ACTAS
Dermo-Sifiliográficas

Full English text available at
www.actasdermo.org

CASE AND RESEARCH LETTHER

[Translated article] On Using the Bayes Factor in Dermatology Research

El método del factor Bayes para la investigación en Dermatología

To the Editor:

On December 6, this journal published an important article reporting a statistically significant association ($P < .05$) between a favorable attitude toward using sunscreen and previous sunburn among ultramarathon runners in the Gran Trail Aneto-Posets race.[1] The association was measured using the odds ratio (OR).

Replication of health science research based on significance testing is recommended in order to generate more reliable evidence in clinical investigation and in the subdiscipline of dermatology.

Such an approach is made possible via Bayesian inference, which enables us to reanalyse the significant finding reported by García-Malinis et al.,[1] who referred to the Bayes factor method as probability under one hypothesis with respect to the other (null hypothesis vs. alternative hypothesis).[2,3] In other words, the Bayes factor quantifies the degree of evidence on which the data support both the null hypothesis and the alternative hypothesis to enable contrast beyond the dichotomous interpretation of rejection or acceptance of the null hypothesis.[2,3] This method provides information that goes further than the dichotomous explanation of rejection or acceptance of the null hypothesis, whose interpretation is based on Jeffrey's[4] classification of values as weak, moderate, strong, very strong, and extreme (Table 1).

The aim of the present letter was to report a simple example of how Bayesian reanalysis can be applied to specify the strength of proof of statistical hypotheses. Therefore, we first considered the conversion of the OR value (1.57) to a correlation effect size (r) using the online calculator of Lenhard et al.[5] We reported an r of 0.124, which, in addition to the sample size (657), is considered essential for the replication of the Bayes factor.[2] This method allows 2 interpretations: BF_{10} (in favor of the alternative hypothesis of significance) and BF_{01} (in favor of the null hypothesis), with a 95% confidence interval.[6] The results obtained for the Bayes factor are $BF_{10} = 7.7$ and $BF_{01} = 0.13$, with a 95% CI of 0.048–0.198. These findings support the significant finding reported by García-Malinis et al.,[1] with a substantial degree of evidence (7-fold) in favor of the alternative statistical hypothesis (correlation).

Table 1 Values for the Quantifiable Interpretation 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
0.3–0.9	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

Source: Adapted from the evidence categories for the Bayes factor according to Jeffreys.[4]

The maximum Bayes factor is also reported ($maxBF_{10} = 35.19$) to determine the stability of the results, whose maximum size strengthens the estimation of Bayesian reassessment.

The Bayes factor is extremely useful in other statistical analyses and reanalyses based on significance testing.[7,8] Consequently, the use of this method and the interpretation of its findings should be disseminated in the clinical field of dermatology. Furthermore, this approach strengthens systematic quantitative research that applies these statistical tests to ensure greater credibility for the conclusions of meta-analyses.

The interpretation of significance levels in hypothesis testing has been called into question because of the misuse of P values resulting from erroneous interpretations, one of the most common being the inverse probability fallacy, which is the mistaken belief that the P value refers to the probability of a true null hypothesis (H_0), and the effect size fallacy, which links statistical significance to effect size.

DOI of original article:
https://doi.org/10.1016/j.ad.2021.01.014

https://doi.org/10.1016/j.ad.2021.01.027

Thus, small significant *P* values are interpreted as large effects, even though they provide no information on the size of the effect.[9] Similarly, the criteria for interpreting these values differ depending on the field of health science in question owing to a range of factors, such as statistical power or the clinical measures used.[10]

Therefore, we hope that the present letter goes some way to extending the use of the Bayes factor to consolidate the reproducibility of clinical research data beyond the notion of statistical significance. The inclusive contribution to methodology of this approach is essential for future articles in this journal.

References

1. García-Malinis AJ, Gracia-Cazaña T, Zazo M, Aguilera J, Rivas-Ruiz F, de Troya Martín M, et al. Hábitos y conocimientos sobre fotoprotección y factores de riesgo para quemadura solar en corredores de maratones de montaña. Actas Dermosifiliogr. 2020, http://dx.doi.org/10.1016/j.ad.2020.11.003.
2. Ly A, Raj A, Etz A, Gronau QF, Wagenmakers EJ. Bayesian reanalyses from summary statistics: a guide for academic consumers. Adv Meth Pract Psychol Sci. 2018;1:367–74, http://dx.doi.org/10.1177/2515245918779348.
3. Marsmamn M, Wagenmakers EJ. Bayesian benefits with JASP. Eur J Dev Psychol. 2017;14:545–55, http://dx.doi.org/10.1016/j.pmrj.2017.11.003.
4. Jeffreys H. Theory of probability. Oxford: Oxford University Press; 1961.
5. Lenhard W, Lenhard A. Calculation of effect sizes. Dettelbach; 2016. Available from: https://www.psychometrica.de/effect_size.html
6. Goss-Sampson MA. Bayesian inference in JASP: a guide for students. University of Amsterdam: JASP team; 2020, http://dx.doi.org/10.17605/OSF.IO/CKNXM.
7. Kelter R. Bayesian alternatives to null hypothesis significance testing in biomedical research: a non-technical introduction to Bayesian inference with JASP. BMC Med Res Methodol. 2020;20:1–12, http://dx.doi.org/10.1186/s12874-020-00980-6.
8. Ramos-Vera CA. Replicación bayesiana: cuán probable es la hipótesis nula e hipótesis alterna. Educ Méd. 2020;02Dic, http://dx.doi.org/10.1016/j.edumed.2020.09.014.
9. Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, et al. Statistical tests *P* values, confidence intervals, and power: a guide to misinterpretations. Eur J Epidemiol. 2016;31:337–50, http://dx.doi.org/10.1007/s10654-016-0149-3.
10. Brydges CR. Effect size guidelines sample size calculations, and statistical power in gerontology. Innov Aging. 2019;3:igz036, http://dx.doi.org/10.1093/geroni/igz036.

C. Ramos-Vera

Área de investigación, Facultad de Ciencias de la Salud, Universidad Cesar Vallejo, Lima, Peru

E-mail address: cristony_777@hotmail.com