

Bayesian Posterior Predictive Probability Happiness

Gabriela Rodríguez-Hernández, Galileo Domínguez-Zacarías*, Carlos Juárez Lugo

Universidad Autónoma del Estado de México, Ecatepec, Mexico

Email: *grodriguez@uaemex.mx, *gadoza3@gmail.com

Received 15 March 2016; accepted 17 May 2016; published 20 May 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

We propose to determine the underlying causal structure of the elements of happiness from a set of empirically obtained data based on Bayesian. We consider the proposal to study happiness as a multidimensional construct which converges four dimensions with two different Bayesian techniques, in the first we use the Bonferroni correction to estimate the mean multiple comparisons, on this basis it is that we use the function t and a z -test, in both cases the results do not vary, so it is decided to present only those shown by the t test. In the Bayesian Multiple Linear Regression, we prove that happiness can be explained through three dimensions. The technical numerical used is MCMC, of four samples. The results show that the sample has not atypical behavior too and that suitable modifications can be described through a test. Another interesting result obtained is that the predictive probability for the case of sense positive of life and personal fulfillment dimensions exhibit a non-uniform variation.

Keywords

Bayesian Inference, Posterior Predictive Distribution, MCMC, Happiness

1. Introduction

Bayesian networks emerged about three decades ago as alternatives to conventional systems-oriented decision-making and forecasting under uncertainty in probabilistic terms [1]. A Bayesian network is a statistical tool that represents a set of associated uncertainties given conditional independence relationships established between them [2] [3].

The rule Bayes is a rigorous method for interpreting evidence in the context of previous experience or knowledge. The Bayes rule has recently emerged as a powerful tool with a wide range to applications which include:

*Corresponding author.

genetics, image processing, ecology, physics and engineering. The essential characteristics of Bayesian methods are their explicit use of probability for quantifying uncertainty in inferences based on statistical data analysis. A Bayesian probability interval for an unknown quantity of interest can be directly regarded as having a high probability of containing the unknown quantity, in contrast to a frequentist confidence interval which may strictly be interpreted only in relation to a sequence of similar inferences that might be made in repeated practice.

Using Bayesian networks as a tool for data analysis is not widespread in the context of Psychology. However, the benefits of using this tool in all areas of psychology are identified in several ways: on the economic front with Bayesian networks could develop systems to make judgments appropriate chance to improve diagnosis and psychological treatment. On the scientific level, Bayesian networks cannot be overlooked if psychology strives to clarify the mechanisms by which people evaluate, decide and make inferences; as they may serve as analytical and theoretical reference in the development of models of reasoning, learning and perception of uncertainty [3].

In psychology, the area's most prolific work in the use of Bayesian networks has been the causal learning [4]-[7], where they test the hypothesis that people represent causal knowledge similarly to as it does a Bayesian network. Other theoretical developments have made possible to extend the use of Bayesian network outside the domain of causal knowledge, for instance, on inductive generalization of concepts and learning words [8]. Another facet of Bayesian networks which have proved useful tools is in market research [9]. Others found that a Bayesian network classifies customers of a company relative to the perspective of long-term purchases [10]. On the other hand, it analyzed the nature of human emotions and the negative behavior resulting from overcrowding during mass events. Utilizing the Bayesian network, his model shows the dependence structure between different emotions and negative behaviors of pilgrims in the crowd [11].

Finally, by its shaping power, the Bayesian networks could be used to generate more and better models of how organizations, groups or social aggregate which is the subject of study for psychology. Thus the aim of this study was to determine the underlying causal structure of the elements of happiness from a set of empirically obtained data. It is considered a Bayesian statistical parameter, which is inferred as an uncertain event, in this study the knowledge about happiness is not accurate and is subject to uncertainty, therefore happiness can be described by a probability distribution. Whereby the Metropolis algorithm used is a specific type of process Monte Carlo, which generates a random way so that every step along the way is completely independent of the previous steps of the current position and generates Markov chains. Process in each step has not memory of the previous states. This is known as Markov Chain Monte Carlo (MCMC)

Happiness has been defined as an entity that can be described by a specific set of measures [12], [13], a mental state that people can gain control over in a cognitive way to perceive and conceive both themselves and their world as an experience of joy, satisfaction or positive welfare [14]. Unfortunately, terms like happiness have been used frequently in daily discourse and may now have vague and somewhat different meanings.

The difficulty of defining happiness has led pioneer psychologists in the study of happiness propose the term subjective well-being (SWB). SWB refers to people's evaluations of their own lives and encompasses both cognitive judgments of satisfaction and affective appraisals of moods and emotions. This conceptualization emphasizes the subjective nature of happiness and holds individual human beings to be the best judges of their own happiness [15]-[17].

There is empirical evidence indicating that well-being is a much broader construct than stability of emotions and subjective judgment about life satisfaction; e.g. situational models consider that the sum of happy moments in life results in the satisfaction of people [18], that is, a person exposed to a greater amount of happy events will be more satisfied with his or her life. People briefly react to good and bad events, but in a short time they return to neutrality. Thus, happiness and unhappiness are merely short-lived reactions to changes in people [19] and it depends strongly on intentional activity [20], [21]. But also it has been identified the temperament is suggested to influence happiness [22], [23] and other personality traits such as optimism and self-esteem [24]-[26], the self-determination [27].

Scholars have noticed that happiness is not a single thing, but it can be broken down into its constituent elements. Considering this information, Alarcón [28] proposes to study happiness as a multidimensional construct which converges satisfaction of what has been achieved, positive attitudes toward life, (experiences that reflect positive feelings concerning one's self and life) personal fulfillment, and joy of living.

Happiness Scale of Lima (HSL) [28], consists of 27 items and reported item-scale correlations were highly

significant and high internal consistency ($\alpha = 0.92$). The factorial analysis of principal components and varimax rotation, revealed that happiness is a multidimensional behavior, consisting of four dimension: X_1 Sense positive of life, X_2 Satisfaction with life, X_3 Personal fulfillment and X_4 Joy of living.

2. The Model

Definition 1

The σ -algebra is used for measured definition. A probability measure is a mapping $\mathbb{P} : \mathcal{F} \rightarrow [0,1]$ that assigns a probability $\mathbb{P}(A) \in [0,1]$ to any event, with the properties $\mathbb{P}(\Omega) = 1$ and $\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$ whenever $A_k \cap A_l = \emptyset$. A real value random variable is a mapping $X : \Omega \rightarrow \mathbb{R}$ and $\omega \mapsto X(\omega)$

Definition 2

A random variables es given by $X : \Omega \rightarrow \mathbb{R}$ where A is measurable and subset of \mathbb{R} , we denote by $\{X \in A\}$ then the event is $\{X \in A\} = \{\omega \in \Omega : X(\omega) \in A\}$.

Definition 3

The random samples, say $X_1 = \{X_{1i}, \dots, X_{ni}\}$, $X_2 = \{X_{2i}, \dots, X_{2i}\}$, $X_3 = \{X_{3i}, \dots, X_{3i}\}$ y $X_4 = \{X_{4i}, \dots, X_{4i}\}$, have the same underlying distribution.

We assume that we have four data samples $\{X_1, X_2, X_3, X_4\}$. If the data have the same distributions, then $X_1 \triangleq X_2 \triangleq X_3 \triangleq X_4$. Uncertainty about the true value of parameter it described by a measurement Π conditional observed datasets. The posterior predictive distribution, is defined as

$$P_X(A \in \Omega) = \sum_{\Theta} P_{\Theta}(A \in \Omega) \Pi_X \tag{1}$$

Lema 1.

Let \mathbb{P} be a probability measure on $(\Omega_1 \times \Omega_2 \times \Omega_3 \times \Omega_4, \mathcal{F}_1 \otimes \mathcal{F}_2 \otimes \mathcal{F}_3 \otimes \mathcal{F}_4)$ such that

$$\mathbb{P}(\{\omega_1, \omega_2, \omega_3, \omega_4\} : I_i(\omega_i) = 1) = \mathbb{P}_i(I_i = 1) \quad i = 1, 2, 3, 4$$

The probability for $X_1 \triangleq X_2 \triangleq X_3 \triangleq X_4$ denoted $\mathbb{P}(X_1 \triangleq X_2 \triangleq X_3 \triangleq X_4)$ is equal to the joint probability $\mathbb{P}(I_1 = 1, I_2 = 1, I_3 = 1, I_4 = 1)$ for any fixed $\|P_{X_1} \wedge P_{X_2} \wedge P_{X_3} \wedge P_{X_4}\|$, the probability is bounded by

$$\max\{0, 4\|P_{X_1} \wedge P_{X_2} \wedge P_{X_3} \wedge P_{X_4}\| - 1\} \leq \mathbb{P}(X_1 \triangleq X_2 \triangleq X_3 \triangleq X_4) \leq \|P_{X_1} \wedge P_{X_2} \wedge P_{X_3} \wedge P_{X_4}\| \tag{2}$$

Consider the problem of selecting independent samples from several populations for the purpose of between-group comparisons, either through hypothesis testing or estimation of mean differences. A companion problem is the estimation of within-group mean levels. Together, these problems form the foundation for the very common analysis of variance framework, but also describe essential aspects of stratified sampling, cluster analysis, empirical Bayes, and other settings. Procedures for making between-group comparisons are known as multiple comparisons methods. The goal of determining which groups have equal means requires testing a collection of related hypotheses [29].

Consider independent samples from l normally distributed populations with equal variances (3a y 3b)

$$X_{11}, X_{12}, \dots, X_{1n_1} \sim N(\mu_1, \sigma^2) \tag{3a}$$

$$X_{l1}, X_{l2}, \dots, X_{ln_l} \sim N(\mu_l, \sigma^2) \tag{3b}$$

Then $H^{(a,b)} \rightarrow \mu_a = \mu_b$ for each (a,b) then there is $l(l-1)/2$ tests. For example, for two sample there is 1 test, for three sample there is 3 test, and 4 there is 6 test. When two or more means are taken as equal, we merely combine all relevant samples into one.

Let Y_n denote the observed data. Assume that Y_n is to be described using a model M_k selected from a set of candidate models $\{M_1, \dots, M_l\}$. Assume that each M_k is uniquely parameterized by θ_k , an element of the

parameter space $\Theta(k)$. In the multiple comparisons problem, the class of candidate models consists of all possible mean level clustering. Each candidate model is parameterized by the mean vector $\mu = (\mu_1, \dots, \mu_l)$ and the common variance σ^2 , with the individual means restricted by the model defined clustering of equalities. That is, each model determines a corresponding parameter space where particular means are taken as equal.

Let $L_{X_1}(\theta_1 | Y_1), \dots, L_k(\theta_k | Y_n)$, where θ take values in a d-dimensional parameter space Θ , be likelihood [30]-[33] functions associated with the samples and $\pi(k)$ denote a prior density on Θ over the model $\{M_1, \dots, M_l\}$. Then $g(\theta_k | k)$ denote a prior on θ_k given the model M_k . The posterior probability for M_k and θ_k can be written as

$$\mathbb{P}(M_n | Y_n) = \frac{\pi(k)g(\theta_k | k)L_k(\theta_k | Y_n)}{\sum_{k=1}^n \mathbb{P}(M_k) p_{M_k}(Y_k)} \tag{4}$$

$$B_k = -\ln L(\hat{\theta} | Y_n) + \dim(\theta_k) \ln(n) \tag{5}$$

You can write the posterior probability (4) y (5) as,

$$\pi_k(k | Y_n) = h(Y_n)^{-1} \pi(k) \int_{\Theta_k} g(\theta_k | k) L(\theta_k | Y_n) d\theta_k \tag{6}$$

We use a uniform prior for $\pi_k = \frac{1}{L} \forall k$ (6) it can be written as

$$\pi_k(k | Y_n) = h(Y_n)^{-1} \int_{\Theta_k} L(\theta_k | Y_n) d\theta_k$$

$$\pi_k(k | Y_n) = \int_{\Theta_k} \frac{L(\theta_k | Y_n)}{h(Y_n)} d\theta_k$$

we using equations (5,6) and the previous development, we write the following relationship

$$-2 \ln \pi_k(k | Y_n) = -2 \ln \int L(\theta_k | Y_n) d\theta_k + c_n \tag{7}$$

Definition 4

The Taylor series of a real or complex-valued function $f(x)$ that is infinitely differentiable at a real or complex number a is the power series

$$f(a) = \sum_{n=0}^{\infty} \frac{d^n f}{n!} (x-a)^n$$

Then, applying the definition 4 to (7), we obtain

$$-\int \left\{ L(\hat{\theta}_k | Y_n) + \frac{\partial \ln L(\hat{\theta}_k | Y_n)}{\partial \theta_k} + \frac{\partial^2 \ln L(\hat{\theta}_k | Y_n)}{\partial \theta_k \partial \hat{\theta}_k} + \dots \right\} d\theta_k + c_n \tag{8}$$

The Fisher information [30] is a way of measuring the amount of information that an observable random variable X carries out unknown parameter θ .

$$E \left[\frac{\partial}{\partial \theta} \log f(X; \theta) | \theta \right] = 0 \tag{9}$$

Applying (9) to (8), is obtained

$$\int L(\theta_k | Y_n) d\theta_k \approx L(\theta_k | Y_n) \int \exp \left\{ -\frac{n}{2} (\theta_k - \hat{\theta}_k) \left[\frac{\partial^2 \ln L(\hat{\theta}_k | Y_n)}{\partial \theta_k \partial \hat{\theta}_k} \right] (\theta_k - \hat{\theta}_k) \right\} d\theta_k$$

$$\int L(\theta_k | Y_n) d\theta_k \approx L(\hat{\theta}_k | Y_n) (2\pi)^{\dim(\theta_k/2)} \left\| n \frac{\partial^2 \ln L(\hat{\theta}_k | Y_n)}{\partial \theta_k \partial \hat{\theta}_k} \right\|^{1/2} \tag{10}$$

taken

$$-2 \ln \pi(k | Y_n) = -\ln L(\hat{\theta}_k | Y_n) + \dim(\theta_k) \ln(n)$$

$$B_k = -\ln L(\hat{\theta}_k | Y_n) + \dim(\theta_k) \ln(n)$$

Then

$$\pi(k | Y_n) \approx \exp\left(-\frac{B_k}{2}\right) \tag{11}$$

With respect to the candidate model class $\{M_1, M_2, \dots, M_l\}$, we obtain, the posterior model probabilities

$$\pi(k | Y_n) \cong \frac{\exp\left(-\frac{B_k}{2}\right)}{\sum_{l=1}^L \exp\left(-\frac{B_l}{2}\right)} \tag{12}$$

Many clever methods have been devised for constructing and sampling from arbitrary posterior distributions. Markov chain simulation (also called Markov chain Monte Carlo, or MCMC) is a general method based on drawing values of θ from approximate distributions and then correcting those draws to better approximate the target posterior distribution, $p(\theta | y)$ [34] [35]. The sampling is done sequentially, with the distribution of the sampled draws depending on the last value drawn; hence, the draws form a Markov chain. (As defined in probability theory, a Markov chain is a sequence of random variables $\theta_1, \theta_2, \dots$, for which, for any t, the distribution of θ_{t-1} given all previous θ 's depends only on the most recent value) The key to the method's success, however, is not the Markov property but rather that the approximate distributions are improved at each step in the simulation, in the sense of converging to the target distribution [35].

A z-statistic should be calculated when the standard deviation of the population(s) is known. If the standard deviation is not known, then the standard error must be estimated using the standard deviation of the sample(s). Due to this estimation, we must use the t-distribution which is thicker in the tails to account for estimating the standard error with the sample standard deviation [34].

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \tag{13}$$

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \tag{14}$$

Until now, we have built the theory to apply to the case of four sample data, now what we will do, will be a particular case study, and see that Lema and definitions 1, 2 y 3 are applied naturally like the Equation (14)

Example particular case

Suppose $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$. Select a random sample of size n_i . From i th group ($i = 1, 2, 3, 4$) so sample sizes are (n_1, n_2, n_3, n_4)

$$\bar{x}_i = \sum_{j=1}^{n_i} \frac{x_{ij}}{n_i}$$

$\bar{\bar{x}}$ is the sample mean for the observations in all group combined

$$\bar{\bar{x}} = \frac{\sum_{i=1}^4 \sum_{j=1}^{n_i} x_{ij}}{n_1 + n_2 + n_3 + n_4}$$

$$\bar{\bar{x}} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2 + n_3\bar{x}_3 + n_4\bar{x}_4}{n_1 + n_2 + n_3 + n_4}$$

Variability in the data, the deviation of an individual observation

$$x_{ij} - \bar{\bar{x}} = (x_{ij} - \bar{x}_i) + (\bar{x}_i - \bar{\bar{x}}) = (x_{ij} - \bar{x}_i) + (\bar{x}_i - \bar{\bar{x}})$$

$$\sum_{i=1}^4 \sum_{j=1}^{n_i} (x_{ij} - \bar{\bar{x}})^2 = \sum_{i=1}^4 \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 + \sum_{i=1}^4 \sum_{j=1}^{n_i} (\bar{x}_i - \bar{\bar{x}})^2$$

To test the null hypothesis that the population means are all the same, us the test statistic

$$t = \frac{\left[n_1 (\bar{x}_1 - \bar{\bar{x}})^2 + n_2 (\bar{x}_2 - \bar{\bar{x}})^2 + n_3 (\bar{x}_3 - \bar{\bar{x}})^2 + n_4 (\bar{x}_4 - \bar{\bar{x}})^2 \right] [n_1 + n_2 + n_3 + n_4 - 4]}{3 \left[(n_1 - 1) s_1^2 + (n_2 - 1) s_2^2 + (n_3 - 1) s_3^2 + (n_4 - 1) s_4^2 \right]}$$

Under H_0 , this statistic has t distribution with $k-1$ and $n-k$ degree freedom.

Now four three

Suppose $H_0 : \mu_1 = \mu_2 = \mu_3$. Select a random sample of size n_i . From i th group ($i=1,2,3$) so sample sizes are (n_1, n_2, n_3)

$$\bar{x}_i = \sum_{j=1}^{n_i} \frac{x_{ij}}{n_i}$$

$$\bar{\bar{x}} = \frac{\sum_{i=1}^3 \sum_{j=1}^{n_i} x_{ij}}{n_1 + n_2 + n_3}$$

$$\bar{\bar{x}} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2 + n_3\bar{x}_3}{n_1 + n_2 + n_3}$$

$$\sum_{i=1}^3 \sum_{j=1}^{n_i} (x_{ij} - \bar{\bar{x}})^2 = \sum_{i=1}^3 \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 + \sum_{i=1}^3 \sum_{j=1}^{n_i} (\bar{x}_i - \bar{\bar{x}})^2$$

$$t = \frac{\left[n_1 (\bar{x}_1 - \bar{\bar{x}})^2 + n_2 (\bar{x}_2 - \bar{\bar{x}})^2 + n_3 (\bar{x}_3 - \bar{\bar{x}})^2 \right] [n_1 + n_2 + n_3 - 3]}{2 \left[(n_1 - 1) s_1^2 + (n_2 - 1) s_2^2 + (n_3 - 1) s_3^2 \right]}$$

Then $H_0 : \mu_1 = \mu_2$

$$t = \frac{\left[n_1 (\bar{x}_1 - \bar{\bar{x}})^2 + n_2 (\bar{x}_2 - \bar{\bar{x}})^2 \right] [n_1 + n_2 - 2]}{\left[(n_1 - 1) s_1^2 + (n_2 - 1) s_2^2 \right]}$$

and

$$H_0 : \mu_1$$

$$t = \frac{\left[n_1 (\bar{x}_1 - \bar{\bar{x}})^2 \right] [n_1 - 1]}{\left[(n_1 - 1) s_1^2 \right]}$$

The principle of Bayes model is to compute posteriors bases on specified priors and the likelihood function of data, the four groups of size 1110.

We began with a descriptive model of data from four groups, wherein the parameters were meaningful measures of central tendency, variance, and normality. Bayesian inference reallocates credibility to parameter values that are consistent with the observed data. The posterior distribution across the parameter values gives complete information about which combinations of parameter values are credible. In particular, from the posterior distri-

bution we can assess the credibility of specific values of interest, such as zero difference between means, or zero difference between standard deviations. We can also decide whether credible values of the difference of means are practically equivalent to zero, so that we accept the null value for practical purposes.

The Bayesian posterior distribution can also be used as a complete hypothesis for assessing power, that is, the probabilities of achieving research goals such as rejecting a null value, accepting a null value, or reaching a desired precision of estimation. The power estimation incorporates all the information in the posterior distribution by integrating across the credible parameter values, using each parameter-value combination to the extent it is credible. **Figure 1** shows histograms of data that are labeled with X_i on their abscissas, and these data are fixed at their empirically observed values.

3. Bayesian Multiple Linear Regression

(x_{i1}, \dots, x_{ik}) now, we need to find an adjustment function for the above data. A linear regression model where more than one variable involved is called multiple regression model [36]-[38]

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon \tag{15}$$

where, we have k regressors, parameters β they are regression coefficients, then the Approach Bayesian Multiple Linear Regression is [36] [37]

$$(Y | X, \beta, \sigma^2) \sim N(X, \beta, \sigma^2 \delta_n) \tag{16}$$

The prior distribution of (β, σ^2) is NIG (Normal-Inverse-Gamma) and it is given by (b_0, B_0, n_0, S_0) thus b_0, B_0, n_0, S_0 they are hyperparameters.

$$(\beta | \sigma^2, Y, X) \sim N(b_1, \sigma^2 B_1) \tag{17}$$

$$(\beta | Y, X) \sim t_{n1}(b_1, S_1 B_1).$$

$$(\sigma^2 | \beta, Y, X) \sim IG\left(\frac{n_1}{2}, \frac{n_1 S_{11}}{2}\right).$$

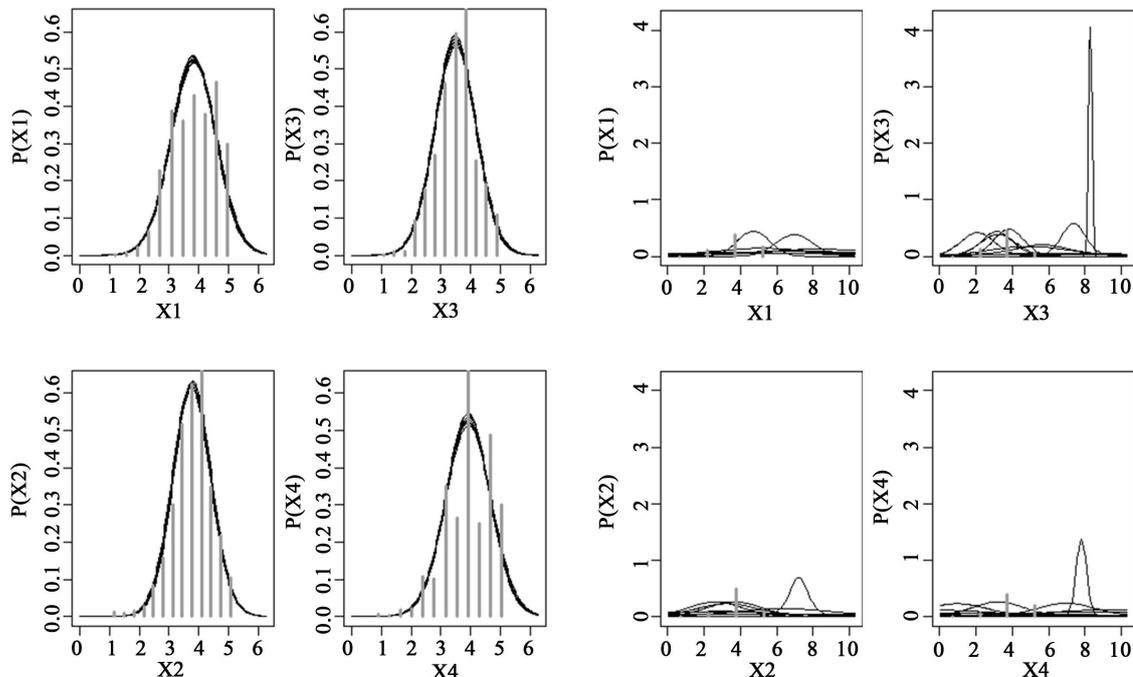


Figure 1. Posterior predictive probability dimensions.

$$(\sigma^2 | Y, X) \sim IG\left(\frac{n_1}{2}, \frac{n_1 S_{11}}{2}\right).$$

where IG (Inverse Gaussian)

If we denote by $\omega = (Y, X)$ y $\theta = (\mu, \delta, \beta, \gamma)$

We can express the density function as

$$f(w) = g(\theta)h(w)\exp\left(\sum_{j=1}^4 \varphi_j(\theta)t_j\right). \tag{18}$$

The conjugate prior distribution will be given by

$$p(\theta) \approx g(\theta)^{a_0} \exp\left(\sum_{j=1}^4 \varphi_j(\theta)a_j\right). \tag{19}$$

In the scheme MCMC

$$Y | \theta_1 \sim N_n(A_1\theta_1, C_1). \tag{20}$$

where A_1 is a matrix ($n \times p$) and C_1 is the covariance matrix and θ_1 is a vector dimension p of regression coefficients to do Bayesian estimation. Assume the prior distribution of θ_1 is

$$\theta_1 \sim N_p(A_2\theta_2, C_2). \tag{21}$$

where A_2 and C_2 are matrices representing our beliefs about average and covariance of prior distribution. Take $A_2 = I$ and using Bayesian approach, we obtain

$$\theta_1 | Y \sim N_p(Dd, D). \tag{22}$$

where

$$d = A_1' C_1^{-1} Y + C_2' A_2 \theta_2.$$

$$D^{-1} = A_1' C_1^{-1} A_1 + C_2^{-1}.$$

We can see that both matrices have two terms, one that only it depends on the prior and other that only it depends of data. This is very useful because in each iteration. We have to update only the last term. The question to be dealt is the choice of hyperparameters C_2 y θ_2

$$y_i | X_i, z_i \sim N(X_i \mathbf{b} + \beta_{k+1} z_i, z_i).$$

$$z_i \sim IG(\gamma, \delta)$$

Then, we can write the equation as follows [38]

$$y(i) = \beta_0 + \sum_{j \neq [i]}^n \beta_{[j]} x_{[j]}(i) + \beta_{cov.x_{cov}}(i) \tag{23}$$

Figures 2-5 show the range HDI, which means Higher Density Interval. Values inside HDI have a greater probability density (credibility) that values outside this. Therefore, the 95% HDI includes the most incredible parameters values. There is a way that the posterior 95% HDI could exclude zero even when the data have a frequency of zero. It can happen if the prior already excludes zero. This interval is useful as a summary of the distribution and decision tool. The decision rule is simple. Any value outside of the 95% HDI is rejected [38].

In all four cases the focus is on assessing if the predictors were differentially predictive, for which we examine the posterior distribution of the differences standardized regression coefficients, given that the comparison is based in the normalization using the single sample.

As shown in **Figure 2**, the data to $X_1 \sim f(X_2, X_3, X_4)$, none of the coefficients are within the supposed. We observed than for X_1 (sense positive of life) for differences in β are out of interval to be credible, indicating than $X_1 \sim f(X_2, X_3, X_4)$, cannot be a linear combination of the other dimensions.

In **Figure 4**, the outcomes for dataset $X_2 \sim f(X_1, X_3, X_4)$ where X_2 (satisfaction with life) $\beta_2 - \beta_3 = 0.301$,

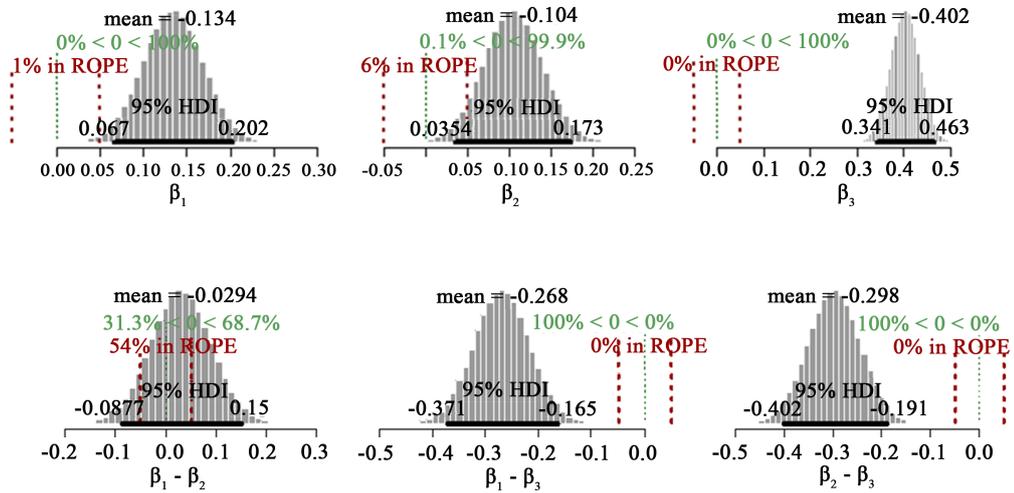


Figure 2. Posteriori distribution for X_1 .

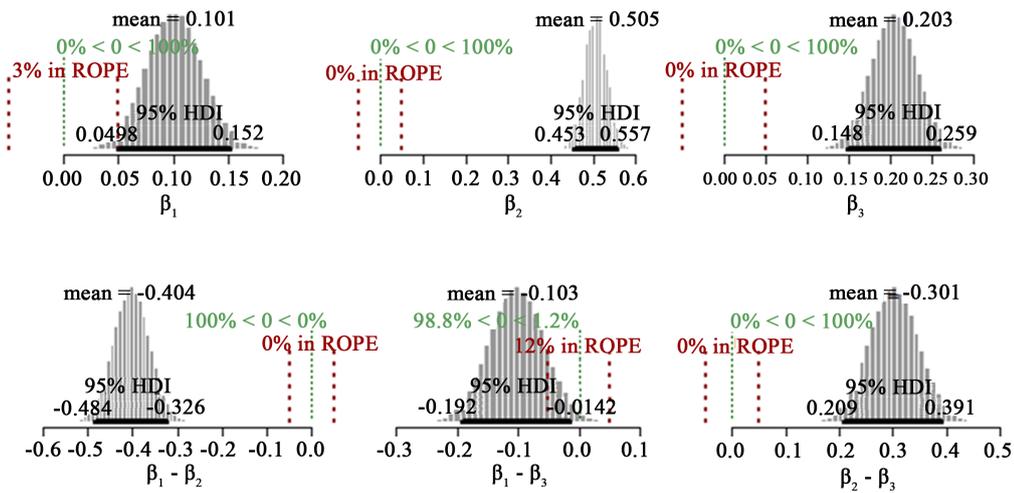


Figure 3. Posteriori distribution for X_2 .

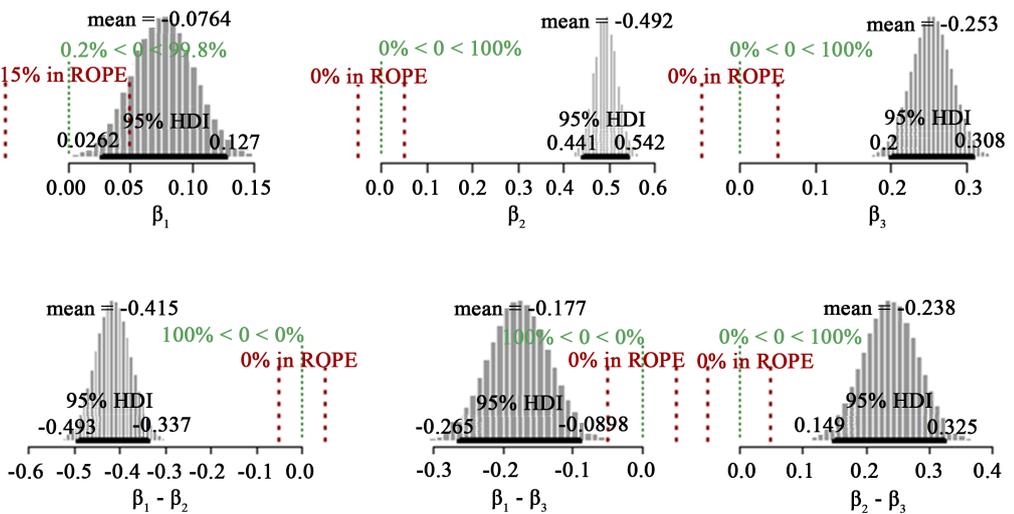


Figure 4. Posteriori distribution for X_3 .

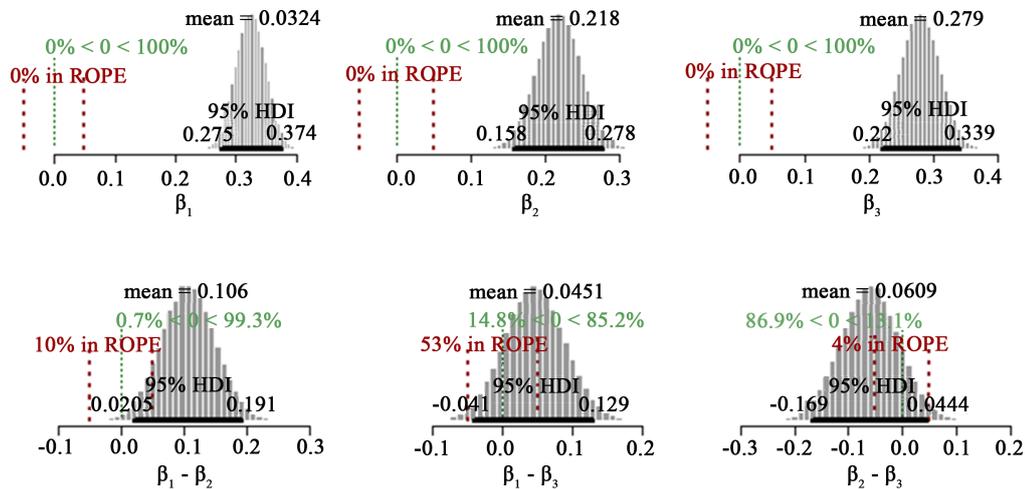


Figure 5. Posteriori distribution for X_4 .

HDI goes from 0.209 to 0.391. Therefore X_1 and X_3 are equally costly and if we want to avoid double cost of measuring both, then, is probable to be more effective assess X_2 than X_3 .

In Figure 5 the outcomes for dataset $X_3 \sim f(X_1, X_2, X_4)$, where X_3 (personal fulfillment) $\beta_2 - \beta_3 = 0.238$, HDI goes from 0.149 to 0.325. Therefore X_1 and X_2 are equally costly and if we want to avoid double cost of measuring both, then, is probable to be more effective assess X_1 than X_4 .

Figure 5 shows for dataset $X_4 \sim f(X_1, X_2, X_3)$, where X_4 (joy of living) $\beta_1 - \beta_2 = 0.106$, HDI goes from 0.0205 to 0.191. Therefore X_1 and X_2 are equally costly, and if we want to avoid double cost of measuring both, then, is probable to be more effective assess X_1 than X_2 .

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire journals, and not as an independent document. Please do not revise any of the current designations.

4. Conclusions

Bayesian methods have been developed as a tool for reasoning quantitatively in situations where arguments cannot be made with certainty. The focus recent developments of Markov chain Monte Carlo algorithms, in many situations the only way to integrate over the parameter space. The use of posterior predictive distributions makes the method robust to the choice of priors on the model parameters and enables the use of improper priors even when only very few observations are available. To measure the agreement between posterior predictive distributions, we derive a measure which has an intuitive probabilistic interpretation.

Within the discussion of the results, we found that the sample has not atypical behavior, too, and that suitable modifications can be described through a test. Another interesting result obtained is that the predictive probability for the case of X_1 (sense positive of life) and X_3 (personal fulfillment) dimensions exhibit a non-uniform variation, while other factors are uniform distribute.

The hypotheses of work, was that if through sample analysis could infer that happiness, only one is affected by three dimensions X_2 (sense positive of life), X_3 (personal fulfillment) and X_4 (satisfaction with life). In this context, we note that the hypothesis was tested, the marked tendency on distributions in recent factors was sufficient to support this theory, on the other hand, through the Multilinear Regression Bayesian, also tested this hypothesis.

Due to recent revolutionary advances in Bayesian posterior computation via computer-intensive MCMC simulation techniques, difficulties with posterior computations can be overcome. A Bayesian state-space model is readily implemented using standard Bayesian software such as JAGS, BUGS, NIMBLE and STAN. One can therefore avoid writing one-off programs in a low-level language. Any modifications, such as different prior

distributions, applications to different data sets, or the use of different sampling distributions, require the change of just a single line in the code.

References

- [1] Pearl, J. (2001) Bayesian Networks, Causal Inference and Knowledge Discovery (Tech. Rep. R-281). University of California, Los Angeles.
- [2] Cowell, R.G., Dawid, A.P., Lauritzen, S.L. and Spiegelhalter, D.J. (1999) Probabilistic Networks and Expert Systems. Springer, Harrisonburg.
- [3] Edwards, W. (1998) Hailfinder. Tools for and Experiences with Bayesian Normative Modelling. *American Psychologist*, **53**, 416-428. <http://dx.doi.org/10.1037/0003-066X.53.4.416>
- [4] Gopnik, A., Glymour, C., Sobel, D.M., Schulz, L.E., Kushnir, T. and Danks, D. (2004) A Theory of Causal Learning in Children: Causal and Bayes Nets. *Psychological Review*, **111**, 3-32. <http://dx.doi.org/10.1037/0033-295X.111.1.3>
- [5] Gopnik, A. and Schulz, L. (2004) Mechanisms of Theory Formation in Young Children. *Trends in Cognitives Sciences*, **8**, 371-377. <http://dx.doi.org/10.1016/j.tics.2004.06.005>
- [6] Sobel, D.M., Tenenbaum, J.B. and Gopnik, A. (2004) Children's Causal Inferences from Indirect Evidence: Backwards Blocking and Bayesian Reasoning in Preschoolers. *Journal of Experimental Psychology*, **130**, 380-400. http://dx.doi.org/10.1207/s15516709cog2803_1
- [7] Waldman, M.R. and Hagmayer, Y. (2005) Seeing versus Doing: Two Models of Accessing Causal Knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **31**, 216-227. <http://dx.doi.org/10.1037/0278-7393.31.2.216>
- [8] Tenenbaum, J.B., Griffiths, T.L. and Kemp, C. (2006) Theory-Based Bayesian Models of Inductive Learning and Reasoning. *Trends in Cognitive Science*, **10**, 309-318. <http://dx.doi.org/10.1016/j.tics.2006.05.009>
- [9] Nadkarni, S. and Shenoy, P.P. (2001) A Bayesian Network Approach to Making inferences in Causal Maps. *European Journal of Operational Research*, **128**, 479-498. [http://dx.doi.org/10.1016/S0377-2217\(99\)00368-9](http://dx.doi.org/10.1016/S0377-2217(99)00368-9)
- [10] Baesens, B., Verstraeten, G., Van den Poel, D., Egmont-Petersen, M., Van Kenhove, P. and Vanthienen, J. (2004) Bayesian Classifiers for Identifying the Slope of the Customer Life Cycle of Long-Life Customer. *European Journal of Operational Research*, **156**, 508-523. [http://dx.doi.org/10.1016/S0377-2217\(03\)00043-2](http://dx.doi.org/10.1016/S0377-2217(03)00043-2)
- [11] Ramli, N., Abdul G., Ahmad, H., Mohd, H., Sulong, J., Mohd, M., Abd, R. and Mat, S. (2014) Bayesian Network Model of Crowd Emotion and Negative Behavior. *AIP Conference Proceeding*, **163**, 867. <http://dx.doi.org/10.1063/1.4903685>
- [12] Seligman, M. (2011) Flourish. A Visionary New Understanding of Happiness and Well-Being. Free Press, New York.
- [13] Seligman, M., Steen, T., Park, N. and Peterson, C. (2005) Positive Psychology Progress: Empirical Validation of Interventions. *American Psychologist*, **60**, 410-421. <http://dx.doi.org/10.1037/0003-066X.60.5.410>
- [14] Lyubomirsky, S. (2008) La ciencia de la Felicidad. Un método probado para conseguir el bienestar. Ediciones Urano, Santiago.
- [15] Diener, E. (2000) Subjective Well-Being: The Science of Happiness and Proposal for a National Index. *American Psychologist*, **55**, 34-43. <http://dx.doi.org/10.1037/0003-066X.55.1.34>
- [16] Diener, E. (2009) Subjective Well-Being. In: Diener, E., Ed., *The Science of Well-Being: The Collected Works of Ed Diener, Social Indicators Research Series 37*, Springer Science + Bussiness Media B. V, 11-58. http://dx.doi.org/10.1007/978-90-481-2350-6_2
- [17] Diener, E., Sandvik, E. and Pavot, W. (1991) Happiness Is the Frequency, Not the Intensity, of Positive versus Negative Affect. In: Strack, F., Argyle, M. and Schwarz, N., Eds., *Subjective Well-Being: An Interdisciplinary Perspective*, Pergamon, New York, 119-139.
- [18] Veenhoven, R. (1994) Is Happiness a Trait? *Social Indicators Research*, **32**, 101-160. <http://dx.doi.org/10.1007/BF01078732>
- [19] Lyubomirsky, S., Sheldon, K.M. and Schkade, D. (2005) Pursuing Happiness: The Architecture of Sustainable Change. *Review of General Psychology*, **9**, 111-131. <http://dx.doi.org/10.1037/1089-2680.9.2.111>
- [20] Sheldon, K.M. and Lyubomirsky, S. (2006) Achieving Sustainable Gains in Happiness: Change Your Actions, Not Your Circumstances. *Journal of Happiness Studies*, **7**, 55-86. <http://dx.doi.org/10.1007/s10902-005-0868-8>
- [21] Diener, E., Helliwell, F.B. and Kahneman, D. (2010) International Differences in Well-Being. Oxford University Press, New York. <http://dx.doi.org/10.1093/acprof:oso/9780199732739.001.0001>
- [22] Diener, E. and Lucas, R. (2008) Personality and Subjective Well-Being. In: John, O., Robins, R. and Pervin, L., Eds., *Handbook of Personality*, 3rd Edition, Guilford, New York, 795-814.

- [23] Lucas, R. and Fujita, F. (2000) Factors Influencing the Relation between Extraversion and Pleasant Affect. *Journal of Personality and Social Psychology*, **79**, 1039-1056.
- [24] Lucas, R., Diener, E. and Suh, E. (1996) Discriminant Validity of Well-Being Measures. *Journal of Personality and Social Psychology*, **71**, 616-628. <http://dx.doi.org/10.1037/0022-3514.71.3.616>
- [25] Peterson, C. (2000) The Future of Optimism. *American Psychologist*, **55**, 44-55. <http://dx.doi.org/10.1037/0003-066X.55.1.44>
- [26] Schimmack, U. and Diener, E. (2003) Predictive Validity of Explicit and Implicit Self-Esteem for Subjective Well-Being. *Journal of Research in Personality*, **37**, 100-106.
- [27] Ryan, R.M. and Deci, E. (2000) Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist*, **55**, 68-78. <http://dx.doi.org/10.1037/0003-066X.55.1.68>
- [28] Alarcón, R. (2006) Desarrollo de una Escala Factorial para medir Felicidad. *Interamerican Journal of Psychology*, **40**, 99-106.
- [29] Blomstedt, P., Gauriot, R., Viitala, N., Reinikainen, T. and Corander, J. (2014) Bayesian Predictive Modeling and Comparison of Oil Samples. *Journal of Chemometrics*, **28**, 52-59. <http://dx.doi.org/10.1002/cem.2566>
- [30] Yin, Y. and Li, B. (2014) Analysis of the Behrens-Fisher Problem Based on Bayesian Evidence. *Journal of Applied Mathematics*, **2014**, Article ID: 978691. <http://dx.doi.org/10.1155/2014/978691>
- [31] Myung, J.I. (2003) Tutorial on Maximum Likelihood Estimation. *Journal of Mathematical Psychology*, **47**, 90-100. [http://dx.doi.org/10.1016/S0022-2496\(02\)00028-7](http://dx.doi.org/10.1016/S0022-2496(02)00028-7)
- [32] Mossel, E. and Tamuz, O. (2010) Iterative Maximum Likelihood on Networks. *Advances in Applied Mathematics*, **45**, 36-49. <http://dx.doi.org/10.1016/j.aam.2009.11.004>
- [33] Mengersen, K.L., Pudlo, P. and Robert, P.C. (2013) Bayesian Computation via Empirical Likelihood. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 1321-1326. <http://dx.doi.org/10.1073/pnas.1208827110>
- [34] Lensen, J.L. (1987) A Note on Asymptotic Expansions for Markov Chains Using Operator Theory. *Advances in Applied Mathematics*, **8**, 377-392. [http://dx.doi.org/10.1016/0196-8858\(87\)90016-9](http://dx.doi.org/10.1016/0196-8858(87)90016-9)
- [35] D'Angeli, D. and Donno, A. (2013) The Lumpability Property for a Family of Markov Chains on Poset Block Structures. *Advances in Applied Mathematics*, **51**, 367-391. <http://dx.doi.org/10.1016/j.aam.2013.04.007>
- [36] Mitchell, T.J. and Beauchamp, J.J. (1988) Bayesian Variables Selection in Linear Regression. *Journal of the American Statistical Association*, **83**, 1023-1032. <http://dx.doi.org/10.1080/01621459.1988.10478694>
- [37] Hoeting, J., Raftery, A.E. and Madigan, D. (1996) A Method for Simultaneous Variables Selection and Outlier Identification in Linear Regression. *Computational Statistics & Data Analysis*, **22**, 251-270. [http://dx.doi.org/10.1016/0167-9473\(95\)00053-4](http://dx.doi.org/10.1016/0167-9473(95)00053-4)
- [38] Kruschke, J.K. (2015) *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. Academic Press/Elsevier, Waltham.