

B. I. T. – Development of a dynamic visualization tool for Bayesian inference on various types of normal distributions for medical decision-making and education –

Hiroki Inoue¹, Hachiro Uchiyama², Kensuke Moriwaki³, Shinichiro Yanagisawa⁴, Kunihiro Nishimura⁵

¹Department of Health Informatics, Niigata University of Health and Welfare, Niigata, Japan

²Center for General and Liberal Education, Doshisha University, Kyoto, Japan

³Department of Medical Statistics, Kobe Pharmaceutical University, Kobe, Japan

⁴Division of Medical Economics, Himeji Dokkyo University School of Pharmacy, Himeji, Japan

⁵Departments of Preventive Medicine and Epidemiology, National Cerebral and Cardiovascular Center, Suita, Japan

Keywords: visualization tool, intuitive operation, normal distribution, medical decision-making, Bayesian inference

Received: 22 March 2019 / Accepted: 8 May 2019

Abstract

Bayesian inference Tool (B. I. T.) is a software program developed to support medical professionals and students in the application of Bayesian inference to medical decision-making and education. This software program makes it possible to visualize the posterior probability distributions of Bayesian inference. In addition, its interface aims to facilitate the intuitive understanding of the operative procedures. For many cases of Bayesian inference with normal distributions, the graphical representation of the posterior probability distribution changes dynamically when the input value is assigned by the user with the use of a computer mouse; this procedure also enables the dynamic calculation of the posterior probability. When operating this software, it is desirable for the user to have basic knowledge and understanding of the presuppositions as well as applications of Bayesian inference.

Introduction

In outcomes research of health and medical sciences, Bayesian inference is frequently used in meta-analysis [1], estimation of morbidity prevalence rate [2], as well as calculation of public health indicators of a certain region [3]. The application of Bayesian inference not only facilitates the understanding in terms of the tendencies of the overall patient group, but also helps decide whether or not to accept the difference between two or more groups, with a certain degree of confidence. Moreover, with Bayesian inference, it is feasible to infer whether there is a difference in the data collected from different healthcare facilities. In probability and statistics, the tool in information technology—with features for statistical calculation, interactive graphical user interface, computer simulation, data analysis and visualization—has been recognized to be effective in terms of raising students' examination scores, and level of satisfaction, and so on [4]. When applying Bayesian statistics in daily life, it would be

Corresponding author: Hiroki Inoue

Department of Health Informatics, Niigata University of Health and Welfare, 1398 Shimami-cho, Kita-ku, Niigata 950-3198, Japan
TEL/FAX: +81-25-257-4532, E-mail: inoue@nuhw.ac.jp

time-consuming and labor-intensive for the user to calculate and make inferences with the given formulae on each occasion. It is considered that the feature for displaying the dynamic changes in graphical curves of the probability distributions is highly suitable for expressing the Bayesian updating [5] that is characteristic of Bayesian statistics. In the future, when Bayesian inference is applied in the contexts of medical decision-making, user-friendly software programs that can not only be operated intuitively via the graphical user interface (GUI) but also whose visualized results can be understood easily by the user are deemed highly useful and practical, particularly for healthcare professionals and students. In this study, the word “intuitive” is defined as needing no complex procedures for manipulating the data input as well as interpreting the resulting output. Despite the usefulness of the Bayesian approach, a want of user-friendly software programs for healthcare professionals and students is a non-negligible issue. MRBAYES [6] is an already known software program dedicated to Bayesian inference. However, despite its capabilities to handle intricate statisti-

cal models, MRBAYES may not be suitable for general medical practitioners and students who wish to gain a deeper understanding of Bayesian statistics (because of its rather complicated operational procedures). The purpose of the software program—Bayesian inference Tool (B. I. T.) that the authors of this paper have developed—is to support general clinicians and medical students in learning Bayesian inferences for normal distributions that are rudimentary and highly applicable; the software program, B. I. T., supports the user in his/her learning process by providing an environment where dynamic and intuitive operations are possible.

Materials and Methods

The innovative aspects of B. I. T.’s functions that should be achieved in development are shown in Table 1. The essential requirements for realizing the innovative functions of the software program are listed in Table 2. The required items for inferences and posterior probability distributions are shown in Table 3. Each item in Table 3 was selected, in accordance with the studies on

Table 1. The innovative aspects of B. I. T.’s functions to be achieved.

Requirement 1.	As many operations as possible shall be performed on one operation screen.
Requirement 2.	User’s operations are executed at ease through visual means.
Requirement 3.	When the observed data required for Bayesian inference are assigned by entering the data, the graphical curve of the posterior distribution is immediately calculated and displayed.
Requirement 4.	As the initial values for Bayesian inference are replaced with newly observed values, the graphical curve of the posterior distribution is re-calculated and re-displayed.
Requirement 5.	By continuously performing Requirements 3 and 4, the graphical curves are displayed in animation simultaneously.
Requirement 6.	When the observed data required for Bayesian inference are entered, the prior probability and the posterior probability are immediately calculated and displayed.
Requirement 7.	At the time of replacing the data on observed values for Bayesian inference from the initial value(s), the prior probability and the posterior probability shall be re-calculated and re-displayed.
Requirement 8.	Updating in graphical curves and probabilities are automatically reflected in synchronization with the execution of data-updating. There is no need to update the display of graphical curves and probabilities with buttons or other procedures.

Table 2. The essential requirements for improving this software to realize the innovative functions of B. I. T.

Requirement 9.	It can be operated only with a computer mouse; i.e., the use of a PC keyboard is not required.
Requirement 10.	Users need not enter command statements when operating.
Requirement 11.	The program can be controlled by users with no expertise in computers.
Requirement 12.	The option with/without prior information needs to be added (exclusion when using approximations).
Requirement 13.	The feature by which the user can print every operation screen shall be added to the software program.
Requirement 14.	The facility in operations for downloading, installing and launching is ensured.

Table 3. Bayesian inferences Available in B. I. T.

(i) Bayesian inferences on one normal population $N(\mu, \sigma^2)$ (μ : mean, σ^2 : variance)
(i-1) Inference of μ with a known σ^2
(i-2) Inference of σ^2 with a known μ
(i-3) Inference of μ with unknown μ and σ^2
(i-4) Inference of σ^2 with unknown μ and σ^2
(ii) Bayesian inferences on two populations, $N_1(\mu_1, \sigma_1^2)$ and $N_2(\mu_2, \sigma_2^2)$ (μ_1, μ_2 : mean, σ_1^2, σ_2^2 : variance)
(ii-1) Inference of $\mu_1 - \mu_2$ (where $\sigma_1^2 = \sigma_2^2 = \sigma^2$ and σ^2 is known)
(ii-2) Inference of $\mu_1 - \mu_2$ (where $\sigma_1^2 = \sigma_2^2 = \sigma^2$ and σ^2 is unknown)
(ii-3) Inference of $\mu_1 - \mu_2$ (where $\sigma_1^2 \neq \sigma_2^2$ and both σ_1^2 and σ_2^2 are unknown)
(ii-4) Inference of variance ratio on two normal populations
(iii) Bayesian inferences on three normal populations ; $N_1(\theta_1, \sigma_1^2)$, $N_2(\theta_2, \sigma_2^2)$ and $N_3(\theta_3, \sigma_3^2)$
(iii-1) Inference of $\theta = (\theta_1, \theta_2, \theta_3)$ (where $\sigma_1^2, \sigma_2^2, \sigma_3^2$ are unknown, $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma^2$ can be assumed).
(iii-2) Comparisons of the means of two groups with three normal populations respectively. Three normal populations' variances are unknown and equal.
(iii-3) Evaluation of the equality of variances among three normal populations, $\sigma_1^2, \sigma_2^2, \sigma_3^2$
(iii-4) Inference of the variance of three normal populations (where the three groups' variances are unknown but their equality is known)
(iv) Bayesian inference of bivariate normal distributions: $N\left(\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}, \begin{bmatrix} \sigma_{11}^2 & \sigma_{12}^2 \\ \sigma_{21}^2 & \sigma_{22}^2 \end{bmatrix}\right)$
(iv-1) Inference on the correlation coefficient of bivariate normal distributions
(iv-2) Inference of two random variables with bivariate normal distributions: $\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}$
(iv-3) Inference of variance-covariance matrix with bivariate normal distributions: σ_{11}^2 and σ_{22}^2
(v) Bayesian inference on two bivariate normal distributions, where: $N\left(\begin{bmatrix} \mu_{1X} \\ \mu_{1Y} \end{bmatrix}, \Sigma_1\right)$
and $N\left(\begin{bmatrix} \mu_{2X} \\ \mu_{2Y} \end{bmatrix}, \Sigma_2\right)$, (Σ_1, Σ_2 ; variance-covariance matrix)
(v-1) Inference of the difference in means of two normal distributions: $\begin{bmatrix} \mu_{1X} - \mu_{2X} \\ \mu_{1Y} - \mu_{2Y} \end{bmatrix}$ ($\Sigma_1 = \Sigma_2 = \Sigma$)

Bayesian statistics by Gelman et al. [7], Lee [8], Lindley [9], and Zellner [10]. For more details on each of the formulae, see [7-10].

The software program we have developed has been written with Microsoft(R) Visual Basic(R).net 2017(TM). The software's intended function-ability has been confirmed with a computer with the Windows(R) 10 operating system.

Results

One of the window screens is shown in Figure 1 as an example. The items listed in Table 3 have all been implemented in this software program. A dedicated operation screen is provided for each Bayesian inference. Some of the formulae for calculating posterior probability distributions in this software program utilize various approximations. (ii-3) in Table 3 uses Patil's approximation [11]. The formula in (iii-3) is approximated by that of Bartlett test [12]. The formula in (iv-1), likewise, is an approximation of a complicated formula of integration [13]. The formula of integration used to calculating posterior probability distribution uses a simple method of quadrature by parts.

The following details pertinent to B. I. T. show that the requirements of Table 1 have been met. The display is divided into the entire left section where the user can assign and change observed values, and the right section which is, in turn, split into the upper-right section for displaying graphs, and the lower-right section for showing calculated values (Corresponding to Requirement 1). As needed, the user can adjust or change the range of average values via manipulating the slide bar. Moreover, the x-axis of each curve in the graph can be adjusted according to the user's needs (Requirement 2).

This software program automatically calculates posterior probability distribution upon the user's assignment of various observed values required for Bayesian inference, and the user can visually comprehend prior and posterior probability distributions with dynamic animations (Requirements

3, 4, and 5).

The posterior probability of the credible interval within the range specified by the user can also be approximated and displayed (Requirement 6).

The re-calculation of credible interval can be done if the user instructs the program to do so by changing the range of the interval, which is then followed by re-calculation and graphical re-display (Requirement 7). Users can not directly change the display of graphical curves and probabilities (Requirement 8).

B. I. T. implements the requirements of Table 2 as detailed below. With a simple use of a computer mouse, all of the operations shown in Table 3 can be executed. There is no need for the user to write and enter a computer command, and it can be operated without a keyboard. B. I. T. does not necessitate the user to have specialized knowledge of computers (Corresponding to Requirements 9, 10, and 11).

When using this feature, the user can select either a locally uniform distribution with an objective perspective, or a natural conjugate prior distribution with an emphasis on subjective and experiential prior information (Requirement 12).

Upon clicking on the "Instructions" button, a manual in PDF format opens and the user can refer to information on theorems and formulae of each inference along with information on how to operate. "Print" button prints the screen (Requirement 13). B. I. T. can be downloaded from the following website (URL: <https://upload.umin.ac.jp/fileshare/registant.cgi>). B. I. T. is distributed as Windows Installer package and is easy to install and launch (Requirement 14).

Examples 1-5 show selected samples of this software program's applications carried out by the developers. The following are examples where the use of Bayesian statistics for medical decision-making in daily clinical practice is assumed. The operation screens that suit the given situations have been selected. Figures 2-9 are examples of graphical curves dynamically changed

by operating this software program.

Example 1:

When the trial material assuming an unknown normal distribution was used in the test, the observed values were as follows:

- the mean, 45;
- the variance, 17.8;
- and the sample size, 10.

Provisional experiments conducted beforehand resulted in the mean of 40, the variance of 22.2, and the sample size of 10.

Using the analyses in (i-3) and (i-4), 95% credible interval of the marginal posterior distribution of the mean turned out to be 42.5 ± 10.307 as shown in Figure 2. As shown in figure 3, the marginal posterior distribution at 95% credible interval resulted in the range of variance between 0 and 48.

Example 2:

This example has the following scenario: There is a hypothetical case wherein both a new medicine and an old medication are known to have the effect of lowering Alanine Aminotransferase (ALT) values (IU/L). If the new medicine lowers ALT values by more than 20 (IU/L) as compared to the older (conventional) medication, one may judge that the new one is more effective than the other one. Administering these medicines to two groups (one group composed of 17 subjects given the new medicine, and the other made up of 12 subjects given the old drug) without any difference in terms of severity resulted in the following results in ALT values: Group 1 with the mean of 45, and the variance of 25, and Group 2 with the mean of 68 and the variance of 54.5. No prior information was utilized. According to (ii-4), the p-value of the posterior distribution of these two samples being within the range between 0.5 and 2 is 0.43, as shown in Figure 4. This confidence estimate is not robust enough to ascertain equal variance or homoskedasticity. Without consider-

ing the possibility of equal variance, the posterior probability distribution of the means, as shown in (ii-3), has the mean value, 23, the variance, 82, and the degree of freedom, 17, which approximate the generalized t-distribution. The confidence estimate for the new drug being more effective than the prior one is about 64 percent (Figure 5).

Example 3:

At a medical facility X, three groups of patients were selected from three clinical departments, each of which was named, A, B, and C. Subsequently, ALT values were measured for the three groups, and the results indicated that the mean, the variance, and the sample size were respectively as follow:

A (45, 158.3, 7), B (52, 175, 5), and C (60, 160, 6).

Each group's ALT values follow a normal distribution, whose variance is unknown.

No prior information was utilized. Assessing the homogeneity of the variance for each of the three groups through approximating chi-square distribution with the degree of freedom of 2 (as shown in iii-3), it turned out that the variances were contained within the 99% highest density region. As a result, it was determined that the homoskedasticity could not be rejected (Figure 6). In the analysis as seen in (iii-2), it was inferred at the 99% credible interval that the variance assumes the value between 0 and 476 (Figure 7).

Example 4:

Based on the patient groups as illustrated in Example 3, the posterior probability distribution that the mean value of the ALT measurements in each of the patient groups assumes is a three-dimensional multivariate t-distribution. Hence, each group's marginal posterior probability distribution follows the generalized t-distribution, and the 95% credible intervals for the three groups turned out to be the following:

A: 45 ± 8.8 ; B: 52 ± 9.0 ; and C: 60 ± 8.8 (See Figure 8).

At another medical facility called Y, ALT values were measured for the three groups D, E, and F, whose mean values resulted in 54, 68, and 76, respectively.

The secondary form of the three-dimensional-multivariate t-distribution follows the F distribution, $F(3,15)$. The F value for the mean value at the facility Y was 6.9, which exceeds the 99% highest density region (iii-3). In consequence, with the credible level of 99%, it was inferred that the mean ALT values between patient groups at the two facilities were different (See Figure 9).

Example 5:

Two kinds of clinical tests G and H, each of which with a bivariate normal distribution, were conducted with a subject group composed of 10 patients. Without prior information, sample correlation coefficient calculated from the observed values resulted in 0.34. The analyses in (iv-1) revealed that the marginal posterior distribution of the correlation coefficient assumed an asymmetric curve, and the credible level (p-value) that the correlation coefficient for G and H falling somewhere between 0.4 and 0.6 was 26 percent (Figure 10).

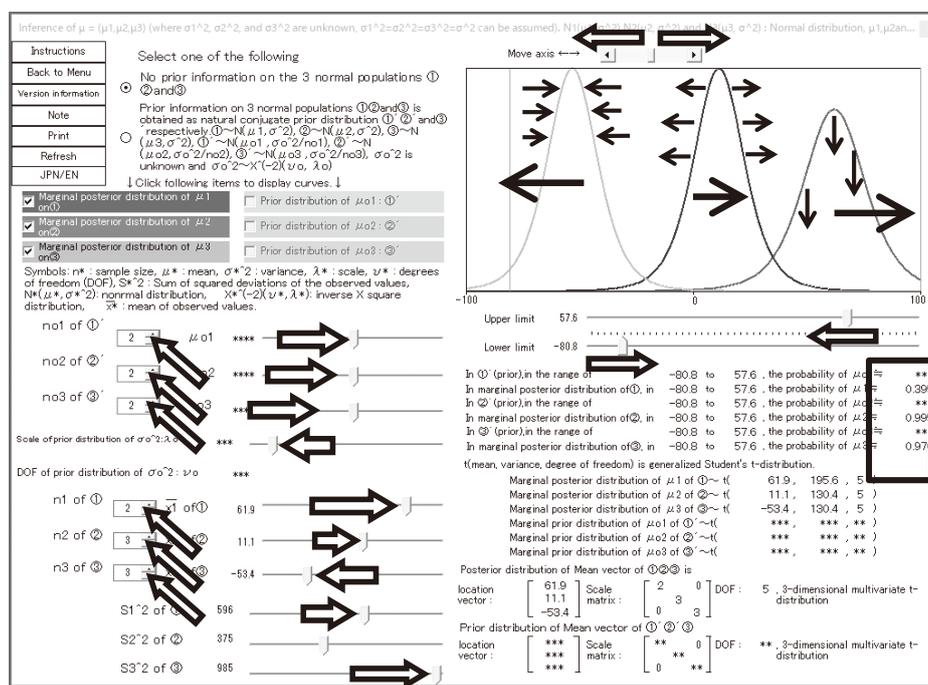


Figure 1. Example of Data Entry Interface. 16 types of Bayesian inferences are available in B.I.T. When one Bayesian inference is selected, the screen corresponding to the selected Bayesian inference is launched. The design of the interfaces is individually different for each Bayesian inference. The screen size is up to 1035 × 768 pixels. “JPN/EN” button changes the screen notation Japanese to English. As numerical values are set (⇨), the graph curves of the posterior or prior probability distribution moves and deforms (→) sequentially, automatically, and the posterior or prior probability is displayed (□) simultaneously.

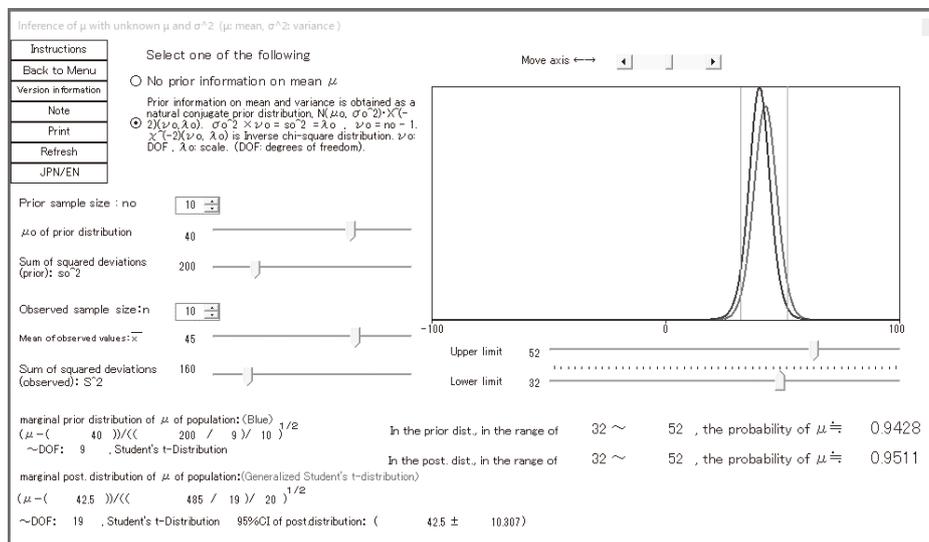


Figure 2. Marginal posterior probability distribution of sample means with an unknown normal distribution.

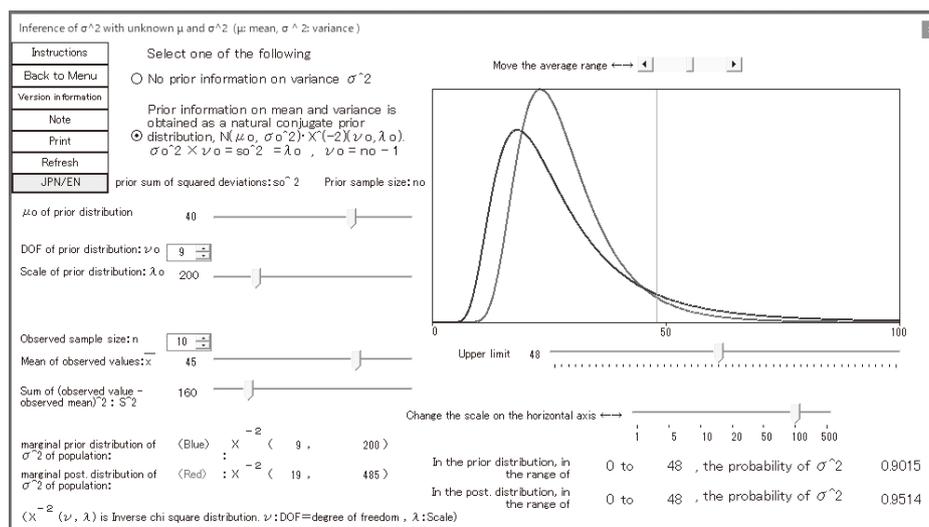


Figure 3. Marginal posterior probability distribution of sample variance with an unknown normal distribution.

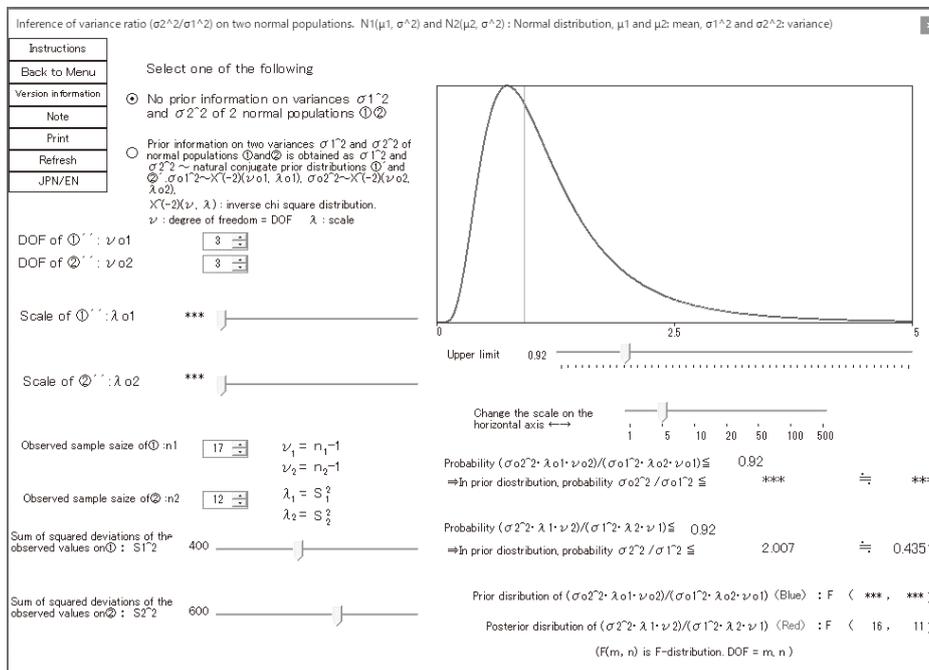


Figure 4. Assessing the homogeneity of variance from the posterior distribution of the variance ratio between a new drug and an old drug.

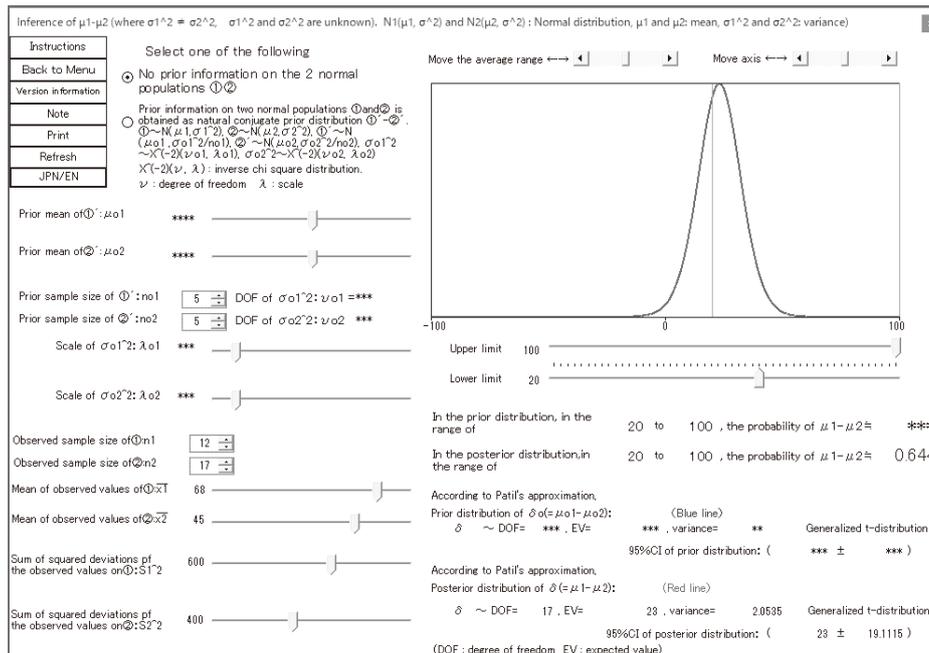


Figure 5. Posterior probability distribution of the difference of the means of two groups, not assuming the possibility of equal variance.

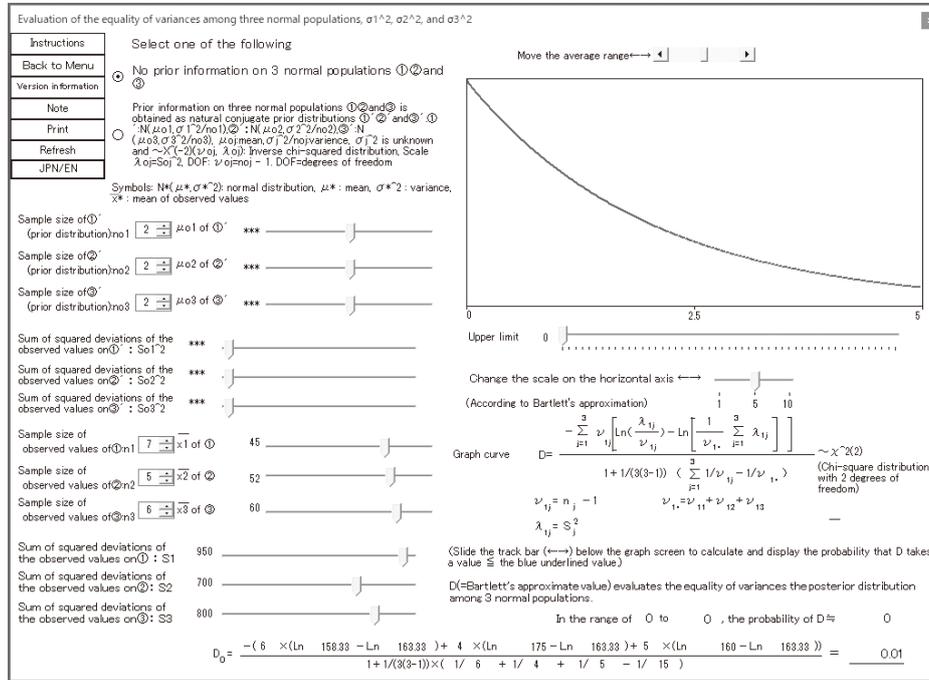


Figure 6. Making an inference about the homogeneity of variance among the three groups (A, B, C) composed of patients at three clinical departments.

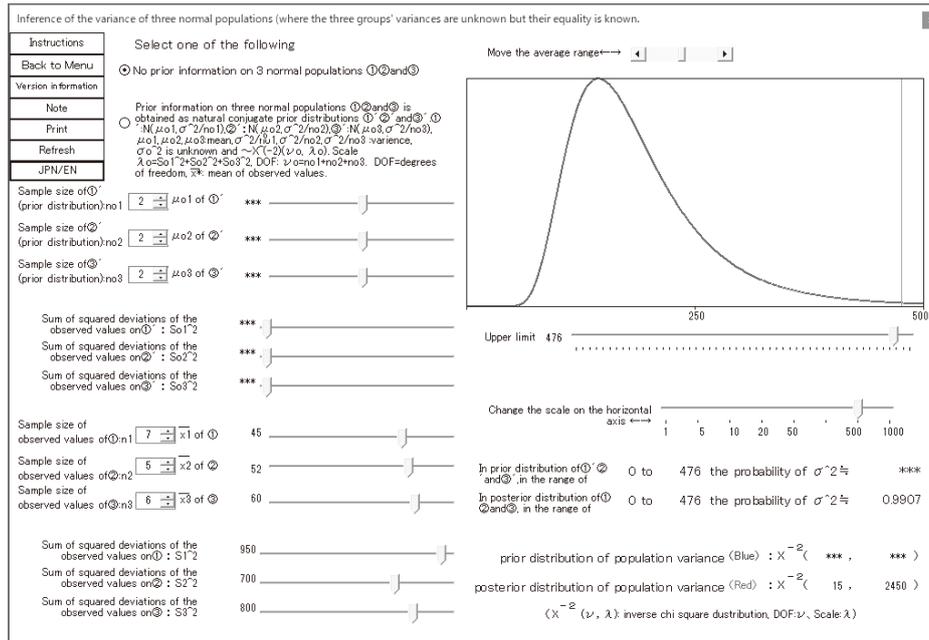


Figure 7. Posterior probability distribution of the variances, assuming the variances of three groups (A, B, C) being homogeneous.

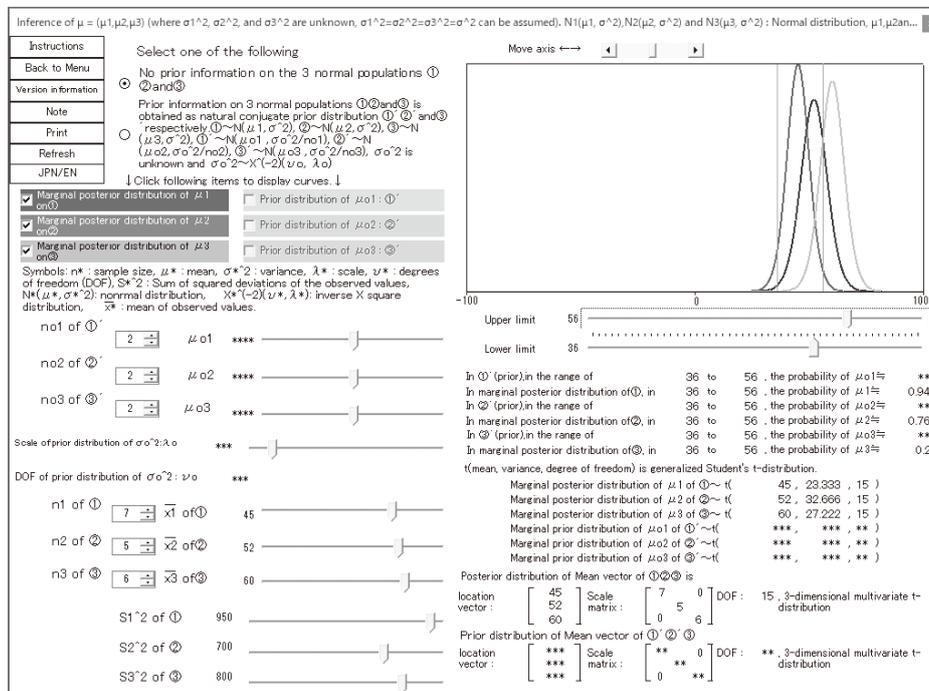


Figure 8. Inference of the marginal posterior distributions of means of three groups (A, B, C). The marginal posterior probability of group A is measured in this figure. Since the three groups (A, B, C) curves are not the same, it is impossible to measure the marginal posterior probabilities of A, B and C simultaneously.

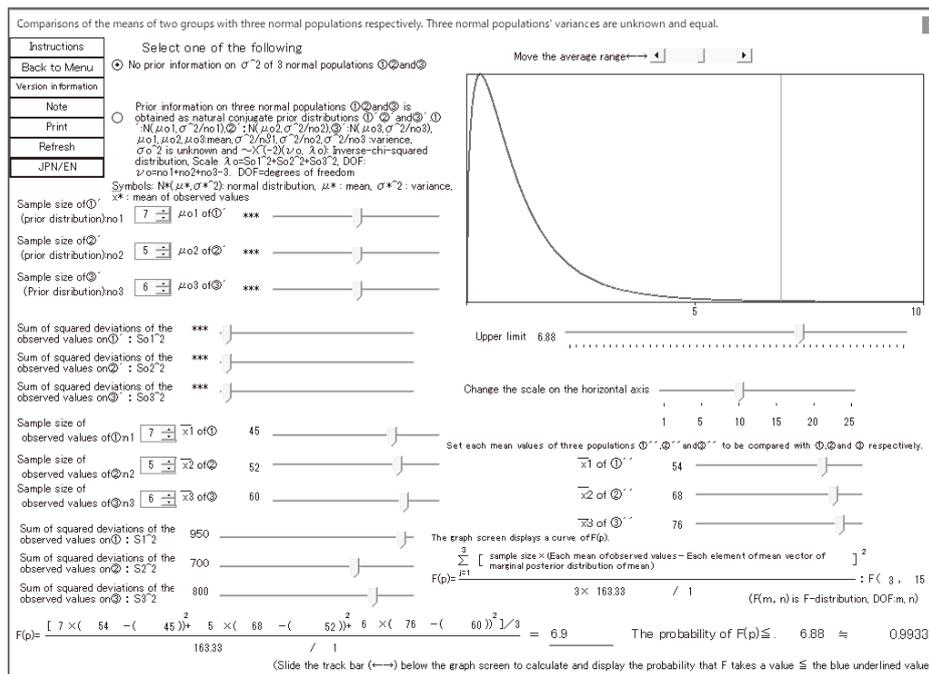


Figure 9. Making an inference about the mean values of two groups (A, B, C) and (D, E, F).

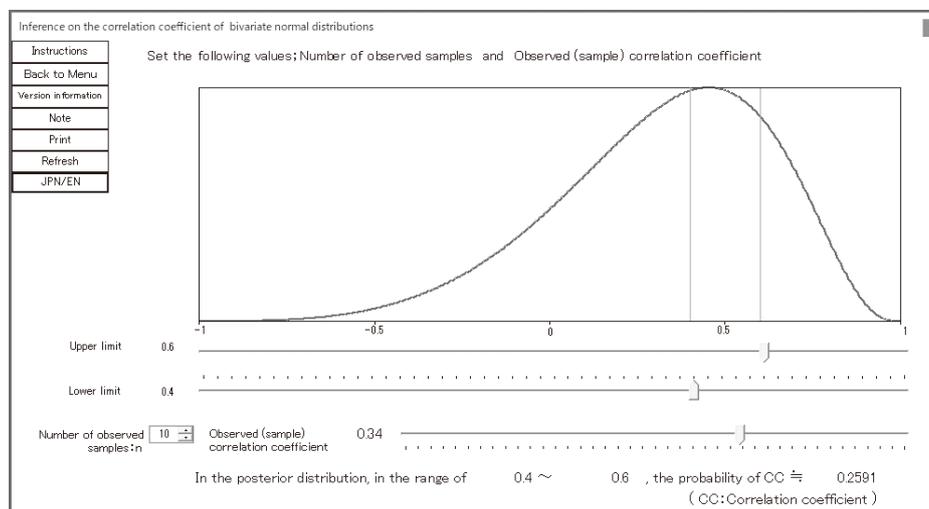


Figure 10. Making an inference about correlation coefficient of two variables that follow bivariate normal distributions.

Discussion

1. Characteristic features of B. I. T. and comparison with other visualization systems for probability distributions

First Bayes can be cited as a software program that performs Bayesian inference on many probability distributions including normal distributions [14]. With it, it is also possible to display graphical curves on the PC screen. However, unlike B. I. T., it does not have a function that allows the user to dynamically change the graphical curves that have already been displayed. Furthermore, First Bayes has a complicated installation procedure. B. I. T. can easily be installed through interactive means.

SOCR [15] is an example of educational computer programs that enable the user to visualize probability distributions. SOCR has a function called “SNAPSHOT” by which the user can print out the images on the monitor screen. Similarly, B. I. T. has a “SNAPSHOT” function which is supported by the “Print” button on the screen. However, as of 2018, the visualization function of Bayesian statistics has not been added to SOCR.

Although First Bayes and SOCR have a function for the operator to import data sets, it is considered that the data import function is difficult to express the Bayesian updating since, in Bayesian statistics, observed data are constantly updated. B. I. T. enables the user to enter numerical values quickly with, e.g., a slide bar and is user-friendly. The dynamic change of graphical curves in B. I. T. facilitates the user to get an intuitive grasp of Bayesian updating.

2. Availability of B. I. T.

It is essential to heed that the applicability of Bayesian inference has prerequisite conditions: that is, each observed value must be independent of one another, or the condition of exchangeability must be fulfilled. In addition, if the sample size were large enough to allow statistical tests based on non-Bayesian probability and/or frequentist theories, the need for Bayesian inference would be diminished. As above descriptions illustrate, it is critical to comprehend the prerequisites for Bayesian inference, before applying the approach with this software program. In biostatistics, nor-

mal distributions are assumed in many cases, [16] and, similarly, the majority of statistical analyses in many fields assume normal distributions. Nevertheless, as Tango [17] reported, of all clinical tests, many blood tests in particular are known to assume normal or lognormal distributions. Although this software program does not have a function for lognormal distributions [18], the possibility for adding this feature to the program in the future is certainly recognized by the authors. Furthermore, when learning Bayesian statistics with this software, it is required that the user acquire knowledge in terms of the principles of Bayesian inference and concomitant formulae, along with learning how to operate the interface.

3. Issues and future plans

Limitations in graphical displays

Since displaying joint probability density functions is rather intricate and entails difficulties, B. I. T. allows the user to display each peripheral probability distribution. It is important to note that the issue to be dealt with in the future—if B. I. T. were to incorporate the function to display joint probability density function curves—is to make the graphical displays as easily comprehensible as possible. For example, the possibility for a three-dimensional display of a graphical curve might be considered.

4. Numerical calculation methods

As seen in the results, the software program uses approximations for some of the probability distributions and integral calculations. With regard to these, the program would require improvement in precision and accuracy in numerical calculations, which is an area in which future progress is expected.

5. This software's possibility of expansion in the future

The programming language used to develop this software has been extended in recent years

to enable the use of other languages such as R and Python. It is expected that it will be possible in the future to use the interface of this software program in conjunction with other languages for applied numerical calculations.

6. Evaluation of the effectiveness of B. I. T.

B. I. T. requires continued improvements so that its effectiveness in facilitating the comprehension of Bayesian inference ameliorates alongside. For example, by collecting user feedback and conducting a questionnaire survey on learners' opinions, the program can be evaluated in areas of how the software has improved in ease of use, the level of advancement in terms of Bayesian statistics, and the level of comprehension in mathematical formulae as well as applications. They will be reflected in the improvement of B. I. T.

Conflicts of Interest

The authors declare they have no conflicts of interest.

References

1. Carlin JB. Meta-analysis for 2 x 2 tables: a Bayesian approach. *Stat Med.* 1992; 11: 141-158.
2. Zeger SL, See LC, Diggle PJ. Statistical methods for monitoring the AIDS epidemic. *Stat Med.* 1989; 8: 3-21.
3. Clayton D, Kaldor J. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics.* 1987; 43: 671-681.
4. Dinov I, Sanchez J, Christou N. Pedagogical utilization and assessment of the statistic online computational resource in introductory probability and statistics courses. *Comput Educ.* 2008; 50: 284-300.
5. Parmigiani G. Modeling in medical decision making: A bayesian approach (Statistics in practice). Chichester: John Wiley & Sons;

- 2002; 40-46. 1971; 66: 327-330.
6. Huelsenbeck JP, Ronquist F. MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics*. 2001; 17: 754-755.
 7. Gelman A, Carlin JB, Stern H. S, et al. *Bayesian data analysis*. 2nd ed. Boca Raton: Chapman & Hall/CRC; 2003; 46-88.
 8. Lee PM. *Bayesian statistics: An introduction*. 3rd ed. London: A Hodder Arnold Publication; 2004; 31-71.
 9. Lindley DV. *Introduction to probability and statistics from a Bayesian viewpoint: Inference Pt.2*. 1st ed. London: Cambridge University Press; 1965; 1-121.
 10. Zellner A. *An Introduction to Bayesian inference in econometrics*. 1st ed. Hoboken: Wiley-Interscience; 1996; 363-399.
 11. Patil VH. Approximations to the Behrens-Fisher distribution. *Biometrika*. 1965; 52: 267-271.
 12. Box G. E. P, Tiao GC. A Bayesian approach to the importance of assumptions applied to the comparison of variances. *Biometrika*. 1964; 51: 153-167.
 13. Box G. E. P, Tiao GC. *Bayesian inference in statistical analysis*. 1st ed. Canada: Wiley-Interscience; 1992; 465-470.
 14. First Bayes. Tony O'Hagan-First Bayes software; 2017. Available from: <http://tonyohagan.co.uk/1b/> (accessed December 17, 2018).
 15. Dinov ID, Christou N. Statistics online computational resource for education. *Teach Stat*. 2009; 31: 49-51.
 16. Spiegelhalter DJ, Myles JP, Jones DR, et al. Bayesian methods in health technology assessment: a review. *Health Technol Assess*. 2000; 4: 121.
 17. Tango T. Normal distribution or log-normal distribution?. *Rinsho Byori*. 1979; 27: 1151-1155.
 18. Zellner A. Bayesian and non-Bayesian analysis of the log-normal distribution and log-normal regression. *J Am Stat Assoc*.