JP3I (Jurnal Pengukuran Psikologi dan Pendidikan Indonesia), 10(2), 2021, 99-117
DOI: http://dx.doi.org/10.15408/jp3i.v10i2.20185
http://journal.uinjkt.ac.id/index.php/jp3i

Bayesian Statistics in Psychological Research

Edwin Adrianta Surijah[1,2], I Made Feby Anggara[2]

The Queensland University of Technology, Faculty of Health, School of Psychology and Counselling, Brisbane, Qld, Australia[1]

Universitas Dhyana Pura, Program Studi Psikologi, Badung, Bali, Indonesia[2]

edwin@adrianta.com

Abstract

One of the key developments in psychological data analysis is the Bayesian implementation. This article aims to introduce Bayesian statistics application in psychological research. A data set of Marital Satisfaction and Positive Affect (n = 200) became an example to compare the regression results based on frequentist and Bayesian statistics. The data analysis examined the influence of positive affect on marital satisfaction. Based upon the prior information and observed data, results suggest that the average of the distribution of the posterior coefficient of positive affect is .31, with a deviation standard of .01 and a credible interval ranging from .30 to .33. The study's results present the unique approach in interpreting the Bayesian result. This article also outlines diagnostic steps to obtain a robust Bayesian result and avoid misuse of Bayesian statistics. Finally, discussions cover the probability principle in Bayesian analysis and how to interpret its result to encourage Indonesian psychological scientists to implement Bayesian as an alternative to data analysis.

Keywords: Bayesian, probability, regression, statistics.

Abstrak

Salah satu perkembangan analisis data dalam penelitian psikologi adalah penggunaan analisis Bayesian. Tulisan ini bertujuan memperkenalkan implementasi Bayesian statistics dalam konteks penelitian psikologi. Sebuah set data Kepuasan Pernikahan dan Afek Positif (n = 200) menjadi contoh untuk menunjukkan perbandingan hasil analisis regresi berdasarkan statistika frekuentis dan Bayesian. Analisis data menguji kontribusi afek positif terhadap kepuasan pernikahan. Berdasarkan informasi terdahulu dan data hasil pengamatan, rerata distribusi probabilitas koefisien afek positif adalah .31 dengan simpangan baku .01 dan selang kepercayaan (credible interval) dari .30 hingga .33. Hasil studi menunjukkan keunikan memaknai hasil analisis Bayesian. Lebih dari itu, tulisan ini juga memaparkan beberapa langkah diagnostik untuk mendapatkan hasil analisis Bayesian yang handal dan mencegah penyalahgunaan analisis data. Diskusi mengenai prinsip probabilitas dalam analisis Bayesian dan cara memaknai hasil analisis menjadi upaya untuk mendorong peneliti psikologi di Indonesia tertarik menggunakan analisis Bayesian sebagai perspektif yang berbeda.

***Kata Kunci**: Bayesian, probabilitas, regresi, statistika.*

Introduction

Psychological research, in general, uses frequentist statistics, although, in the last 25 years, the use in such research of the Bayesian approach has increased (van de Schoot, Winter, Ryan, Zondervan-Zwijnenburg, & Depaoli, 2017). A meta-analysis has shown that there have been at least 1,579 pieces of psychological research using Bayesian analysis between 1990 and 2015, and 46.9% of these were in the form of research-based upon regression (van de Schoot, Winter, Ryan, Zondervan-Zwijnenburg, & Depaoli, 2017)Thus, in general, trends indicate that Bayesian analysis in psychological research is increasing (Andrews & Baguley, 2013).

In Indonesia, the attention paid to Bayesian analysis in psychological research is also increasing. In examinations of the literature, the Bayesian approach has been used to determine types of clinical psychological disturbances (Nurmansyah & Hartati, 2013), or the design of personality tests (Buaton & Astuti, 2014). What is of interest is, these studies use psychological variables, but they are not reported by psychological researchers nor published by psychological journals. This indicates that there is space for introducing the implementation of Bayesian analysis in psychological research for psychologists in Indonesia.

Why would Bayesian analysis be of benefit for psychological research in Indonesia? A review article has comprehensively explained the advantages of Bayesian analysis for psychological researchers (Wagenmakers, *et al.*, 2018). One of the superiorities of Bayesian analysis is the use of previous knowledge in data analysis (Dienes, 2011). By way of illustration, if a researcher intends to test reaction-time scores using frequentist statistics, they will conduct a parametric test of differences (e.g., t-test for an independent sample) of the reaction-time scores, based upon the observed data obtained. In Bayesian analysis, previous knowledge, for instance, the experience of previous researchers, previous research results, or results of meta-analysis, can become part of the Bayesian data analysis, in addition to the results of the observed data obtained by the current researcher.

Another advantage of Bayesian analysis is that the parameters obtained (θ) take the form of distributions. These distributions are known by the terms "*credible intervals*" and "*highest probability densities*" (HPD) (Chen & Shao, 1999). The advantages of these distributions are that they assist the researcher in making more accurate analysis result decisions. The results of *frequentist statistics* analysis often take the form of estimations and *confidence intervals*. In Bayesian analysis, *credible intervals* are in the form of not only value intervals but also distributions. For example, previous knowledge indicates that the results of intelligence testing indicated data distribution in Gaussian form, with an average score of 100. Suppose the results of Bayesian analysis indicate the highest average probability (HPD) of the average parameters of intelligence testing is 100. In that case, however, the form of its *credible intervals* is *uniform* or *exponential.* Therefore, these analysis results may not be trusted as intellectual score probabilities. This indicates the accuracy of the details of Bayesian analysis.

The final advantage of Bayesian analysis is to be found in the matter of the testing of hypotheses. The results of *frequentist statistics* analysis generally accomplish the rejection or acceptance of a *null* hypothesis through an analysis of the size of the *p-values.* However, conventional data analysis cannot show the degree of invariance between observed data and a *null* hypothesis (Rouder, Speckman, Sun, Morey, & Iverson, 2009). Bayesian analysis makes it possible for the researcher to compare probability distributions, on the basis of a *null* hypothesis, and probability distributions based upon an alternative hypothesis. The result of such a comparison is termed the *Bayes factor* (Penny, Mattout, & Trujillo-Barreto, 2007). When the results of a *frequentist statistics* analysis can report the values of the *effect size*, then the analysis can perform a quantification of the support for a *null*, or an alternative, hypothesis.

http://journal.uinjkt.ac.id/index.php/jp3i

Another advantage of Bayesian analysis is the use of simulation for *data sampling* to be able to overcome the weakness inherent in a small sample size (van de Schoot & Miočević, 2020). The Bayesian analysis uses computerized sampling methods to obtain the distribution of results (*posterior*). One of the sampling approaches used is Markov Chain Monte-Carlo (van Ravenzwaaij, Cassey, & Brown, 2018). With this sampling, Bayesian analysis allows the researcher to formulate a simulation model by using the sample from just one distribution (van Ravenzwaaij, Cassey, & Brown, 2018). Because of this, the number of samples in a piece of research depends not only upon those samples resulting from observation but also upon samples resulting from the simulation.

With the various advantages of Bayesian analysis, this article aims to introduce the application of such analysis in the context of psychological research. Furthermore, the article is intended to focus on examples of Bayesian analysis based upon regression as one of the analytical approaches frequently used in psychological research. This study shows and compares the results of regression analysis using the *frequentist statistics* approach and the Bayesian. Through this study, it is hoped that the reader may obtain a picture of the implementation of Bayesian analysis in the context of psychological research. For this reason, this article is limited in that it does not explain the basics of Bayesian principles, nor does it introduce Bayesian analytical software (*e.g.*, JASP). The authors invite the readers to give their attention in greater depth to the book "*Doing Bayesian Data Analysis*" (Kruschke, 2014), to be able to understand Bayesian principles.

Because this article aims to give an introduction or guide, the writers have attempted to present the Bayesian concept, based upon regression analysis, in simple language and draw parallels with the concepts of *frequentist statistics*, known previously. If the reader meets with difficulty in understanding the terminology in this article, a list of terms, with their simple explanations, has been provided in Appendix D. As a beginning, a regression equation in *frequentist statistics* analysis is generally written as follows:

$$y_i = \beta_0 + \beta_1 * x_i + \varepsilon_i \quad (1)$$

with y_i as the outsider/observed variable as a function of the *intercept* (β_0), the *slope* (β_1) the predictor (x_i), and the residual variables (ε). Meanwhile, in Bayesian analysis, this regression equation is understood in the context of probabilities and written using *Bayes' Rule* (Etz & Vandekerckhove, 2018). The *Bayes' Rule* referred to here is as follows:

$$Posterior = \frac{Likelihood \times Prior}{Evidence} \quad (2)$$

$$P(model|data) = \frac{P(data|model) \times P(model)}{P(data)} \quad (3)$$

The *posterior* is the predicted model probability, based on the researcher's data at the time and previous knowledge. *The likelihood* is the probability of a piece of data, based upon the model of the regression equation, as determined by the writers, or, in other words, *likelihood* is the probability of Equation 1. The *prior* probability is that of the model, based upon previous knowledge. *Finally, evidence* is the probability of the data held by the writers/researchers resulting from the measurements y_i and x_i. Based upon this *Bayes' Rule* formula, psychological researchers will later be able to create a general *posterior* regression analysis model, consisting of:

$$y \sim N(\beta^T X . \delta^2 I) \quad (4)$$

This means that the value y is obtained from the normal average/*mean* distribution (N) and its variance. This average or *mean* is a transposition of the matrix of parameter β/*beta* (*intercept* and *slope*),

times the matrix of the predictor variable, and the variance is the square of the deviation standard, multiplied by the Identity matrix (*e.g.* for the identity of the participants). To make the comparison easier, Equation 4 can be understood to mean that the output variable (y) is obtained from the normal distribution of all parameters in Equation 1.

Regression analysis is always compiled from the *intercept, slope,* and *residual* probabilities. In regression analysis, in psychological research, it is generally of interest to know the *slope,* the degree of inclination of the regression equation line, or the impact of the predictor (x) on the external variables (y). By combining Equations 3 and 4, psychological researchers may formulate the *posterior* probability of the research model, as follows:

$$P(\beta|y.x) = \frac{P(y.x|\beta)\times P(\beta)}{P(y.x)} \quad (5)$$

The *posterior* is the probability of the parameter β (intercept and slope) value, based upon data resulting from measurements y and x. Because of this, Bayesian analysis activity involves not only the calculation of the data held (*likelihood*), but also the calculation of the probabilities of a number of the components in Equation 5. To make it easier for the reader to comprehend this sequence, an example is given in the Method section, using y and x, based upon the variables generally found in psychological research.

The variable of marital satisfaction is chosen as the focus and, at the same time, the example in this study to represent a psychological construct. Studies related to marriage are important because marriage is related to psychological welfare (Becker, Kirchmaier, & Trautmann, 2019), and the health of an individual (Kim, Lee, & Park, 2018). In Indonesian, as well as in cross-cultural studies, research into marital satisfaction has also been conducted previously (Az Zahra & Caninsti, 2017; Sorokowski, *et al*., 2017; Surijah & Prakasa, 2020). For this reason, the examples chosen are relevant constructs for the reader or psychological researcher in Indonesia.

This article is also aimed at showing the diagnostic steps resulting from Bayesian analysis. A general guide, such as looking at the values of *R-hat*, *trace plot*, and *density plot* (Abdelkader, 2020; Law, 2019; Gabry & Modrák, 2021), together with the application of the WAMBS-v2 (*when to Worry and how to Avoid the Misuse of Bayesian Statistics*) *checklist* (van de Schoot, *et al*., 2021) were used in this study. The WAMBS *checklist* is a crucial guide for researchers using the Bayesian approach in order to be able to pay critical attention to the results of Bayesian analysis to obtain an accurate interpretation. The stages of Bayesian analysis, in this article, are based upon the general recommendations for such analysis and the recommendations expounded in the WAMBS *checklist* (see Table 1).

Table 1. The Stages of Bayesian Analysis and the Diagnostic Guide

	Stages of the Analysis
1	Construct a model with *non-informative prior*
2	Compare the model with *informative prior*
3	Examine the values of *R-hat* and *trace plot*
4	Run and compare the models with different numbers of iteration and thinning
5	Examine the *effective sample size*
6	Examine the *posterior* distribution by looking at the *density plot*
7	Examine auto-correlation
8	Examine the convergence using the *Gelman-Rubin Convergence diagnostic tool*
9	Determine the values of the *credible intervals*
10	Calculate the *effect sizes* using the *Bayes Factor* or the ROPE
11	Interpret the results using a Bayesian model

http://journal.uinjkt.ac.id/index.php/jp3i

Methods

The details of the information used as an example or illustration of the application of the Bayesian approach in psychological research in Indonesia are related in this section. Further, the sequence for data analysis is explained in brief. A more detailed elaboration concerning the stages of data analysis is explained together with the clarification of the Results. This aside, a simple description of the terms used in Bayesian analysis may be found in Appendix D.

Data and Design

Examples and illustrations in this article use existing data previously held by the authors. These data form an extract from the *daily diary study* research data of husband and wife couples in Indonesia. The design of the original research was that husband and wife couples were to perform *self-reporting* for 30 days. The number of husband and wife couples participating in the research was (n = 200). Research participants reported on the quality of their marriages, their emotional states), their levels of happiness, and other marital relationship indicators. This *self-reporting* used the instrument commonly found in psychological research, *i.e.*, *Likert* scale questionnaires, to which the subjects responded. The principle design of this research had been examined and approved by an ethics committee.

For this article, extracts from these data were taken, in the form of the data from the participants on Marital Satisfaction and Positive Affect on the first day of the diary study. Marital Satisfaction was made the external variable (y), and Positive Affect was made the predictor variable (x). Demographic data and other variables from the original research were not included as *covariates*. For this reason, the *dummy* research question in this study was as follows: "*Does positive affect contribute positively to marital satisfaction?"*

Instruments

Marital Satisfaction was measured using the *Satisfaction with Married Life* scale. This scale is a modified version of the Satisfaction with Life Scale (Diener, Emmons, Larsen, & Griffin, 1985). In the main study, data collection was performed using a scale that had previously been translated into Indonesian (Surijah & Prakasa, 2020). The scale comprised five points, and the participants then allocated values ranging from 1 (*Strongly Disagree*) to 7 (*Strongly Agree*). An example of a question on this scale is: *I am satisfied with my married life.* The reliability coefficient of the original scale was α = .92 (Johnson, Zabriskie, & Hill, 2006), and the reliability coefficient of the scale in the Indonesian language translation was α = .82 (Surijah & Prakasa, 2020).

The main research measured Positive Affect, using the Positive and Negative Affect Schedule (PANAS) scale. This scale consists of 20 items or lists of human emotion/affect components(Zevon & Tellegen, 1982). More particularly, there are ten points that measure Positive Affect, such as *fervency* or *enthusiasm.* The research participants then evaluated how frequently they experienced positive feelings on that particular day, from 1 (*Very Rarely/Not at All*) to 2 (*Rarely*), 3 (*Occasionally*), 4 (*Often*), and 5 (*Very Frequently*). The coefficient of reliability of this scale is α = .86. The original scale was written in English, then later, the principal researchers undertook a translation, using translation standards for research (Beaton, Bombardier, Guillemin, & Ferraz, 2000).

Data Analysis

Data analysis was originally performed using the *frequentist statistics* approach, followed by using Bayesian analysis. This was aimed at facilitating the reader's ability to make a comparison between these two approaches. Later on, inferences based upon the results of the Bayesian analysis (Yau, n.d.) were made, following the guidance in the WAMBS *checklist*. *Frequentist statistics* regression analysis uses the basic function of *R*, *i.e.,* 'lm', whilst Bayesian analysis uses OpenBUGS software (Lunn, Thomas, Best,

& Spiegelhalter, 2000). The term OpenBUGS (an acronym drawn from *Bayesian inference Using Gibbs Sampling*) uses a *sampling* method known as Gibbs *sampling* and may be performed in *R* by using the 'R2OpenBUGS' packet (Sturtz, Ligges, & Gelman, 2005).

R is used as an implementation to assist data analysis by the evaluation of the transparency of the data analysis process. The code used in this data analysis process will be included in the appendix so that the reader can use that code independently. Besides this, OpenBUGS is also a *software* that is often used by Bayesian researchers and bears a resemblance to JAGS (Depaoli, Clifton, & Cobb, 2016), so that the reader can easily choose an alternative approach appropriate to his or her preferences. The reader may use a different approach, such as Stan (Gelman, Lee, & Guo, *Stan: A probabilistic programming language for Bayesian inference and optimization*, 2015), or JASP (Marsman & Wagenmakers, 2017).

Results and Discussion

Results

This section sets forth the results of regression analysis using both *frequentist statistics* and the Bayesian approach. Besides this, the results of Bayesian analysis are tested by diagnostic stages to obtain an accurate result. In addition to the relating of results, explanations related to the rationale for the use of diagnostic stages are also explained in brief in this section.

Table 2. Descriptive Statistics of the Study Variables

	Mean (SD)	**Range**
Marital Satisfaction	5.89 (.44)	1-7
Positive Affect	3.89 (.49)	1-5

Table 2 shows the descriptive statistical results (averaged and deviation standard) of the results of measurements made related to the 200 participants in the research. After that, the process is carried forward, with the relating of the results of linear regression analysis, as follows:

Table 3. Regression Parameter of Marital Satisfaction as a Function of Positive Affect

	β (SE)	**t**	***p-value***	**LL (2.5%)**	**UL (97.5%)**
Intercept	17.61 (.94)	18.71	.00	15.75	19.47
Slope	.30 (.02)	12.68	.00	.25	.35

The results of the regression analysis (see Table 3) indicate that Positive Affect makes a positive contribution to Marital Satisfaction, with an effect size of. 33 ($F = 160.8$, $df = 198$, $R^2 = .44$, $p = .00$). The diagnostic plot (e.g., the residual normality) is also included in the appendices section in order to enable careful attention to be paid to the accuracy of the results of this regression analysis. By using the results of this regression analysis, psychological researchers may be able to answer research questions and reject *null* hypotheses.

In accord with the analysis stages in Table 1, further stages begin to enter Bayesian analysis, with non-informative prior. Non-informative prior (Model 1) means the research has not determined the scale of the parameter values based upon prior knowledge. A regressive model, written using BUGS, may be seen in Appendix B. The results of Bayesian analysis show a *posterior* distribution. This means that the analysis in Table 4 shows a probability distribution of the values of a combination of the parameters of the data held (*likelihood*) and previous (*prior*) knowledge. Model 1 shows a "value" similar to those of the results of *frequentist statistics*. However, before making too quick a conclusion, Bayesian analysis is continued by analysis using informative prior (Model 2).

http://journal.uinjkt.ac.id/index.php/jp3i

Table 4. Posterior Distribution of the Regression Parameter with Bayesian Statistics

	Mean (SD)	*R-hat*	*n.eff*	LL (2.5%)	UL (97.5%)
(1) Non-informative					
Intercept	18.1 (.80)	1.1	65	16.5	19.7
Slope	.30 (.00)	1.1	84	.30	.30
(2) Informative					
Intercept	17.2 (.40)	1	81	16.48	17.91
Slope	.31 (.00)	1	87	.27	.35
(3) Informative					
Intercept	17.08 (.43)	1	89	16.31	17.89
Slope	.31 (.01)	1	88	.29	.34
(4) Informative					
Intercept	17.11 (.40)	1	350	16.33	17.9
Slope	.31 (.01)	1	360	.29	.33
(5) Informative					
Intercept	17.11 (.40)	1	5,400	16.84	17.91
Slope	.31 (.01)	1	6,700	.30	.33

One of the ways to obtain previous knowledge to compile an informative *prior* model is by using assumptions based upon the previous experience of the researcher. If a researcher has no previous experience, a meta-analysis may be conducted first as the most accurate way to discover such information (Van de Schoot, 2015). This study used the investigation of previous research to form an assumption regarding the size of the relationship between Marital Satisfaction and Positive Affect. Table 5 shows the results of the investigations made by the writers. On the basis of the information in Table 5, it was determined that *prior* information for the *slope* parameter was .20 and for the *intercept* parameter was 17.00

Table 5. Investigation of Previous Research as an Organizational Basis for Informative Prior

Study	**Variable**	***Effect Size*/*Beta*/Correlation**
(Johnson, *et al.*, 2005)	Average for Marital Satisfaction ~ Positive Affect - Husband	.17
	Average for Marital Satisfaction ~ Positive Affect - Wife	.24
(Otero, *et al.*, 2020)	Marital Satisfaction ~ Positivity Resonance	r = .25 β = .23
	Marital Satisfaction ~ Positive Affect (joint)	r = .22
	Marital Satisfaction ~ Positive Affect - Husband	r = .06
	Marital Satisfaction ~ Positive Affect - Wife	r = .05

Table 4 shows that Model 2 has a value resembling the results of previous analyses. However, the value of *R-hat* = 1 (the closer to 1, the greater the indication of convergence). This means that the use of informative prior increases the reliability of the results of the analysis. Although this is the case, the value of the *effective sample size* (ESS) of Model 2 is classified as low. Generally, the value of the ESS increasing

will increasingly indicate the availability of a quantity of independent information, and one of the *threshold* values which may be used is ≥ 1,000 (Bürkner, 2017). Because of this, the final conclusion regarding the results of the Bayesian analysis could not yet be obtained.

One of the steps that may be taken after that to increase the reliability of the analysis results is to add the total of the *iterations* of the *sampling* process. The two previous models performed Gibbs *sampling* 1,000 times, preceded by the warming up/practice *sampling* performed 500 times. For that, the number of *samplings* was increased to 2,000 (Model 3). Analysis results indicated no drastic change in the values of the parameter distribution nor the ESS values. After that, the number of samplings was again increased to 5,000 (Model 4). The results of Model 4 indicated that the number of ESS was increased, although its value was lower than 1,000. Because of this, the analyses of the *trace plots* between Models 2 and 4 were then compared.

The *R-hat* and the *trace plot* are diagnostic tools used to indicate the convergence, or uniformity, of the results of each *chain.* Models 1 to 4 use three *chains* (three instances of the Gibbs *sampling* process). In Model 4, the Gibbs sampling process was undertaken on each occasion and performed 5,000 times (*iterations*). The *R-hat* value approaching 1.00 indicates that there is some convergence of results between the three *chains*. Meanwhile, the *trace plot* is a visual diagnostic tool, which provides the size of the value of the *R-hat*. The Bayesian analysis produces a *trace plot* for each parameter. As an example, the appearance of the *trace plot slope* is as follows:

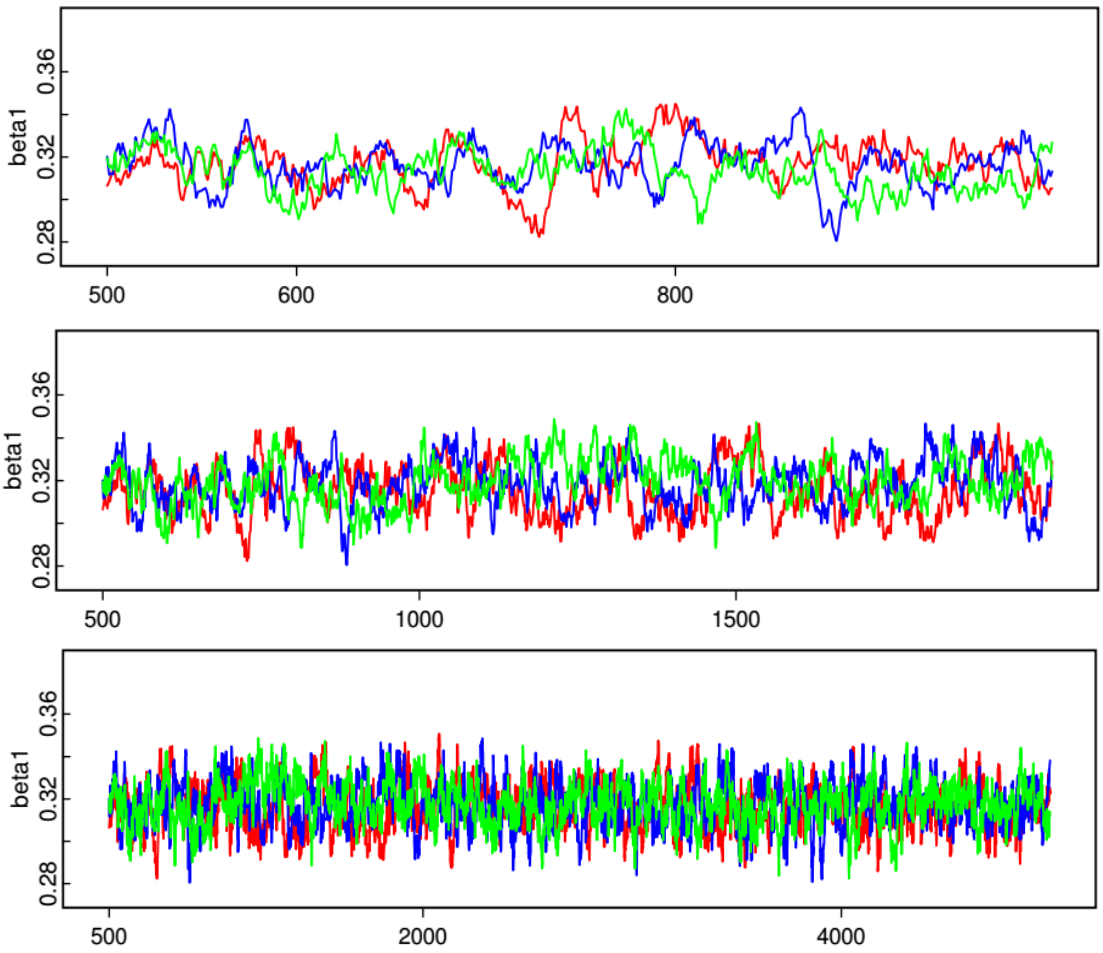


Figure 1. *Example of Trace Plot for Parameter Slope.* The *trace plot* of Model 2 (1,000 *iterations*) is uppermost, and that of Model 4 (5,000 *iterations*) is the one at the very bottom. The *Y-axis* shows the probability of parameter "*beta* 1" (β_1 = *slope*).

http://journal.uinjkt.ac.id/index.php/jp3i

Figure 1 shows that the *trace plot* of Model 2 is moving in the direction of that of Model 4, and the further it goes, the closer it gets. This closeness indicates that the *sampling* results in each *chain* resemble each other so that the size of the β_1 value becomes increasingly trustworthy. This means that, although Models 2 to 4 have a value of *R-hat* = 1.00, nonetheless, the *trace plots* indicate that Model 4 has a more accurate convergence of results. Because of this, Model 4 becomes the reference model for further diagnostic processes.

The Bayesian analysis process, which uses BUGS, is performed using Gibbs *sampling*. For this reason, further diagnostic sampling needs to be adapted to the *sampling* technique which is generally performed for the Bayesian approach, that being that of *Markov Chain Monte Carlo* (MCMC). This adaptation process uses a 'coda' packet on *R* (Plummer, Best, Cowles, & Vines, 2006). The results of the analysis using Gibbs *sampling* may be examined using analysis of the same type as that of the MCMC *sampling* technique. The *R* codes for using the 'coda' packet and the diagnostic analysis can be seen in Appendix C.

The further diagnostic process is related to auto-correlation and *thinning*. Each time the *sampling* process is conducted, there is the possibility of auto-correlation occurring because the *sampling* process is performed thousands of times. One of the ways to reduce auto-correlation is by performing a reduction in the number of samples by conducting *sub-sampling,* known as *thinning*. Repeated testing of Model 4 was performed by conducting *sub-sampling* through *thinning* of every fourth *sampling* (*n.thin* = 4). This model, which was accompanied by *sub-sampling*, did not show significant differences in the parameter values, but the values of the ESS surged above 1,000 (Model 5). This indicated that the researchers had increasingly achieved trustworthy Bayesian analysis models. The specifications of the Gibbs *sampling* for Model 5 may be seen in Appendix C.

The next diagnostic stage is diagnostic auto-correlation analysis. The analysis results show that, in line with the *sampling* process, the auto-correlation of each parameter was convergent on 0 (zero). This indicates that the information obtained was independent and was not caused by any *time-series* effect in the *sampling* process.

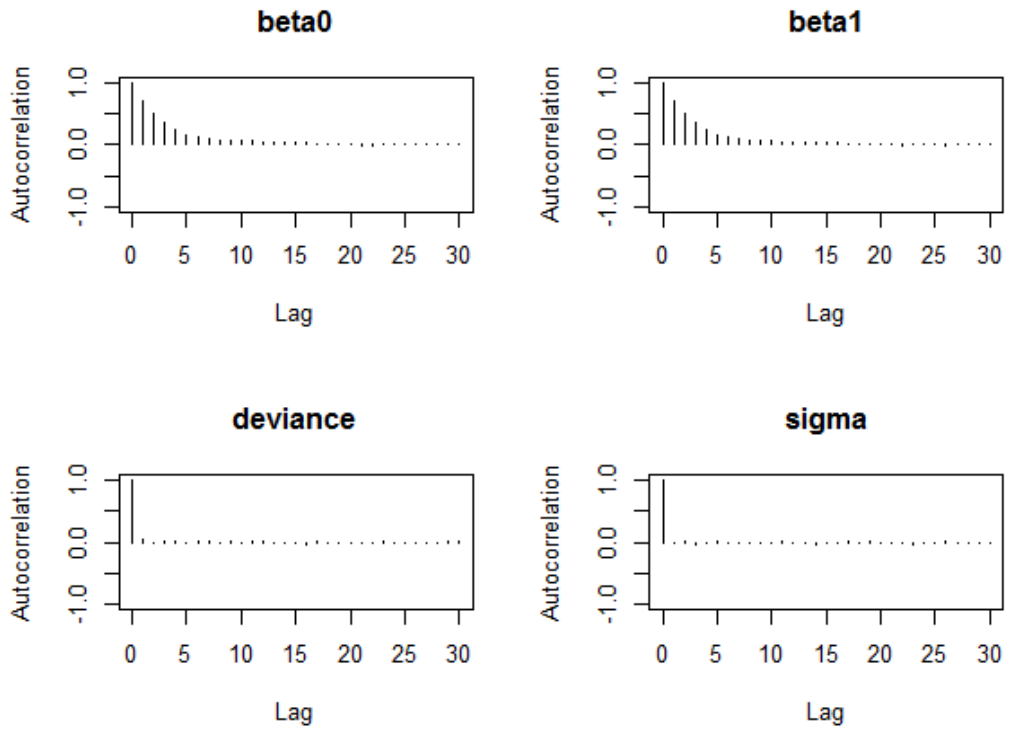


Figure 2. Auto-correlation Diagnostics. The auto-correlation in each parameter increasingly approaches 0 (zero).

The next analysis was that of the *Density Plots*. The Bayesian analysis produces a *density plot for each parameter.* However, this example uses the *density plot* "*beta* 1" as an abbreviated illustration. At the time of the compilation of the *prior*, the Gaussian (normal) distribution for the slope distribution was expected. The results of the *density plot* indicated that the results of the analysis (*posterior*) would also be in Gaussian form and reach the highest peak of probability at the value of .30. This would increasingly support the value of the parameter distribution of Model 5 in Table 4.

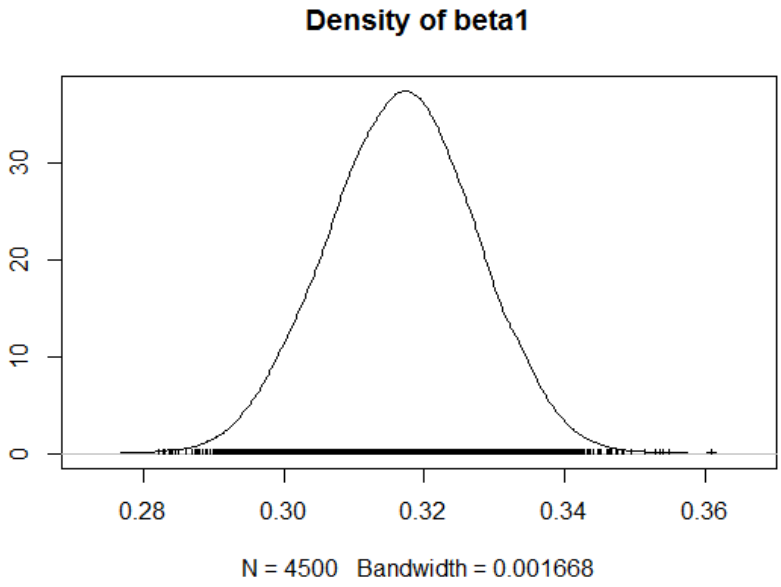


Figure 3. Examples of the *Density Plot* for the *Slope* Parameters

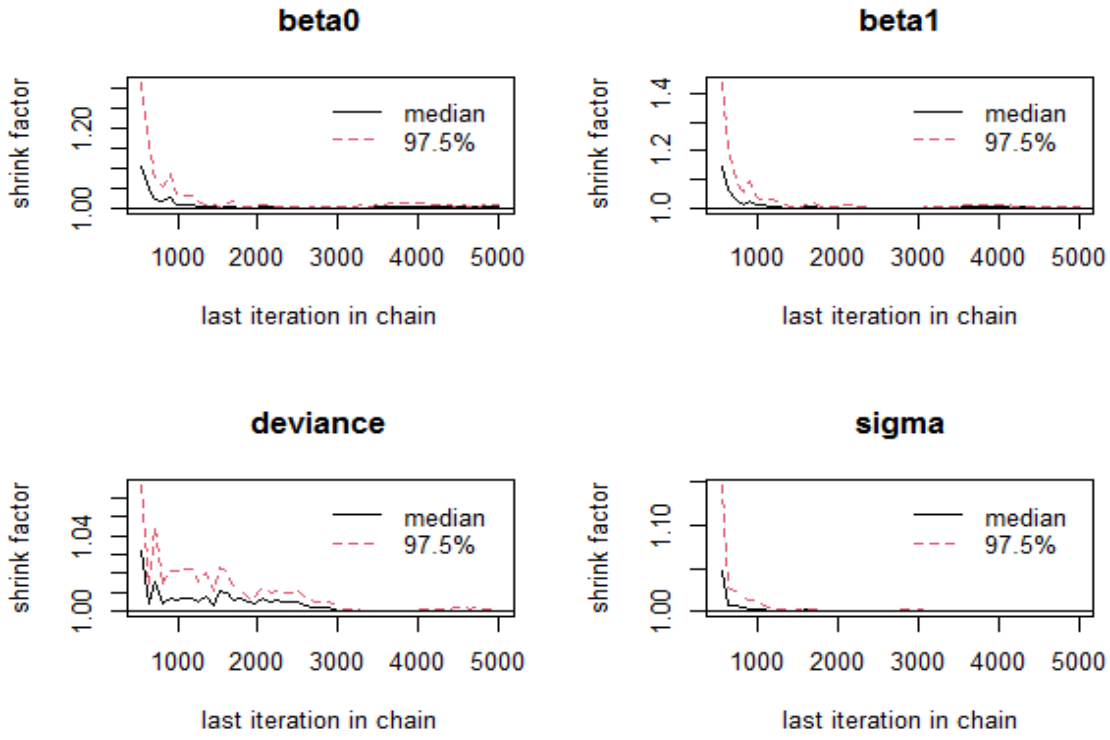


Figure 4. The Gelman-Rubin diagnostic. The Gelman-Rubin diagnostic shows that the variance in every parameter is shrinking (is convergent). The convergence values are stated by the *shrink factors,* which approach 1.00.

http://journal.uinjkt.ac.id/index.php/jp3i

The Gelman-Rubin analysis is the next diagnostic tool. The Gelman-Rubin Diagnostic calculates the variance of the differences of each *chain* and that between *chains* (Gelman & Rubin, 1992). Variances that are concentrated, once upon the other, show that the Bayesian analyses have achieved convergence. Figure 4 shows that the results of the analyses have converged for each parameter. This finding is supportive of the information concerning the *trace plots*, auto-correlation, sizes of the ESS, values of the *R-hat*, and the *density plots*, previously obtained. Because of this, the *posterior* distribution of each parameter, presented in Table 4, is trustworthy.

The next step which may be taken as supplementation in the Bayesian analysis is related to hypothesis testing. In Bayesian analysis, the next analytical step compatible with the *p-values* is to calculate the *Bayes Factor* (compare the probabilities of the distribution of the sample, on the bases of H_0 and of H_1), or the *Region of Practical Equivalence* (ROPE). In this example, the ROPE is used to test the hypothesis because it is similar to *effect size*, which has frequently been used.

ROPE is an analysis to determine the probabilities of the results of Bayesian analysis, being in the *region of practically no effect*. A high percentage of the *credible interval* in the *region of practically no effect* indicates that a *null* hypothesis is accepted. The *R* packet 'bayestestR' is used to calculate the probability of being in this *region of practically no effect*. (Makowski, Ben-Shachar, & Lüdecke, 2019). This test uses the equivalent of interval *effect size*. What is normally known is that an interval *effect size* of -.10 to +.10 is a weak *effect size* (*no effect*). The analysis results show that the percentage of *credible interval* in the *region of practically no effect* is 0%. This means that the *null* hypothesis is rejected.

The final step in all of the analysis process is performing an interpretation and making conclusions from the results of the Bayesian analysis. The peculiarity of the Bayesian analysis is the use of the principle of probability in the data analysis, together with the externals, in the forms of distribution. This is at variance with the *frequentist statistics* perspective, which stresses the values of significance and point estimation. For that reason, reports of the results of Bayesian analysis need to show unabridged distribution. Table 6 shows an example of one sentence from a report on analysis results, using two differing approaches.

Table 6. *Examples of the Writing of Reports on the Results of Analysis.*

Results of Linear Regression Analysis (*frequentist statistics*)	**Results of Bayesian Regression Analysis**
Positive Affect contributes positively to Marital Satisfaction (β = .30 [.25, .35], *SE* = .02, t = 12.68, p < .01).	Based upon the prior data and information held, the average of the distribution of the *posterior* coefficient of Positive Affect is .31, with a deviation standard of .01 and a *credible interval* ranging from .30 to .33.

Discussion

This study aimed to support psychological researchers interested in Bayesian analysis and shows a viewpoint different from that of the *frequentist statistics* approach. Bayesian analysis is clarified not through the estimation of one value but instead by using the principle of probability. The results from Bayesian analysis always take the form of *posterior* distributions for each parameter. This *posterior* distribution depends upon the distribution of the data held (*likelihood*) and previously held knowledge (*prior*). Thus, Bayesian analysis has the flexibility to adjust findings in line with the development of knowledge (*prior*). Because of this, this approach is fitting to be used for the analysis of *big data* or longitudinal research, facilitating researchers to pay close attention to data in line with the development of the data population held (Ekström, Lunn, Jackson, Best, & Thomas, 2013).

The study results demonstrate the Bayesian approach to the analysis of data, generally by using a *frequentist* approach to psychological research. At a cursory glance, the results of regression analysis using a *frequentist statistics* approach resemble Bayesian analysis results. It is hoped that this article will not give rise to debates about whether one approach is better than/superior to the other. For this reason, the examples used do not show strikingly different numerical indicators. The article shows that Bayesian analysis is an alternative data analysis in response to various challenges in psychological research.

One of these challenges in psychological research is to overcome bias in the reporting of research results. Psychological research is currently striving to relieve the dependency on the *p-value* only when reporting statistical analysis. However, the *p-value* requires accompanying information (Altman & Krzywinski, 2017) to reduce the possibility of its misuse (Krawczyk, 2015). One of the necessary information is the *a priori* calculation of *power* and the reporting of *effect size.* In general, *power* depends upon a sufficient number of samples, and often psychological researchers cannot achieve the sample numbers representative of the research population.

Bayesian analysis can assist psychological researchers in overcoming limited sample numbers (van de Schoot & Miočević, 2020), by using *sampling* algorithms (*e.g.* MCMC). However, this does not mean that Bayesian analysis is not sensitive to the number of participants because this analysis still requires data from participants (*likelihood*). Therefore, to get accurate results, psychological researchers who cannot obtain adequate sample sizes need to have precise prior information and conduct diagnostics, as shown in Table 1.

The bias in data analysis related to *p-values is p-hacking,* manipulating the data analysis process to obtain the hoped-for value of *p* (Head, Holman, Lanfear, Kahn, & Jennions, 2015). Bayesian analysis is also susceptible to misuse, such as with the augmentation of the *iterations*, the amount of *thinning*, and the number of *chain*s to obtain convergent results.. The reader needs to pay careful attention to the fact that the analyses of Models 1 to 5 are not aimed solely at discovering convergent results. For example, the amount of *thinning* is increased in Model 5 so that the value of the *effective sample size* (ESS) is > 1,000. On the other hand, the increase in the amount of *thinning* actually reduces the number of *iterations* in each *chain*. Thus, the reader needs to comprehend that the process in Models 1 to 5 aims to discover an adequate number of samplings to obtain convergent results.

The results of data analysis using *Bayesian statistics* are influenced by the number of iterations or chains that are performed. Therefore, the clarification of the analysis results is susceptible to bias. Because of this, *prior* information plays a large role in determining things external to an accurate Bayesian analysis (Van de Schoot, 2015). Besides this, the complete performance of the diagnostic steps may also determine that the results of Bayesian analysis are accepted as transparent and accurate. Several indicators, such as the *R-hat*, *trace plots*, ESS, *density plots*, auto-correlation, and the Gelman-Rubin *plot,* are used as sources of supportive information.

The entirety of the sources of supportive evidence is intended to indicate convergence (shrinking, reducing/*stationary*). The Bayesian analysis uses computerized simulations with specific algorithms, such as MCMC, to find constant *posterior* distribution (St. John, Strutz, Broadbelt, Tyo, & Bomble, 2019). This means that Bayesian analysis compares the results of simulations from a number of different *chains*. Results that contract to one distribution are the most *probable*. The steps described in Table 1 comprise a small section of the ways to show convergence. Researchers may use other approaches, for instance, the convergence analysis of each *chain*, by dividing each *chain* into two and comparing the *R-hat* values (van de Schoot, *et al.*, 2021).

http://journal.uinjkt.ac.id/index.php/jp3i

Conclusion

As an introduction, this study has several weaknesses. The first is that the study has not initially set out the history and basic principles of the Bayesian approach, such as the Bayesian theorems. The second is that terminology frequently found in Bayesian analysis has not been explained in detail. There are so many things to be examined and dissected through the article; however, writing one article capable of covering all of the Bayesian analysis elements is not easily achieved. The introductory book on Bayesian analysis (Kruschke, 2014; Bolstad & Curran, 2017) is a source of information worthy of study to assist the reader in comprehending Bayesian analysis in greater depth.

Psychological research in Indonesia using Bayesian analysis will increasingly develop in the future. This study used an assistance tool, *i.e.* OpenBUGS, which requires a basic understanding of the *coding* process and the model writing of the examples in the article. Nonetheless, psychological researchers can use a variety of alternatives to facilitate analyzing data through the Bayesian approach. Furthermore, Bayesian analysis assistance tools made specifically for psychological research are also increasingly being developed, for example, 'bayes4psy' (Demšar, Repovš, & Štrumbelj, 2020) so that Bayesian analysis is increasingly accessible for psychological researchers in Indonesia.

References

Abdelkader, M. (2020, April 25). *Bayesian linear regression*. Retrieved March 4, 2021, from https://modelingwithr.rbind.io/bayes/Bayesian-linear-regression/

Altman, N., & Krzywinski, M. (2017). Interpreting P values. *Nature Methods, 14*, 213-214. https://doi.org/10.1038/nmeth.4210

Andrews, M., & Baguley, T. (2013). Prior approval: the growth of Bayesian methods in psychology. *British Journal of Mathematical and Statistical Psychology, 66*(1), 1-7. https://doi.org/10.1111/bmsp.12004

Az Zahra, S., & Caninsti, R. (2017). Hubungan antara kepuasaan pernikahan dengan spiritualitas pada istri bekerja yang berada dalam tahap pernikahan Families With School Children. *Jurnal Psikogenesis, 4*(2), 215-223. https://doi.org/10.24854/jps.v4i2.351

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *SPINE, 25*(24), 3186-3191. https://doi.org/10.1097/00007632-200012150-00014

Becker, C., Kirchmaier, I., & Trautmann, S. T. (2019). Marriage, parenthood and social network: Subjective well-being and mental health in old age. *PLOS ONE, 14*(7), e0218704. https://doi.org/10.1371/journal.pone.0218704

Bolstad, W. M., & Curran, J. M. (2017). *Introductions to Bayesian statistics* (3rd ed.). Hoboken, New Jersey: John Wiley & Sons, Inc. https://doi.org/10.1002/9781118593165

Buaton, R., & Astuti, S. (2014). Perancangan sistem pakar tes kepribadian dengan menggunakan metode Bayes. *METHODA, 4*(1), 1-16.

Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 80*, 1-28.

Chen, M., & Shao, Q. (1999). Monte Carlo estimation of Bayesian credible and HPD intervals. *Journal of Computation and Graphical Statistics, 8*(1), 69-92. https://doi.org/10.2307/1390921

Demšar, J., Repovš, G., & Štrumbelj, E. (2020). bayes4psy—An open source R package for Bayesian statistics in psychology. *Frontiers in Psychology, 11*, 947. https://doi.org/10.3389/fpsyg.2020.00947

Depaoli, S., Clifton, J. P., & Cobb, P. R. (2016). Just Another Gibbs Sampler (JAGS): Flexible software for MCMC implementation. *Journal of Educational and Behavioral Statistics, 41*(6), 628-649. https://doi.org/10.3102/1076998616664876

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessments, 49*(1), 71-75.

Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on? *Perspectives on Psychological Science, 6*, 274-290.

Ekström, R. J., Lunn, D., Jackson, C., Best, N., & Thomas, A. (2013). *The BUGS Book: A practical introduction to Bayesian analysis.* New York: CRC Press.

Etz, A., & Vandekerckhove, J. (2018). Introduction to Bayesian inference for psychology. *Psychonomic Bulletin and Review, 25*, 5-34. https://doi.org/10.3758/s13423-017-1262-3

Gabry, J., & Modrák, M. (2021, January 11). *Visual MCMC diagnostics using the bayesplot package*. Retrieved March 4, 2021, from https://mc-stan.org/bayesplot/articles/visual-mcmc-diagnostics.html

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science, 7*, 457-472.

Gelman, A., Lee, D., & Guo, J. (2015). Stan: A probabilistic programming language for Bayesian inference and optimization. *Journal of Educational and Behavioral Statistics, 40*(5), 530-543. https://doi.org/10.3102/1076998615606113

Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. (2015). The extent and consequences of p-hacking in science. *PLoS Biology, 13*(3), e1002106. https://doi.org/10.1371/journal.pbio.1002106

Johnson, H. A., Zabriskie, R. B., & Hill, B. (2006). The contributions of couple leisure involvement, leisure time, and leisure satisfaction to marital satisfaction. *Marriage & Family Review, 40*(1), 69-91. https://doi.org/10.1300/J002v40n01_05

Johnson, M. D., Cohan, C. L., Davila, J., Lawrence, E., Rogge, R. D., Karney, B. R., . . . Bradbury, T. N. (2005). Problem-solving skills and affective expressions as predictors of change in marital satisfaction. *Journal of Consulting and Clinical Psychology, 73*(1), 15-27. https://doi.org/10.1037/0022-006X.73.1.15

Kim, A., Lee, J. A., & Park, H. S. (2018). Health behaviors and illness according to marital status in middle-aged Koreans. *Journal of Public Health, 40*(2), e99-e106. https://doi.org/10.1093/pubmed/fdx071

Krawczyk, M. (2015). The search for significance: A few peculiarities in the distribution of p values in experimental psychology literature. *PLoS ONE*, e0127872. https://doi.org/10.1371/journal.pone.0127872

Kruschke, J. K. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan.* New York, NY: Academic Press.

Law, J. (2019, June 14). *Bayesian statistics and functional programming: Bayesian linear regression with Gibbs sampling in R*. Retrieved March 4, 2021, from https://jonnylaw.rocks/posts/2019-06-14-Bayesian-linear-regression/

Lunn, D., Thomas, A., Best, N., & Spiegelhalter, D. (2000). WinBUGS — a Bayesian modelling framework: Concepts, structure, and extensibility. *Statistics and Computing, 10*, 325-337.

Makowski, D., Ben-Shachar, M. S., & Lüdecke, D. (2019). bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. *Journal of Open Source Software, 4*(40), 1541. https://doi.org/10.21105/joss.01541

Marsman, M., & Wagenmakers, E.-J. (2017). Bayesian benefits with JASP. *European Journal of Developmental Psychology, 14*(5), 545-555. https://doi.org/10.1080/17405629.2016.1259614

http://journal.uinjkt.ac.id/index.php/jp3i

Nurmansyah, W., & Hartati, S. (2013). Prototipe sistem pakar penentu jenis gangguan psikologi klinis menggunakan Forward Chaining dan Formula Bayes. *Berkala MIPA, 23*(2), 149-160.

Otero, M. C., Wells, J. L., Chen, K.-H., Brown, C. L., Connelly, D. E., Levenson, R. W., & Fredrickson, B. L. (2020). Behavioral indices of positivity resonance associated with long-term marital satisfaction. *Emotion, 20*(7), 1225-1233. https://doi.org/10.1037/emo0000634

Penny, W. D., Mattout, J., & Trujillo-Barreto, N. (2007). CHAPTER 35-Bayesian model selection and averaging. In K. Friston, J. Ashburner, S. Kiebel, T. Nichols, & W. Penny (Eds.). Academic Press. https://doi.org/10.1016/B978-012372560-8/50035-8

Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). CODA: Convergence Diagnosis and Output Analysis for MCMC. *R News, 6*(1), 7-11.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*(2), 225-237. https://doi.org/10.3758/PBR.16.2.225

Sorokowski, P., Randall, A., Groyecka, A., Frackowiak, T., Cantarero, K., Hilpert, P., . . . Sorokowska, A. (2017). Marital satisfaction, sex, age, marriage duration, religion, number of children, economic status, education, and collectivistic values: Data from 33 countries. *Frontiers in Psychology, 8*, 1199. https://doi.org/10.3389/fpsyg.2017.01199

St. John, P. C., Strutz, J., Broadbelt, L. J., Tyo, K. E., & Bomble, Y. J. (2019). Bayesian inference of metabolic kinetics from genome-scale multiomics data. *PLOS Computational Biology, 15*(11), e1007424. https://doi.org/10.1371/journal.pcbi.1007424

Sturtz, S., Ligges, U., & Gelman, A. (2005). R2WinBUGS: A package for running WinBUGS from R. *Journal of Statistical Software, 12*(3), 1-16.

Surijah, E. A., & Prakasa, D. Y. (2020). Kepuasan pernikahan dan bias harapan sosial. *Jurnal Ilmiah Psikologi MIND SET, 11*(01), 14-25. https://doi.org/10.35814/mindset.v11i01.1269

Van de Schoot, R. (2015). Latent trajectory studies: The basics, how to interpret the results, and what to report. *European Journal of Psychotraumatology, 6*, 27514. http://dx.doi.org/10.3402/ejpt.v6.27514

van de Schoot, R. D., Kramer, B., Märtens, K., Tadesse, M. G., Vannucci, M., Gelman, A., . . . Yau, C. (2021). Bayesian statistics and modelling. *Nature Reviews Methods Primers, 1*, 1. https://doi.org/10.1038/s43586-020-00001-2

van de Schoot, R., & Miočević, M. (Eds.). (2020). *Small sample size solutions: A guide for applied researchers and practitioners* (1 ed.). Routledge. https://doi.org/10.4324/9780429273872

van de Schoot, R., Winter, S. D., Ryan, O., Zondervan-Zwijnenburg, M., & Depaoli, S. (2017). A systematic review of Bayesian articles in psychology: The last 25 years. *Psychological Methods, 22*(2), 217-239. http://dx.doi.org/10.1037/met0000100

van Ravenzwaaij, D., Cassey, P., & Brown, S. (2018). A simple introduction to Markov Chain Monte–Carlo sampling. *Psychonomic Bulletin and Review, 25*, 143-154. https://doi.org/10.3758/s13423-016-1015-8

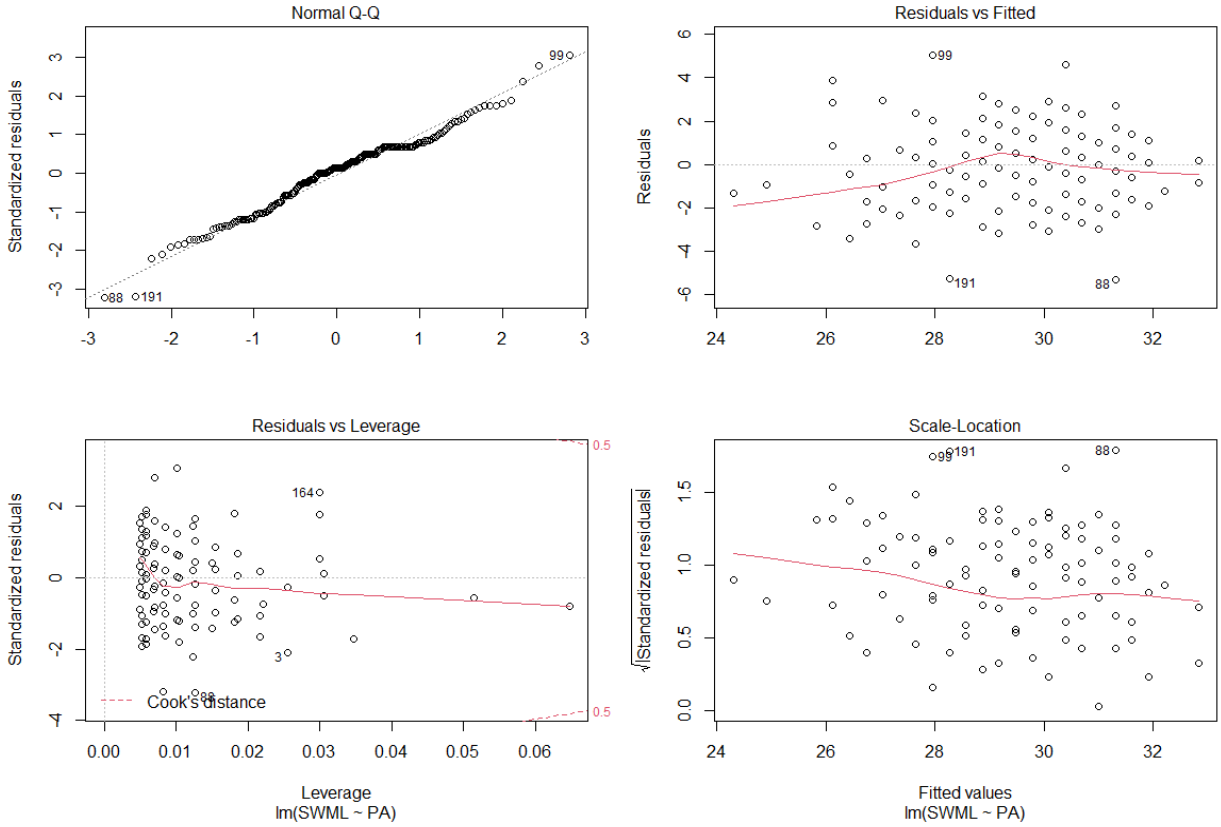
Wagenmakers, E. J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., . . . Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review, 25*, 35-57. https://doi.org/10.3758/s13423-017-1343-3

Watson, D., Clark, L. A., & Carey, G. (1988). Positive and negative affectivity and their relation to anxiety and depressive disorders. *Journal of Abnormal Psychology, 97*(3), 346-353. https://doi.org/10.1037/0021-843X.97.3.346

Yau, C. (n.d.). *Bayesian inference using OpenBUGS*. Retrieved March 4, 2021, from R Tutorial: http://www.r-tutor.com/Bayesian-statistics/openbugs

Zevon, M. A., & Tellegen, A. (1982). The structure of mood change: An idiographic/nomothetic analysis. *Journal of Personality and Social Psychology, 43*, 111-122.

Appendix A. Diagnostic Plots for the Linear Regression Analysis



http://journal.uinjkt.ac.id/index.php/jp3i

Appendix B. Non-Informed Bayesian Regression Model

```
model {
for (i in 1:N) {
y[i] ~ dnorm(mu[i], tau)
mu[i] <- beta0 + beta1*x[i]
}
beta0 ~ dnorm(0, 0.00001)#define Prior
beta1 ~ dnorm(0, 100)
tau <- 1/pow(sigma, 2)
sigma ~ dunif(0, 100)}

After we write the model, we prepare the R environment:
library(R2OpenBUGS)
setwd("C:/Users/ThinkPad/Documents/Project 2020/Bayes Intro")

Set the Data and Parameters:
y <- dataset$SWML
x <- dataset$PA
N <- length(y)
the_data <- list("y" = y, "x" = x, "N" = N)
parameters <- c("beta0", "beta1", "sigma")
set.seed(123)
posterior <- bugs(data = the_data, model.file="trial001.txt",
                  n.chains = 3,inits=NULL,
                  parameters = parameters,
                  n.iter = 1000,
                  n.burnin = 500,n.thin=1,codaPkg=FALSE,debug=TRUE)
```

Appendix C. Gibbs Sampling Specification for Model 5

```
set.seed(123)
posterior <- bugs(data = the_data, model.file="trial001.txt",
                  n.chains = 3,inits=NULL,
                  parameters = parameters,
                  n.iter = 5000,
                  n.burnin = 500,n.thin=4,codaPkg=TRUE,debug=TRUE)
library(coda)
check<-read.bugs(posterior)
traceplot(check)
densplot(check)
autocorr.plot(check)
gelman.plot(check)
summary(check, quantiles = c(0.25, 0.975))
```

Appendix D. Specific Terms in Layman's Terms

Term	Simple Explanation
Bayes Factor	Likelihood ratio of a hypothesis when compared with another hypothesis
Burn-in	Numerical information that indicates the number of warming-up simulated before the iteration is run
Chain	Numerical information that indicates how many times each parameter is simulated
ESS	Effective Sample Size; The number of independent samples that gives information to the posterior distribution
Gelman-Rubin Diagnostic	Visual information of variance convergence of a parameter in each chain and between-chains
Iterations	Numerical information that indicates the number of sampling on each chain
Likelihood	The model's probability based on the acquired/observed data
MCMC	Computerized sampling algorithms Markov Chain Monte Carlo
Posterior	The model's probability based on the observed data and prior knowledge
Prior	The model's probability based on the prior knowledge
R-hat	Numerical information that indicate a parameter convergence based on the between-chains simulations; R-hat = 1.00 indicates convergent results
ROPE	The value of a probability that the Bayesian result falls within an interval of small effect
Trace Plot	Visual information of a parameter convergence based on the between-chains simulations

.