Journal of Medical and Health Studies
ISSN: 2710-1452
DOI: 10.32996/jmhs
Journal Homepage: www.al-kindipublisher.com/index.php/jmhs

JMHS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Bayesian Approach: Adding Clinical Edge in Interpreting Medical Data

R.Vijayaragunathan[1] ✉ Kishore K John[2] and M.R.Srinivasan[3]
[12]Indira Gandhi College of Arts and Science, Kathirkamam, Puducherry, India
[3]School of Mathematical Sciences, University of Hyderabad, Hyderabad, India
Corresponding Author: R.Vijayaragunathan, **E-mail**: rvijayaragunathan@gmail.com

| ABSTRACT
In frequentist tests, the significance testing framework for null hypothesis permits dichotomous conclusions alone, and such tests do not quantify the strength of the evidence supporting the null hypothesis. Under the Bayesian approach, probability reflects their uncertainty or degree of belief, that is, how scientific belief should be modified by data. This paper attempts to demonstrate the advantages of the Bayes factor in hypothesis testing that can quantify evidence in favour of the null hypothesis and how the prior specification is used for statistical tools, such as independent t-test and Analysis of Variance (ANOVA). Despite the advantages of the Bayesian approach, the use of conventional tests that rely on inference by p-values is ubiquitous in medical research. The adoption of the Bayesian approach may be seriously hindered by the absence of formulae, algorithms, etc. Furthermore, we have attempted to validate our argument by interpreting the application of both the Frequentist and Bayesian approaches for dietary intake of calcium mg/day with the help of JASP software.

| KEYWORDS

Independent t-test, Analysis of Variance, Bayes Factor, Jeffrey-Zellner-Siow Prior, JASP Software.

| ARTICLE DOI: 10.32996/jmhs.2022.3.1.9

1. Introduction

The practice of using frequentist tests as a primary choice has become ubiquitous among health and social science researchers and scientists over the years. Of late, Bayesian Inference has been extensively used in research under a wide range of disciplines. The early development of Bayesian hypothesis testing and the concept of Bayes Factors are discussed in the book "*Theory of Probability*" by Jeffrey [3]. In brief, Bayes Factors have many strengths and limitations under the Bayesian approach. The chief limitations of Bayes Factors are their sensitivity to the assumptions in a parametric model and the choice of priors (Kass and Raftery 1995). The basic idea of Bayes factors, choice of priors, the asymptotic approximation to find the posterior, and various examples are explained by (Kass and Raftery 1995). Bayesian tests of two or more independent groups and some selected models are described by (Gelman 2005), (Rouder et al. 2009), (Wang et al. 2016), (Morey et al. [7], Wetzels et al. 2012), and others. The main objective of this paper is to demonstrate the superiority of the Bayesian Hypotheses Tests when applied in health science research and its ability to enhance the distinctiveness of such research work.

2. Literature Review

Suppose a model data D, assume that the two hypotheses H_0 and H_1 according to a probability density $P(D/H_0)$ and $P(D/H_1)$. The prior probabilities are $P(H_0)$ and $P(H_1)$, and the posterior probabilities to the models are $P(H_0/D)$ and $P(H_1/D)$ = 1-$P(H_0/D)$. In general, using Bayes Theorem, the posterior probability to the data is

$$P(H_i/D) = \frac{P(H_i)P(D/H_i)}{P(H_0)P(D/H_0) + P(H_1)P(D/H_1)} \quad (1)$$

These are the quantities that are conditional on observed data. Therefore, an appropriate statistic for comparing hypotheses is the posterior odds

$$\frac{P(H_0/D)}{P(H_1/D)} = \frac{P(H_0)P(D/H_0)}{P(H_1)P(D/H_1)} \tag{2}$$

From this, we see that the transformation of the prior odds into the posterior odds is a matter of multiplying by the Bayes factor

$$B_{01} = \frac{P(D/H_0)}{P(D/H_1)} \tag{3}$$

Note that the Bayes factor is equal to the posterior odds when the equal prior probability is given to each hypothesis. It has on occasion been suggested that the value of the Bayes factor can be interpreted independently of the prior odds.

The Bayesian Approach, which quantifies evidence instead of forcing an all-or-none decision (in frequentist, either accept or reject the null hypothesis), can distinguish between "data support H_0" and "data are not diagnostic." Interpretations of the Bayes factor for the null and alternative hypothesis are given in the below table. The results are quantified evidence in Table 1 as four-component each for both the hypothesis, which is discussed elaborately by Jeffery (1961), Kass, and Raftery (1995).

Table: 1 interpretation for Bayes Factor

log (BF_{01})	**Interpretation of Bayes Factor**
<-2	Decisive evidence for H_1
-2 to -1	Strong evidence for H_1
-1 to -0.5	Substantial evidence for H_1
-0.5 to 0	Poor evidence for H_1
0 to 0.5	Poor evidence for H_0
0.5 to 1	Substantial evidence for H_0
1 to 2	Strong evidence for H_0
>2	Decisive evidence for H_0

3. Methodology

In this section, we will discuss the technique used by Jeffreys-Zellner-Siow prior to the selected statistical tools, such as independent two samples t-tests and analysis of variance one-way model.

3.1 Jeffreys-Zellner-Siow Bayes Factor for Independence two-sample t-test:

JZS Bayes Factor for the independent two-sample t-test was developed by (Rouder et al. 2012). Let x_i and y_i denote the i[th] observations in the first and second groups, respectively.

$$x_i \overset{iid}{\sim} N(\mu - \frac{\alpha}{2},\ \sigma^2),\ i = 1, 2, \cdots, n_1 \text{ and } y_i \overset{iid}{\sim} N(\mu + \frac{\alpha}{2},\ \sigma^2),\ i = 1, 2, \cdots, n_2 \tag{4}$$

Where μ and α denote the grand mean and total effect respectively and, $n_1$and n_2 denote the sample size for the first and second groups, respectively. The null hypothesis corresponds to $\alpha = 0$. In the parameterization, the null hypothesis corresponds to effect size $\delta = 0$ and the Jeffrey-Zellner-Siow prior for $\delta \sim Cauchy$ under the alternative hypothesis. Priors needed for μ andσ^2, these parameters are common to both models, the resulting BF is relatively robust to the choice. The Jeffreys noninformative prior onσ^2, i.e., $p(\sigma^2) = \frac{1}{\sigma^2}$ appropriate. A noninformative prior may also be placed on μ, this prior denoted $p(\mu) = 1$. Jeffreys-Zellner-Siow Bayes Factor for two-sample t-test (Rouder et al. 2012) is in equation (5).

$$BF_{01} = \frac{\left(1 + \frac{t^2}{v}\right)^{-(v+1)/2}}{\int (1+Ng)^{-\frac{1}{2}} (2\pi)^{-\frac{1}{2}} g^{-3/2} e^{-1/2g}\, dg} \tag{5}$$

where t = two-sample t value, $N = \frac{n_1 \times n_2}{n_1 + n_2}$ is effect size and $v = n_1 + n_2 - 2$ degrees of freedom.

3.2 Jeffreys-Zellner-Siow Bayes Factor for Analysis of Variance One-way model:

JZS Bayes Factor for ANOVA was discussed elaborately by (Wetzels et al. 2012). The extending tending Jeffrey's suggestion to variable selection in the regression model, (Zellner and Siow 1980) proposed a multivariate Cauchy prior on regression coefficients and a flat prior on the common intercept. (Liang et al. 2008) developed the JZS prior as a mixture of g- priors, i.e., a $Inverse\ Gamma\left(\frac{1}{2}, \frac{n}{2}\right)$ prior on g and Jeffreys' prior on the precision ∅

$$
\begin{aligned}
p(\phi) &\propto \frac{1}{\phi} \\
p(\beta \mid \phi, g, \mathrm{X}) &\propto \int N(0, \frac{g}{\phi}(X^T X)^{-1})\, p(g)\, dg \\
p(g) &= \frac{(n/2)^{1/2}}{\Gamma(1/2)}\, g^{-3/2}\, e^{-n/2g}
\end{aligned}
\tag{6}
$$

In the JZS approach, the Bayes Factor comparing the full model to the null model is in equation (7). A drawback of the JZS prior is that the Bayes Factor is not analytically available but not vulnerable to the Jeffreys-Lindley- Bartlett paradox nor to the information paradox.

$$BF_{01} = \frac{(n/2)^{1/2}}{\Gamma(1/2)} \int_0^\infty (1+g)^{(n-k-1)/2}\, [1 + g(1-R^2)]^{-(n-1)/2} g^{-3/2} e^{-n/2g}\, dg \tag{7}$$

4. Results and Discussion

We discuss an interpretation of both frequentist and Bayesian approaches of independent t-test and Analysis of Variance with illustrations.

4.1 *Illustration for Bayesian Independent t-test*

A clinical trial was conducted at the gynaecology unit of a major hospital to determine the effectiveness of drug A in preventing premature birth. In the trial, 30 pregnant women were studied, 15 in a treatment group to receive drug A and 15 in a control group to receive a placebo. The patients were to take a fixed dose of each drug on a one-time-only basis between the 24th and 28th weeks of pregnancy. The patients were assigned to groups based on computer-generated random numbers, where for every two patients eligible for the study, one was random to the treatment group and the other to the control group. This example from Bernard Rosner [8] was regarding the birth weights in a clinical trial to test a drug for preventing low-birth-weight deliveries of the treatment group and control group. Both the treatment and placebo groups were the same sample size 15 pregnant women, the treatment group's mean and standard deviation of baby weights (lb) are 7.08 and 0.899, and the placebo group's mean and standard deviation of baby weights (lb) are 6.26 and 0.961 respectively.

The test statistic is 2.414, and the corresponding p-value is 0.023, which is less than 0.05, so we reject the null hypothesis. Hence we conclude that there is a significant difference between the treatment and placebo group's baby weights (lb). It means that the treatment group's baby birth weight is more than the placebo group's baby birth weight.

Table: 2 Descriptive statistics for treatment and placebo group

Group	N	Mean	SD	SE
1	15	7.080	0.899	0.232
2	15	6.260	0.961	0.248

Table: 3 JASP output for two independent samples of treatment and placebo group

t	df	P	Mean Difference	SE Difference	Cohen's d
2.414	28.000	0.023	0.820	0.340	0.881

Table: 4 JASP output of Bayesian t-test.

	BF01	error %
Data	0.358	8.283e-4

In the Bayesian approach, the Bayes factor BF_{01} value for two independent t-tests is 0.358, i.e., $\log_{10} BF_{01}$=0.4461, now we refer the table 1, Bayes factor value lies 0 to -0.5, it would be poor evidence to support the alternative hypothesis. In other words, BF_{10}=1/0.358=2.79 the data support the alternative hypothesis more than 2.79 times the null hypothesis.

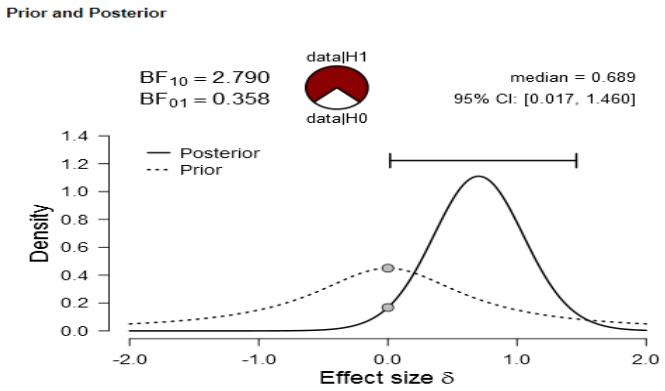

Figure 1: JASP output shows the prior and posterior distributions.

In the figure, the dotted line represents the prior distribution, and the smooth line represents the posterior distribution indicated by graphical for Bayesian independent t-test and also indicates 95% credibility interval for mean differences from 0.017 to 1.460, which means that 95% confident that the mean difference lies in that interval and Bayes Factor BF_{01}=0.358.

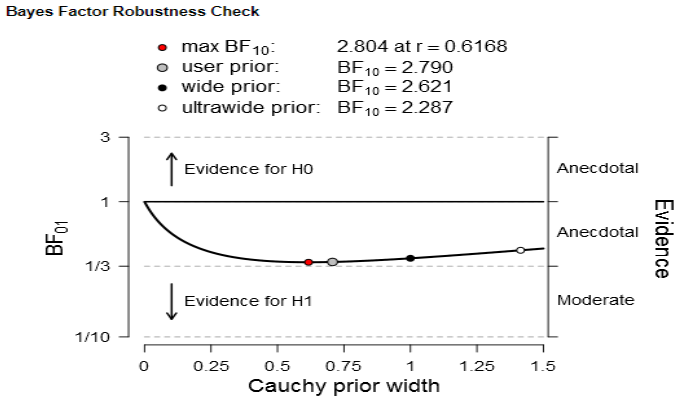

Figure 2: JASP output for Bayes factor for Robustness check for the different prior values.

If we increase the prior (Cauchy prior) values, the data supports the alternative hypothesis only. If the prior is 0, then the data does not support any hypothesis, but it (prior) would be an increase wider the data somewhat support the alternative hypothesis.

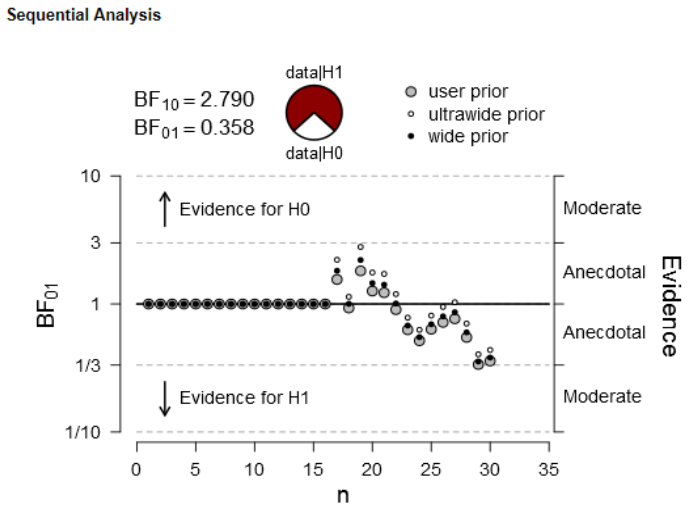

Figure 3: JASP output for sequential Analysis for an independent t-test.

Interestingly, if the sample size is small, the data does not support any hypothesis. If the sample size is lies between 15 and 20, the data is found to somewhat support the null hypothesis, but finally, the Bayes Factor value falls in the Poor Evidence region of the alternative hypothesis.

4.2 Illustration for Bayesian Analysis of Variance

The dietary intake of calcium was investigated by (Ilich-Ernst et al. 2002) among a cross-section of 113 healthy women ages group between 20 and 88. The researchers formed four age groupings as follows: 20.0–45.9 years in Group A; 46.0–55.9 years in group B; 56.0–65.9 years in group C; and over 66 years in group D. Calcium was measured in mg/day from the food intake. Using JASP software, the ANOVA output for calcium intake from food among the different age groups of women was obtained, which is shown in table 5. The test statistic value was 9.359, and the corresponding significant value was <0.001. Subsequently, the null hypothesis was rejected. Hence, it was concluded that there is a significant difference in the average intake of calcium among the different age groups of women. Furthermore, we applied the Post-Hoc test to find which pair of treatment means were significantly different.

Table:5 JASP output for ANOVA one-way classification

Cases	Sum of Squares	Df	Mean Square	F	p
Group	5.931e+6	3	1.977e+6	9.359	< 0.001
Residuals	2.303e+7	109	211252.293		

Table: 6 JASP output for Post Hoc Comparisons-Group

Treatments	Mean Difference	SE	t	P-value
A Vs B	455.721	157.136	2.900	0.023
A Vs C	574.536	129.950	4.421	<0.001
A Vs D	596.634	118.336	5.042	<0.001
B Vs C	118.815	149.579	0.794	0.857
B Vs D	140.914	139.609	1.009	0.744
C Vs D	22.098	108.100	0.204	0.997

The post-hoc test shows that the average calcium intake of group A (20 to 45.9 years) is different from all other age groups of women, such as 46 to 55.9 years, 56 to 65.9 years, and > 66 years. Thus, the average intake of calcium (mg/day) among the above 46 years of women does not make much difference, but this group is different.

Table: 7 JASP output for Bayesian Analysis of Variance

Models	P(M)	P(M|data)	BFM	BF10	error %
Group	0.500	0.999	1463.727	1.000	
Null model	0.500	6.827e-4	6.832e-4	6.832e-4	6.407e-5

In the Bayesian ANOVA table 7, P(M) indicates prior model probabilities, P(M/data) indicates the updated probabilities after the data, and BF_{10} Mindicates the degree to which the data have changed the prior model odds. The Bayes Factor value is 0.006832, i.e., $\log_{10}$ (0.0006832)=-3.17, and it indicates that Decisive evidence supports the alternative hypothesis. In other words, the data supporting the alternative hypothesis is found to be almost 3.17 times that of the null hypothesis. It provides additional support to correctly interpret the results, whereas researchers can detect the tendency of the data instead of accepting or rejecting the hypothesis completely.

5. Conclusion

In this article, we tried to outline and apply a default Bayesian approach to two simple and widely used statistical tools, such as independent two-sample t-test and Analysis of Variance with the help of JASP – freely downloadable statistical software. The interpretations of frequentist tests are mostly familiar to researchers than the Bayesian tests. We have demonstrated how the Bayesian approach would differ when compared to the frequentist test in a two-sample t-test and ANOVA. Under the Bayesian approach, researchers may be using different priors to their study, and we preferred the Jeffreys-Zellner-Siow prior for both the statistical tests as it directly follows the regression framework (Wetzels et al. 2012). The Bayesian approach to the test offers practical advantages over the frequentist approach as it quantifies evidence in favour of the null hypothesis. As observed in the illustration, under the frequentist independent two-sample t-test, the treatment group's baby birth weight differs from the placebo group's baby birth weight. Whereas, in the Bayesian approach, the data supported the alternative hypothesis by more than three times that of the null hypothesis. In ANOVA, under frequentist, there was a significant difference among the average calcium (mg/day) intake from the food for various age groups of women, and further, the Post-hoc test showed that calcium intake per day of group 'A' women are different from the other age groups of women. However, under the Bayesian approach, it was found that the data supported the alternative hypothesis by almost 3.17 times than the null hypothesis. From the illustrations, we detected Bayesian inferences that there was *poor evidence to support the alternative hypothesis* in the two-sample t-test and *decisive evidence to support the alternative hypothesis* in ANOVA design. Such inferences would serve as additional support to interpret the results whereby medical researchers can practically observe the tendency of the data. This would lead them to more precise conclusions instead of accepting or rejecting the hypothesis altogether when they exclusively depend upon frequentist tests. In the future, medical and paramedical researchers may apply Bayesian concepts to draw findings with instructive conclusions.

Disclosure statement: No potential conflict of interest was reported by the authors.

Funding: No funding grant from any funding agencies/research institutions.

Ethical approval: We used secondary data to apply Bayesian concepts. Therefore, no need to get approval.

Informed consent: Not required

References

[1] Daniel W. W, and Cross C.L. (2013). Biostatistics: A foundation for analysis in the Health Sciences. *Tenth edition, John Wiley & Sons, Inc*.
[2] Gelman A, and Carlin J.B, Stren HS, Rubin DB. (2004). Bayesian Data Analysis. *Chapman & Hall/CRC*.
[3] Jeffreys H. (1961). Theory of Probability. *Oxford University Press*.
[4] Kass R.E, and Raftery A.E. (1995). Bayes Factors. *Journal of the American Statistical Association*. 90, 773-795.
[5] Liang F, Paulo R, Molina G, Clyde MA and Berger JO. (2008). Mixtures of g priors for Bayesian Variable Selection. *Journal of the American Statistical Association*. *103*(481).
[6] Marsman M, and Wagenmakers E. (2017). Bayesian benefits with JASP. *European Journal of Developmental Psychology*. *14(*5) 545-555.
[7] Morey R.D, Romeijn J, and Rouder J.N. (2016). The philosophy of Bayes Factors and the quantification of statistical evidence, *Journal of Mathematical Psychology*, 72, 6-18.
[8] Rosner B. (2011). Fundaments of Biostatistics, *Nelson Education Ltd, Canada*.
[9] Rouder J.N, Morey R.D, Speckman P.L, and Province J.M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology 171*: 251-266.
[10] Rouder J.N, Speckman P.L, Sun D, and Morey R.D. (2009). Bayesian t-tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*. *16*(2): 225-237.
[11] Wang M, and Liu G. (2016). A Simple Two-Sample Bayesian t-Test for Hypothesis Testing. *The American Statistician*. *70*(2), 195-201.

[12] Wetzels R, and Wagenmakers E. (2012). A default Bayesian hypothesis test for correlations and partial correlations. *Psychonomic Bulletin & Review*. 19: 1057-1064.

[13] Wetzels R, Grassman RPPP, and Wagenmakers E. A. (2012). default Bayesian hypothesis test for ANOVA Designs. *The American Statistician*. *66(*2): 104-110.

[14] Zellner A, and Siow A. (1980). Posterior Odds Ratio for Selected Regression Hypothesis", in *Bayesian Statistics: Proceeding of the First International Meeting held in Valencia (Spain), Valencia: University Press*. 585-603.