entropy

MDPI

Article

From *p*-Values to Posterior Probabilities of Null Hypotheses

Daiver Vélez Ramos [1,*], Luis R. Pericchi Guerra [2] and María Eglée Pérez Hernández [2]

[1] Faculty of Business Administration, Statistical Institute and Computerized Information Systems, Río Piedras Campus, University of Puerto Rico, 15 AVE Universidad STE 1501, San Juan, PR 00925-2535, USA
[2] Faculty of Natural Sciences, Department of Mathematics, Río Piedras Campus, University of Puerto Rico, 17 AVE Universidad STE 1701, San Juan, PR 00925-2537, USA; luis.pericchi@upr.edu (L.R.P.G.); maria.perez34@upr.edu (M.E.P.H.)
* Correspondence: daiver.velez@upr.edu

Abstract: Minimum Bayes factors are commonly used to transform two-sided *p*-values to lower bounds on the posterior probability of the null hypothesis, in particular the bound $-e \cdot p \cdot \log(p)$. This bound is easy to compute and explain; however, it does not behave as a Bayes factor. For example, it does not change with the sample size. This is a very serious defect, particularly for moderate to large sample sizes, which is precisely the situation in which *p*-values are the most problematic. In this article, we propose adjusting this minimum Bayes factor with the information to approximate an exact Bayes factor, not only when p is a *p*-value but also when p is a pseudo-*p*-value. Additionally, we develop a version of the adjustment for linear models using the recent refinement of the Prior-Based BIC.

Keywords: *p*-value calibration; Bayes factor; linear model; pseudo-*p*-value; adaptive levels

check for updates

Citation: Vélez Ramos, D.; Pericchi Guerra, L.R.; Pérez Hernández, M.E. From *p*-Values to Posterior Probabilities of Null Hypothesis. *Entropy* **2023**, *25*, 618. https://doi.org/10.3390/e25040618

Academic Editors: Carlos Alberto De Bragança Pereira, Adriano Polpo, Agatha Rodrigues and Débora Corrêa

Received: 15 February 2023
Revised: 28 March 2023
Accepted: 30 March 2023
Published: 6 April 2023

1. Introduction

By now, it is well known by practitioners that *p*-values are not posterior probabilities of a null hypothesis, which is what science would need to declare a scientific finding. So *p*-values, and particularly the threshold of 0.05, need to be recalibrated. Two widespread practical attempts are (i) the so-called Robust Lower Bound on Bayes factors $BF \geq -e \cdot p \cdot \log(p)$ [1] and (ii) the replacement of the ubiquitous $\alpha = 0.05$ by $\alpha^* = 0.005$ [2]. These suggestions, which are an improvement of usual practice, fall short of being a real solution, mainly because the dependence of the evidence on the sample size is not considered. Still, the Robust Lower Bound is useful since it is valid from small sample sizes and onward and only depends on the *p*-value. It is known that the evidence of a *p*-value against a point null hypothesis depends on the sample size. In [3], they consider *p*-values in linear models and propose new monotonic minimum Bayes factors that depend on the sample size and converge to $-e \cdot p \cdot \log(p)$ as the sample size approaches infinity, which implies it is not consistent, as Bayes factors are. It turns out that the maximum evidence for an exact two-tailed *p*-value increases with decreasing sample size. There are several proposals in the literature, and most do not depend on the sample size, while those that do continue to be Robust Lower Bounds; however, neither behaves like a real Bayes factor. In this article, we propose to adjust the Robust Lower Bound $-e \cdot p \cdot \log(p)$ so that it behaves in a similar or approximate way to actual Bayes factors for any sample size. A further complication arises, however, when the null hypotheses are not simple, that is, when they depend on unknown nuisance parameters. In this situation, what is usually called *p*-values are only pseudo-*p*-values [4] (p. 397). So, we first need to extend the validity of the Robust Lower Bound to pseudo-*p*-values. The effect of adjusting this minimum Bayes factor with the sample size is shown in a simulation in Section 5.1.

The outline of the article is as follows: In Section 2 we define pseudo-*p*-values using the *p*-value definition of [4] (p. 397) and extend for them the validity of the Robust Lower

Bound. In Section 3, we present the adaptive significance levels that will be used for incorporating the sample size in the lower bound: the general adaptive significance level presented in [5] and the refined version for linear models developed in [6]; in both cases, we use versions calibrated using the Prior-Based BIC (PBIC) [7]. In Section 4, we derive adaptive approximate Bayes factors and apply them to pseudo-p-values in Section 5. We close in Section 6 with some final comments.

2. Valid p-Values and Robust Lower Bound

Under the null hypotheses, p-values are well known to have Uniform(0, 1); in [4] (p. 397), a more general definition is given.

Definition 1. *A p-value $p(\boldsymbol{X})$ is a statistic satisfying $0 \leq p(\boldsymbol{x}) \leq 1$ for every sample point $\boldsymbol{x}$. Small values of $p(\boldsymbol{X})$ give evidence that $H_1 : \theta \in \Theta_0^c$ is true, where Θ_0 is some subset of the parameter space and Θ_0^c is its complement. A p-value is valid if, for every $\theta \in \Theta_0$ and every $0 \leq \alpha \leq 1$,*

$$P_\theta(p(\boldsymbol{X}) \leq \alpha) \leq \alpha.$$

Based on this definition, we can say that there are valid p-values that are Uniformly Distributed in (0, 1), that is,

$$P_\theta(p(\mathbf{X}) \leq \alpha) = \alpha \text{ for every } \theta \in \Theta_0 \text{ and every } 0 \leq \alpha \leq 1, \tag{1}$$

and others that are not, that is, when there is at least one α, such that

$$P_\theta(p(\mathbf{X}) \leq \alpha) < \alpha \text{ for every } \theta \in \Theta_0. \tag{2}$$

Remark 1. *We consider any valid p-value complying with (2) a pseudo-p-value.*

The "Robust Lower Bound" (*RLB*), as we call it here and proposed by [1], is

$$B_L(p) = \begin{cases} -e \cdot p \cdot \log(p) & p < e^{-1} \\ 1 & \text{otherwise} \end{cases}$$

The authors consider that under the null hypothesis, the distribution of the p-value, $p(\mathbf{X})$, is Uniform(0, 1). Alternatives are typically developed by considering alternative models for $\mathbf{X}$, but the results then end up being quite problem-specific. An attractive approach is instead to directly consider alternative distributions for p itself. In effect, they consider that, under H_1, the density of p is $f(p|\xi)$, where ξ is an unknown parameter. So, consider testing

$$H_0 : p \sim \text{Uniform}(0, 1) \text{ versus } H_1 : p \sim f(p|\xi)$$

If the test statistic (T) has been appropriately chosen so that large values of $T(\mathbf{X})$ would be evidence in favor of H_1 , then the density of p under H_1 should be decreasing in p. A class of decreasing densities for p that is very easy to work with is the class of Beta$(\xi, 1)$ densities, for $0 < \xi \leq 1$, given by $f(p|\xi) = \xi p^{\xi-1}$. The uniform distribution (i.e., H_0) arises from the choice $\xi = 1$ [1]. The expression $B_L(p) = \inf_{all\ \pi} B_\pi(p)$, where $B_\pi(p)$ is the Bayes factor of H_0 to H_1 for a given prior density $\pi(\xi)$ on this alternative.

Note that this calibration has already been proposed in [8]. Another class of decreasing densities is Beta$(1, \xi)$ with $\xi > 1$. This leads to the "$-e \cdot q \cdot \log(q)$" calibration, where $q = 1 - p$ see [9].

In contrast with Remark 1, if we consider $p(\mathbf{X})$ a pseudo-p-value under H_0, that is,

$$p \sim \text{Beta}(\xi_0, 1) \quad \text{with} \quad \xi_0 > 1, \quad \text{fixed but arbitrary,}$$

under the test

$$H_0 : p \sim \text{Beta}(\xi_0, 1) \quad \text{vs.} \quad H_1 : p \sim f(p|\xi)$$

with $f(p|\xi) \sim \text{Beta}(\xi, 1)$ for $0 < \xi \leq \xi_0$, then a generalized Robust Lower Bound RLB_{ξ_0} can be defined as

$$B_L(p, \xi_0) = \begin{cases} -e \cdot \xi_0 \cdot p^{\xi_0} \log(p) & p < e^{-\frac{1}{\xi_0}} \\ 1 & \text{otherwise} \end{cases} \tag{3}$$

where ξ_0 has to be estimated or calculated theoretically (see [10] for a proposal when extending for multiple testing). Any value $\xi_0 \neq 1$ corresponds to a pseudo-p-value.

On the other hand, since $f(p|\xi) = \xi p^{\xi-1}$ has its maximum in $\xi = -\frac{1}{\log(p)} < 1$ with $p < e^{-1}$, then $f(p|\xi)$ is decreasing for $\xi > -\frac{1}{\log(p)}$, thus for any Bayes factor B_{01}

$$B_{01} \geq B_L(p) > B_L(p, \xi_0) \quad \text{with} \quad \xi_0 > 1 \tag{4}$$

see Figure 1.

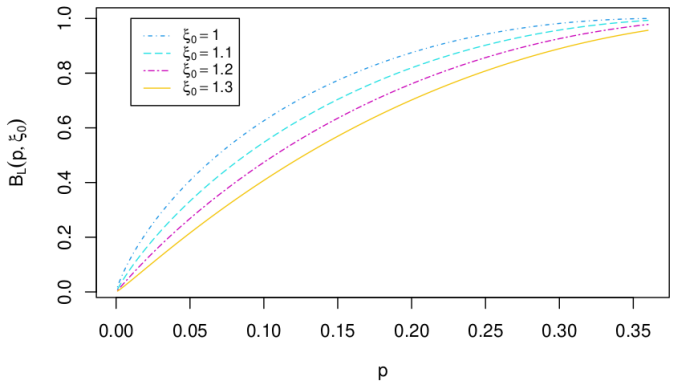

Figure 1. Extended Robust Lower Bound RLB_{ξ_0} as a function of p for different values of ξ_0.

In the following, we calibrate RLB_{ξ_0} such that $RLB_{\xi_0} \approx B_{01}$.

Lemma 1. $B_L(p_{val}, \xi) = -e \cdot \xi \cdot p_{val}^{\xi} \cdot \log(p_{val}) \geq e \cdot \xi \cdot p_{val}^{\xi} > p_{val}^{\xi}$, *for,* $0 < p_{val} < e^{-1}$ *and* $\xi \geq 1$. *Note that* $B_L(p_{val}, 1) = B_L(p_{val})$

Proof. Appendix A. □

Theorem 1. *The* RLB_{ξ} *is a valid p-value for* $\xi \geq 1$, *that is,*

$$P(B_L(p, \xi) \leq \alpha | p \sim f(p|\xi)) \leq \alpha, \quad \textit{for each} \quad 0 \leq \alpha \leq 1.$$

Proof. Appendix A. □

3. Adaptive α with PBIC Strategy

The Bayesian literature has been criticizing for several decades the implementation of hypothesis testing with fixed significance levels and, in particular, the use of the scale p-value < 0.05. An adaptive α allows us to adjust the statistical significance with the amount of information; see [5,11,12]. The adaptive values we work with in this section were calculated so that they allow to arrive to results equivalent to those obtained with a Bayes factor. In [5], the authors present an adaptive α based on BIC as

$$\alpha_n(q) = \frac{[\chi^2_\alpha(q) + q \log(n)]^{\frac{q}{2}-1}}{2^{\frac{q}{2}-1} n^{\frac{q}{2}} \Gamma\left(\frac{q}{2}\right)} \times C_\alpha, \tag{5}$$

where C_α is a calibration constant, and strategies for calculating it are presented in [5]. It yields a consistent procedure; it alleviates the problem of the divergence between prac-

tical and statistical significance; and it makes it possible to perform Bayesian testing by computing intervals with the calibrated α-levels.

An adaptive α is also presented in [6], but this time it is a version refined to nested linear models with calibration based on the Bayesian information criterion based on Prior PBIC [7],

$$\alpha_{(b,n)}(q) = \frac{[g_{n,\alpha}(q) + \log(b) + C]^{\frac{q}{2}-1}}{b^{\frac{n-j}{2(n-1)}} \cdot \left(\frac{2(n-1)}{n-j}\right)^{q/2-1} \Gamma\left(\frac{q}{2}\right)} \times \exp\left\{-\frac{n-j}{2(n-1)}(g_{n,\alpha}(q) + C)\right\}. \quad (6)$$

Here, $b = \frac{|\mathbf{X}_j^t \mathbf{X}_j|}{|\mathbf{X}_i^t \mathbf{X}_i|}$ and $\mathbf{X}_i, \mathbf{X}_j$ are design matrices and

$$C = 2\sum_{m_i=1}^{q_i} \log \frac{(1 - e^{-v_{m_i}})}{\sqrt{2}v_{m_i}} - 2\sum_{m_j=1}^{q_j} \log \frac{(1 - e^{-v_{m_j}})}{\sqrt{2}v_{m_j}},$$

$v_{m_l} = \frac{\hat{\xi}_{m_l}}{[d_{m_l}(1+n^e_{m_l})]}$ with $l = i, j$ corresponding to each model. Here, $n^e_{m_l}$, with $l = i, j$, refers to The Effective Sample Size (called TESS) corresponding to that parameter; see [7].

The adaptive α in (5) can also be presented using the PBIC strategy (this strategy was not considered in [5]), and the following expression is obtained

$$\alpha_n(q) = \frac{[\chi^2_\alpha(q) + q\log(n) + C]^{\frac{q}{2}-1}}{n^{\frac{q}{2}} 2^{\frac{q}{2}-1} \Gamma\left(\frac{q}{2}\right)} \times \exp\left\{-\frac{1}{2}\left(\chi^2_\alpha(q) + C\right)\right\}. \quad (7)$$

Note that this adaptive α is still of BIC structure, since the expression $\chi^2_\alpha(q) + q\log(n)$ remains.

Example: Binomial Models

Consider comparing two binomial models $S_1 \sim$ binomial(n_1, p_1) and $S_2 \sim$ binomial(n_2, p_2) via the test

$$H_0 : p_1 = p_2 \text{ vs. } H_1 : p_1 \neq p_2.$$

Defining $n = n_1 + n_2$ and $\hat{p}$, the MLE from $p_1 - p_2$, then (7) gives

$$\alpha_n = \left[\frac{2}{n\pi(\chi^2_\alpha(1) + \log(n) + C)}\right]^{1/2} \times \exp\left\{-\frac{1}{2}\left(\chi^2_\alpha(1) + C\right)\right\}, \quad (8)$$

here, $\chi^2_\alpha(1)$ is the quantile α from chi-square with $df = 1$, $C = -2\log\frac{(1-e^{-v})}{\sqrt{2}v}$, $v = \hat{p}^2/[d(1+n^e)]$, $d = \left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$, $n^e = \max\left\{\frac{n_1^2}{\sigma_1^2}, \frac{n_2^2}{\sigma_2^2}\right\} d$.

Table 1 shows the behavior of this adaptive α_n for $\alpha = 0.05$ and different values of n_1 and n_2.

Table 1. Adaptive α via PBIC in (8) for testing equality of two proportions for different sample sizes when $\alpha = 0.05$.

		Adaptive α via PBIC (α_n)
n_1	n_2	$n = n_1 + n_2$
10	10	0.0068
25	25	0.0040
50	50	0.0027
100	50	0.0021
50	100	0.0021
100	100	0.0018

4. Adjusting RLB_{ξ} Using Adaptive α

In this section, we combine (3) with the formulas for adaptive α in (6) and (7) for adjusting RLB_{ξ} and obtaining an approximation to an objective Bayes factor. Indeed, we adjust the RLB_{ξ} through the expression $B(\alpha) = B_L(\alpha, \xi_0) \cdot g(\cdot)$, where g is determined in such a way that when $B(\alpha)$ is evaluated in (6) or (7), it converges to a constant (this allows us to obtain equivalent results from the Frequentist and Bayesian point of view, that is, the decision does not change).

Substituting p in (3) by the adaptive α value in (7) results in the following expression.

$$B(\alpha, q, n, \xi_0) = -\alpha^{\xi_0} \log(\alpha) \Gamma(q/2)^{\xi_0} n^{\frac{\xi_0 q}{2}} \left[\frac{2}{\chi^2_{\alpha}(q) + q \cdot \log(n) + C} \right]^{\frac{\xi_0 q}{2} - (\xi_0 - 1)}. \tag{9}$$

For a Uniform$(0,1)$ p-value with $\xi_0 = 1$, this expression simplifies to

$$B(\alpha, q, n) = -\alpha \log(\alpha) \Gamma(q/2) n^{\frac{q}{2}} \left[\frac{2}{\chi^2_{\alpha}(q) + q \cdot \log(n) + C} \right]^{\frac{q}{2}}. \tag{10}$$

The refined version of this calibration for linear models is obtained when (3) is evaluated in (6)

$$B(\alpha, q, n, b) = -\alpha \log(\alpha) \Gamma(q/2) b^{\frac{n-j}{2(n-1)}} \left[\frac{2(n-1)}{(g_{n,\alpha}(q) + \log(b) + C)(n-j)} \right]^{\frac{q}{2}} \tag{11}$$

in this case, we only consider $\xi_0 = 1$.

Balanced One-Way Anova

Suppose we have k groups with r observations each, for a total sample size of kr, and let $H_0 : \mu_1 = \cdots = \mu_k = \mu$ vs. H_1 : At least oneμ_i different. Then, the design matrices for both models are

$$\mathbf{X}_1 = \begin{pmatrix} 1 \\ 1 \\ . \\ \vdots \\ 1 \end{pmatrix}, \mathbf{X}_k = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 1 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}, b = \frac{|\mathbf{X}_k^t \mathbf{X}_k|}{|\mathbf{X}_1^t \mathbf{X}_1|} = k^{-1} r^{k-1},$$

and the adaptive α for the linear model in accordance with what was presented in [6] is

$$\alpha(k, r) = \frac{[g_{r,\alpha}(k-1) - \log(k) + (k-1)\log(r) + C]^{\frac{k-3}{2}}}{\left(k^{-1} r^{k-1}\right)^{\frac{r-1}{2(r-1/k)}} \left(\frac{2(r-1/k)}{r-1}\right)^{\frac{k-3}{2}} \Gamma\left(\frac{k-1}{2}\right)} \times \exp\left\{ -\frac{r-1}{2(r-1/k)} (g_{r,\alpha}(k-1) + C) \right\}.$$

Here, the number of replicas r is The Effective Sample Size (TESS). Therefore, the approximate Bayes factor for this test calculated with (8) is

$$B(\alpha, k, r) = -\alpha \log(\alpha) \Gamma((k-1)/2) \left(k^{-1} r^{k-1}\right)^{\frac{r-1}{2(r-1/k)}} \left[\frac{2(r-1/k)}{(g_{r,\alpha}(k-1) - \log(k) + (k-1)\log(r) + C)(r-1)} \right]^{\frac{k-1}{2}}$$

A very important case arises when $k = 2$. For this situation, the last formula simplifies to

$$B(\alpha, r) = -\alpha \log(\alpha) \left(\frac{r}{2}\right)^{\frac{r-1}{2r-1}} \left[\frac{2(r-1)\pi}{(g_{r,\alpha}(1) - \log\left(\frac{r}{2}\right) + C)(r-1)}\right]^{\frac{1}{2}} \tag{12}$$

5. Obtaining Bounds for $P(H_0|Data)$

In this section, we use (9) and (11) to produce bounds for the posterior probability of the null hypothesis H_0.

Since for any Bayes factor B_{01}

$$B_{01} \geq B_L(p, \xi_0) \quad \text{with} \quad \xi_0 \geq 1, \text{ fixed but arbitrary,}$$

a lower bound for the posterior probability of the null hypothesis can be obtained as

$$\min P(H_0|Data) = \left[1 + \frac{1}{B_L(p, \xi_0)}\right]^{-1}. \tag{13}$$

Figure 2 shows these posterior probabilities (called $P_{RLB_{\xi_0}}$) for different values of ξ_0. To simplify the use of these Bayes factors, we call BFG_{ξ_0} the Bayes factor of Equation (9), BFG the Bayes factor of Equation (10), and BFL the Bayes factor of Equation (11).

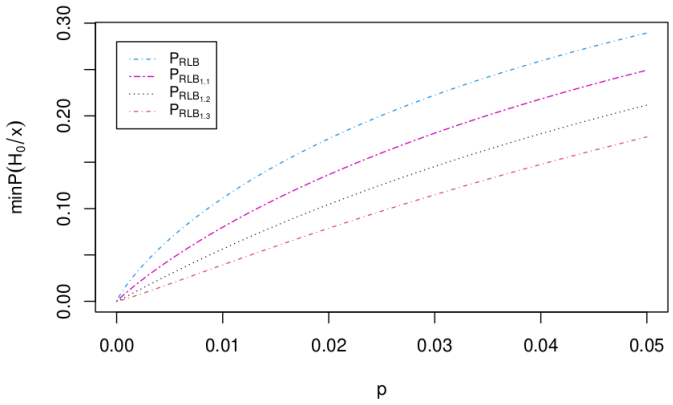

Figure 2. Lower bound for posterior probability for the null hypothesis H_0 (in (13)) for $\xi_0 = 1, \xi_0 = 1.1, \xi_0 = 1.2, \xi_0 = 1.3$.

5.1. Testing Equality of Two Means

Consider comparing two normal means via the test

$$H_0 : \mu_1 = \mu_2 \text{ versus } H_1 : \mu_1 \neq \mu_2,$$

where the associated known variances, σ_1^2 and σ_2^2, are not equal.

$$\mathbf{Y} = \mathbf{X}\mu + \epsilon = \begin{pmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \epsilon_{11} \\ \vdots \\ \epsilon_{2n_2} \end{pmatrix},$$

$$\times \epsilon \sim N(\mathbf{0}, \text{diag}\{\underbrace{\sigma_1^2, \ldots, \sigma_1^2}_{n_1}, \underbrace{\sigma_2^2, \ldots, \sigma_2^2}_{n_2}\})$$

Defining $\nu = (\mu_1 + \mu_2)/2$ and $\zeta = (\mu_1 - \mu_2)/2$ places this in the linear model comparison framework,

$$\mathbf{Y} = \mathbf{B}\begin{pmatrix}\nu\\ \zeta\end{pmatrix} + \epsilon$$

with

$$\mathbf{B} = \begin{pmatrix}1 & 1\\ \vdots & \vdots\\ 1 & 1\\ 1 & -1\\ \vdots & \vdots\\ 1 & -1\end{pmatrix}$$

where we are comparing $M_0 : \zeta = 0$ versus $M_1 : \zeta \neq 0$.
So, for BFG and BFL,

$$C = -2\log\frac{(1-e^{-v})}{\sqrt{2}v}$$

$$v = \frac{\hat{\zeta}^2}{d(1+n^e)}, d = \left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right), n^e = \max\left\{\frac{n_1^2}{\sigma_1^2}, \frac{n_2^2}{\sigma_2^2}\right\}\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right).$$

A special case is the standard test of equality of means when $\sigma_1^2 = \sigma_2^2 = \sigma^2$. Then,

$$n^e = \min\left\{n_1\left(1 + \frac{n_1}{n_2}\right), n_2\left(1 + \frac{n_2}{n_1}\right)\right\}.$$

On the other hand, considering $\mu = \mu_1 - \mu_2$ with $\sigma_1^2 = \sigma_2^2 = \sigma^2$:

- $H_0 : \mu_1 = \mu_2 \longleftrightarrow \mu = 0$;
- $H_1 : \mu_1 \neq \mu_2 \longleftrightarrow \mu \neq 0$.

Assuming priors:

- $\mu|\sigma^2, H_1 \sim Normal(0, \sigma^2/\tau_0), \tau_0 \in (0, \infty)$;
- $\pi(\sigma^2) \propto 1/\sigma^2$ for both H_0 and H_1.

The Bayes factor is

$$BF_{01} = \left(\frac{n+\tau_0}{\tau_0}\right)^{1/2}\left(\frac{t^2\frac{\tau_0}{n+\tau_0} + l}{t^2 + l}\right)^{\frac{l+1}{2}} \tag{14}$$

where

$$t = \frac{|\bar{\mathbf{Y}}|}{s/\sqrt{n}}$$

a t-statistic with degrees of freedom $l = n - 1$ and $n = n_1 + n_2$; see [13].

Figure 3 shows the posterior probability for the null hypothesis H_0 when $n = 50$ and $n = 100$ for the Robust Lower Bound with $\xi_0 = 1$ (called P_{RLB}), the Bayes factor BFL (called P_{BFL}), the Bayes factor BFG (called P_{BFG}), and the Bayes factor BF_{01} (called $P_{BF_{01}}$). Note that the posterior probability with BF_{01} when $\tau_0 = 6$ looks very similar to the result obtained using the Bayes factors BFL and BFG.

We now present a simulation that shows how our adjustment, or calibration, to RLB_ξ works quite similarly to an exact Bayes factor. We perform the following experiment: We simulate r data points from each of the two normal distributions, $N(\mu_1, \sigma)$ and $N(\mu_2, \sigma)$. We reproduce this K times. For all K simulations, $\mu_1 - \mu_2 = 0$. For all K replicates, we test the hypotheses $H_0 : \mu_1 = \mu_2$ vs. $H_1 : \mu_1 \neq \mu_2$, and then we count how many of the p-values lie between $0.05 - \varepsilon$ and 0.05. Note that all of these p-values would be considered sufficient to reject H_0 if $\alpha = 0.05$ is selected. Finally, we determine the proportion of these "significant" p-values obtained from samples where H_0 is true.

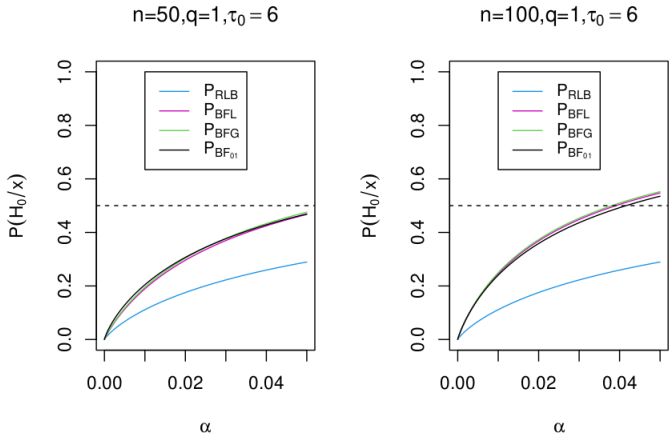

Figure 3. Posterior probability for the null hypothesis H_0 for $n = 50$ and $n = 100$ using the Bayes factor RLB_{ξ_0} with $\xi_0 = 1$, the Bayes factor BF_{01}, and the Bayes factor BFL and BFG.

Table 2 presents the mean percentage of these significant p-values coming from samples, where H_0 is true for 100 iterations of the simulation scheme with $K = 8000$, $\sigma = 1$, and $\varepsilon = 0.05$ for $r = 10, 50, 100, 500$, and 1000. As expected, the distribution of the p-values behaved Uniform(0, 1) under H_0, since H_0 was assumed true in the K replicates. Table 2 also presents the proportion of posterior probability of H_0 greater than or equal to 0.5 (50%) when using the RLB_{ξ}, when corrected according to the method suggested in this document (Equations (10) and (11)), and when an exact Bayes factor (Equation (14)) is used. It is clear that the method suggested here behaves very similarly to an exact Bayes factor.

Table 2. Mean percentage of p-values less than 0.05 (considered significant) coming from data generated under the null hypothesis for 100 experiments, where $K = 8000$ testing problems are generated under $H_0 : \mu_1 = \mu_2$. This experiment is performed for different groups with sample sizes r. Corrected and uncorrected Bayes factors are considered, as well as an exact Bayes factor.

		% Of Samples with $P(H_0|x) \geq 0.5$			
r	**% Of Samples with $p < 0.05$**	RLB_{ξ}	BFG	BFL	BF_{01}
10	5%	0%	58%	66%	75%
50	5%	0%	81%	86%	87%
100	5%	0%	86%	89%	91%
500	5%	0%	94%	96%	96%
1000	5%	0%	95%	96%	97%

5.2. Fisher's Exact Test

This is an example where the p-value is a pseudo-p-value (see the example 8.3.30 in [4]). Let S_1 and S_2 be independent observations with $S_1 \sim$ binomial(n_1, p_1) and $S_2 \sim$ binomial(n_2, p_2). Consider testing $H_0 : p_1 = p_2$ vs. $H_1 : p_1 \neq p_2$.
Under H_0, if we let p be the common value of $p_1 = p_2$, the joint pmf of (S_1, S_2) is

$$f(s_1, s_2|p) = \binom{n_1}{s_1}\binom{n_2}{s_2} p^{s_1+s_2}(1-p)^{n_1+n_2-(s_1+s_2)}$$

and the conditional pseudo-p-value is

$$p(s_1, s_2) = \sum_{j=s_1}^{\min\{n_1, s\}} f(j|s), \tag{15}$$

the sum of hypergeometric probabilities, with $s = s_1 + s_2$.

Remark 2. *It does not seem to be simple to estimate the appropriate ξ_0 that best fits the pseudo-p-value in (15), in Figure 4 some arbitrary possibilities are given.*

It is important to note that in Bayesian tests with a point null hypothesis, it is not possible to use continuous prior densities, because these distributions (as well as posterior distributions) will grant zero probability to $p = (p_1 = p_2)$. A reasonable approximation will be to give $p = (p_1 = p_2)$, a positive probability π_0, and to $p \neq (p_1 = p_2)$ the prior distribution $\pi_1 g_1(p)$, where $\pi_1 = 1 - \pi_0$ and g_1 proper. One can think of π_0 as the mass that would be assigned to the real null hypothesis, $H_0 : p \in ((p_1 = p_2) - b, (p_1 = p_2) + b)$ if it had not been preferred to approximate by the null point hypothesis. Therefore, if

$$\pi(p) = \begin{cases} \pi_0 & p = (p_1 = p_2) \\ \pi_1 g_1(p) & p \neq (p_1 = p_2) \end{cases}$$

then

$$\begin{aligned} m(s) &= \int_\Theta f(s|p)\pi(p)dp \\ &= f(s|(p_1 = p_2))\pi_0 + \pi_1 \int_{p \neq (p_1 = p_2)} f(s|p)g_1(p)dp \\ &= f(s|(p_1 = p_2))\pi_0 + (1 - \pi_0)m_1(s) \end{aligned}$$

where $m_1(s) = \int_{p \neq (p_1 = p_2)} f(s|p)g_1(p)dp$ is the marginal density of $(S = S_1 + S_2)$ with respect to g_1.

So,

$$\pi((p_1 = p_2)|s) = \frac{\pi_0 f(s|(p_1 = p_2))}{m(s)}$$

thus

$$\begin{aligned} \text{posterior odds} &= \frac{\pi((p_1 = p_2)|s)}{1 - \pi((p_1 = p_2)|s)} \\ &= \frac{f(s|(p_1 = p_2))\pi_0}{m(s)(1 - \frac{f(s|(p_1 = p_2))\pi_0}{m(s)})} \\ &= \frac{f(s|(p_1 = p_2))\pi_0}{m(s) - f(s|(p_1 = p_2))\pi_0} \\ &= \frac{f(s|(p_1 = p_2))\pi_0}{(1 - \pi_0)m_1(s)} \\ &= \frac{\pi_0 f(s|(p_1 = p_2))}{\pi_1 m_1(s)} \\ &= \text{prior odds} \cdot \frac{f(s|(p_1 = p_2))}{m_1(s)} \end{aligned}$$

and the Bayes factor is

$$B_{01} = \frac{f(s|(p_1 = p_2))}{m_1(s)}.$$

Now, if we take $g_1(p) = \text{Beta}(a, b)$ such that $E(p) = \frac{a}{a+b} = (p_1 = p_2)$, then

$$BF_{Test} = \frac{B(a, b)}{B(s + a, n_1 + n_2 - s + b)} p^s (1 - p)^{n_1 + n_2 - s}.$$

Figure 4 shows the posterior probability for the null hypothesis H_0 when $n = n_1 + n_2 = 50$ and 100, for the Robust Lower Bound, the Bayes factor BFG_{ξ_0} (called $P_{BFG_{\xi_0}}$), the Bayes factor BFG (called P_{BFG}), and the Bayes factor BF_{Test} (called $P_{BF_{Test}}$). We can note that all the $P_{BFG_{\xi_0}}$ are comparable, even though in the case $\xi_0 = 1$ (P_{BFG}) it is a p-value and not a pseudo-p-value.

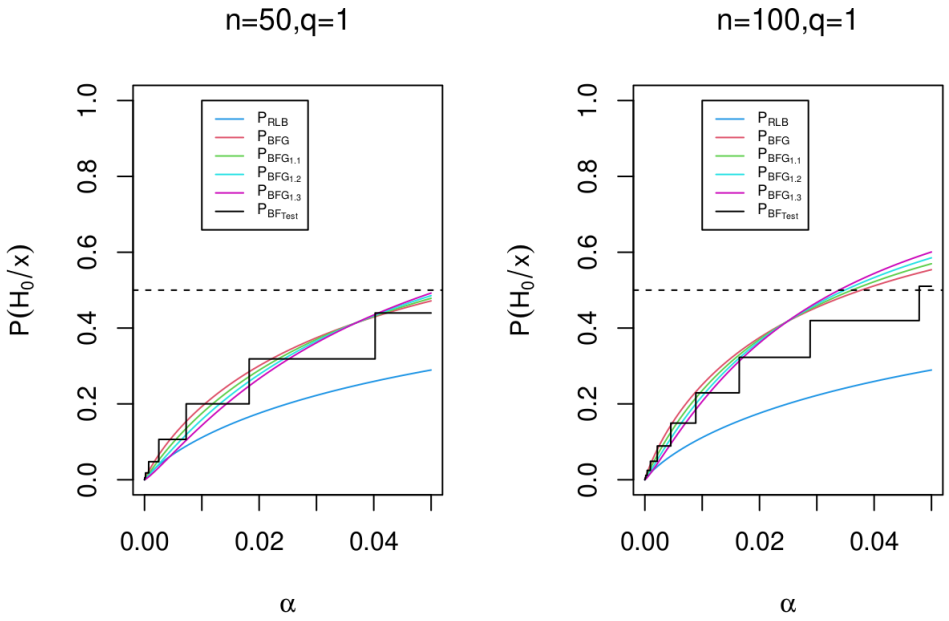

Figure 4. Posterior probability for the null hypothesis H_0 for $n = 50$ and $n = 100$ using the Bayes factor RLB_{ξ_0} with $\xi_0 = 1$, the Bayes factor BF_{Test}, the Bayes factor BFG_{ξ_0}, and the Bayes factor BFG.

5.3. Linear Regression Models

Consider comparing two nested linear models $M_3 : y_l = \lambda_1 + \lambda_2 x_{l2} + \lambda_3 x_{l3} + \epsilon_l$ with $M_2 : y_l = \lambda_1 + \lambda_2 x_{l2} + \epsilon_l$ via the test

$$H_0 : M_2 \text{ versus } H_1 : M_3,$$

with $1 \leq l \leq n$, and the errors ϵ_l are assumed to be independent and normally distributed with unknown residual variance σ^2. According to the Equation (3) in [6,7]

$$b = (n-1)s_3^2(1-\rho_{23}^2),$$

where s_3^2 is the variance x_{v3}, ρ_{23} is the correlation between x_{v2} and x_{v3}, and

$$C = 2\log\frac{(1-e^{-v_2})}{\sqrt{2}v_2} - 2\log\frac{(1-e^{-v_3})}{\sqrt{2}v_3},$$

where $v_2 = \hat{\lambda}_2^2/[d_2(1+n_2^e)]$, $d_2 = \sigma^2/s_{x_{l2}}^2$, $n_2^e = s_{x_{l2}}^2/\max_i\{(x_{l2}-\bar{x}_2)^2\}$ and $v_3 = \hat{\lambda}_3^2/[d_3(1+n_3^e)]$, $d_3 = \sigma^2(\tilde{X}^t\tilde{X})^{-1}$, $n_3^e = \tilde{X}^t\tilde{X}/\max_i\{|\tilde{X}_i|^2\}$ with $\tilde{X} = (\mathbf{I}_n - X^*(X^{*t}X^*)^{-1}X^*)x_{l3}$ and $X^* = (\mathbf{1}_n|x_{l2})$.

As an example, we analyze a data set taken from [14], which can be accessed at http://academic.uprm.edu/eacuna/datos.html (accessed on 13 January 2022). We want to predict the average mileage per gallon (denoted by `mpg`) of a set of $n = 82$ vehicles using four possible predictor variables: cabin capacity in cubic feet (`vol`), engine power (`hp`), maximum speed in miles per hour (`sp`), and vehicle weight in hundreds of pounds (`wt`).

Through the Bayes factors BFG and BFL, we want to choose the best model to predict the average mileage per gallon by calculating the posterior probability of the null hypothesis of the following test

$$H_0 : M_2 : \text{mpg} = \lambda_1 + \lambda_2 \text{wt}_l + \epsilon_l \text{ vs. } H_1 : M_3 : \text{mpg} = \lambda_1 + \lambda_2 \text{wt}_l + \lambda_3 \text{sp}_l + \epsilon_l$$

with $\alpha = 0.05$, $q = 1$, $j = 3$, the posterior probabilities for the null hypothesis H_0 are

$$P_{BFL} = 0.9253192, P_{BFG} = 0.7209449.$$

The use of this posterior probability in both cases will change the inference, since the p-value of the F test is $p = 0.0325$, which is smaller than 0.05.

Findley's Counterexample

Consider the following simple linear model [15]

$$Y_i = \frac{1}{\sqrt{i}} \cdot \theta + \epsilon_i, \text{ where } \epsilon_i \sim N(0,1), i = 1,2,3,\ldots,n$$

and we are comparing the models $H_0 : \theta = 0$ and $H_1 : \theta \neq 0$. This is a classical and challenging counterexample against BIC and the Principle of Parsimony. In [7], the inconsistency of BIC is shown, but the consistency of PBIC is shown in this problem.

Here, we show through the posterior probabilities of the null hypothesis that the Bayes factor BFG (based on BIC) is inconsistent, while the Bayes factor BFL (based on PBIC) is consistent if it is. We perform the analysis in two contexts: First, when n grows and $\alpha = 0.05$ or $\alpha = 0.01$ are fixed. Second, when n is fixed and $0 < \alpha < 0.05$. For calculations

$$C = -2 \log \frac{(1 - e^{-v})}{\sqrt{2}v}, v = \frac{\hat{\theta}^2}{d(1 + n^e)}, d = \left(\sum_{i=1}^{n} \frac{1}{i} \right)^{-1}, n^e = \sum_{i=1}^{n} \frac{1}{i}.$$

Figures 5 and 6 show, through the posterior probability of the null hypothesis H_0, the consistency of the Bayes factor based in PBIC (P_{BFL}), as well as the inconsistency of the Bayes factor based in BIC (P_{BFG}).

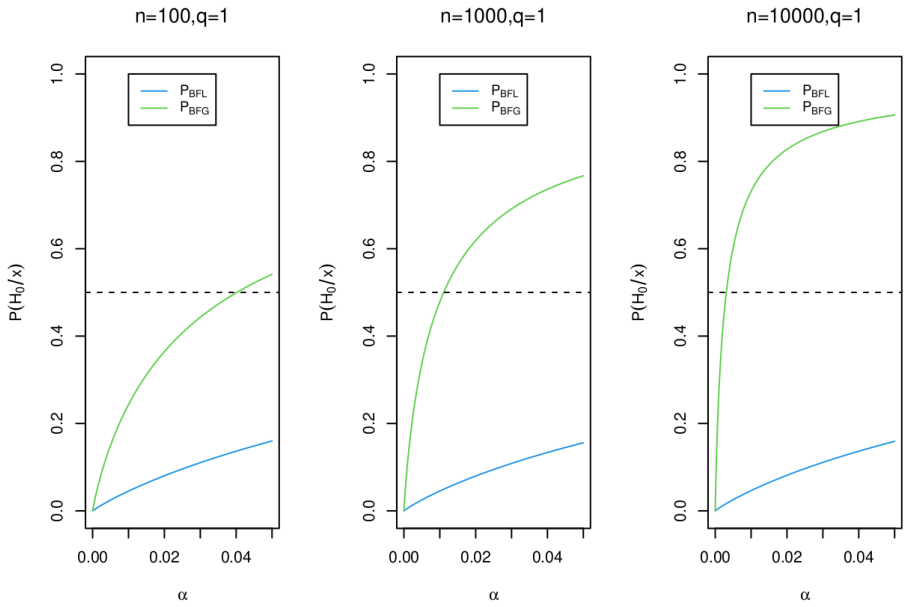

Figure 5. Posterior probability for the null hypothesis H_0 for $n = 100$, $n = 1000$ and $n = 10{,}000$ using the Bayes factors BFL and BFG.

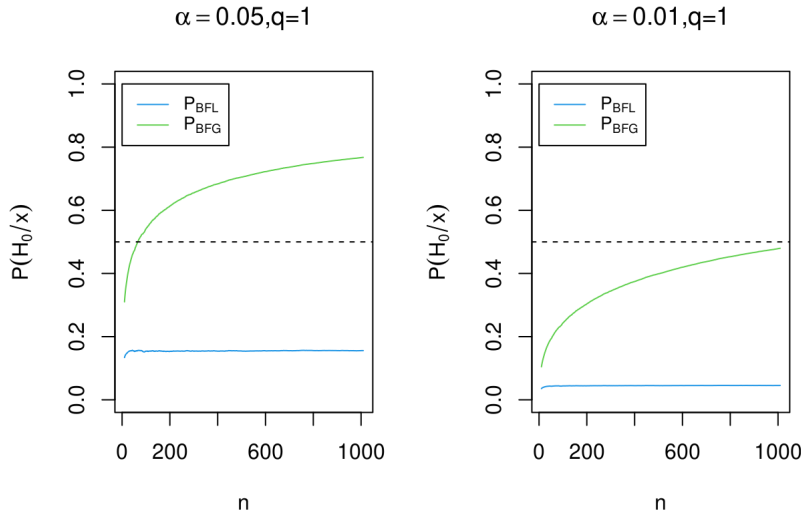

Figure 6. Posterior probability for the null hypothesis H_0 for $\alpha = 0.05$ and $\alpha = 0.01$ using the Bayes factors BFL and BFG when n grows.

6. Discussion and Final Comments

1. Lower bounds have been an important development to give practitioners alternatives to classical testing with fixed α levels. A deep-seated problem with the useful bound $-e \cdot p \cdot log(p)$ is that it depends on the p-value, which it should, but it is static, not a function of the sample size n. This limitation makes the bound of little use for moderate to large sample sizes, where it is arguably the correction to p-values more needed.
2. The approximation develops here as a function of p-values, and sample size has a distinct advantage over other approximations, such as BIC, in that it is a valid approximation for any sample size.
3. The (approximate) Bayes factors (9) and (11) are simple to use and provide results equivalent to the sensitive p-value Bayes factors of hypothesis tests. In this article, we extended the validity of the approximation for "pseudo-p-values," which are ubiquitous in statistical practice. We hope that this development will give tools to the practice of statistics to make the posterior probability of hypotheses closer to everyday statistical practice, on which p-values (or pseudo-p-values) are calculated routinely. This allows an immediate and useful comparison between raw-p-values and (approximate) posterior odds.

Author Contributions: Conceptualization, D.V.R., L.R.P.G. and M.E.P.H.; methodology, D.V.R., L.R.P.G. and M.E.P.H.; software, D.V.R.; validation, D.V.R., L.R.P.G. and M.E.P.H.; formal analysis, D.V.R., L.R.P.G. and M.E.P.H.; investigation, D.V.R., L.R.P.G. and M.E.P.H.; writing—original draft preparation, D.V.R.; writing—review and editing, D.V.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The real datasets are freely available in http://academic.uprm.edu/eacuna/datos.html.

Acknowledgments: The first author gratefully acknowledges financial support from the Faculty of Business Administration of the University of Puerto Rico Río Piedras Campus. The work of L.R Pericchi and M.E Pérez has been partially funded by NIH grant U54CA096300, P20GM103475, and R25MD010399.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Proof of Lemma 1. Let $h(p_{val}) = -e \cdot \xi \cdot \log(p_{val})$, then $\frac{d[h(p_{val})]}{dp_{val}} = -\frac{e \cdot \xi}{p_{val}} < 0$; thus, h is decreasing with minimum at $\xi = e^{-1}$. So, $h(p_{val}) \geq h(e^{-1}) = e \cdot \xi$, which implies $B_L(p_{val}, \xi)/p_{val}^{\xi} = h(p_{val}) \geq e \cdot \xi$, so $B_L(p_{val}, \xi) \geq e \cdot \xi \cdot p_{val}^{\xi} > p_{val}^{\xi}$ □

Proof of Theorem 1. First of all, it can be seen that $B_L(p, \xi) = -e \cdot \xi \cdot p^{\xi} \cdot \log(p)$ is well-defined, since $0 \leq B_L(p, \xi) \leq 1$.

Let $\alpha \in [0, 1]$ and denote by D_B the subset of R_p (range of p), such that

$$-e \cdot \xi \cdot p^{\xi} \cdot \log(p) \leq \alpha,$$

then

$$(B_L(p, \xi) \leq \alpha) = [-e \cdot \xi \cdot p^{\xi} \cdot \log(p) \leq \alpha] = (p \in D_B)$$

where $(p \in D_B)$ is the event that consists of all the result x, such that the point $p(x) \in D_B$. Therefore,

$$\begin{aligned} F_B(\alpha) = P(B_L(p, \xi) \leq \alpha | p \sim f(p|\xi)) &= P(-e \cdot \xi \cdot p^{\xi} \cdot \log(p) \leq \alpha | p \sim f(p|\xi)) \\ &= P(p \in D_B | p \sim f(p|\xi)) \\ &= \int_{D_B} f_p(p) dp \\ &= \int_0^{\rho} \xi p^{\xi - 1} dp \\ &= \rho^{\xi} \end{aligned}$$

where ρ is determined such that

$$0 < \rho < \frac{1}{e} \text{ and } \alpha = -e \cdot \xi \cdot \rho^{\xi} \cdot \log(\rho)$$

as shown in the Figure A1 for the case when $\xi = 1$.

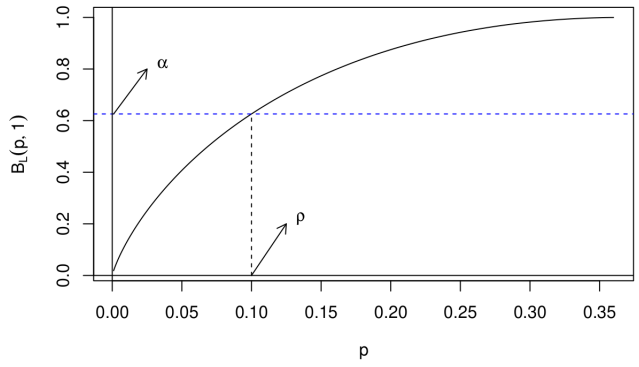

Figure A1. Proof of Theorem 1: graph of the generalized Robust Lower Bound for $\xi = 1$ ($B_L(p, 1)$), identifying the value ρ where $-e \cdot \rho \cdot \log(\rho) = \alpha$.

□

Now, by Lemma 1 $F_B(\alpha) = \rho^{\xi} < -e \cdot \xi \cdot \rho^{\xi} \cdot \log(\rho) = \alpha$.

Appendix B. Codes

```
I=seq(1,n1+n2,1)
y=I
for (i in I) {
y[i]=1
}
return(y)
}
Y=function(n1=10,n2=10){
I=seq(1,n1,1)
y=rep(-1,n1+n2)
for (i in I) {
y[i]=1
}
return(y)
}
ml=function(n1=10,n2=10){return(lm(X(n1,n2)~Y(n1,n2)))}
sigma=function(n1=10,n2=10){
return(as.numeric(summary(ml(n1,n2))$sigma^2))}
d=function(n1=10,n2=10){return(sigma(n1,n2)*(1/n1+1/n2))}
ne=function(n1=10,n2=10){return(min(n1*(1+n1/n2),n2*(1+n2/n1)))}
beta.=function(n1=10,n2=10){
return(as.numeric(ml(n1,n2)$coefficients[2]^2))}
v=function(n1=10,n2=10){
return(beta.(n1,n2)/(d(n1,n2)*(1+ne(n1,n2))))}
C=function(n1=10,n2=10){
return(-2*log((1-exp(-v(n1,n2)))/(sqrt(2)*v(n1,n2))))}
# Adaptive alpha eq.8
alphabinom=function(n1,n2,alpha){
sqrt(2/((n1+n2)*pi*(qchisq(alpha, df=1, lower.tail=F)
+log(n1+n2)
+C(n1,n2))))*exp(-(qchisq(alpha, df=1, lower.tail=F)
+C(n1,n2))/2)
}
```

```
# RLB_xi
RLB=function(a,b){
  -exp(1)*b*a^b*log(a)}
pval=seq(0.001,0.36,0.00001)
plot(pval,RLB(pval,1),col=4,lty=4,
ylab=expression(paste(B[L](p,xi[0]))),
xlab=expression(paste(p)),type="l")
lines(pval,RLB(pval,1.1),col=5,lty=5)
lines(pval,RLB(pval,1.2),col=6,lty=6)
lines(pval,RLB(pval,1.3),col=7,lty=7)
legend(0.01,1,col =c(4,5,6,7),
c(expression(paste(xi[0]==1)),
expression(paste(xi[0]==1.1)),
expression(paste(xi[0]==1.2)),
expression(paste(xi[0]==1.3))),
lty=c(4,5,6,7),cex = 0.8)
```

```
plot(pval,RLB(pval,1),
ylab=expression(paste(B[L](p,1))),
xlab=expression(paste(p)),type="l")
abline(h=RLB(.1,1),lty=2,col="blue")
abline(v=0)
abline(h=0)
segments(0.1,0,0.1,RLB(0.1,1),lty=2)
arrows(0.001,RLB(0.1,1),0.025,0.8,length = 0.1)
arrows(0.1,0,0.125,0.2,length = 0.1)
legend(0.01,0.9,expression(paste(alpha)),bty = "n")
legend(0.11,0.3,expression(paste(rho)),bty = "n")
```

```
alpha=seq(0.000000000001,.05,.00001)
# posterior probability of H_0
pP=function(a){
  1/(1+1/(a))}
# posteriors probability (RLB_xi)
plot(alpha,pP(RLB(alpha,1)),col=4,lty=4,xlab="p",
ylab=expression(paste(minP(H[0]/x))),type = "l")
lines(alpha,pP(RLB(alpha,1.1)),col=6,lty=6)
lines(alpha,pP(RLB(alpha,1.2)),col=9,lty=9)
lines(alpha,pP(RLB(alpha,1.3)),col=10,lty=10)
legend(0,.28,col =c(4,6,9,10),
c(expression(paste(P[RLB])),
expression(paste(P[RLB[1.1]])),
expression(paste(P[RLB[1.2]])),
expression(paste(P[RLB[1.3]]))),
lty=c(4,6,9,10),cex = 0.8)
```

```
Y=function(n1,n2){
c=cbind2(c(rep(1,n1),rep(1,n2)))
return(c)}
Y1=function(n1,n2){
set.seed(2)
a=rnorm(n1+n2,0,.05)
c=cbind2(c(rep(1,n1),rep(3,n2))+a)
return(c)
}
X1=function(n1,n2){
c=cbind2(c(rep(1,n1),rep(-1,n2)))
return(c)
}
X=function(n1,n2){
return(cbind2(Y(n1,n2),X1(n1,n2)))
}
b=function(n1,n2){
return(abs(det(t(X(n1,n2))%*%X(n1,n2))/det(t(Y(n1,n2))%*%
Y(n1,n2))))}
l.model=function(n1,n2){return(lm(Y1(n1,n2)~X1(n1,n2)))}
beta=function(n1,n2){as.numeric(l.model(n1,n2)$coefficient[2])}
d=function(n1,n2){return(2/n1+2/n2)}
ne=function(n1,n2){return(min(n1^2,n2^2)*(1/n1+1/n2))}
```

```
v=function(n1,n2){return(beta(n1,n2)^2/(d(n1,n2)*(1+ne(n1,n2))))}
C=function(n1,n2){return(-2*log((1-exp(-v(n1,n2)))/(sqrt(2)*
v(n1,n2))))}
# Bayes Factor Linear Version (Eq.8)
BFL=function(alpha,q,n,b,C,j){
-alpha*log(alpha)*gamma(q/2)*b^((n-j)/(2*(n-1)))*
((2*(n-1))/((qgamma(alpha,shape=q/2,rate=(n-j)/
(2*(n-1)),lower.tail = FALSE)
+log(b)+C)*(n-j)))^(q/2)
}
# Bayes Factor General (E.q 9)
BFG=function(alpha,q,n,C){
-alpha*log(alpha)*gamma(q/2)*n^(q/2)*
(2/(qchisq(alpha,q,lower.tail=FALSE)+q*log(n)+C))^(q/2)
# Bayes Factor $BF_{01}$ (means)

BF=function(t,n1,n2,alpha){
  n=n1+n2
  l=n-1
  return(((n+t)/t)^(1/2)*((qt(alpha,l,lower.tail=FALSE))^2*
  (t/(n+t))+l)/((qt(alpha,l,lower.tail = FALSE))^2+l))^((l+1)/2))
}
# Plot posteriors probability
par(mfrow=c(1,2))
plot(alpha,pP(RLB(alpha,1)),col=4,
xlab=expression(paste(alpha)),
ylab=expression(paste(P(H[0]/x))),
main =expression(paste("n=50,","q=1,",
tau[0]==6)),type="l",ylim = c(0,1))
lines(alpha,pP(BFL(alpha,1,50,b(25,25),C(25,25),2)),
col=6)
lines(alpha,pP(BFG(alpha,1,50,C(25,25))),col=3)
lines(alpha,pP(BF(6,25,25,alpha)),col=9)
legend(0.01,1,col =c(4,6,3,9),
c(expression(paste(P[RLB])),
expression(paste(P[BFL])),
expression(paste(P[BFG])),
expression(paste(P[BF["01"]]))),
 lty=c(1,1,1,1),cex = 0.9)
abline(.5,0,lty=2)
plot(alpha,pP(RLB(alpha,1)),col=4,
xlab=expression(paste(alpha)),
ylab=expression(paste(P(H[0]/x))),
main = expression(paste("n=100,","q=1,",tau[0]==6)),
type="l",ylim = c(0,1))
lines(alpha,pP(BFL(alpha,1,100,b(50,50),
C(50,50),2)),col=6)
lines(alpha,pP(BFG(alpha,1,100,C(50,50))),col=3)
lines(alpha,pP(BF(6,50,50,alpha)),col=9)
legend(0.01,1,col =c(4,6,3,9),
 c(expression(paste(P[RLB])),
expression(paste(P[BFL])),
expression(paste(P[BFG])),
expression(paste(P[BF["01"]]))),
```

```
 lty=c(1,1,1,1),cex = 0.9)
abline(.5,0,lty=2)
```

```
# Bayes factor Fisher's␣Exact␣Test

B_01=function(p,a,b,alpha,n){
p^(qbinom(alpha,n,p,lower.tail␣=␣FALSE))*
(1-p)^(n-qbinom(alpha,n,p,lower.tail␣=␣FALSE))*
beta(a,b)/beta(qbinom(alpha,n,p,lower.tail␣=␣FALSE)+a,
n-qbinom(alpha,n,p,lower.tail␣=␣FALSE)+b)
}
z=B_01(.7,7,3,alpha,50)
x=B_01(.7,7,3,alpha,100)
#␣Posteriors␣probability
par(mfrow=c(1,2))
plot(alpha,pP(RLB(alpha,1)),col=4,
xlab=expression(paste(alpha)),
ylab=expression(paste(P(H[0]/x))),
main␣=␣expression(paste("n=50,","q=1")),type␣=␣"l",ylim␣=␣c(0,1))
lines(alpha,pP(BFG(1,alpha,25,25,1)),col=2)
lines(alpha,pP(BFG(1,alpha,25,25,1.1)),col=3)
lines(alpha,pP(BFG(1,alpha,25,25,1.2)),col=5)
lines(alpha,pP(BFG(1,alpha,25,25,1.3)),col=6)
lines(alpha,pP(z),col=9)
legend(0.01,1,col␣=c(4,2,3,5,6,9),
c(expression(paste(P[RLB])),
expression(paste(P[BFG])),
expression(paste(P[BFG[1.1]])),
expression(paste(P[BFG[1.2]])),
expression(paste(P[BFG[1.3]])),
expression(paste(P[BF[Test]]))),
␣lty=c(1,1,1,1,1,1),cex␣=␣0.6)
abline(.5,0,lty=2)
plot(alpha,pP(RLB(alpha,1)),col=4,
xlab=expression(paste(alpha)),
ylab=expression(paste(P(H[0]/x))),
main␣=␣expression(paste("n=100,","q=1")),
type␣=␣"l",ylim␣=␣c(0,1))
lines(alpha,pP(BFG(1,alpha,80,20,1)),col=2)
lines(alpha,pP(BFG(1,alpha,80,20,1.1)),col=3)
lines(alpha,pP(BFG(1,alpha,80,20,1.2)),col=5)
lines(alpha,pP(BFG(1,alpha,80,20,1.3)),col=6)
lines(alpha,pP(x),col=9)
legend(0.01,1,col␣=c(4,2,3,5,6,9),
␣c(expression(paste(P[RLB])),
expression(paste(P[BFG])),
expression(paste(P[BFG[1.1]])),
expression(paste(P[BFG[1.2]])),
expression(paste(P[BFG[1.3]])),
expression(paste(P[BF[Test]]))),
␣lty=c(1,1,1,1,1,1),cex␣=␣0.6)
abline(.5,0,lty=2)
```

```
# C and b

Y=function(n){
c=cbind2(rep(1,n))
return(c)}

X1=function(n){
I=seq(1,n,1)
x=I
for (i in I) {
x[i]=1/i
}
return(as.matrix(x))
}
Y1=function(n){
set.seed(4)
a=rnorm(n,0,1)
return(a+X1(n)*0.5)
}
X=function(n){
return(cbind2(Y(n),X1(n)))
}
b=function(n){
return(abs(det(t(X(n))%*%X(n))/det(t(Y(n))%*%Y(n))))}
l.model=function(n){return(lm(Y1(n)~X1(n)))}
theta=function(n){as.numeric(l.model(n)$coefficient[2])}
d=function(n){return(1/apply(X1(n),2,sum))}
ne=function(n){return(apply(X1(n),2, sum))}
v=function(n){return(theta(n)^2/(d(n)*(1+ne(n))))}
C=function(n){return(-2*log((1-exp(-v(n)))/(sqrt(2)*v(n))))}

# plot posteriors probability in function of alpha.

par(mfrow=c(1,3))
plot(alpha,pP(BFL(alpha,1,100,b(100),C(50),2)),
col=4,xlab=expression(paste(alpha)),
ylab=expression(paste(P(H[0]/x))),
main =expression(paste("n=100,","q=1")),
type="l",ylim = c(0,1))
lines(alpha,pP(BFG(alpha,1,100,C(100))),col=3)
legend(0.01,1,col =c(4,3),
 c(expression(paste(P[BFL])),
expression(paste(P[BFG]))),
 lty=c(1,1),cex = 0.9)
abline(.5,0,lty=2)

plot(alpha,pP(BFL(alpha,1,1000,b(1000),C(1000),2)),
col=4,xlab=expression(paste(alpha)),
ylab=expression(paste(P(H[0]/x))),
main =expression(paste("n=1000,","q=1")),
type="l",ylim = c(0,1))
lines(alpha,pP(BFG(alpha,1,1000,
C(1000))),col=3)
legend(0.01,1,col =c(4,3),
```

```
c(expression(paste(P[BFL])),
expression(paste(P[BFG]))),
 lty=c(1,1),cex = 0.9)
abline(.5,0,lty=2)

plot(alpha,pP(BFL(alpha,1,10000,b(10000),C(10000),2)),
col=4,xlab=expression(paste(alpha)),
ylab=expression(paste(P(H[0]/x))),
main =expression(paste("n=10000,","q=1")),
type="l",ylim = c(0,1))
lines(alpha,pP(BFG(alpha,1,10000,
C(10000))),col=3)
legend(0.01,1,col =c(4,3),
c(expression(paste(P[BFL])),
expression(paste(P[BFG]))),
lty=c(1,1),cex = 0.9)
abline(.5,0,lty=2)
```

```
# plot posteriors probability in function of n.

I=seq(1,1000,1)
BL=I
BL1=I
BG=I
BG1=I
for (n in I) {
   i=9+n
 BL[n]=BFL(0.05,1,i,b(i),C(i),2)
 BL1[n]=BFL(0.01,1,i,b(i),C(i),2)
 BG[n]=BFG(0.05,1,i,C(i))
 BG1[n]=BFG(0.01,1,i,C(i))
}

m=seq(10,1009,1)
par(mfrow=c(1,2))
plot(m,pP(BL),col=4,
xlab=expression(paste("n")),
ylab=expression(paste(P(H[0]/x))),
main =expression(paste(alpha==0.05,",","q=1")),
type="l",ylim = c(0,1))
lines(m,pP(BG),col=3)
legend(0.01,1,col =c(4,3),
 c(expression(paste(P[BFL])),
expression(paste(P[BFG]))),
lty=c(1,1),cex = 0.8)
abline(.5,0,lty=2)
plot(m,pP(BL1),col=4,
xlab=expression(paste("n")),
ylab=expression(paste(P(H[0]/x))),
main =expression(paste(alpha==0.01,",","q=1")),
type="l",ylim = c(0,1))
lines(m,pP(BG1),col=3)
legend(0.01,1,col =c(4,3),
```

```
c(expression(paste(P[BFL])),
expression(paste(P[BFG]))),
 lty=c(1,1),cex = 0.8)
abline(.5,0,lty=2)
```

References

1. Sellke, T.; Bayarri, M.J.; Berger, J.O. Calibration of p values for testing precise null hypotheses. *Am. Stat.* **2001**, *55*, 62–71. [CrossRef]
2. Benjamin, D.; Berger, J.; Johannesson, M.; Nosek, B.; Wagenmakers, E.-J.; Berk, R.; Bollen, K.; Brembs, B.; Brown, L.; Camerer, C.; et al. Redefine statistical significance. *Nat. Hum. Behav.* **2018**, *2*, 6–10. [CrossRef] [PubMed]
3. Held, L.; Ott, M. How the Maximal Evidence of *p*-Values Against Point Null Hypotheses Depends on Sample Size. *Am. Stat.* **2016**, *70*, 335–341. [CrossRef]
4. Casella, G.; Berger, R. *Statistical Inference*, 2nd ed.; Duxbury Resource Center: Belmont, CA, USA, 2017.
5. Pérez, M.E.; Pericchi, L.R. Changing statistical significance with the amount of information: The adaptive alfa significance level. *Stat. Probab. Lett.* **2014**, *85*, 20–24. [CrossRef] [PubMed]
6. Vélez, D.; Pérez, M.E.; Pericchi, L.R. Increasing the replicability for linear models via adaptive significance levels. *Test* **2022**, *31*, 771–789. [CrossRef]
7. Bayarri, M.J.; Berger, J.O.; Jang, W.; Ray, S.; Pericchi, L.R.; Visser, I. Prior-based bayesian information criterion. *Stat. Theory Relat. Fields* **2019**, *3*, 2–13. [CrossRef]
8. Vovk, V. A logic of probability, with application to the foundations of statistic. *J. R. Stat. Soc. Ser. B* **1993**, *55*, 31–351. [CrossRef]
9. Held, L.; Ott, M. On *p*-values and bayes factors. *Annu. Rev. Stat. Appl.* **2018**, *5*, 393–419. [CrossRef]
10. Cabras, S.; Castellanos, M. *p*-value calibration in multiple hypotheses testing. *Stat. Med.* **2021**, *36*, 2875–2886. [CrossRef] [PubMed]
11. Patiño Hoyos, A.E.; Fossaluza, V.; Esteves, L.G.; Bragança Pereira, C.A.d. Adaptive Significance Levels in Tests for Linear Regression Models: The *e*-Value and *p*-Value Cases. *Entropy* **2023**, *25*, 19. [CrossRef] [PubMed]
12. Luis, P.; Pereira, C. Adaptative Significance Levels Using Optimal Decision Rules: Balancing by Weighting the Error Probabilities. *Braz. J. Probab. Stat.* **2015**, *29*, 70–90.
13. Roger, S.Z.; Sarkar, A.; Carroll, R.J.; Mallick, B.K. A powerful bayesian test for equality of means in high dimensions. *J. Am. Stat. Assoc.* **2018**, *113*, 1733–1741.
14. Acuna, E. *Regresion Aplicada Usando R*; Universidad de Puerto Rico en Mayagüez, Departamento de Ciencias Matemáticas: Mayagüez, Puerto Rico, 2015
15. Findley, D.F. Counterexamples to parsimony and BIC. *Ann. Inst. Stat. Math.* **1991**, *43*, 505–514. [CrossRef]