Statistics and Probability Letters 92 (2014) 121–124

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Simple relation between Bayesian order-restricted and point-null hypothesis tests

Richard D. Morey[a,*], Eric-Jan Wagenmakers[b]

[a] *University of Groningen, Department of Psychometrics and Statistics, Grote Kruisstraat 2/1, 9712 TS Groningen, The Netherlands*

[b] *University of Amsterdam, Department of Psychological Methods, Weesperplein 4, 1018 XA Amsterdam, The Netherlands*

ARTICLE INFO

Article history:
Received 4 October 2013
Received in revised form 14 May 2014
Accepted 14 May 2014
Available online 24 May 2014

Keywords:
Bayes factor
Order-restriction
Sign-restriction
One-sided test
t test
Hypothesis testing

ABSTRACT

One of the main challenges facing potential users of Bayes factors as an inferential technique is the difficulty of computing them. We highlight a useful relationship that allows certain order-restricted and sign-restricted Bayes factors, such as one-sided Bayes factor tests, to be computed with ease.

1. Introduction

Consider an encompassing model $\mathcal{M}_e$ with nuisance parameters θ and parameter of interest $\boldsymbol{\delta}$ of length K with marginal prior distribution $p(\boldsymbol{\delta})$. Two restrictions of $\mathcal{M}_e$ can be considered: the null hypothesis $\mathcal{M}_0$ states that $\boldsymbol{\delta} = \mathbf{0}$, and $\mathcal{M}_r$ is an order-restricted hypothesis that the $\boldsymbol{\delta}$ parameters have a specific ordering. If R is the set of all vectors $\boldsymbol{\delta}$ that meet the specified restriction, then $\mathcal{M}_r$ states that $\boldsymbol{\delta} \in R$. If $K = 1$ and δ is a scalar parameter, then $\mathcal{M}_r$ is a sign hypothesis that δ is either positive or negative. We use the general term "order-restriction" to refer both the $K = 1$ case and the $K > 1$ case. Suppose that $p(\boldsymbol{\delta})$ is such that all orderings are equally-likely a priori, as will occur if the prior distributions on the $K\delta$ parameters are identical and mutually conditionally independent. The Bayes factor $B_{r0} = p(\mathbf{y} \mid \mathcal{M}_r)/p(\mathbf{y} \mid \mathcal{M}_0)$ quantifies the evidence that the data $\mathbf{y}$ provide for $\mathcal{M}_r$ versus $\mathcal{M}_0$ (Jeffreys, 1961; Kass and Raftery, 1995). This Bayes factor is of practical interest because researchers often have strong prior expectation about the direction of an effect or the ordering of means under the assumption that the null hypothesis is false. Unfortunately, B_{r0} is often not available in closed form because almost all tests have been developed for the two-sided scenario B_{e0}. In addition, the computation of B_{r0} is made difficult by the fact that the prior and posterior distributions under model $\mathcal{M}_r$ are bounded at 0 and therefore may not be members of familiar families of distributions. Hence, the calculation of $p(\mathbf{y} \mid \mathcal{M}_r)$ can be a non-trivial task that appears to require general procedures such as reversible jump Markov chain Monte Carlo (Green, 1995) that applied researchers may find challenging to implement.

However, Pericchi et al. (2008) proposed a general and simple solution to the computation of the one-sided Bayes factor B_{r0}, avoiding the need for integration over the parameter space when the two-sided Bayes factor B_{e0} is already in hand.

* Corresponding author.
E-mail addresses: r.d.morey@rug.nl (R.D. Morey), ej.wagenmakers@gmail.com (E.-J. Wagenmakers).

http://dx.doi.org/10.1016/j.spl.2014.05.010

Theorem 1. *Let L be*

$$L = \begin{cases} 2 & K = 1 \\ K! & K > 1. \end{cases}$$

Then

$$B_{r0} = Lp(\boldsymbol{\delta} \in R \mid \mathbf{y}, \mathcal{M}_e)B_{e0}.$$

Proof of Theorem 1. There are L specific order-restricted hypotheses on $\boldsymbol{\delta}$. Under a proper prior $p(\boldsymbol{\delta})$ in which all orderings on K means are equally likely, each ordering has an a priori probability of $1/L$. Because the total prior probability of all orderings is 1, the prior odds of $\mathcal{M}_r$ against the encompassing model $\mathcal{M}_e$ are thus $(1/L)/1 = 1/L$. The corresponding posterior odds are $p(\boldsymbol{\delta} \in R \mid \mathbf{y}, \mathcal{M}_e)$ (Klugkist et al., 2005). Because the Bayes factor is the ratio of the posterior odds to the prior odds,

$$\begin{aligned} B_{re} &= \frac{p(\boldsymbol{\delta} \in R \mid \mathbf{y}, \mathcal{M}_e)}{1/L} \\ &= Lp(\boldsymbol{\delta} \in R \mid \mathbf{y}, \mathcal{M}_e). \end{aligned}$$

Bayes factors are ratios of the corresponding marginal likelihoods, and thus

$$\begin{aligned} B_{r0} &= \frac{p(\mathbf{y} \mid \mathcal{M}_r)}{p(\mathbf{y} \mid \mathcal{M}_0)} \\ &= \frac{p(\mathbf{y} \mid \mathcal{M}_r)}{p(\mathbf{y} \mid \mathcal{M}_e)} \times \frac{p(\mathbf{y} \mid \mathcal{M}_e)}{p(\mathbf{y} \mid \mathcal{M}_0)} \\ &= B_{re}B_{e0} \end{aligned}$$

and the result follows. ∎

The term B_{e0} is from the familiar two-sided test, and the term B_{re} equals the ratio between the marginal posterior and the marginal prior mass consistent with the restriction (Klugkist et al., 2005). If B_{re} is not available analytically, it can be easily obtained to any desired degree of approximation using numerical methods such as Markov chain Monte Carlo (e.g., Morey et al., 2011).

One application of Theorem 1 is the one-sided tests that arise when $K = 1$. In such one-sided tests, Theorem 1 implies that the one-sided test B_{r0} equals the two-sided test B_{e0} only when the posterior $p(\delta \mid \mathbf{y}, \mathcal{M}_e)$ is symmetric around 0. In addition, the use of a one-sided test can increase the evidence against $\mathcal{M}_0$ by a factor of 2 at most, which happens when almost the entire posterior distribution is consistent with the order-restriction. When the data are inconsistent with the sign-restriction $\delta > 0$ this means that $p(\delta > 0 \mid \mathbf{y}, \mathcal{M}_e)$ is lower than 0.5, and the use of a one-sided test increases the evidence for $\mathcal{M}_0$. In fact, when the data are wildly inconsistent with the order-restriction it may happen that B_{r0} is extremely low (indicating that $\mathcal{M}_0$ should be retained) and that, at the same time, B_{e0} is extremely high (indicating that $\mathcal{M}_0$ should be rejected). This underscores the relative nature of the Bayes factor as a measure of evidence.

The relevance of order-restricted tests is particularly acute for the replication research and for clinical trials, where compelling evidence for $\mathcal{M}_0$ may be obtained when the effect goes in the direction opposite to what was expected. The effects become more pronounced when more parameters are subject to test. Suppose $\mathcal{M}_e$ is a one-way model with $K = 4$ condition means for which the analyst has a strong a priori commitment to the orderings of the K means, if the null hypothesis were false. For instance, if the conditions arose from a manipulation of a single independent variable, such as dosage or difficulty, then the analyst may wish to test the specific ordering that implies a monotone relationship. If the posterior probability $p(\boldsymbol{\delta} \in R \mid \mathbf{y}, \mathcal{M}_e)$ in favor of the restriction is maximal, then increase in the evidence from B_{e0} to B_{r0} from properly restricting the test will be $4! = 24$, a substantial change in the evidence.

In special cases where the posterior probability can be easily approximated by the p value, such as in one- and two-sample tests, the correction factor can be easily computed using the output of a standard classical analysis. In the one-sample case, $L = 2$ and the correction needed to obtain the sign-restricted test equals $2 \times p(\delta > 0 \mid \mathbf{y}, \mathcal{M}_e)$ if the desired sign restriction is that $\delta > 0$. Exploiting the fact that for the test of location parameters the classical one-sided p value approximates $p(\delta < 0 \mid \mathbf{y}, \mathcal{M}_e) = 1 - p(\delta > 0 \mid \mathbf{y}, \mathcal{M}_e)$ (Casella and Berger, 1987; Lindley, 1965; Pratt, 1965), we obtain:

$$B_{r0} \approx \begin{cases} (2 - p) \times B_{e0} & \text{if } \hat{\delta} > 0, \\ p \times B_{e0} & \text{if } \hat{\delta} \leq 0, \end{cases} \tag{1}$$

where p is two-sided, and $\hat{\delta} > 0$ indicates that the observed effect is consistent with the sign-restriction. When $\hat{\delta} > 0, B_{r0} > B_{e0}$, with a maximum of $B_{r0} = 2 \times B_{e0}$ when $p \to 0$. When $\hat{\delta} < 0$ (i.e., the observed effect goes in the opposite direction), $B_{r0} < B_{e0}$. In sum, (1) shows how the sign-restricted Bayes factor can be approximated by the product of two familiar terms, one involving the two-sided p value, and one involving the two-sided Bayes factor.

The p value approximation is particularly useful when the posterior probability $p(\delta < 0 \mid \mathbf{y}, \mathcal{M}_e)$ is not immediately available. For instance, not all methods of estimating Bayes factors involve MCMC chains that can be used to estimate the required posterior probability, and even when they do the software may not report the chains. The widely-used JZS Bayes factor web calculator (Rouder et al., 2009; http://pcl.missouri.edu/bayesfactor), for instance, does not return posterior probabilities.