様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　５ 年　 ６ 月　 ６ 日現在

機関番号：１５４０１
研究種目：研究活動スタート支援
研究期間：2021～2022
課題番号：２１Ｋ２０２１１

研究課題名（和文）Thorough overview of the use of the Bayes factor in applied psychological research

研究課題名（英文）Thorough overview of the use of the Bayes factor in applied psychological research

研究代表者
ＮＵＮＥＳ・ＴＥＮＤＥＩＲＯ　ＪＯＲＧＥ（Tendeiro, Jorge）

広島大学・学術・社会連携室・教授

研究者番号：９０７８６３１０

交付決定額（研究期間全体）：（直接経費）　　1,100,000円

研究成果の概要（和文）：本プロジェクトは、ベイズ統計的仮説検定を行う際に統計的応用を主とする社会学者がベイズファクター(BF)をどのように利用しているかについて理解を深めることを目的とした。BFは、古典的仮説検定のp-値に対する「ベイズ統計的代替品」とみなされることが多いが、応用を主とする研究者にとっては比較的新しいものであり、研究論文でどの程度正しく使用されているのかは不明であった。そこで、国際的な研究者チームと協力し、ベイズ統計による仮説検定の誤用を示唆する証拠を発見し、その改善のために取り組むべき問題をいくつか特定した。現在、私はBFの正しい理解を促進することを目的としたツールの開発に従事している。

研究成果の学術的意義や社会的意義
Performing responsible research is a mandate that should guide any scientist, because we must use research funds wisely. The research performed in this project will benefit all those interested in performing (null) hypothesis testing optimally, in particular under the Bayesian framework.

研究成果の概要（英文）：In this project I aimed at getting a better understanding into how social applied researchers use the Bayes factor in order to conduct hypothesis testing. The Bayes factor, often considered the Bayesian replacement of the p-value, is still relatively new to most practitioners. As such, it was unclear how well it has been used in the literature. In this project I worked together with an international team of researchers. I found evidence suggesting that there is severe misunderstanding about how Bayesian hypothesis testing should be conducted. I have identified several issues that need to be addressed in order to improve the current state of affairs. I am currently working on tools aiming at easing the understanding of the Bayes factor.

研究分野：Statistics

キーワード：Bayes factor

様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９（共通）

１．研究開始当初の背景

The most popular statistical inferential framework in use nowadays is frequentist, or classical, statistics. Specifically for hypotheses testing, null hypothesis significance testing (NHST) and the *p*-value are heavily used. Even though these tools have been around for about 100 years, there is ample evidence in the literature showing that practitioners do not understand NHST nor the *p*-value very well. This is extremely problematic because much of the decision-making in science relies on these tools. A poor understanding of hypothesis testing may be detrimental towards establishing solid scientific foundations at best, and may easily mislead scientists into making egregious errors.

This state of affairs has motivated researchers, and statisticians in particular, to look into possible solutions and alternatives. In this regard, Bayesian statistics has emerged as one solid viable alternative to frequentist statistics. Concerning hypothesis testing there is the Bayes factor, which can be thought of as the Bayesian counterpart to the *p*-value. The Bayes factor has been gaining popularity since the mid 1990s with the advent of MCMC algorithms. The big problem is that most researchers did not receive a good (or any) preparation in Bayesian statistics. As such, using a new inferential tool such as the Bayes factor is not an easy task. Knowing how often researchers make mistakes using well-established frequentist statistical tools, I could not help but wonder whether researchers have also been making mistakes while using the Bayes factor in their research.

２．研究の目的

We had two main purposes with this research project: To identify existing problems related to the use of the Bayes factor in practice, and to offer solutions to improve matters. Concerning the first purpose, I studied how the Bayes factor has been (mis)understood and interpreted. Together with my research team, we also reasoned on why the identified errors do come about. Based on these findings, we addressed our second purpose by developing new materials aiming at improving the education and understanding of Bayesian hypothesis testing. In particular, an interactive Shiny app is currently entering deployment stage.

３．研究の方法

(1) A large-scale literature study was conducted. I used various citation databases to find all empirical psychology papers published since 2010 that report at least one Bayes factor in their results. After intensive selection and screening, a final set of 167 papers was selected. The criteria for rating the adequacy of each use of the Bayes factor in each paper was determined by my research team. This is an international team that I coordinated and that included the following researchers: Myself, Prof. Henk Kiers and Dr. Rink Hoekstra (University of Groningen), Dr. Richard Morey (Cardiff University), and Tsz Keung Wong (Tilburg University). Together, we assessed potential misconceptions in uses of the Bayes factor. We curated a list of criteria according to which we judged each of the papers in our sample; we refer to these criteria as QRIPs (questionable reporting and interpreting practices). Afterwards, we read all papers in our sample and marked all QRIP occurrences. This provided us with a broad picture of the type and severity of problems that exist. At this point, my team met several times to discuss the results. More than just looking at the total counts for each QRIP, we studied many text passages and tried to reason why such errors occurred. Based on this, we speculated on various possible causes for the identified problems. Furthermore, we elaborated on a series of suggested solutions to address the problems that we found. We created a scheme relating each cause to each proposed solution.

(2) Based on the previous project, we followed up by working on new education materials. A Shiny app has been created for this purpose. This app is essentially a modern HTML page which includes means to compute the Bayes factor in one very simple situation,

raise awareness about the various errors that are important to avoid, and finally go through a practice test in order to check whether the contents were well learned. We are in the process of finalizing the app and the supporting paper, which will be submitted to an international peer-reviewed journal.

４．研究成果

(1) The incidence of the various QRIP criteria in our papers sample is shown in Figure 1. It can be seen that the percentages are in some cases rather high. This is indeed worrisome and unfortunately confirms our worse suspicions that led us to work on this project to start with.

Criterion	Count (Percentage)
#1	22 (13.2)
#3a	18 (10.8)
#3b	50 (29.9)
#3c	10 (6.0)
#4	104 (62.3)
#5	59 (35.3)
#6	34 (20.4)
#7	7 (4.2)
#9	6 (3.6)
#10	9 (5.4)
A	59 (35.3)
B	27 (16.2)
C	30 (18.0)

Figure 1: Count (percentage) of papers displaying the corresponding criterion.

As a result of dissecting the identified problems, we offered some suggestions for solutions. Figure 2 summarizes both the purported causes and proposed solutions.

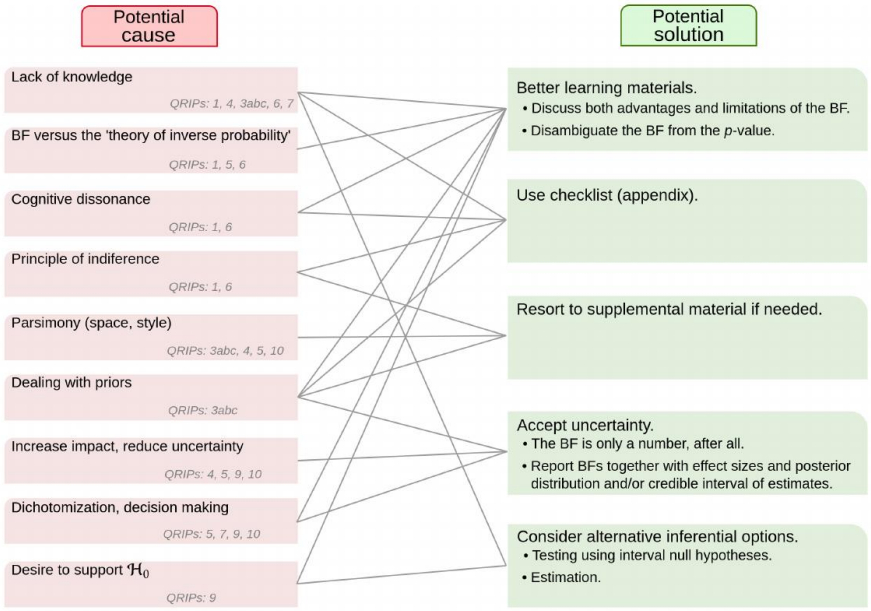

Figure 2: Summary of the potential causes for the problems identified in the literature study (left) and suggestions for potential solutions (right).

The results from this project were presented at an international conference (CSP 2023, San Francisco, 03 February 2023) and we submitted a paper for publication (preprint here: https://psyarxiv.com/du3fc/; current stage: resubmitted after revision).

(2) A Shiny app is currently in an advanced stage of development. The weblink is currently https://statsedge.org/StatsEdgeShinyApps/LearnBF/, but this is still subject to change. Most elements of the app are in place, I am finalizing the more educational parts of the tool. This project suffered a small delay due to problems related to securing my one server and website to make sure this app stays functional for years to come. A paper supporting this app is also under development. I hope to submit the paper to an international journal in the coming few weeks.

(3) I further proceeded with my research with the Bayes Factor itself. Together with colleagues from Europe, I co-authored two papers revolving around the Bayes factor. In Linde, Tendeiro, Wagenmakers, and van Ravenzwaaij (2023; in press), we wrote a rebuttal to Campbell and Gustafson on the merits of the Bayes factor in the realm of equivalence testing. In Linde, Tendeiro, and van Ravenzwaaij (2023; under revision), we developed new Bayes factor for use with Cox regression in a two-group design.

＜引用文献＞

① Tendeiro, J. N., Kiers, H. A. L., Hoekstra, R., Wong, T. K., and Morey, R. D. (2023). *Diagnosing the misuse of the Bayes factor in applied research*. Submitted for publication. Preprint: https://psyarxiv.com/du3fc/

② Tendeiro, J. N. (2023). *Misuse of Bayesian inference in applied research: Current status and constructive feedback*. Presentation at the CSP 2023 conference, San Francisco, February 2023.
Slides: https://www.jorgetendeiro.com/talk/2023_csp/

③ Linde, M., Tendeiro, J. N., Wagenmakers, E.-J., and van Ravenzwaaij (2023). Practical implications of equating equivalence tests: Reply to Campbell and Gustafson (2022). In press, *Psychological Methods*.
Preprint: https://psyarxiv.com/xk43y

④ Linde, M., Tendeiro, J. N., and van Ravenzwaaij, D. (2023). *Bayes factors for two-group comparisons in Cox regression*. Submitted for publication. Preprint: https://www.medrxiv.org/content/10.1101/2022.11.02.22281762v1

⑤ Tendeiro, J. N. (2023). Learn about the Bayes factor! Software app (https://statsedge.org/StatsEdgeShinyApps/LearnBF/). In late development stage.

５．主な発表論文等

〔雑誌論文〕　計3件（うち査読付論文　3件／うち国際共著　3件／うちオープンアクセス　2件）

１．著者名 Tendeiro, J. N., Kiers, H. A. L., Hoekstra, R., Wong, T. K., Morey, R. D.	４．巻 -
２．論文標題 Diagnosing the misuse of the Bayes factor in Applied Research	５．発行年 2023年
３．雑誌名 Advances in Methods and Practices in Psychological Science (submitted)	６．最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.31234/osf.io/du3fc	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

１．著者名 Linde, M., Tendeiro, J. N., Wagenmakers, E.-J., and van Ravenzwaaij	４．巻 -
２．論文標題 Practical implications of equating equivalence tests: Reply to Campbell and Gustafson (2022)	５．発行年 2023年
３．雑誌名 Psychological Methods (in press)	６．最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.31234/osf.io/xk43y	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する

１．著者名 Linde, M., Tendeiro, J. N., and van Ravenzwaaij, D.	４．巻 -
２．論文標題 Bayes Factors for Two-group Comparisons in Cox Regression	５．発行年 2023年
３．雑誌名 Journal of Applied Statistics (submitted)	６．最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.1101/2022.11.02.22281762	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する

〔学会発表〕　計1件（うち招待講演　0件／うち国際学会　1件）

１．発表者名 Tendeiro, J. N.
２．発表標題 Misuse of Bayesian inference in applied research: Current status and constructive feedback.
３．学会等名 Conference on Statistical Practice 2023（国際学会）
４．発表年 2023年

〔図書〕　計0件

〔産業財産権〕

〔その他〕

Tendeiro, J. N. (2023). Learn about the Bayes factor! Software app (https://statsedge.org/StatsEdgeShinyApps/LearnBF/). In late development stage.

６．研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考

７．科研費を使用して開催した国際研究集会

〔国際研究集会〕　計0件

８．本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関