Journal of Statistical Theory and Applications
Vol. 19(4), December 2020, pp. 540–546
DOI: https://doi.org/10.2991/jsta.d.201230.001; ISSN 1538-7887
https://www.atlantis-press.com/journals/jsta

Bayes Factors for Comparison of Two-Way ANOVA Models

R. Vijayaragunathan[1,*], M. R. Srinivasan[2]

[1] Assistant Professor, Department of Statistics, Indira Gandhi College of Arts and Science, Puducherry, India
[2] Professor and Head, Department of Statistics, University of Madras, Chennai, India

ARTICLE INFO

Article History

Received 19 Oct 2020
Accepted 28 Dec 2020

Keywords

ANOVA model
Bayes factor
Zellner's *g* prior
Jeffrey-Zellner-Siow prior
Hyper-*g* prior
Simulation data

ABSTRACT

In the traditional two-way analysis of variance (ANOVA) model, it is possible to identify the significance of both the main effects and their interaction based on the *P* values. However, it is not possible to determine how much data supports the model when these effects are incorporated into the model. To overcome this practical difficulty, we applied Bayes factors for hierarchical models to check the intensity of the effects (both main and interaction). The objective is to identify the impact of the main and interaction effects based on a comparison of Bayes factors of the hierarchical ANOVA models. The application of Bayes factors enables to observe which model strengthens more while including or eliminating the effects in the model. Consequently, this paper proposes three priors such as Zellner's *g*, Jefferys-Zellner-Siow, and Hyper-*g* priors, to compute the Bayes factor. Finally, we extended this procedure to the simulation data for the generalization of the Bayesian results.

1. INTRODUCTION

The Bayesian method is widely applied in a range of disciplines such as engineering, medicine, biology, economics, etc. Jeffrey [1] developed the Bayesian approach to solve hypothesis testing problems. He also developed a methodology for quantifying the evidence in favor of a scientific theory. Kass and Raftery [2] discussed the Bayesian viewpoints and several techniques for computing Bayes factor. Bayesian approach to the analysis of variance (ANOVA) model was discussed by many authors (Liang *et al.* [3]; Maruyama and George [4]; Wetzels *et al.* [5], etc.). In this study, we aim to identify the influence of the effects when we incorporate eliminating from the ANOVA two-way models by the Bayesian approach. The main advantage of this approach is that the evidence is quantified, instead of forcing an all or none decision as in the frequentist approach. The Bayes factor results are quantified in Table 1. Several priors are discussed to find the Bayes factor by many authors, Liang *et al.* [3], Wang and Sun [6], Kass and Raftery [2], Clyde *et al.* [7].

1.1. Zellner's *g* Prior

Zellner's *g* priors are most commonly used prior for Bayesian hypothesis testing. The limit of inverse-gamma distribution with shape and scale used as prior density. Many authors, Kass and Raftery [2], Zellner [8], and others discussed extensively this prior. Its modified version was introduced by Liang *et al.* [3]. That is, the g prior is assigned to all the regression coefficients except the intercept. We consider two priors by setting the value of *g*, (i) *Unit Information Prior (UIP)* if $g = n$ and (ii) *Risk Inflation Criterion (RIC)* if $g = k^2$, where n = number of observations and k = the number of predictors in the regression model. The Bayes factor for the full model to the null model is

$$BF = (1+g)^{(n-k-1)/2} \left[1 + g\left(1 - R^2\right)\right]^{-(n-1)/2} \tag{1}$$

1.2. Jeffrey-Zellner-Siow Prior

Jeffrey-Zellenr-Siow (JZS) prior is a mixture of priors, such that, Jeffreys' prior on the intercept and an inverse gamma with ½ and $n/2$ prior on g. Further, estimate g from the data and the Bayes factor is not analytically available, Liang *et al.* [3], Ley et al. [9]. The Bayes Factor for

*Corresponding author. Email: rvijayaragunathan@gmail.com

Table 1 Interpretations of Bayes factor.

Data Support the Model M_j if $\log(B_{ij})$ Value	Data Support the Model M_i if $\log(B_{ij})$ Value	Evidence for Respective Model
< -2	> 2	Decisive
−2 to −1	1 to 2	Strong
−1 to −0.5	0.5 to 1	Substantial
−0.5 to 0	0 to 0.5	Poor

the full model to the null model is

$$BF = \frac{(n/2)^{1/2}}{\Gamma(1/2)} \int_0^\infty (1+g)^{(n-k-1)/2} \left[1 + g\left(1 - R^2\right)\right]^{-(n-1)/2} g^{-3/2} e^{-n/2g} \, dg \tag{2}$$

1.3. Hyper-*g* Prior

Hyper-g prior is a family of prior distributions on g and this hyper-g approach $p(g) = \frac{a-2}{2}(1+g)^{-a/2}, g > 0$. The range of values for a lies in between 2 and 4. It produces different behavior of hyper-g prior, generally, its proposed values are $a = 3$ and/or $a = 4$, apart from that the issue of arbitrary constants of proportionality leads to indeterminate Bayes factors. The Bayes factor for the full model to the null model is

$$BF = \frac{a-2}{2} \int_0^\infty (1+g)^{(n-k-1-a)/2} \left[1 + g\left(1 - R^2\right)\right]^{-(n-1)/2} dg \tag{3}$$

To avoid the ambiguous conclusions of the Bayesian approach, we propose three different priors such as Zellner's g, Jeffreys-Zellner-Siow, and Hyper-g priors, Wetzels *et al.* [5], Perrakis and Ntzoufras [10], Vijayaragunathan and Srinivasan [11]. The level of interpretations for the Bayes factor for the null and alternative model are in Table 1. The results are quantified evidence of data support the respective model, which was discussed extensively by Jeffreys [1] and Kass and Raftery [10].

2. ANOVA TWO-WAY MODELS AND BAYESIAN APPROACH

Factor A has a levels, a factor B has b levels, and these factors are arranged in a factorial design each replicate of the experiment contains all ab treatment combinations. An ANOVA two-factor experimental design is

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk} \quad i = 1, 2, \cdots, a; j = 1, 2, \cdots, b; k = 1, 2, \cdots, n \tag{4}$$

where μ is the overall mean effect, τ_i is the effect of ith level of factor A, β_j is the effect of jth level of factor B, $(\tau\beta)_{ij}$ is the effect of the interaction between A and B, and ε_{ijk} is a random error component.

The two-way ANOVA model is converted into a regression framework model for our convenience in applying the Bayesian technique to different priors Zellner *et al.* [13]. To test the outcome of main effects A, and B and the interaction effects of A and B, by comparing four possible ANOVA designs. An example of a two-way ANOVA model is discussed in the next section. The main effect A is plate material has three levels P_1, P_2, and P_3; the second main effect B is the temperature also has three levels T_1, T_2, and T_3; and its interaction effect has nine possible combinations. The regression framework for the two-way ANOVA experiment is as follows:

$$Y = X\beta \tag{5}$$

where X- is a design matrix formed according to the factors by using dummy codes. The design matrix is not necessarily a full column rank. We adopt the sum-to zero constraints, so the regression coefficient of the last level of each factor is equal to the sum of the other regression coefficients. Therefore, the two-factor ANOVA experiment full model contains grand mean μ, two levels of main effect A, and two levels of the main effect B and it's interaction effects (the third level of A and B omitted because of sum-to-zero constraint) is

$$Model\,1: \; Y = 1_n\mu + X_{p_1}\beta_{p_1} + X_{p_2}\beta_{p_2} + X_{t_1}\beta_{t_1} + X_{t_2}\beta_{t_2} + X_{p_1xt_1}\beta_{p_1xt_1} + X_{p_1xt_1}\beta_{p_1xt_1} + X_{p_2xt_1}\beta_{p_2xt_1} + X_{p_2xt_1}\beta_{p_2xt_1} \tag{6}$$

The two-way ANOVA model containing only the main effects, without interaction effects, is

$$Model\;2: \; Y = 1_n\mu + X_{p_1}\beta_{p_1} + X_{p_2}\beta_{p_2} + X_{t_1}\beta_{t_1} + X_{t_2}\beta_{t_2} \tag{7}$$

The two-way ANOVA model containing only the main effect A is

$$Model\ 3: Y = 1_n\mu + X_{p_1}\beta_{p_1} + X_{p_2}\beta_{p_2} \tag{8}$$

The two-way ANOVA model containing only the main effect B is

$$Model\ 4: Y = 1_n\mu + X_{t_1}\beta_{t_1} + X_{t_2}\beta_{t_2} \tag{9}$$

The two-way ANOVA model without main and interaction effects (the null model) is

$$Model\ 0: (Null)\ Y = 1_n\mu \tag{10}$$

Let $BF_{M_i:M_0}$ or BF_{i0}, $i = 1, 2, 3,$ and 4, be the Bayes factor for the ith model to the null model Rouder *et al.* [5]. The ANOVA Model 1 (full model) is a larger model to other models; it means that the reduced models are nested to the full model. We compare reduced models with the larger model to test the cause of the respective factors. To test an interaction effect, we want to find BF_{12} that is Bayes factor for a larger model (model 1) to the reduced model (model 2). We have to compute the Bayes factor BF_{10} and BF_{20} to find $BF_{12,}$ then the ratio of these Bayes factors is

$$\frac{BF_{10}}{BF_{20}} = BF_{10} \times BF_{02} = BF_{12} \tag{11}$$

because Bayes factors are inverse and transitivity. Similarly, the ratio of Bayes factors BF_{23} and BF_{24} are found by using BF_{20} & BF_{30} and BF_{20} & BF_{40} respectively to test the effect of main effects. It does not give any assertion about the impact of main and interaction effects in the model based on a single prior. Therefore, we proposed three different priors to find the Bayes factor. In the next session, we will discuss on Bayes factors of these priors. In general, it couldn't take any legitimate decision based on the simple application. Further, we decided to find the number of Bayes factors from the simulation datasets, which is discussed in the fourth section.

3. APPLICATIONS

Consider an engineer is designing a battery for use in a device that will be subject to some extreme variations in temperature, and he has three possible choices of plate material [14]. The engineer has no control over the temperature extremes that the device will encounter, and he knows from experience that temperature will probably affect the effectiveness of battery life. He decides that all three plate materials at three temperature levels (15∘F, 70∘F, and 125∘F) because these temperature levels are consistent with the product end-use environment. Four batteries were tested at each combination of plate material and temperature and all 36 tests are run in random order. The F ratio and corresponding P values of four ANOVA models have presented in Table 2. The main effect of A (plate material) is significant in models 1 and 2, but in the single treatment model 3, there is no significant difference among the average battery life of different plate material. Thus, the effect of plate material would be significant only if it is with the other effect. The main effect of B (temperature) is strongly significant in all models when it exists. Also, the interaction effect between the plate material and temperature is moderately significant.

In the first step of the Bayesian approach, we have computed the Bayes Factor for all ANOVA models with the null model, to know how the data support the respective models. From Table 3, the evidence of data supports "decisively" for the two-way ANOVA models such as model 1 (full model), model 2 (except interaction effects model), and model 4 (only main effect B). But, the Bayes factor for model 3 (only the main effect A) does not provide impressive results, such that two priors are "poorly" support the null model and the other three priors are almost "poorly" support model 3.

The two-way ANOVA with interaction effect model (full model) is a larger model, other models are nested models to it, and so on. Now, to test the effect of interaction by comparing the full model and without the interaction model. In the same way, we used the without interaction model and respective main effect models to test the main effects. In general, the Bayes factors for the larger model to the nested model have been computed to know the strength of the effect which was eliminated in nested models.

The ratio of Bayes factors of the full model and without interaction models indicate that there is an intensity of the interaction effect in the model. The Bayes factor values are close to one for BF(12) and BF(24), which gives the impression that the data support "strongly" for both interaction and main effect B (temperature) of battery life. All the priors support the respective models invariably for these effects. But, for the main effect, A (plate material) model provides notable results, the UIP and JZS priors are "strongly" support model 3 and the other three

Table 2 F ratio for four ANOVA models (* represents P values).

ANOVA	Main Effect A	Main Effect B	Interaction	R Square
Model 1	7.9114 (0.001976*)	28.967 (1.9×10^{-7}*)	3.5595 (0.018611*)	0.7652
Model 2	5.947 (0.00651*)	21.776 (1.24×10^{-6}*)	–	0.6414
Model 3	2.633 (0.0869*)	–	–	0.1376
Model 4	–	16.75 (9.51×10^{-6}*)	–	0.5038

priors are "strongly" support model 2. Because, the main effect A is significant when it is together with the main effect *B* and nonsignificant when it is alone in the ANOVA model, are perhaps the reason for the variation in results. We cannot trust the applications with a single Bayes factor value, therefore, in the next section, we will demonstrate several Bayes factors in the simulation study to generalize the interpretations.

4. SIMULATION STUDY

In an examination of the effects of the two-way ANOVA model, we constructed the different datasets by varying different error variance in the simulation. To find the Bayes Factors for all priors, we have to simulate datasets to obtain reliable conclusions. Firstly, we simulated 1000 different data sets with variance 1 to compute the 1000 Bayes factor values for each prior. Similarly, we simulated the other two datasets with the error variances 10 and 25 to compute Bayes factors. These results are shown in Figures 1–5 and also display the summary statistics for the Bayes factor of respective simulation data to the ANOVA model are shown in Table 4–9. The overall conclusion of all the priors do not change in BF(12) and BF(24), but it varies in BF(23). Thus, our simulation study provides broad knowledge about the different priors of various datasets. Consequently, it may be useful to make decisions for inclusion or elimination of the interaction effect in the two-way ANOVA model as well as the influence of the main effects in the model.

Table 3 Bayes factor for respective two-way ANOVA model with the null model.

Prior	BF_{10}	BF_{20}	BF_{30}	BF_{40}
UIP	4.0987	4.2990	−0.4759	3.5521
RIC	3.3896	4.3936	0.1874	3.2220
JZS	4.1908	4.1087	−0.5450	3.3363
Hyper-*g* (*a* = 3)	4.2165	3.9933	0.0615	3.1474
Hyper-*g* (*a* = 4)	4.0314	3.7685	0.1466	2.9049

UIP, Unit Information Prior; RIC, Risk Inflation Criterion; JZS, Jeffrey-Zellenr-Siow.

Table 4 Ratio of Bayes factors for the relevant ANOVA models.

Prior	BF_{12}	BF_{23}	BF_{24}
UIP	0.9534	−9.0334	1.2103
RIC	0.7715	23.4450	1.3636
JZS	1.0200	−7.5389	1.2315
Hyper-*g* (*a* = 3)	1.0559	64.9317	1.2688
Hyper-*g* (*a* = 4)	1.0698	52.6239	1.2973

UIP, Unit Information Prior; RIC, Risk Inflation Criterion; JZS, Jeffrey-Zellenr-Siow.

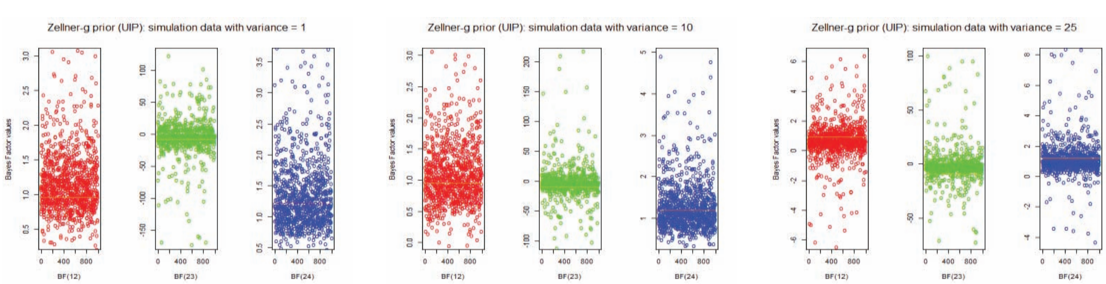

Figure 1 Bayes factors of simulation data sets for different variances to the Zellner's *g* prior (unit information prior).

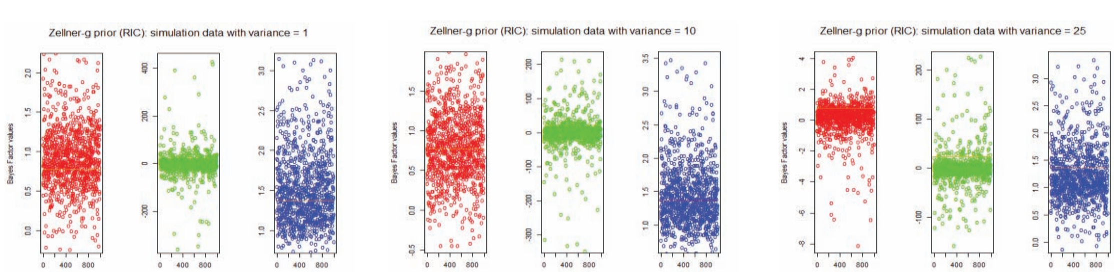

Figure 2 Bayes factors of simulation data sets for different variances to the Zellner's *g* prior (risk inflation criterion).

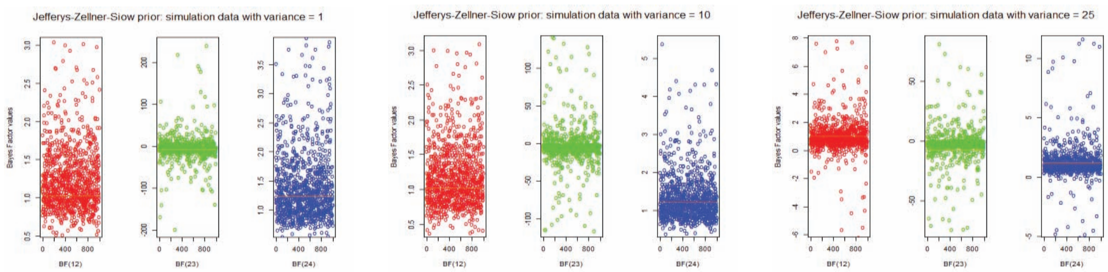

Figure 3 Bayes factors of simulation datasets for different variances to the Jeffreys-Zellner-Siow prior.

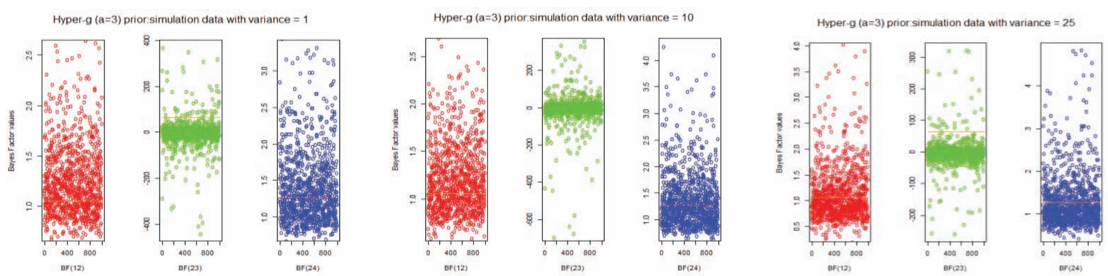

Figure 4 Bayes factors of simulation datasets for different variances to the hyper-g ($a = 3$) prior.

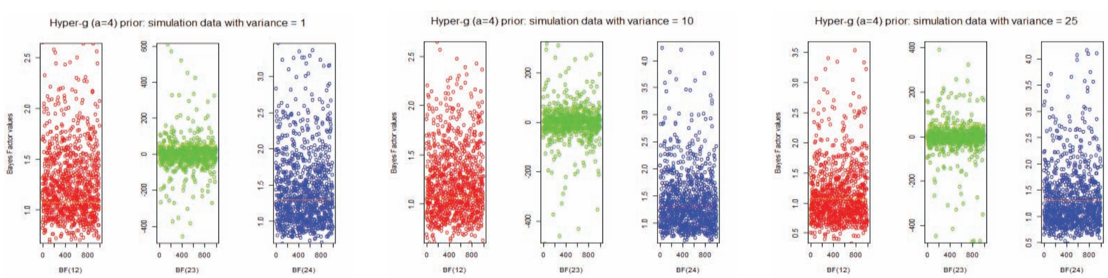

Figure 5 Bayes factors of simulation datasets for different variances to the hyper-g ($a = 4$) prior.

Table 5 Summary of Bayes factors for Zellner's g prior (unit information prior).

Error Variance	**Descriptive**	BF_{12}	BF_{23}	BF_{24}
	Actual	**0.95**	**−9.03**	**1.21**
1	Mean (SD)	1.19 (0.49)	−7.10 (172.93)	1.38 (0.61)
10	Mean (SD)	1.05 (2.20)	−1.88 (77.78)	1.51 (5.94)
25	Mean (SD)	0.73 (3.91)	1.51 (5.94)	1.05 (2.83)

Table 6 Summary of Bayes factors for Zellner's g prior (risk inflation criterion).

Error Variance	**Descriptive**	BF_{12}	BF_{23}	BF_{24}
	Actual	**0.77**	**23.45**	**1.36**
1	Mean (SD)	0.92 (0.43)	−132.39 (4261.64)	1.53 (0.48)
10	Mean (SD)	0.79 (0.49)	−1.70 (173.45)	1.48 (0.60)
25	Mean (SD)	0.22 (3.99)	−5.83 (530.68)	1.15 (1.80)

5. CONCLUSIONS

An ANOVA is one of the most popularly used statistical tools in research. We considered a two-way ANOVA model consisting of two factors such as plate material (A) and temperature (B) related to battery life. In the classical approach of the ANOVA two-way model, we examined and discussed the significance of main and interaction effects as usual. We also examined the contribution level of effect(s) in the two-way ANOVA model by the Bayes factor of choice of different priors. In Bayesian Approach, it is found that the Bayes factors for all four ANOVA

Table 7 Summary of Bayes factors for Jeffreys-Zellner-Siow prior.

Error Variance	Descriptive	BF_{12}	BF_{23}	BF_{24}
	Actual	**1.02**	**−7.54**	**1.23**
1	Mean (SD)	1.26 (0.46)	1.24 (0.84)	1.43 (0.69)
10	Mean (SD)	1.24 (0.84)	−21.61 (417.78)	1.28 (3.14)
25	Mean (SD)	1.02 (2.17)	−4.55 (75.76)	1.64 (8.01)

Table 8 Summary of Bayes factors for hyper-*g* ($a = 3$) prior.

Error Variance	Descriptive	BF_{12}	BF_{23}	BF_{24}
	Actual	**1.06**	**64.93**	**1.27**
1	Mean (SD)	1.27 (0.38)	−9.88 (364.94)	1.43 (0.54)
10	Mean (SD)	1.23 (0.41)	−36.57 (733.52)	1.42 (0.76)
25	Mean (SD)	1.17 (0.85)	−16.17 (1.33)	1.33 (1.01)

Table 9 Summary of Bayes factors for hyper-*g* ($a = 4$) prior.

Error Variance	Descriptive	BF_{12}	BF_{23}	BF_{24}
	Actual	**1.07**	**52.62**	**1.30**
1	Mean (SD)	1.28 (0.37)	32.36 (750.47)	1.46 (0.55)
10	Mean (SD)	1.25 (0.40)	8.64 (281.79)	1.45 (0.71)
25	Mean (SD)	1.18 (0.67)	−6.58 (309.83)	0.95 (14.07)

models with the null model to know the strength of interaction and main effects in the respective models. From the simulation study, the Bayes factor results provide ample evidence for support of the data "strongly" both the models (with the interaction effect and main effect B (temperature)). For these effects, all the priors support the respective models invariably. Nevertheless, the main effect A (plate material) model provides "strongly" support to model 2 as well as model 3. We may deduct that the main effect A is significant only when it is together with the main effect B and nonsignificant when it is alone in the ANOVA model. All priors generate almost similar average Bayes factor values with different standard deviations for the models consisting of interaction and the main effect of temperature. It is observed that the same priors do not generate the same results in the model consisting of the main effect of plate material. Thus, we conclude that there is no single correct prior distribution and if we use different priors it may provide more instructive conclusions. The simulation technique may be highly rewarding to the researchers if properly applied in their respective areas.

CONFLICTS OF INTEREST

The authors declare that they have no conflict of interest.

AUTHORS' CONTRIBUTIONS

Both the authors contributed equally and approved the final manuscript.

ACKNOWLEDGMENTS

The authors would like to thank the editor and the reviewers for their constructive comments and suggestions which highly improved the paper. Also, state that this research did not receive any specific grant from any funding agencies.

REFERENCES

1. H. Jeffreys, Theory of Probability, third ed., Oxford University Press, Oxford, UK, 1961.
2. R.E. Kass, A.E. Raftery, J. Am. Stat. Assoc. 90 (1995), 773–795.
3. F. Liang, R. Paulo, G. Molina, M.A. Clyde, J.O. Berger, J. Am. Stat. Assoc. 103 (2008), 410–423.
4. Y. Maruyama, E.I. George, Ann. Stat. 39 (2011), 2740–2765.
5. R. Wetzels, R.P.P.P. Grasman, E.-J. Wagenmakers, Am. Stat. 66 (2012), 104–111.

6. M. Wang, X. Sun, Commun. Stat. Theory Methods. 43 (2014), 5072–5090.
7. M.A. Clyde, J. Ghosh, M.L. Littman, J. Comput. Graph. Stat. 20 (2011), 80–101.
8. A. Zellner, Statistician. 32 (1983), 23–34.
9. E. Ley, M.F.J. Steel, J. Econom. 171 (2012), 251–266.
10. K. Perrakis, I. Ntzoufras, WIREs Comput. Stat. 10 (2018), e1442.
11. R. Vijayaragunathan, M.R. Srinivasan, Adalya J. 9 (2020), 898–903.
12. A. Zellner, A. Siow, in Bayesian Statistics: Proceeding of the First International Meeting, University of Valencia Press, Valencia, Spain, 1980, pp. 585–603.
13. J.N. Rouder, R.D. Morey, P.L. Speckman, J.M. Province, J. Math. Psychol. 171 (2012), 251–266.
14. D.C. Montgomery, Design and Analysis of Experiments, John Wiley & Sons, Inc., Hoboken, NJ., 2013.