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 6

Unknown Number of Exponentials

The unknown number of exponentials package estimates amplitudes and decay rate constants of an unknown number of exponential signal components in data that are known to contain sums of exponential signals. The data analyzed by this package are Ascii data files and may be input as Ascii files, peak picks, the amplitudes from a Bayes Analyze analysis run or extracted from an image stack. The unknown number of exponentials package is accessed by selecting the "Exponential" package on the package menu. When the Exponential package is selected, the interface shown in Fig. 6.1 is displayed. This interface is used in both the given and unknown number of exponentials packages. The unknown number of exponential package is selected by setting the model order to "Unknown". Note when the model order is set to unknown, the "Constant" widget is grayed out, i.e., becomes inactive because the unknown number of exponentials package computes, among other things, the posterior probability a constant is present. To use the unknown number of exponentials package, you must do the following:

Select the exponential package from the Package menu.

Load one or more Ascii data sets using the Files menu.

Set the "Unknown" model order in the model order selection menu.

Check the Find Outliers check box if you suspect outliers are present in the data.

Review the prior probabilities assignment for the decay rate constants 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.

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 6.1: The Unknown Exponential Interface

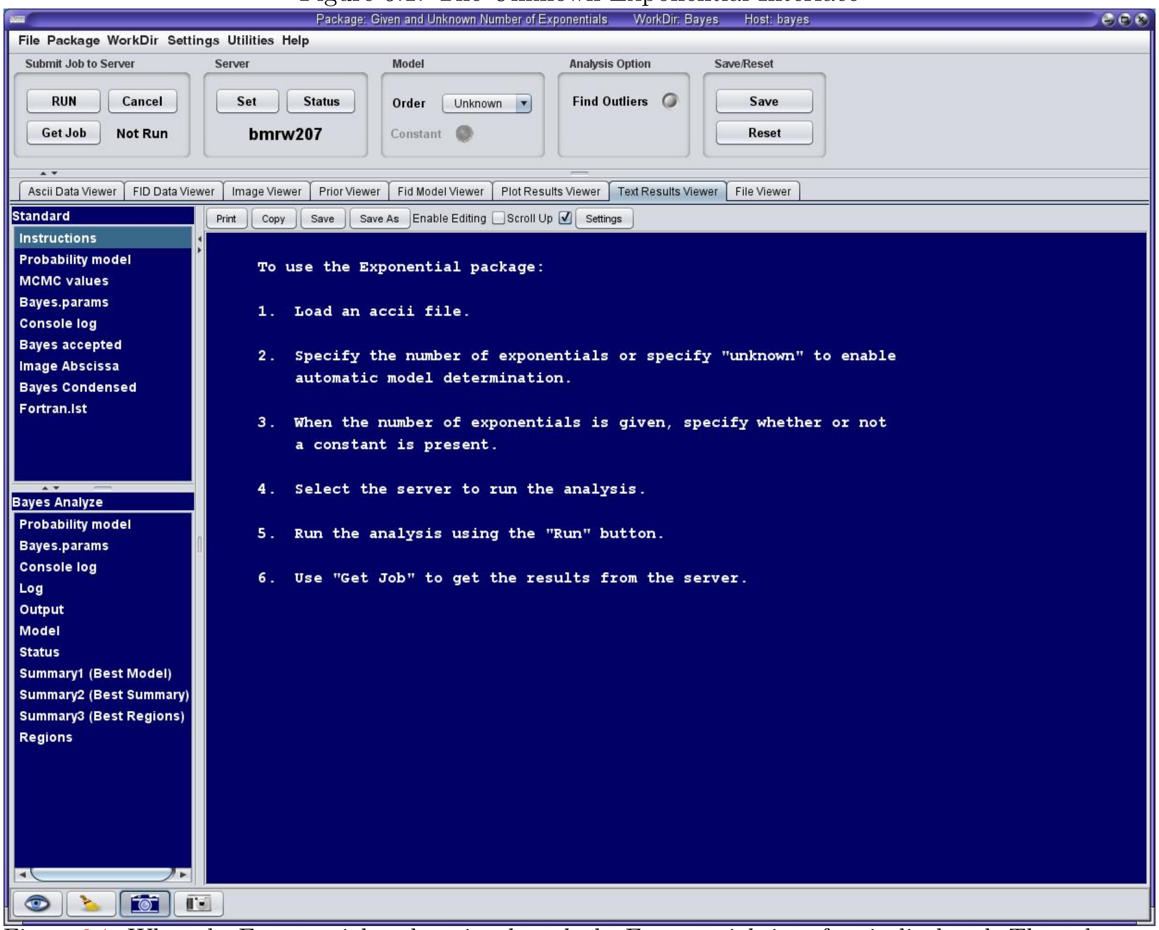

Figure 6.1: When the Exponential package is selected, the Exponentials interface is displayed. The unknown number of exponentials package is selected by setting the model order to "Unknown". The prior viewer can be used to set the prior probability for the decay rate constants. Note in this package the posterior probabilities are marginal posterior probabilities, so there are no adjustable prior probabilities for the amplitudes.

6.1 The Bayesian Calculations

The Bayesian calculation is for the joint posterior probability for the number of exponentials and whether or not a constant offset is present in the data. As with all Bayesian calculations, one begins by relating the data to the parameters of interest. In this case the model which relates the unknown number of exponentials to the data is of the form

$$d_{ij} = C_j + \sum_{l=1}^{m} A_{lj} \exp\{-\alpha_l t_i\} + n_{ij} \tag{6.1}$$

where d_{ij} is the ith data value in the jth data set containing N_j data values. Note that the number of data values is data set dependent and may differ from one data set to the next, C_j is a constant offset in the jth data set and part of the calculation is to determine if this constant is present, m is the unknown number of exponentials, A_{lj} is the amplitude or intensity of the lth exponential in the jth data set, α_l is the lth exponential decay rate constant, and n_{ij} is noise in the ith data value in the jth data set. Except for the lack of prior information about the number of exponentials m, and whether or not the constant C is present, this is the given number of exponentials model, Eq. (5.1). However, now the Bayesian calculations are more substantial because both m and whether or not C_j are present must be part of the Bayesian calculation.

The calculation begins by first ranked the models from the simplest to the most complicated and then numbering them from 1 to the total number of models. In the program that implements this calculation the maximum number of exponentials is set to 4, consequently there are a total of 10 models. The model number is designate as u and when $u = 1$, the model is no exponential and no constant, $u = 2$ is no exponential and a constant, $u = 3$ is one exponential and no constant, etc., up to model 10, 4 exponentials and a constant. The model indicator, u, is a discrete hypotheses that must be incorporated into the Bayesian calculation and its numerical value indicates both the number of exponentials and whether or not a constant offset is present.

To determine the model, i.e., estimates u, one computes the posterior probability for the model indicator given the data and the prior information, $P(u|DI)$. This posterior probability is computed by an application of Bayes' theorem [1],

$$P(u|DI) = \frac{P(u|I)P(D|uI)}{P(D|I)} \tag{6.2}$$

where $P(u|I)$ is the prior probability for the model indicator, $P(D|uI)$ is the marginal direct probability for the data given the model indicator, and the denominator is

$$P(D|I) = \sum_u P(Du|I) = \sum_u P(u|I)P(D|uI) \tag{6.3}$$

is the normalization constant needed to ensure the total probability for the model indicator sums to one. If this posterior probability is normalized at the end of the calculation, then

$$P(u|DI) \propto P(u|I)P(D|uI). \tag{6.4}$$

The two terms on the right-hand side of this equation are the prior probability for the model indicator, $P(u|I)$, which can be assign, and the direct probability for the all of the data given the model indicator, $P(D|uI)$, which is much too complicated to see how it should be assigned.

However, this direct probability can be computed from the joint probability for all of the data and the parameters given the model indicator, $P(D\{A\}\{\alpha\}\{C\}\{\sigma\}|uI)$, by application of the sum rule:

$$P(u|DI) \propto P(u|I) \int d\{A\}d\{C\}d\{\alpha\}d\{\sigma\}P(D\{A\}\{C\}\{\alpha\}\{\sigma\}|uI) \tag{6.5}$$

where all of the parameters except the model indicator have been removed by marginalization. The notation $\{\cdot\}$, means all of the parameters of the indicated type. So for example $\{A\}$ means all of the amplitudes appearing in the model. If there are 3 exponentials and 5 data sets there would be 15 total amplitudes. Also note that the parameters represented by $\{C\}$ may or may not occur in any given model depending on the model indicator. So if the model indicator is even then $\{C\}$ is present, otherwise its not. The right-hand side of this equation may be factored, one obtains

$$P(u|DI) \propto P(u|I) \int d\{A\}d\{\alpha\}d\{C\}d\{\sigma\}P(\{A\}\{\alpha\}\{C\}\{\sigma\}|uI)P(D|u\{A\}\{\alpha\}\{C\}\{\sigma\}I) \tag{6.6}$$

where $P(\{A\}\{\alpha\}\{C\{\sigma\}|uI)$ is the joint prior probability for all of the parameters appearing in model u and $P(D|u\{A\}\{\alpha\}\{C\{\sigma\}I)$ is the direct probability for all of the data given all of the parameters and the model indicator. Both of these probabilities are too complicated to see how they might be assigned, but the rules of probability theory can be used to factor these probabilities into increasingly simple terms: Assuming logical independence, i.e., changing the prior information about the amplitudes would not change the prior information about the decay rate constants, standard deviations, etc., then the joint prior probability for all of the parameters, $P(\{A\}\{\alpha\}\{C\}\{\sigma\}|I)$ can be factored into a series is into a series of independent prior probabilities for each parameter:

$$P(\{A\}\{\alpha\}\{C\}\{\sigma\}|I) = \left[\prod_{j=1}^{m}\prod_{k=1}^{n} P(A_{kj}|I)\right]\left[\prod_{j=1}^{m} P(\alpha_j|I)\right]\left[\prod_{k=1}^{n} P(C_k|I)\right]\left[\prod_{k=1}^{n} P(\sigma_k|I)\right] \tag{6.7}$$

where m is the number of exponentials in the model indicated by u, n is the number of data sets, $P(A_{kj}|I)$ is the prior probability for amplitude k in data set j, $P(\alpha_j|I)$ is the prior probability for the jth decay rate constant, $P(C_k|I)$ is the prior probability for the constant offset in the kth data set, and $P(\sigma_k|I)$ is the prior probability for the standard deviation of the noise in the kth data set. Note, that the prior probability for the constant offset, $P(C_k|I)$, is only present when the model indicator is even. Also note that the dependence on the model indicator has been dropped in these priors, because knowing the number of exponentials and whether or not a constant is present would not change the prior information about the amplitudes and decay rate constants. Substituting the prior, Eq. (6.7) back into the posterior, Eq. (6.6) to obtain

$$\begin{aligned}
P(u|DI) \quad &\propto P(u|I) \quad \int d\{A\}d\{\alpha\}d\{C\}d\{\sigma\} \\
&\times \quad \left[\prod_{j=1}^{m}\prod_{k=1}^{n} P(A_{kj}|I)\right] \\
&\times \quad \left[\prod_{j=1}^{m} P(\alpha_j|I)\right] \\
&\times \quad \left[\prod_{k=1}^{n} P(C_k|I)\right] \\
&\times \quad \left[\prod_{k=1}^{n} P(\sigma_k|I)\right] \\
&\times \quad \left[\prod_{k=1}^{n} P(D_k|uA_{1k}\dots A_{mk}\{\alpha\}C_k\sigma_k I)\right]
\end{aligned} \tag{6.8}$$

as the posterior probability for the model indicator, where $P(D_k|uA_{1k}\dots A_{mk}\{\alpha\}C_k\sigma_k I)$ is the direct probability for data set D_k given both the parameters that occur in all data sets, the $\{\alpha\}$, and the parameters specific to data set D_k. These data set specific parameters are the constant offset C_k and the amplitudes of the m exponentials in the kth data set. Here these amplitudes have been designated as $A_{1k}\dots A_{mk}$.

If you examine the integrand of this posterior, you will discover that it is the joint posterior probability for the parameters given the model indicator. Consequently, this integrand is exactly the same as the integrand in the joint posterior probability given the model order, Eq. (5.3), and in what follows the same prior probabilities will be assigned in both problems. This will have the effect of making the parameter estimates from the unknown number of exponentials and the given number of exponentials essentially identical when the model indicators are the same.

Equation. (6.8) cannot be further simplified and the only recourse is to assign the various prior probabilities and evaluate the integrals. The prior probability for the model indicator, $P(u|I)$, was assigned using a uniform prior probability:

$$P(u|I) \propto \begin{cases} \dfrac{1}{10} & \text{if } 1 \leq u \leq 10 \\ 0 & \text{otherwise} \end{cases} . \tag{6.9}$$

The prior probability for the noise standard deviation, $P(\sigma_k|I)$, was assigned a Jeffreys' prior

$$P(\sigma_k|I) \propto \frac{1}{\sigma_k}. \tag{6.10}$$

Note that the Jeffreys' prior has not been bounded and normalized because the exact same normalization constant appears in all models tested by this calculation. The prior probabilities for the amplitudes, $P(A_{kj}|I)$, are assigned broad Gaussian prior probability:

$$P(A_{jk}|I) = \left(\frac{2\pi\sigma_k^2}{\delta^2}\right)^{-\frac{1}{2}} \exp\left\{-\frac{\delta^2}{2\sigma^2}A_{jk}^2\right\} \tag{6.11}$$

where δ/σ_k plays the part of the standard deviation of this prior probability and cannot be changed by the user. In the program that implements this calculation $\delta = 0.01$.

The prior probabilities for the decay rate constants are under use control and the user may select the type of prior used. However, if the user does not set the prior probability for the decay rate constants, then the same default prior probability described in Chapter 5 Eq. (5.4) is used.

The model equation is symmetric under relabeling of the amplitudes and decay rate constants. This symmetry causes the joint posterior probability for the decay rate constants to be symmetric in the sense that if there is a peak at $\alpha_1 = \beta$ and $\alpha_2 = \gamma$, then there is also a peak at $\alpha_1 = \gamma$ and $\alpha_2 = \beta$; this symmetry is caused because the model does not tell us which exponential signal corresponds to to which model component. Consequently, a convention must be introduced which brakes this symmetry by identifying a model component with a signal component. This is done by forcing the decay rate constants to be ordered.

The full Bayesian calculation and the assignment of the prior probabilities is discussed in reference [16] and this paper is available in pdf by activating this link. Additionally, much more about exponential parameter estimation is contained in [15, 17]. The first paper describes the problem when the exponential model is given and is a simpler version of this problem, while the second paper discusses how the accuracy of the parameter estimates depends on the number of data values, signal-to-noise level and the rate of decay of the sample.

6.2 Outputs From The Unknown Number of Exponentials Package

The Text outputs from the unknown number of exponentials package consist of: “Bayes.prob.model,” “BayesExpGiven.mcmc.values,” “Bayes.params,” “Console.log,” “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.

When the unknown number of exponentials package is running, the distribution of simulations having a given model indicator u is in constant flux. The unknown number of exponentials program writes the number of simulations having $u = 1$, $u = 2$, etc. to the console log after it completes each annealing step. This display serves as a kind of visual picture of the calculation of the posterior probability for the model indicator. A sample of this output is shown in Fig. 6.2. This listing was generated from a data set that contains three exponential plus a constant. When the annealing parameter is zero, the distribution of simulations reflects only the prior probability for the model. Because that prior probability is uniform, the simulations are essentially evenly distributed across all models. As the program runs, the annealing parameter is increased and the distribution of simulations begins to change to reflect the increasing importance of the likelihood. Typically the simulations begin to cluster around one or two models. This clustering continues until of the simulations have settled into the posterior probability for the model indicator.

Figure 6.2: The Distribution Of Models

Phase	Annl Param	<Likelihood>	<StdDevLike>	0	0+C	1	1+C	2	2+C	3	3+C	4	4+C
Annealing	0.000	-7.2157E+02	1.5195E+02	12	6	7	8	12	13	9	11	10	12
2	0.005	-5.6842E+02	1.6292E+02	7	5	4	4	2	21	4	24	3	26
3	0.011	-4.8783E+02	1.1564E+02	1	0	0	2	0	14	13	28	16	26
4	0.018	-4.3429E+02	8.3983E+01	0	0	1	0	0	18	2	29	17	33
5	0.026	-3.9971E+02	5.2703E+01	0	0	0	0	0	9	4	26	21	40
6	0.038	-3.8265E+02	3.4063E+01	0	0	0	0	0	8	2	37	21	32
7	0.054	-3.7660E+02	3.6781E+01	0	0	0	0	0	7	2	34	24	33
8	0.069	-3.6507E+02	1.7327E+01	0	0	0	0	0	4	0	39	24	33
9	0.090	-3.6031E+02	1.3435E+01	0	0	0	0	0	1	0	44	24	31
10	0.113	-3.5629E+02	8.9340E+00	0	0	0	0	0	1	0	54	21	24
.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.
Phase	Annl Param	<Likelihood>	<StdDevLike>	0	0+C	1	1+C	2	2+C	3	3+C	4	4+C
Annealing	0.745	-3.4957E+02	1.5147E+00	0	0	0	0	0	0	0	97	2	1
32	0.777	-3.4935E+02	1.2221E+00	0	0	0	0	0	0	0	94	1	5
33	0.809	-3.4936E+02	1.1754E+00	0	0	0	0	0	0	0	99	0	1
34	0.841	-3.4917E+02	9.8822E-01	0	0	0	0	0	0	0	94	1	5
35	0.873	-3.4893E+02	8.6573E-01	0	0	0	0	0	0	0	99	1	0
36	0.906	-3.4926E+02	1.3015E+00	0	0	0	0	0	0	0	97	1	2
37	0.938	-3.4910E+02	9.9978E-01	0	0	0	0	0	0	0	97	0	3
38	0.970	-3.4902E+02	9.6721E-01	0	0	0	0	0	0	0	97	2	1
39	1.000	-3.4903E+02	9.9140E-01	0	0	0	0	0	0	0	100	0	0

Figure 6.2: While the unknown number of exponentials 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 model order. In this example the data were three exponentials plus a constant. Notice that as soon as the annealing parameter begins to increase the simulations quickly move to high order models and eventually the all end up in the three exponential plus a constant model.

In addition to these text reports there is one new plot result, the posterior probability for the model order, Fig. 6.3. The abscissa is labeled with the model order. These model orders, 1, 2, 3, etc. are the number of exponentials in the model. When a constant is present this, the model order is changed to 1+C, 2+C, 3+C etc. to indicate the presence of one, two or three exponentials plus a constant. The vertical axis, is the posterior probability for the model indicator. Each probability is computed by counting the number of simulations having the indicated model divided by the total number of simulations. The “Plot Results Viewer” can be used to view the probability for the model indicator and the output probability density functions for the parameters in a given model order. 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.

Figure 6.3: The Posterior Probability For Exponential Model

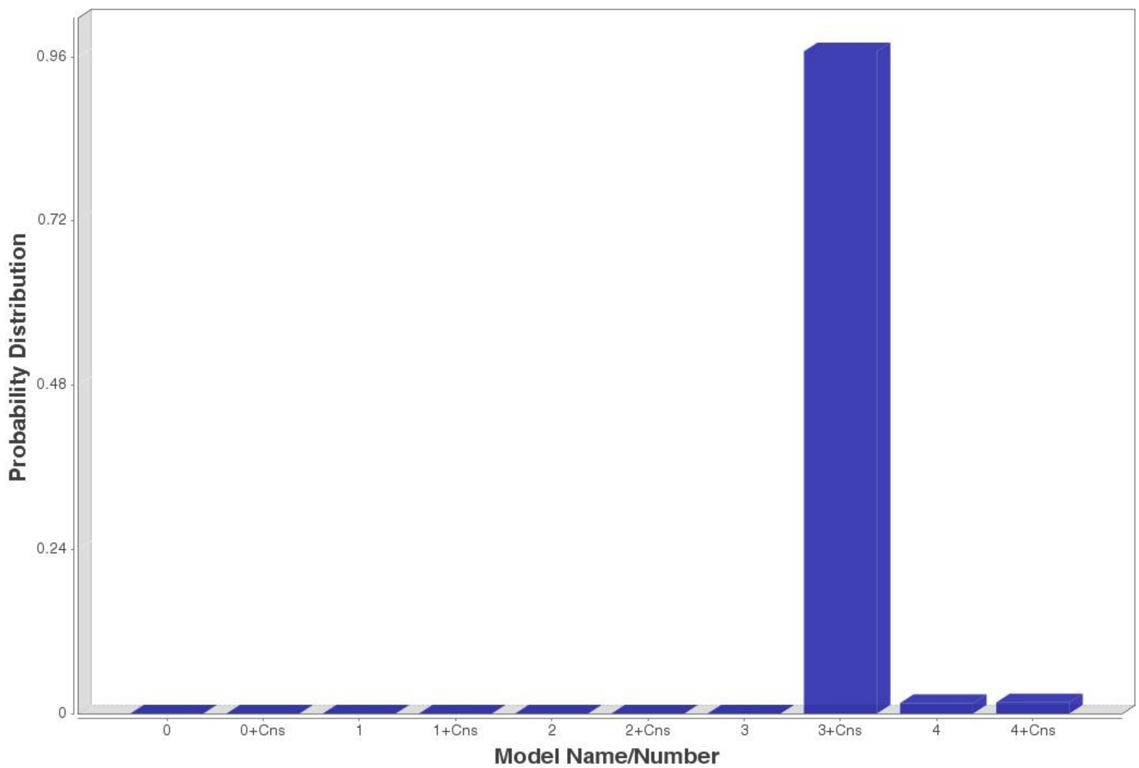

Figure 6.3: When the Unknown number of Exponentials package runs a new output plot is produced. The plot is of the posterior probability for the model order. The abscissa is labeled with the model order. These model orders are arbitrated as 1, 2, 3, etc. when the model is one, two or three exponentials without a constant. When a constant is present this label is changed to 1+C, 2+C, 3+C etc. to indicate the presence of one, two or three exponentials plus a constant.

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.