d_i = \sum_{j=0}^{m} A_j G_j(t_i) + n_i \tag{16.1}$$

with

$$G_j(t_i) = t_i^j \tag{16.2}$$

where A_j is the amplitude of the jth polynomial, n_i represents noise in the ith data value and we have written these polynomials as $G_j(t_i)$ for notational convenience. We will think of these polynomials as functions of time, but in the analysis which follows t is simply a single column abscissa and may be any quantity. In the problem we are describing here, m is the given order of the polynomial. Additionally, the assumption that these are polynomials is unimportant in the following discussions, the $G_j(t_i)$ could be any set of functions.

16.1.1 Gram-Schmidt

The Bayesian calculation is implemented using Markov chain Monte Carlo with simulated annealing to draw samples for the joint posterior probability for the parameters. Before we do this calculation we introduce a change of function and a change of variables. The reason for this is simply that computing polynomials of the form $\sum_{j=0}^{m} A_j t_i^j$ is computationally vary unstable in the sense that only orders up to about 8 can be computed using single precision arithmetic. You can get to higher orders only by using numerical procedures that avoid differencing large numbers, for example the remainder theorem. However, here we take a different approach by transforming the problem to something that is computationally more stable. Using Gram-Schmidt, the polynomials are transformed to a set of orthogonal polynomials. We then solve the problem using these orthogonal polynomials and finally, we transform the derived amplitudes back to the A_j given in Eq. (16.1). If we designate the Gram-Schmidt polynomials as $H_j(t_i)$ and the expansion coefficients as B_j, Eq. (16.1) becomes:

$$d_i = \sum_{j=0}^{m} B_j H_j(t_i) + n_i. \tag{16.3}$$

We chose the Gram-Schmidt polynomials because they can be computed trivially from the t_i^j, they preserve the notation of the order of the polynomial, and each polynomial depends only on the lower orders, not the higher orders. This change of variables and change of function is an identity, i.e., the polynomial expansions in Eq. (16.1) and Eq. (16.3) are exactly equal to each other. Finally, the amplitudes, B_j and A_j, are linearly related to each other through an lower triangular matrix, and consequently, the conversion back and forth between these representations is very easy to program.

As a reminder to those unfamiliar with Gram-Schmidt polynomials, the normalized Gram-Schmidt polynomials $H_j(t_i)$, are generated recursively from the $G_j(t_i)$ using:

$$H_j(t_i) = \frac{1}{C_{jj}} \left[G_j(t_i) - \sum_{\ell=0}^{j-1} C_{j\ell} H_\ell(t_i) \right] \tag{16.4}$$

where the sum is not present for the first polynomial, and

$$C_{j\ell} = \sum_{i=1}^{N} G_j(t_i)H_\ell(t_i) \qquad (0 \leq j \leq \ell). \tag{16.5}$$

Gram-Schmidt polynomials have the property

$$\sum_{i=1}^{N} H_j(t_i)H_\ell(t_i) = \delta_{j\ell} \tag{16.6}$$

where $\delta_{j\ell}$ is zero if $j \neq \ell$ and one if $j = \ell$.

To derive the relationship between the A_j and the B_j, note that the expansions given by Eq. (16.1) and Eq. (16.3) are identities, so can write

$$\sum_{k=0}^{m} A_k G_k(t_i) = \sum_{j=0}^{m} B_j H_j(t_i) \tag{16.7}$$

where we changed the summation index on the left-hand side just to remind people that these summations are independent of each other. Multiplying this equation by $H_\ell(t_i)$, and summing over time:

$$\sum_{i=1}^{N}\sum_{k=0}^{m} A_k G_k(t_i)H_\ell(t_i) = \sum_{i=1}^{N}\sum_{j=0}^{m} B_j H_j(t_i)H_\ell(t_i). \tag{16.8}$$

The right-hand side of this equating is zero unless $j = \ell$ and then one obtains

$$\sum_{i=1}^{N}\sum_{k=0}^{m} A_k G_k(t_i)H_\ell(t_i) = B_j \tag{16.9}$$

and the sum over time on the left-hand side of this equation can be written as

$$\sum_{k=0}^{m} A_k \left[\sum_{i=1}^{N} G_k(t_i)H_\ell(t_i)\right] = B_j. \tag{16.10}$$

The quantity in big square brackets is just the right-hand side of Eq. (16.5), so this equation becomes

$$\sum_{k=0}^{m} A_k C_{k\ell} = B_\ell. \tag{16.11}$$

The matrix $C_{k\ell}$ is a lower triangular matrix, so inverting it is trivial and one can use this equation solve for the nonorthogonal expansion coefficients, the A_k, from the orthogonal expansion coefficients, the B_j.

16.1.2 The Bayesian Calculation

The Bayesian calculation is for the joint posterior probability for the amplitudes, B_j, given the data and the prior information. This joint probability, denoted by $P(B_0B_1\ldots B_m|DI)$, is computed by application of Bayes' theorem

$$P(B_0B_1\ldots B_m|DI) \propto P(B_0B_1\ldots B_m|I)P(D|B_0B_1\ldots B_mI) \tag{16.12}$$

where $P(B_0 B_1 \ldots B_m|I)$ is the joint prior probability for the amplitudes, and $P(D|B_0 B_1 \ldots B_m I)$ is the likelihood. Because each polynomial is orthogonal, we will factor the joint prior probability for the amplitudes, $P(B_0 B_1 \ldots B_m|I)$, into a series of independent prior probabilities:

$$P(B_0 B_1 \ldots B_m|I) = \prod_{j=0}^{m} P(B_j|I) \tag{16.13}$$

and each of the $P(B_j|I)$ will be assigned a unbound Gaussian. The posterior probability for the B_j, is thus given by:

$$P(B_0 B_1 \ldots B_m|DI) \propto \left[\prod_{j=0}^{m} P(B_j|I)\right] P(D|B_0 B_1 \ldots B_m I). \tag{16.14}$$

We want the Markov chain to explore the amplitude parameter space, but we don't want it to excessively waist time. All the amplitudes in an orthogonal model are estimated to be $\pm\sqrt{\langle\sigma^2\rangle}$, where $\langle\sigma^2\rangle$ is the mean-square residual given the model. Consequently, if we center the prior probability for an amplitude on the expected amplitude, T_j Eq. (16.16) below, and make the standard deviation of the prior very wide, then the prior probability for the amplitudes will do little more than keep the Markov chain in the physically meaningful region of the parameter space. Here is the prior actually used for the amplitudes:

$$P(B_j|I) = \left(2\pi\delta^2\right)^{-\frac{1}{2}} \exp\left\{-\frac{(T_j - B_j)^2}{2\delta^2}\right\} \tag{16.15}$$

where the expected amplitude, T_j, is given by

$$T_j \equiv \sum_{i=1}^{N} d_i H_j(t_i) \tag{16.16}$$

where N is the total number of data values in the data set. The standard deviation of this prior, δ, is 10 times larger than the expected root mean-square residual:

$$\delta = 10\sqrt{\langle\sigma^2\rangle} \tag{16.17}$$

with

$$\sqrt{\langle\sigma^2\rangle} = \sqrt{\frac{\overline{d^2} - \overline{h^2}}{N}}. \tag{16.18}$$

The quantity, $\overline{d^2} - \overline{h^2}$, is the total-squared residual given the polynomial. So the square root is the root mean-square residual given the polynomial order. The sufficient statistic, $\overline{h^2}$, is the total-squared projection of the data onto the polynomial and is defined as

$$\overline{h^2} \equiv \sum_{k=0}^{m} T_k^2. \tag{16.19}$$

Having assigned the prior probabilities, we must now assign the direct probability. The direct probability, $P(D|B_0 B_1 \ldots B_m I)$, is a marginal probability and is computed from the joint probability for the data and the standard deviation of the noise

$$P(D|B_0 B_1 \ldots B_m I) = \int P(\sigma D|B_0 B_1 \ldots B_m I) d\sigma \tag{16.20}$$

which we factor as

$$P(D|B_0B_1\dots B_mI) = \int P(\sigma|I)P(D|\sigma B_0B_1\dots B_mI)d\sigma. \tag{16.21}$$

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

$$P(B_0B_1\dots B_m|DI) \propto \left[\prod_{j=0}^{m} P(B_j|I)\right] \int \frac{1}{\sigma}\left(2\pi\sigma^2\right)^{-\frac{N}{2}} \exp\left\{-\frac{Q}{2\sigma^2}\right\} d\sigma \tag{16.22}$$

where we have left the prior probabilities in their symbolic form. Evaluating the integral over σ and substituting the prior probability for the amplitudes, Eq. (16.15), into Eq. (16.22), one obtains:

$$P(B_0B_1\dots B_m|DI) \propto \left[\prod_{j=0}^{m} \exp\left\{-\frac{(T_j - B_j)^2}{2\delta^2}\right\}\right]\left[\frac{Q}{2}\right]^{-\frac{N}{2}} \tag{16.23}$$

where we dropped some constant terms that cancel when this probability is normalized. The function Q is defined as

$$\begin{aligned} Q &\equiv \sum_{i=1}^{N}\left(d_i - \sum_{j=0}^{m} B_jH_j(t_i)\right)^2 \\ &= N\overline{d^2} - 2\sum_{j=0}^{m} B_jT_j + \sum_{j=0}^{m} B_j^2. \end{aligned} \tag{16.24}$$

One interesting note about the quantity Q, it does not depend on the individual data values, rather it depends on the total squared data value, the $N\overline{d^2}$, and it depends on the projection of the data onto the orthogonal functions, the T_j. Both of these items can be computed at the beginning of the calculation and used throughout with no further reference to the data. Consequently, the Given Polynomial Order runs very quickly. Also, note that the function Q could have been written as a single summation rather than two. If the standard deviation for the noise is known, the posterior probability for the B_j can be factored into a product of probabilities for each amplitude separately, i.e., the amplitudes of the orthogonal polynomials may be estimated separately, they don't have to be estimated jointly. Finally, because each amplitude can be estimated separately, one can simply plot the posterior probability for each amplitude, there is no need to use a Markov chain Monte Carlo simulation to sample the joint posterior. However, joint estimation is required after marginalizing out the standard deviation for the noise. The joint estimation is done using a Markov chain Monte Carlo simulation to sample the joint posterior probability for the amplitudes, Eq. (16.23).

16.2 Outputs From the Given Polynomial Order Package

The Text outputs files from the Given Polynomial Order package consist of: "Bayes.prob.model," "BayesPolGiven.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

Figure 16.2: Given Polynomial Order Scatter Plot

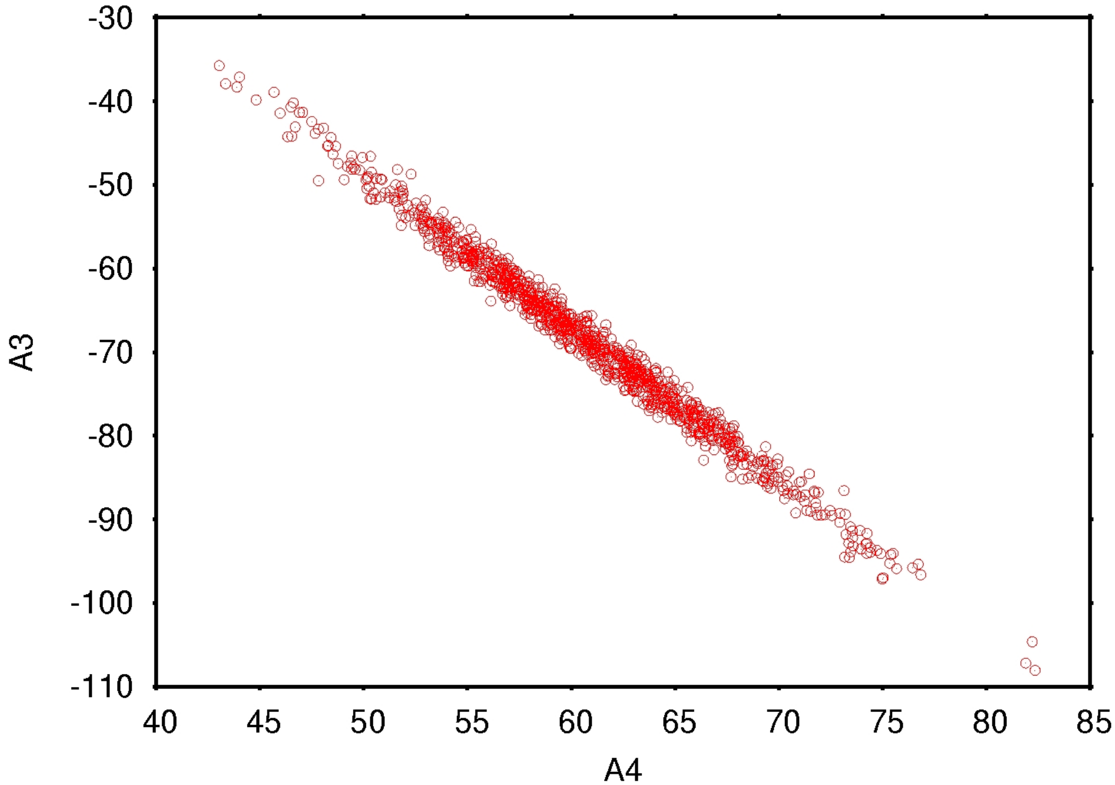

Figure 16.2 The expansion coefficients in a nonorthogonal polynomial expansion tend to be highly correlated. Plotted here is a scatter plot of the 3rd and 4th order expansion coefficients, A_3 and A_4, as generated by a 6th order expansion of the 6th order polynomial test data. This test data can be downloaded using the "Files" menu.

mcmc.values report is discussed in Appendix D and the other reports are discussed in Chapter ??. 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 A_j in the given model. And because estimation of the amplitudes in polynomial models tend to be highly correlated, there are covariance plots to help illustrate these correlations. These covariance plots are scatter plots. The scatter plots are generated from the samples drawn from the Markov chain Monte Carlo simulation. In a typical run, there might be 50 simulations and 30 repeats giving a total of 1500 simulations. Each of these simulations contain the estimated parameters from one Markov chain Monte Carlo simulation. A typical scatter plot just put a dot in the plot at the location of parameter 1 versus parameter 2. In this package that would correspond to plotting A_j versus A_k. In Fig. 16.2, A_3 versus A_4 is plotted. The data used to generate this figure are the 6th order polynomial expansion data available in our test data kit. This test data can be downloaded using the "Files" menu. The covariance plot shown in Fig. 16.2 is the one generated form the A_3 and A_4 expansion coefficients in a 6th order expansion of this data. For these two parameters there is a strong correlation, when A_3 increases the A_4 parameter decreases to counter the effect of changing A_3. Uncorrelated parameters, by contrast, will have elliptical scatter plots with the major and minor axis aligned with coordinate system. The number of possible scatter plots is $m(m+1)$ where m is the order of the polynomial. Because the number of scatter plots can become large very quickly, the package only outputs a representative sample of these plots.

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, Marilyn 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, Marilyn 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] Marilyn 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.

Index