The use of the Bayes factor for statistical inference

CRISTIAN ANTONY RAMOS-VERA[1], ROSELINE OLUWASEUN OGUNDOKUN[2]

[1]*Faculty of Health Sciences, Cesar Vallejo University, Lima, Peru*
[2]*Department of Computer Science, Landmark University Omu Aran, Kwara State, Nigeria*

Dear Editor:

The importance of replicating the results of the "$p < .05$" values in sports research is essential since several studies include specific populations of interest with small sample sizes, the results of which can be controversial due to limited statistical power and higher prevalence of false positives or type I errors (Abt et al., 2020). For example, in correlational studies with less than 250 participants, inflated and unstable coefficients with a higher level of random error are reported, which refer to a higher precision uncertainty of the true effect evidenced in the values of very wide frequent intervals and present bias when it is replicated (Schönbrodt & Perugini, 2013; Ramos-Vera, 2020). Therefore, the practice of replicable findings that generate more credible conclusive evidence in sports science research using Bayesian inference is recommended (Ly, Raj, Etz, Gronau & Wagenmakers, 2018; Marsman & Wagenmakers, 2017).

In 2020, the Journal of Human Sport and Exercise published two important articles that report statistically significant correlations. The first estimated the significant finding between the peak weightlifting data of the season and height in 37 female competitive discus throwers performances (Takanashi, Fujimori & Koikawa, 2020). The other research in this journal reported an association between the number of kilometres run per week and grit-passion in 153 ultramarathon runners (Cousins, Peterson, Christopher, Francis & Betz, 2020). The data from both studies were analysed with the tests of statistical significance ($p < .05$), using Pearson's correlation coefficient.

The use of the Bayes factor is essential for statistical inference in frequent estimation tests (for example, correlation analysis, a statistical test of comparison of student t means, among others). Also, when significant findings are available, the Bayesian model is considered a methodological alternative for statistical replication (Ly et al., 2018), which quantifies the degree of evidence of the alternative hypothesis with the null hypothesis (alternative hypothesis vs null hypothesis). Whose interpretation is based on the Jeffreys (1961) classification scheme: weak, moderate, strong, very strong, and extreme (Table 1).

Corresponding author. *Department of Computer Science, Landmark University Omu Aran, Kwara State, Nigeria.*
E-mail: ogundokun.roseline@lmu.edu.ng
Submitted for publication January 11, 2021.
Accepted for publication March 05, 2021.
Published April 01, 2021.
JOURNAL OF HUMAN SPORT & EXERCISE ISSN 1988-5202

***URI*:** **http://hdl.handle.net/10045/113872**

Table 1. Quantifiable interpretation values of the Bayes factor.

>100	Extreme	Alternative hypothesis
30+100	Very strong	Alternative hypothesis
10+30	Strong	Alternative hypothesis
3.1-10	Moderate	Alternative hypothesis
1.1-3	Weak	Alternative hypothesis
1	0	No evidence
0.3-0.99	Weak	Null hypothesis
0.29-0.1	Moderate	Null hypothesis
0.09-0.03	Strong	Null hypothesis
0.03-0.01	Very strong	Null hypothesis
<0.01	Extreme	Null hypothesis

Note: Own creation according to the Jeffreys classification scale (1961).

The present letter aimed to report two examples of Bayesian reanalysis, based on Pearson correlation values (r = 0.536; r = 0.28) and sample data (37 and 153) from both of the respective studies. The Bayes factor consists of two interpretations: FB10 (in favour of the alternative hypothesis) and FB01 (in favour of the null hypothesis) and the 95% credibility interval (Goss-Sampson, 2020).

The results obtained from the Bayes factor are: BF10 = 56.602 and BF01 = 0.018 and 95% CI [0.241- 0.719], in the investigation of female discus throwers and BF10 = 44.42 and BF01 = 0.22 and 95% CI [0.125 - 0.417], which supported both significant results with very strong evidence in favour of the alternate hypothesis (correlation). In turn, the maximum Bayes factor parameters were reported (maxFB10 = 62.46; maxFB10 = 80.88), whose Bayesian criteria estimate higher magnitude values that strengthen the stability of the results of Bayesian inference.

This alternative method is advised because several investigations in this journal use particular populations of interest with small sample sizes, whose results according to the significance tests can be controversial due to misinterpretations (Abt et al., 2020; Marsman & Wagenmakers, 2017). Therefore, it is advisable to estimate conclusive (strong) or higher evidence (FB10 > 10) for greater certainty of the confirmation of the alternative hypothesis.

The Bayes factor is very useful in other statistical analyses and reanalyses that include significance tests (Kelter, 2020; Ly et al., 2018; Ramos-Vera, 2020), therefore, the dissemination of the use and interpretation of This method in the sports field, additionally allows to reinforce systematic quantitative investigations that use such statistical tests for greater credibility in the conclusions of meta-analytical studies.

Therefore, it is expected that this letter will contribute to the dissemination and use of the Bayes factor that strengthens the reproducibility of sports research beyond the framework of statistical significance, whose inclusive methodological contribution is essential for future articles in this journal.

Keywords: Bayes factor; Statistical influence; Bayesian inference; Sport science.

Cite this article as:
Ramos-Vera, C.A., & Ogundokun, R.O. (2021). The use of the Bayes factor for statistical inference (Letter to the Editor). *Journal of Human Sport and Exercise*, *16*(2), 473-475.

REFERENCES

Abt, G., Boreham, C., Davison, G., Jackson, R., Nevill, A., Wallace, E., & Williams, M. (2020). Power, precision, and sample size estimation in sport and exercise science research. Journal of Sports Sciences, 38(17), 1933-1935. https://doi.org/10.1080/02640414.2020.1776002

Cousins, J.M., Peterson, M.J., Christopher, A.N., Francis, A.P., & Betz, H.H. (2020). Grit-passion and grit-perseverance in ultramarathon runners. Journal of Human Sport and Exercise, in press. https://doi.org/10.14198/jhse.2021.164.18

Goss-Sampson, M. A. (2020). Bayesian Inference in JASP: A Guide for Students. University of Amsterdam: JASP team. Retrieved from: http://dx.doi.org/10.17605/OSF.IO/CKNXM

Jeffreys, H. (1961). Theory of probability (3ra ed.). Oxford, England: Oxford University Press.

Kelter, R. (2020). Bayesian alternatives to null hypothesis significance testing in biomedical research: a non-technical introduction to Bayesian inference with JASP. BMC Medical Research Methodology, 20, 1-12. https://doi.org/10.1186/s12874-020-00980-6

Ly, A., Raj, A., Etz, A., Gronau, Q. F., & Wagenmakers, E.J. (2018). Bayesian reanalyses from summary statistics: a guide for academic consumers. Advances in Methods and Practices in Psychological Science, 1(3), 367-374. https://doi.org/10.1177/2515245918779348

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

Ramos-Vera C. A. (2020). Replicación bayesiana: cuán probable es la hipótesis nula e hipótesis alterna. Educación Médica, in press. https://doi.org/10.1016/j.edumed.2020.09.014

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? Journal of Research in Personality, 47(5), 609-612. https://doi.org/10.1016/j.jrp.2013.05.009

Takanashi, Y., Fujimori, N., & Koikawa, N. (2021). An investigation into the relationship between throw performance and maximum weight in weight training of female discus throwers. Journal of Human Sport and Exercise, 16(1), 226-234. https://doi.org/10.14198/jhse.2021.161.20