

Article

An Exact and Near-Exact Distribution Approach to the Behrens–Fisher Problem

Serim Hong ¹, Carlos A. Coelho ^{2,*}  and Junyong Park ^{3,*}¹ College of Liberal Studies, Seoul National University, Seoul 08826, Korea² NOVA Math (CMA-FCT/UNL) and Mathematics Department, NOVA School of Science and Technology, NOVA University of Lisbon (FCT/UNL), 2829-516 Caparica, Portugal³ Department of Statistics, Seoul National University, Seoul 08826, Korea

* Correspondence: cmac@fct.unl.pt (C.A.C.); junyongpark@snu.ac.kr (J.P.)

Abstract: The Behrens–Fisher problem occurs when testing the equality of means of two normal distributions without the assumption that the two variances are equal. This paper presents approaches based on the exact and near-exact distributions for the test statistic of the Behrens–Fisher problem, depending on different combinations of even or odd sample sizes. We present the exact distribution when both sample sizes are odd and the near-exact distribution when one or both sample sizes are even. The near-exact distributions are based on a finite mixture of generalized integer gamma (GIG) distributions, used as an approximation to the exact distribution, which consists of an infinite series. The proposed tests, based on the exact and the near-exact distributions, are compared with Welch’s *t*-test through Monte Carlo simulations, in particular for small and unbalanced sample sizes. The results show that the proposed approaches are competent solutions to the Behrens–Fisher problem, exhibiting precise sizes and better powers than Welch’s approach for those cases. Numerical studies show that the Welch’s *t*-test tends to be a bit more conservative than the test statistics based on the exact or near-exact distribution, in particular when sample sizes are small and unbalanced, situations in which the proposed exact or near-exact distributions obtain higher powers than Welch’s *t*-test.



Citation: Hong, S.; Coelho, C.A.; Park, J. An Exact and Near-Exact Distribution Approach to the Behrens–Fisher Problem. *Mathematics* **2022**, *10*, 2953. <https://doi.org/10.3390/math10162953>

Academic Editor: Christophe Chesneau

Received: 4 July 2022

Accepted: 13 August 2022

Published: 16 August 2022

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: Behrens–Fisher problem; near-exact distribution; Welch’s *t*-test; generalized integer gamma distribution

MSC: 62F03; 62E15; 62E20

1. Introduction

Let us suppose we have two independent random samples $X_{1i}, i = 1, \dots, n_1$ and $X_{2i}, i = 1, \dots, n_2$, which are drawn from two normal distributions, $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$, respectively. The Behrens–Fisher problem occurs when testing the equality of the two means μ_1 and μ_2 based on random samples like these without the assumption that the two variances, σ_1^2 and σ_2^2 , are equal. [1] showed that a uniformly most powerful test does not exist in this case, and the Behrens–Fisher problem remains one of the unsolved problems of statistics. Many different approaches have been tried to solve this problem. Among those approaches are the fiducial approach proposed by [2,3], which somehow opened the way to the Bayesian approach proposed by [4,5], based on setting independent and locally uniform prior distributions for $\mu_1, \mu_2, \log \sigma_1, \log \sigma_2$. The frequentist approach proposed by [6,7] uses Student’s *t* distribution with $\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2 / \left(\frac{s_1^4}{(n_1-1)n_1^2} + \frac{s_2^4}{(n_2-1)n_2^2}\right)$ degrees of freedom as the approximate distribution of the Behrens–Fisher statistic.

In this paper, we will obtain the exact and near-exact distributions, both the probability density function and cumulative distribution function, of the Behrens–Fisher statistic,

$$T^* = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}, \tag{1}$$

under $H_0 : \mu_1 = \mu_2$, where $\bar{X}_j = \frac{1}{n_j} \sum_{i=0}^{n_j} X_{ji}$ and $S_j^2 = \frac{1}{n_j-1} \sum_{i=0}^{n_j} (X_{ji} - \bar{X}_j)^2$ for $j = 1, 2$, in the form of mixtures of Student's t distributions multiplied by constants. Particularly for the case when both sample sizes are odd, the exact distribution will be derived in a finite closed form without any unsolved integrals or infinite sums by using the GIG (generalized integer gamma) distribution in [8], which is the distribution of the sum of independent gamma variables with integer shape parameters and nonequal rate parameters. For the other cases, that is, when both sample sizes are even or one of them is even and the other one is odd, the near-exact distribution will be obtained by approximating the exact distribution using a finite mixture of GIG distributions to obtain a more manageable cumulative distribution function. Such exact and near-exact distributions include σ_1^2 and σ_2^2 as unknown parameters, which have to be estimated, based on the observed samples, being then the p -values obtained from these exact or near-exact distributions with estimated parameters. The results will be compared with Welch's t -test, one of the most widely used solutions to the problem, through Monte-Carlo simulations for relatively small sample sizes. We will see that the tests based on the exact or near-exact distribution show some advantage in terms of being able to obtain higher power than Welch's t -test, especially when sample sizes are small and unbalanced and variances are also unbalanced.

This paper is organized as follows: Section 2 presents the exact distribution when both sample sizes are odd; Section 3 provides the near-exact distribution when one sample size is even and the other one odd, and Section 4 presents the exact near-exact distribution of the test statistic when both sample sizes are even. Numerical studies are provided in Section 5 to compare the exact and the near-exact distribution approaches with Welch's t test, and concluding remarks are presented in Section 6.

2. The Exact Distribution of the Behrens–Fisher Statistic for Odd-Numbered Sample Sizes

In this section, we present the exact distribution of the Behrens–Fisher statistic in (1), when both n_1 and n_2 are odd. Since $\frac{S_j^2}{n_j} = \frac{\sigma_j^2}{n_j(n_j-1)} \frac{(n_j-1)s_j^2}{\sigma_j^2}$ where $\frac{(n_j-1)s_j^2}{\sigma_j^2} \sim \chi_{n_j-1}^2$ ($j = 1, 2$), $W = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}$ is the sum of two independent gamma variables, each of which follows the distributions $\Gamma\left(\frac{n_1-1}{2}, \frac{n_1(n_1-1)}{2\sigma_1^2}\right)$ and $\Gamma\left(\frac{n_2-1}{2}, \frac{n_2(n_2-1)}{2\sigma_2^2}\right)$, where $\Gamma(r, \lambda)$ indicates a gamma distribution with shape parameter r and rate parameter λ .

Now, we can divide the problem into two cases: one is the case of $\frac{n_1(n_1-1)}{2\sigma_1^2} \neq \frac{n_2(n_2-1)}{2\sigma_2^2}$, and the other is the case $\frac{n_1(n_1-1)}{2\sigma_1^2} = \frac{n_2(n_2-1)}{2\sigma_2^2}$. When $\frac{n_1(n_1-1)}{2\sigma_1^2} \neq \frac{n_2(n_2-1)}{2\sigma_2^2}$, W follows a GIG distribution of depth 2 with shape parameters $r_j = \frac{n_j-1}{2}$ ($j = 1, 2$) and rate parameters $\lambda_j = \frac{n_j(n_j-1)}{2\sigma_j^2}$ ($j = 1, 2$), which are different. Notice that $\frac{n_1-1}{2}$ and $\frac{n_2-1}{2}$ are both integers because of odd-numbered sample sizes. The probability density function for this distribution is

$$f_W(w) = \sum_{j=1}^2 \sum_{k=1}^{r_j} \left(\prod_{j=1}^2 \lambda_j^{r_j} \right) c_{j,k} w^{k-1} e^{-\lambda_j w} I(w > 0)$$

where $c_{j,k}$ are given by (11)–(13) in [8]. In this case,

$$c_{i,r_i} = \frac{1}{(r_i - 1)!} \prod_{j=1, j \neq i}^2 (\lambda_j - \lambda_i)^{-r_j}$$

and

$$c_{i,r_i-k} = \frac{1}{k} \sum_{j=1}^k \frac{(r_i - k + j - 1)!}{(r_i - k - 1)!} \left(\sum_{s=1, s \neq i}^2 r_s (\lambda_i - \lambda_s)^{-j} \right) c_{i,r_i-k+j}$$

for $k = 1, \dots, r_i - 1, i = 1, 2$.

As a matter of fact, this probability density function can be rewritten as

$$f_W(w) = \sum_{j=1}^2 \sum_{k=1}^{r_j} \underbrace{\left(\prod_{j=1}^2 \lambda_j^{r_j} \right) c_{j,k} \frac{\Gamma(k)}{\lambda_j^k}}_{p_{j,k}} \underbrace{\frac{\lambda_j^k}{\Gamma(k)} w^{k-1} e^{-\lambda_j w}}_{\text{probability density function of } \Gamma(k, \lambda_j)} I(w > 0)$$

which is a finite mixture of integer gamma distributions in [9]. Now that we have obtained the exact distribution of W , we can easily get the joint distribution of W and $Y = \bar{X}_1 - \bar{X}_2$, under $H_0 : \mu_1 = \mu_2$. Y follows a normal distribution $N\left(0, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$ under H_0 , and so, given the independence of W and Y , the joint probability density function of these two random variables is

$$f_{W,Y}(w, y) = \sum_{j=1}^2 \sum_{k=1}^{r_j} p_{j,k} \frac{\lambda_j^k}{\Gamma(k)} w^{k-1} e^{-\lambda_j w} \frac{1}{\sigma} \Phi\left(\frac{y}{\sigma}\right) \quad (w > 0, y \in (-\infty, \infty))$$

where $\sigma = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ and Φ is the probability density function of a standard normal distribution.

Since we want to derive the distribution of $T^* = \frac{Y}{\sqrt{W}}$, we need to go further and obtain the joint distribution of T^* and $V = \sqrt{W}$ from $f_{W,Y}(w, y)$ by a simple change of variables. From this process, we obtain

$$f_{T^*,V}(t, v) = \sum_{j=1}^2 \sum_{k=1}^{r_j} p_{j,k} \frac{2\lambda_j^k}{\Gamma(k)} v^{2k} e^{-\lambda_j v^2} \frac{1}{\sigma} \Phi\left(\frac{tv}{\sigma}\right),$$

for $v > 0$ and $t \in (-\infty, \infty)$ and

$$f_{T^*}(t) = \sum_{j=1}^2 \sum_{k=1}^{r_j} p_{j,k} \int_0^\infty 2 \frac{\lambda_j^k}{\Gamma(k)} v^{2k} e^{-\lambda_j v^2} \frac{1}{\sigma} \Phi\left(\frac{tv}{\sigma}\right) dv,$$

where

$$\int_0^\infty 2 \frac{\lambda_j^k}{\Gamma(k)} v^{2k} e^{-\lambda_j v^2} \frac{1}{\sigma} \Phi\left(\frac{tv}{\sigma}\right) dv$$

yields the probability density function of a $\sigma \sqrt{\frac{\lambda_j}{k}} T_{2k}$ random variable, with T_{2k} denoting a Student's t distribution with $2k$ degrees of freedom.

Hence, $f_{T^*}(t)$ is a mixture of probability density functions of $\sigma \sqrt{\frac{\lambda_j}{k}} T_{2k}$ random variables ($k = 1, \dots, r_j \quad j = 1, 2$), with weights $p_{j,k}$. Given that the probability density function of Student's t variable with n degrees of freedom is given by

$$f_{T_n}(t) = \frac{1}{B\left(\frac{n}{2}, \frac{1}{2}\right)} \frac{1}{\sqrt{n}} \left(1 + \frac{t^2}{n}\right)^{-(n+1)/2}$$

for $t \in (-\infty, \infty)$, the probability density function of T^* , under H_0 , can be rewritten as

$$\begin{aligned}
 f_{T^*}(t) &= \sum_{j=1}^2 \sum_{k=1}^{r_j} p_{j,k} f_{\sigma\sqrt{\lambda_j/k}T_{2k}}(t) \\
 &= \left(\prod_{j=1}^2 \lambda_j^{r_j} \right) \frac{1}{\sigma\sqrt{2\pi}} \sum_{j=1}^2 \sum_{k=1}^{r_j} c_{j,k} \Gamma\left(k + \frac{1}{2}\right) \left(\frac{t^2}{2\sigma^2} + \lambda_j\right)^{-(k+\frac{1}{2})}
 \end{aligned}$$

for $t \in (-\infty, \infty)$.

Then, the cumulative distribution function of T^* would also be a mixture of cumulative distribution functions of $\sigma\sqrt{\lambda_j/k}T_{2k}$ ($k = 1, \dots, r_j \quad j = 1, 2$) with weights $p_{j,k}$. For Student's t distribution with even degrees of freedom, the cumulative distribution function is given by

$$\begin{aligned}
 F_{T_{2k}}(t) &= \frac{1}{2} + \frac{t}{B\left(k, \frac{1}{2}\right)} \frac{1}{\sqrt{2k}} {}_2F_1\left(k + \frac{1}{2}, \frac{1}{2}; \frac{3}{2}; -\frac{t^2}{2k}\right) \\
 &= \frac{1}{2} + \frac{t}{B\left(k, \frac{1}{2}\right)} \frac{1}{\sqrt{2k}} {}_2F_1\left(1 - k, 1; \frac{3}{2}; -\frac{t^2}{2k}\right) \\
 &= \frac{1}{2} + \frac{\Gamma\left(k + \frac{1}{2}\right)}{\sqrt{2k}} \frac{t}{2} \left(1 + \frac{t^2}{2k}\right)^{\frac{1}{2}-k} \sum_{i=0}^{k-1} \frac{\left(\frac{t^2}{2k}\right)^i}{\Gamma\left(\frac{3}{2} + i\right)\Gamma(k - i)}
 \end{aligned}$$

for $t \in (-\infty, \infty)$, which is obtained by applying 15.3.3 and 15.4.1 from [10], where ${}_2F_1$ denotes a Gaussian hypergeometric function.

Thus, the cumulative distribution function of T^* , under H_0 , can be expressed as

$$\begin{aligned}
 F_{T^*}(t) &= \sum_{j=1}^2 \sum_{k=1}^{r_j} p_{j,k} F_{T_{2k}}\left(t/\sigma\sqrt{\frac{\lambda_j}{k}}\right) \\
 &= \frac{1}{2} + \sum_{j=1}^2 \sum_{k=1}^{r_j} \frac{\left(\prod_{j=1}^2 \lambda_j^{r_j}\right)}{\sqrt{2\sigma^2}} c_{j,k} \frac{\Gamma(k)\Gamma\left(k + \frac{1}{2}\right)}{\lambda_j} \frac{t}{2} \left(\frac{t^2}{2\sigma^2} + \lambda_j\right)^{\frac{1}{2}-k} \sum_{i=0}^{k-1} \frac{\left(\frac{t^2}{2\sigma^2\lambda_j}\right)^i}{\Gamma\left(\frac{3}{2} + i\right)\Gamma(k - i)}
 \end{aligned}$$

for $t \in (-\infty, \infty)$.

The distribution of T^* is much simpler when $\frac{n_1(n_1-1)}{2\sigma_1^2} = \frac{n_2(n_2-1)}{2\sigma_2^2}$. Since the rate parameters of the two independent gamma variables, $\frac{S_1^2}{n_1}$ and $\frac{S_2^2}{n_2}$, are, in this case, equal, W simply follows the gamma distribution $\Gamma\left(\frac{n_1+n_2-2}{2}, \frac{n_1(n_1-1)}{2\sigma_1^2} = \frac{n_2(n_2-1)}{2\sigma_2^2}\right)$. Therefore, in this case, $T^* \equiv T_{n_1+n_2-2}$. This means that the probability density function and cumulative distribution function of T^* are the same as those of Student's t distribution with $n_1 + n_2 - 2$ degrees of freedom. The probability density function for this distribution is written as

$$f_{T^*}(t) = \frac{1}{B\left(\frac{n_1+n_2-2}{2}, \frac{1}{2}\right)} \frac{1}{\sqrt{n_1 + n_2 - 2}} \left(1 + \frac{t^2}{n_1 + n_2 - 2}\right)^{-(n_1+n_2-1)/2}$$

for $t \in (-\infty, \infty)$.

Additionally, as $n_1 + n_2 - 2$ is an even number when sample sizes are odd-numbered, the cumulative distribution function can be written as

$$F_{T^*}(t) = \frac{1}{2} + \frac{\Gamma\left(\frac{n_1+n_2-1}{2}\right)}{\sqrt{n_1+n_2-2}} \frac{t}{2} \left(1 + \frac{t^2}{n_1+n_2-2}\right)^{\frac{3-n_1-n_2}{2}} \sum_{i=0}^{\frac{n_1+n_2-4}{2}} \frac{\left(\frac{t^2}{n_1+n_2-2}\right)^i}{\Gamma\left(\frac{3}{2}+i\right)\Gamma\left(\frac{n_1+n_2-2}{2}-i\right)}$$

for $t \in (-\infty, \infty)$.

3. The Exact and Near-Exact Distribution of the Behrens–Fisher Statistic for Even-Numbered Sample Sizes

In this section, we present the exact distribution of T^* when sample sizes are both even. The exact distribution consists of an infinite series, but we provide a near-exact distribution based on a finite mixture of GIG distributions, which yields an approximation to the exact distribution.

3.1. The Exact Distribution

Unlike the case where both sample sizes are odd, $W = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}$ does not follow a GIG distribution when both sample sizes are even. This is so because the shape parameters for $\Gamma\left(\frac{n_1-1}{2}, \frac{n_1(n_1-1)}{2\sigma_1^2}\right)$ and $\Gamma\left(\frac{n_2-1}{2}, \frac{n_2(n_2-1)}{2\sigma_2^2}\right)$ are not integers when sample sizes are even. However, we can use the Kummer confluent hypergeometric function to obtain the exact distribution of W for this case.

Given the integral definition of the Kummer confluent hypergeometric function, the probability density function of W , which is the sum of two independent gamma variables with shape parameters

$$r_j = \frac{n_j - 1}{2}, \quad j = 1, 2, \tag{2}$$

and rate parameters

$$\lambda_j = \frac{n_j(n_j - 1)}{2\sigma_j^2}, \quad j = 1, 2, \tag{3}$$

can be expressed as

$$\begin{aligned} f_W(w) &= \frac{\lambda_1^{r_1}}{\Gamma(r_1)} \frac{\lambda_2^{r_2}}{\Gamma(r_2)} e^{-\lambda_2 w} \int_0^w e^{(\lambda_2 - \lambda_1)s} s^{r_1-1} (w-s)^{r_2-1} ds \\ &= \frac{\lambda_1^{r_1} \lambda_2^{r_2}}{\Gamma(r_1 + r_2)} e^{-\lambda_2 w} w^{r_1+r_2-1} {}_1F_1(r_1, r_1 + r_2, (\lambda_2 - \lambda_1)w) \end{aligned}$$

for $w > 0$.

Since

$${}_1F_1(r_1, r_1 + r_2, (\lambda_2 - \lambda_1)w) = \sum_{i=0}^{\infty} \frac{\Gamma(r_1 + i)}{\Gamma(r_1)} \frac{\Gamma(r_1 + r_2)}{\Gamma(r_1 + r_2 + i)} \frac{(\lambda_2 - \lambda_1)^i}{i!} w^i,$$

the probability density function of W can be further written as

$$\begin{aligned} f_W(w) &= \sum_{i=0}^{\infty} \frac{\Gamma(r_1 + i)}{\Gamma(r_1)} \frac{\lambda_1^{r_1} \lambda_2^{r_2}}{\Gamma(r_1 + r_2 + i)} \frac{(\lambda_2 - \lambda_1)^i}{i!} w^{r_1+r_2+i-1} e^{-\lambda_2 w} \\ &= \sum_{i=0}^{\infty} \underbrace{\frac{\Gamma(r_1 + i)}{\Gamma(r_1) i!} \left(\frac{\lambda_1}{\lambda_2}\right)^{r_1} \left(1 - \frac{\lambda_1}{\lambda_2}\right)^i}_{p_i} \underbrace{\frac{\lambda_2^{r_1+r_2+i}}{\Gamma(r_1 + r_2 + i)} w^{r_1+r_2+i-1} e^{-\lambda_2 w}}_{p.d.f \text{ of } \Gamma(r_1+r_2+i, \lambda_2)}, \end{aligned}$$

for $w > 0$, which is the probability density function of an infinite mixture of gamma distributions with weights p_i ($i = 0, 1, \dots$). Now, using a similar approach to the one used in Section 2 for the case of odd-numbered sample sizes, we can obtain the exact probability density function of T^* under H_0 in the form of an infinite mixture of probability density functions of $\sigma\sqrt{\frac{\lambda_2}{r_1+r_2+i}}T_{2(r_1+r_2+i)}$ distributions, with weights p_i , for $i = 0, 1, \dots$. The probability density function of T^* may be then stated as follows:

$$\begin{aligned}
 f_{T^*}(t) &= \sum_{i=0}^{\infty} p_i f_{\sigma\sqrt{\frac{\lambda_2}{r_1+r_2+i}}T_{2(r_1+r_2+i)}}(t) \\
 &= \sum_{i=0}^{\infty} p_i \frac{1}{B\left(r_1+r_2+i, \frac{1}{2}\right)\sqrt{2\sigma^2\lambda_2}} \left(1 + \frac{t^2}{2\sigma^2\lambda_2}\right)^{-(r_1+r_2+i+\frac{1}{2})}
 \end{aligned}$$

for $t \in (-\infty, \infty)$.

Regarding the exact cumulative distribution function of T^* , under H_0 , this cumulative distribution function would also be an infinite mixture of cumulative distribution functions of $\sigma\sqrt{\frac{\lambda_2}{r_1+r_2+i}}T_{2(r_1+r_2+i)}$ distributions, with weights p_i , for $i = 0, 1, \dots$. It can be written as

$$\begin{aligned}
 &F_{T^*}(t) \\
 &= \sum_{i=0}^{\infty} p_i F_{T_{2(r_1+r_2+i)}}\left(t/\sigma\sqrt{\frac{\lambda_2}{r_1+r_2+i}}\right) \tag{4} \\
 &= \frac{1}{2} + \sum_{i=0}^{\infty} p_i \frac{\Gamma\left(r_1+r_2+i+\frac{1}{2}\right)}{\sqrt{2\sigma^2\lambda_2}} \frac{t}{2} \left(1 + \frac{t^2}{2\sigma^2\lambda_2}\right)^{\frac{1}{2}-r_i} \sum_{k=0}^{r_i-1} \frac{\left(\frac{t^2}{2\sigma^2\lambda_2}\right)^k}{\Gamma\left(\frac{3}{2}+k\right)\Gamma(r_1+r_2+i-k)}
 \end{aligned}$$

for $t \in (-\infty, \infty)$, where $r_i = r_1 + r_2 + i$, $p_i = \frac{\Gamma(r_1+i)}{\Gamma(r_1)!} \left(\frac{\lambda_1}{\lambda_2}\right)^{r_1} \left(1 - \frac{\lambda_1}{\lambda_2}\right)^i$, by applying the expression for $F_{T_{2k}}(t)$ obtained in Section 2.

As the exact cumulative distribution function of T^* is expressed as an infinite sum when $\lambda_1 \neq \lambda_2$, it is not much of a manageable cumulative distribution function. In order to obtain numerical values of this cumulative distribution function in a reasonable amount of time, use of an integer upper bound for the summation in i is required. However, the number of terms required in order to obtain a small enough truncation error is often very large. Hence, a near-exact distribution of T^* with a manageable cumulative distribution function needs to be obtained for the case where $\lambda_1 \neq \lambda_2$ so that quantiles and p -values can be computed in a faster and more practical way.

3.2. Near-Exact Distribution

In order to obtain a near-exact distribution for T^* , based on a finite mixture, we will first obtain a near-exact distribution for W and then derive the distribution of T^* from this distribution of W .

Let $r_j^* = \frac{n_j-2}{2}$ ($j = 1, 2$), and let $\lambda_1 \neq \lambda_2$. Then, the exact characteristic function of W is given by

$$\begin{aligned}
 \Phi_W(t) &= \lambda_1^{r_1}(\lambda_1 - \mathbf{i}t)^{-r_1} \lambda_2^{r_2}(\lambda_2 - \mathbf{i}t)^{-r_2} \\
 &= \lambda_1^{r_1^*}(\lambda_1 - \mathbf{i}t)^{-r_1^*} \lambda_2^{r_2^*}(\lambda_2 - \mathbf{i}t)^{-r_2^*} \lambda_1^{0.5}(\lambda_1 - \mathbf{i}t)^{-0.5} \lambda_2^{0.5}(\lambda_2 - \mathbf{i}t)^{-0.5}
 \end{aligned}$$

where \mathbf{i} is the imaginary number with $\mathbf{i}^2 = -1$. Since $\lambda_1^{0.5}(\lambda_1 - \mathbf{i}t)^{-0.5} \lambda_2^{0.5}(\lambda_2 - \mathbf{i}t)^{-0.5}$ is the characteristic function of the sum of two independent gamma variables $\Gamma(0.5, \lambda_1)$ and $\Gamma(0.5, \lambda_2)$, we can make use of the probability density function of W expressed as an infinite mixture of probability density functions of gamma distributions obtained in Section 3.1 to write

$$\lambda_1^{0.5}(\lambda_1 - \mathbf{it})^{-0.5} \lambda_2^{0.5}(\lambda_2 - \mathbf{it})^{-0.5} = \sum_{j=0}^{\infty} p_j \lambda_2^{j+1} (\lambda_2 - \mathbf{it})^{-(j+1)},$$

where $p_j = \frac{\Gamma(0.5+j)}{\Gamma(0.5)j!} \left(\frac{\lambda_1}{\lambda_2}\right)^{0.5} \left(1 - \frac{\lambda_1}{\lambda_2}\right)^j$. This leads to

$$\Phi_W(t) = \lambda_1^{r_1^*} (\lambda_1 - \mathbf{it})^{-r_1^*} \lambda_2^{r_2^*} (\lambda_2 - \mathbf{it})^{-r_2^*} \sum_{j=0}^{\infty} p_j \lambda_2^{j+1} (\lambda_2 - \mathbf{it})^{-(j+1)}.$$

Now, for a given integer m^* , we propose to approximate $\Phi_W(t)$ by

$$\begin{aligned} \Phi_W^*(t) &= \lambda_1^{r_1^*} (\lambda_1 - \mathbf{it})^{-r_1^*} \lambda_2^{r_2^*} (\lambda_2 - \mathbf{it})^{-r_2^*} \sum_{j=0}^{m^*} \pi_j \lambda_2^{j+1} (\lambda_2 - \mathbf{it})^{-(j+1)} \\ &= \sum_{j=0}^{m^*} \pi_j \lambda_1^{r_1^*} (\lambda_1 - \mathbf{it})^{-r_1^*} \lambda_2^{r_2^*+j+1} (\lambda_2 - \mathbf{it})^{-(r_2^*+j+1)} \end{aligned}$$

where $\pi_j (j = 0, \dots, m^* - 1)$ are determined in such a way that satisfies

$$\left. \frac{\partial^h}{\partial t^h} \Phi_W(t) \right|_{t=0} = \left. \frac{\partial^h}{\partial t^h} \Phi_W^*(t) \right|_{t=0} \quad \text{for } h = 1, \dots, m^*,$$

with $\pi_{m^*} = 1 - \sum_{i=0}^{m^*-1} \pi_i$. This approximate characteristic function $\Phi_W^*(t)$ is, in fact, the characteristic function of a finite mixture (with weights π_j) of $m^* + 1$ GIG distributions of depth 2 with integer shape parameters $r_1^*, r_2^* + j + 1 (j = 0, \dots, m^*)$ and rate parameters λ_1 and λ_2 , the first m^* moments of which are the same as those of W .

Hence, a finite mixture of probability density functions of $m^* + 1$ GIG distributions with weights π_j for $j = 0, \dots, m^*$ will then be a near-exact probability density function of W . As the GIG distribution itself is a finite mixture of gamma distributions, as shown in Section 2, this probability density function can be written as

$$\begin{aligned} f_W^*(w) &= \sum_{i=0}^{m^*} \pi_i \sum_{j=1}^2 \sum_{k=1}^{r_{ji}^{**}} \left(\prod_{j=1}^2 \lambda_j^{r_{ji}^{**}} \right) c_{jk,i}^* w^{k-1} e^{-\lambda_j w} \\ &= \sum_{i=0}^{m^*} \pi_i \sum_{j=1}^2 \sum_{k=1}^{r_{ji}^{**}} \underbrace{\left(\prod_{j=1}^2 \lambda_j^{r_{ji}^{**}} \right) c_{jk,i}^*}_{p_{jk,i}^*} \underbrace{\frac{\Gamma(k)}{\lambda_j^k} \frac{\lambda_j^k}{\Gamma(k)} w^{k-1} e^{-\lambda_j w}}_{\text{p.d.f of } \Gamma(k, \lambda_j)} \end{aligned}$$

for $w > 0$, where

$$r_{ji}^{**} = \begin{cases} r_1^* & j = 1 \\ r_2^* + i + 1 & j = 2 \end{cases}$$

for $i = 0, \dots, m^*$ and $c_{jk,i}^*$ are defined in the same way as $c_{j,k}$ in Section 2, except here we use r_{ji}^{**} instead of r_j .

From this near-exact probability density function of W , using the same logic as before just like in Sections 2 and 3.1, we can obtain a near-exact probability density function of T^* in the form of a finite mixture of probability density functions of $\sigma \sqrt{\frac{\lambda_j}{k}} T_{2k}$ with weights $\pi_i p_{jk,i}^*$. Thus, the near-exact probability density function under H_0 is given as follows:

$$\begin{aligned}
 f_{T^*}^*(t) &= \sum_{i=0}^{m^*} \sum_{j=1}^2 \sum_{k=1}^{r_{ji}^{**}} \pi_i p_{jk,i}^* f_{\sigma \sqrt{\frac{\lambda_j}{k}} T_{2k}}(t) \\
 &= \sum_{i=0}^{m^*} \sum_{j=1}^2 \sum_{k=1}^{r_{ji}^{**}} \pi_i \left(\prod_{j=1}^2 \lambda_j^{r_{ji}^{**}} \right) c_{jk,i}^* \frac{\Gamma(k+0.5)}{\sqrt{2\pi\sigma^2}} \left(\lambda_j + \frac{t^2}{2\sigma^2} \right)^{-(k+0.5)}
 \end{aligned}$$

for $t \in (-\infty, \infty)$. Naturally, the corresponding near-exact cumulative distribution function of T^* would be a finite mixture of cumulative distribution functions of $\sigma \sqrt{\frac{\lambda_j}{k}} T_{2k}$ with weights $\pi_i p_{jk,i}^*$, which can be written as

$$\begin{aligned}
 F_{T^*}^*(t) &= \frac{1}{2} + \sum_{i=0}^{m^*} \sum_{j=1}^2 \sum_{k=1}^{r_{ji}^{**}} \left\{ \pi_i \left(\prod_{j=1}^2 \lambda_j^{r_{ji}^{**}} \right) c_{jk,i}^* \frac{\Gamma(k)\Gamma(k+0.5)}{\sqrt{2\sigma^2}} \frac{t}{2\lambda_j} \left(\lambda_j + \frac{t^2}{2\sigma^2} \right)^{0.5-k} \right. \\
 &\quad \left. \times \sum_{s=0}^{k-1} \frac{\left(\frac{t^2}{2\sigma^2 \lambda_j} \right)^s}{\Gamma\left(\frac{3}{2}+s\right)\Gamma(k-s)} \right\} \tag{5}
 \end{aligned}$$

for $t \in (-\infty, \infty)$, by applying $F_{T_{2k}}(t)$ obtained in Section 2.

4. One of the Sample Sizes Is Even and the Other Is Odd

The exact distribution of the Behrens–Fisher statistic T^* for this case is given by the same expressions used in Section 3.1. Just like the case of even-numbered sample sizes, the exact cumulative distribution function is not that manageable for usage when $\lambda_1 \neq \lambda_2$ in this case, too. Hence there is a need to obtain a near-exact distribution of T^* with a manageable cumulative distribution function for faster and more practical computation of quantiles and p -values.

As we did in Section 3.2, we will first obtain a near-exact distribution for W , and then we will derive the near-exact distribution for T^* from the distribution of W . Let, without any loss of generality, n_1 be even and n_2 be odd. Additionally, let $r_1^* = \frac{n_1-2}{2}$, which is equivalent to $[r_1]$, and let $\lambda_1 \neq \lambda_2$. Then, for r_j and λ_j as defined in (2) and (3), the exact characteristic function of W can be written as

$$\begin{aligned}
 \Phi_W(t) &= \lambda_1^{r_1} (\lambda_1 - \mathbf{it})^{-r_1} \lambda_2^{r_2} (\lambda_2 - \mathbf{it})^{-r_2} \\
 &= \underbrace{\lambda_1^{r_1} (\lambda_1 - \mathbf{it})^{-r_1} \lambda_2^{r_2} (\lambda_2 - \mathbf{it})^{-r_2}}_{\Phi_{W_1}(t)} \underbrace{\lambda_1^{\frac{1}{2}} (\lambda_1 - \mathbf{it})^{-\frac{1}{2}}}_{\Phi_{W_2}(t)},
 \end{aligned}$$

where $\Phi_{W_1}(t)$ is the characteristic function of a GIG distribution of depth 2 with integer shape parameters r_1^* and r_2 and rate parameters λ_1 and λ_2 , while $\Phi_{W_2}(t)$ is the characteristic function of a $\Gamma(0.5, \lambda_1)$ distribution. Now, for a given $m^* \in \mathbb{N}$, we propose to approximate $\Phi_{W_2}(t)$ by

$$\Phi_{W_2}^*(t) = \sum_{j=0}^{m^*} \pi_j (2\lambda_1)^{(j+1)} (2\lambda_1 - \mathbf{it})^{-(j+1)}$$

where $\pi_j (j = 0, \dots, m^* - 1)$ are determined in such a way that

$$\left. \frac{\partial^h}{\partial t^h} \Phi_{W_2}(t) \right|_{t=0} = \left. \frac{\partial^h}{\partial t^h} \Phi_{W_2}^*(t) \right|_{t=0} \quad \text{for } h = 1, \dots, m^*$$

with $\pi_{m^*} = 1 - \sum_{i=0}^{m^*-1} \pi_i$.

The characteristic function $\Phi_{W_2}^*(t)$ is the characteristic function of a finite mixture of $m^* + 1$ distributions, which are $\Gamma(j + 1, 2\lambda_1)$ distributions ($j = 0, \dots, m^*$). The first m^* moments of this mixture are the same as those of $\Gamma(0.5, \lambda_1)$.

We should note that the weights π_j do not depend on λ_1 . The h -th non-central moment of $\Gamma(0.5, \lambda_1)$ and of the mixture of $\Gamma(j + 1, 2\lambda_1)$ distributions are, respectively, given by

$$\frac{\Gamma\left(h + \frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\lambda_1^h} \quad \text{and} \quad \sum_{j=0}^{m^*} \pi_j \frac{\Gamma(h + j + 1)}{\Gamma(j + 1)(2\lambda_1)^h},$$

which means π_j are determined in such a way that satisfies

$$\frac{\Gamma\left(h + \frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)} = \sum_{j=0}^{m^*} \pi_j \frac{\Gamma(h + j + 1)}{\Gamma(j + 1)2^h},$$

for $h = 1, \dots, m^*$, definitely not depending on λ_1 .

By using $\Phi_{W_2}^*(t)$ instead of $\Phi_{W_2}(t)$, we can then approximate $\Phi_W(t)$ with

$$\begin{aligned} \Phi_W^*(t) &= \lambda_1^{r_1^*} (\lambda_1 - \mathbf{it})^{-r_1^*} \lambda_2^{r_2^*} (\lambda_2 - \mathbf{it})^{-r_2^*} \sum_{j=0}^{m^*} \pi_j (2\lambda_1)^{(j+1)} (2\lambda_1 - \mathbf{it})^{-(j+1)} \\ &= \sum_{j=0}^{m^*} \pi_j \lambda_1^{r_1^*} (\lambda_1 - \mathbf{it})^{-r_1^*} \lambda_2^{r_2^*} (\lambda_2 - \mathbf{it})^{-r_2^*} (2\lambda_1)^{(j+1)} (2\lambda_1 - \mathbf{it})^{-(j+1)} \end{aligned}$$

which is the characteristic function of a finite mixture of $m^* + 1$ GIG distributions of depth 3 with shape parameters $r_1^*, r_2, j + 1$ ($j = 0, \dots, m^*$) and rate parameters $\lambda_1, \lambda_2, 2\lambda_1$.

Thus, a finite mixture of $m^* + 1$ probability density functions s of GIG distributions of depth 3 with weights π_j can be a near-exact probability density function of W . As GIG distribution itself is a finite mixture of gamma distributions, this near-exact probability density function of W is written as

$$\begin{aligned} f_W^*(w) &= \sum_{i=0}^{m^*} \pi_i \sum_{j=1}^3 \sum_{k=1}^{r_{ji}^*} \left(\prod_{j=1}^3 (\lambda_j^*)^{r_{ji}^*} \right) c_{jk,i}^* w^{k-1} e^{-\lambda_j^* w} \\ &= \sum_{i=0}^{m^*} \pi_i \sum_{j=1}^3 \sum_{k=1}^{r_{ji}^*} \underbrace{\left(\prod_{j=1}^3 (\lambda_j^*)^{r_{ji}^*} \right) c_{jk,i}^*}_{p_{jk,i}^*} \underbrace{\frac{\Gamma(k)}{(\lambda_j^*)^k} \frac{(\lambda_j^*)^k}{\Gamma(k)} w^{k-1} e^{-\lambda_j^* w}}_{\text{p.d.f of } \Gamma(k, \lambda_j^*)} \quad (w > 0) \end{aligned}$$

where

$$r_{ji}^* = \begin{cases} r_1^* & j = 1 \\ r_2 & j = 2 \\ i + 1 & j = 3 \end{cases} \quad (i = 0, \dots, m^*), \quad \lambda_j^* = \begin{cases} \lambda_1 & j = 1 \\ \lambda_2 & j = 2 \\ 2\lambda_1 & j = 3 \end{cases}$$

and $c_{jk,i}^*$ are given by (11)–(13) in [8] using r_{ji}^* and λ_j^* , which are

$$\begin{aligned} c_{j,r_{ji}^*,i}^* &= \frac{1}{(r_{ji}^* - 1)!} \prod_{s=1, s \neq j}^3 (\lambda_s^* - \lambda_j^*)^{-r_{si}^*}, \\ c_{j,r_{ji}^*-k,i}^* &= \frac{1}{k} \sum_{n=1}^k \frac{(r_{ji}^* - k + n - 1)!}{(r_{ji}^* - k - 1)!} \left(\sum_{s=1, s \neq j}^3 r_s (\lambda_j^* - \lambda_s^*)^{-n} \right) c_{j,r_{ji}^*-k+n,i}^* \end{aligned}$$

where $k = 1, \dots, r_{ji}^* - 1$ and $j = 1, 2, 3, i = 0, \dots, m^*$. From this near-exact probability density function of W , using the same logic as before, we can once again obtain a near-exact probability density function of T^* in the form of a finite mixture of probability density functions of $\sigma\sqrt{\frac{\lambda_j^*}{k}}T_{2k}$ distributions, with weights $\pi_i p_{jk,i}^*$. The near-exact probability density function of T^* , under H_0 , is thus given by

$$\begin{aligned}
 f_{T^*}^*(t) &= \sum_{i=0}^{m^*} \sum_{j=1}^3 \sum_{k=1}^{r_{ji}^*} \pi_i p_{jk,i}^* f_{\sigma\sqrt{\frac{\lambda_j^*}{k}}T_{2k}}(t) \\
 &= \sum_{i=0}^{m^*} \sum_{j=1}^3 \sum_{k=1}^{r_{ji}^*} \pi_i \left(\prod_{j=1}^3 (\lambda_j^*)^{r_{ji}^*} \right) c_{jk,i}^* \frac{\Gamma(k + \frac{1}{2})}{\sqrt{2\pi\sigma^2}} \left(\lambda_j^* + \frac{t^2}{2\sigma^2} \right)^{-(k+\frac{1}{2})}
 \end{aligned}$$

for $t \in (-\infty, \infty)$.

Hence, the corresponding near-exact cumulative distribution function of T^* , under H_0 , is also a finite mixture of cumulative distribution functions of $\sigma\sqrt{\frac{\lambda_j^*}{k}}T_{2k}$, with weights $\pi_i p_{jk,i}^*$, which can be written as

$$\begin{aligned}
 F_{T^*}^*(t) &= \frac{1}{2} + \sum_{i=0}^{m^*} \sum_{j=1}^3 \sum_{k=1}^{r_{ji}^*} \left\{ \pi_i \left(\prod_{j=1}^3 (\lambda_j^*)^{r_{ji}^*} \right) c_{jk,i}^* \frac{\Gamma(k)\Gamma(k+0.5)}{\sqrt{2\sigma^2}} \frac{t}{2\lambda_j^*} \left(\lambda_j^* + \frac{t^2}{2\sigma^2} \right)^{0.5-k} \right. \\
 &\quad \left. \times \sum_{s=0}^{k-1} \frac{\left(\frac{t^2}{2\sigma^2\lambda_j^*} \right)^s}{\Gamma(\frac{3}{2}+s)\Gamma(k-s)} \right\}
 \end{aligned} \tag{6}$$

for $t \in (-\infty, \infty)$ by applying $F_{T_{2k}}(t)$ obtained in Section 2.

5. Comparison of the Exact or Near-Exact Distribution and Welch’s t Test

When it is plausible to assume that $\lambda_1 \neq \lambda_2$, that is, that $\frac{n_1(n_1-1)}{2\sigma_1^2} \neq \frac{n_2(n_2-1)}{2\sigma_2^2}$, we can make use of the exact and near-exact distribution of T^* to solve the Behrens–Fisher problem. The exact cumulative distribution function of T^* will be used for computation of p -values when both sample sizes are odd, while the near-exact cumulative distribution functions obtained in Sections 3 and 4 will be used for computing p -values when both sample sizes are even or when one of the sample sizes is even and the other is odd. Because the cumulative distribution functions of T^* include the unknown parameters σ_1^2 and σ_2^2 , these will be estimated by the sample variances S_i^2 ($i = 1, 2$).

We will compare the exact or near-exact distributions and Welch’s t -test by their actual sizes and powers for testing $H_0 : \mu_1 = \mu_2$ versus $H_1 : \mu_1 > \mu_2$. For $T^* = t$, since the hypothesis is for the right-tailed test, the corresponding p -value is computed from $1 - F_{T^*}(t)$ in (5) or $1 - F_{T^*}^*(t)$ in (5) or (6), depending on the parity of the sample sizes, according to the derivations in the previous sections, and with σ estimated by $\hat{\sigma} = \sqrt{S_1^2/n_1 + S_2^2/n_2}$. We used different type I error rates $\alpha \in \{0.1, 0.05, 0.01\}$ and we conducted Monte Carlo experiments under a range of different sets of parameters.

Simulations were conducted for variance and sample size pairs that correspond to

$$\theta = \frac{\sigma_1^2}{n_1} \bigg/ \left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \right) \in \{0.1, 0.3, 0.5\}.$$

For $\mu_1 - \mu_2$, we covered the cases where $\mu_1 - \mu_2$ satisfies

$$\delta = \frac{\mu_1 - \mu_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \in \{0, 1, 2\},$$

with $\delta = 0$ corresponding to the null hypothesis $H_0 : \mu_1 = \mu_2$.

For each combination of parameters, the number of replications was 50,000, and in each subsection, we provide two scenarios: one in which sample sizes were balanced, and another one where they were unbalanced. For each generated sample, we computed t_i for $i = 1, \dots, 50,000$ and then obtained the type I error under $\delta = 0$ and the power under $\delta > 0$ from

$$\frac{1}{50,000} \sum_{i=1}^{50,000} I(p_i \leq \alpha) \tag{7}$$

where $I(\cdot)$ is the indicator function and the p -value $p_i = 1 - F_{T^*}(t_i)$ is computed using $F_{T^*}(\cdot)$ in (5) for the case of both odd sample sizes or $F_{T^*}(\cdot)$ in (5) or (6) when at least one of the sample sizes is even.

All computations were done with the software R, version 4.1.0.

5.1. Odd n_1 and n_2

In Tables 1 and 2, we present the power values for the exact distribution and Welch’s test, represented, respectively, by E and W. We considered for θ and δ the sets of values indicated above.

We may see that when sample sizes are unbalanced, as in Table 2, the exact distribution gives larger values of power than Welch’s t -test, namely for larger values of δ and smaller values of α , while they tend to give similar results when the two sample sizes are homogeneous, still with the exact distribution giving larger power values when the variances are unbalanced. Namely, for the case of $(n_1, n_2) = (15, 3)$ and $(\sigma_1^2, \sigma_2^2) = (15, 27)$ and $\delta = 2$, the near-exact distribution shows a gain of over 30% in power in relation to Welch’s t -test for $\alpha = 0.01$.

5.2. Even n_1 and n_2

Tables 3 and 4 provide numerical results for type I error rates and powers for Welch’s t -test and near-exact distributions that use $m^* = 4$ for cases where both sample sizes are even. In these tables, the near-exact distributions and Welch’s test are, respectively, denoted by NE and W.

Table 1. Power values for $(n_1, n_2) = (5, 5)$.

θ	α	$\delta = 0$		$\delta = 1$		$\delta = 2$	
		E	W	E	W	E	W
0.1	0.1	0.1009	0.1000	0.3578	0.3558	0.7059	0.7036
	$\sigma_1^2 = 5$	0.0518	0.0506	0.2234	0.2199	0.5368	0.5313
	$\sigma_2^2 = 45$	0.0123	0.0114	0.0731	0.0676	0.2291	0.2139
0.3	0.1	0.0971	0.0965	0.3664	0.3648	0.7205	0.7187
	$\sigma_1^2 = 15$	0.0480	0.0473	0.2271	0.2241	0.5528	0.5482
	$\sigma_2^2 = 35$	0.0100	0.0092	0.0636	0.0601	0.2376	0.2263
0.5	0.1	0.0958	0.0953	0.3628	0.3614	0.7221	0.7205
	$\sigma_1^2 = 5$	0.0462	0.0456	0.2243	0.2218	0.5582	0.5545
	$\sigma_2^2 = 5$	0.0094	0.0005	0.0641	0.0606	0.2375	0.2288

Table 2. Power values for $(n_1, n_2) = (15, 3)$.

θ	α	$\delta = 0$		$\delta = 1$		$\delta = 2$	
		E	W	E	W	E	W
0.1	0.1	0.1100	0.1048	0.3542	0.3402	0.6614	0.6441
$\sigma_1^2 = 15$	0.05	0.0660	0.0595	0.2304	0.2090	0.4893	0.4522
$\sigma_2^2 = 27$	0.01	0.0273	0.0053	0.1076	0.0843	0.2609	0.1987
0.3	0.1	0.1052	0.1013	0.3636	0.3538	0.7033	0.6900
$\sigma_1^2 = 45$	0.05	0.0593	0.0557	0.2434	0.2287	0.5414	0.5133
$\sigma_2^2 = 21$	0.01	0.0180	0.0161	0.0990	0.0872	0.2940	0.2533
0.5	0.1	0.1001	0.0976	0.3743	0.3679	0.7263	0.7185
$\sigma_1^2 = 15$	0.05	0.0532	0.0507	0.2390	0.2301	0.5657	0.5488
$\sigma_2^2 = 3$	0.01	0.0133	0.0119	0.0866	0.0788	0.2949	0.2687

Similar to the case of odd sample sizes, we see that the differences in power between the near-exact distribution and Welch’s *t*-test are quite slim when the sample sizes are equal, as shown in Table 3, although with a tendency for the near-exact distributions to exhibit larger powers when variances are unbalanced, while in the unbalanced case $(n_1, n_2) = (12, 4)$ shown in Table 4 the power displayed by the near-exact distributions is quite larger, particularly if the variances are also unbalanced. Namely, for the case $(\sigma_1^2, \sigma_2^2) = (8, 72)$ and $\delta = 2$, the near-exact distribution shows a gain of over 20% in power.

Table 3. Power values for $(n_1, n_2) = (8, 8)$.

θ	α	$\delta = 0$		$\delta = 1$		$\delta = 2$	
		NE	W	NE	W	NE	W
0.1	0.1	0.1002	0.0998	0.3731	0.3722	0.7319	0.7311
$\sigma_1^2 = 8$	0.05	0.0525	0.0519	0.2384	0.2371	0.5775	0.5750
$\sigma_2^2 = 72$	0.01	0.0111	0.0106	0.0755	0.0726	0.2747	0.2672
0.3	0.1	0.1001	0.0992	0.3746	0.3741	0.7392	0.7395
$\sigma_1^2 = 24$	0.05	0.0497	0.0488	0.2373	0.2360	0.5927	0.5922
$\sigma_2^2 = 56$	0.01	0.0105	0.0094	0.0760	0.0744	0.2929	0.2879
0.5	0.1	0.0981	0.0968	0.3775	0.3772	0.7370	0.7385
$\sigma_1^2 = 8$	0.05	0.0511	0.0495	0.2403	0.2393	0.5942	0.5947
$\sigma_2^2 = 8$	0.01	0.0103	0.0087	0.0757	0.0739	0.2911	0.2890

Table 4. Power values for $(n_1, n_2) = (12, 4)$.

θ	α	$\delta = 0$		$\delta = 1$		$\delta = 2$	
		NE	W	NE	W	NE	W
0.1	0.1	0.1061	0.1020	0.3554	0.3458	0.6925	0.6823
$\sigma_1^2 = 12$	0.05	0.0578	0.0539	0.2293	0.2129	0.5233	0.4974
$\sigma_2^2 = 36$	0.01	0.0175	0.0148	0.0894	0.0725	0.2568	0.2082
0.3	0.1	0.1031	0.1009	0.3666	0.3612	0.7186	0.7120
$\sigma_1^2 = 36$	0.05	0.0545	0.0523	0.2361	0.2266	0.5657	0.5492
$\sigma_2^2 = 28$	0.01	0.0131	0.0117	0.0864	0.0769	0.2781	0.2498
0.5	0.1	0.0991	0.0976	0.3730	0.3701	0.7332	0.7308
$\sigma_1^2 = 12$	0.05	0.0511	0.0492	0.2388	0.2340	0.5802	0.5718
$\sigma_2^2 = 4$	0.01	0.0117	0.0103	0.0785	0.0732	0.2841	0.2683

5.3. n_1 Is Even and n_2 Is Odd

Table 5 displays the case of similar sample sizes, and it shows that once again, the near-exact distribution and Welch’s test are fairly similar to each other in terms of Type I error rates and values of power. On the other hand, Table 6 presents the results for the unbalanced sample size case, and it shows that the near-exact distribution and Welch’s test can control well the type I error rates, but that the near-exact distribution can obtain larger powers than Welch’s test. In particular, when $(\sigma_1^2, \sigma_2^2) = (12, 27)$ and $\delta = 2$, the near-exact distribution shows a gain of almost 30% in power in relation to Welch’s t -test for $\alpha = 0.01$. The near-exact distributions and Welch’s test are, respectively, denoted by NE and W in these tables.

Table 5. Power values for $(n_1, n_2) = (8, 7)$.

θ	α	$\delta = 0$		$\delta = 1$		$\delta = 2$	
		NE	W	NE	W	NE	W
0.1	0.1	0.1015	0.1012	0.3684	0.3674	0.7279	0.7269
$\sigma_1^2 = 8$	0.05	0.0504	0.0497	0.2377	0.2354	0.5669	0.5638
$\sigma_2^2 = 63$	0.01	0.0117	0.0111	0.0744	0.0712	0.2642	0.2549
0.3	0.1	0.0980	0.0971	0.3736	0.3730	0.7366	0.7367
$\sigma_1^2 = 24$	0.05	0.0496	0.0486	0.2354	0.2341	0.5874	0.5857
$\sigma_2^2 = 49$	0.01	0.0115	0.0104	0.0747	0.0722	0.2876	0.2820
0.5	0.1	0.1015	0.0995	0.3711	0.3706	0.7389	0.7405
$\sigma_1^2 = 8$	0.05	0.0494	0.0475	0.2410	0.2393	0.5897	0.5896
$\sigma_2^2 = 7$	0.01	0.0116	0.0089	0.0737	0.0701	0.2844	0.2806

Table 6. Power values for $(n_1, n_2) = (12, 3)$.

θ	α	$\delta = 0$		$\delta = 1$		$\delta = 2$	
		NE	W	NE	W	NE	W
0.1	0.1	0.1099	0.1045	0.3495	0.3355	0.6590	0.6418
$\sigma_1^2 = 12$	0.05	0.0632	0.0574	0.2311	0.2099	0.4807	0.4455
$\sigma_2^2 = 27$	0.01	0.0254	0.0206	0.1066	0.0849	0.2586	0.1994
0.3	0.1	0.1074	0.1035	0.3632	0.3523	0.6966	0.6817
$\sigma_1^2 = 36$	0.05	0.0582	0.0546	0.2417	0.2261	0.5417	0.5120
$\sigma_2^2 = 21$	0.01	0.0157	0.0137	0.0929	0.0808	0.2825	0.2428
0.5	0.1	0.1006	0.0986	0.3671	0.3611	0.7185	0.7102
$\sigma_1^2 = 12$	0.05	0.0503	0.0486	0.2351	0.2263	0.5664	0.5493
$\sigma_2^2 = 3$	0.01	0.0115	0.01058	0.0817	0.0755	0.2826	0.2581

5.4. Brief Study of Power Evolution for Increasing Sample Sizes

With the aim of showing the evolution of the values of power for increasing sample sizes, the plots of power curves for increasing values of the sample sizes are shown in Figures 1–3, for given $\sigma_1^2, \sigma_2^2, \mu_1$ and μ_2 . As expected, it is clear that the power curves are increasing with increasing values of the sample sizes. Figures 1, 2 and 3 represent, respectively, cases of $(n_1, n_2) = (odd, odd), (even, even)$ and $(even, odd)$. It is also clear from these Figures that, also as expected, although the use of the exact or near-exact distributions lead to an increase in the values of power, the power values from Welch’s test approach asymptotically those obtained when using the exact or near-exact distributions for increasing sample sizes. In Figures 1–3, we use $(\mu_1, \mu_2) = (\sqrt{40}, 0)$ and present the values of (σ_1^2, σ_2^2) and (n_1, n_2) in each figure. In addition, the corresponding values of δ are also presented.

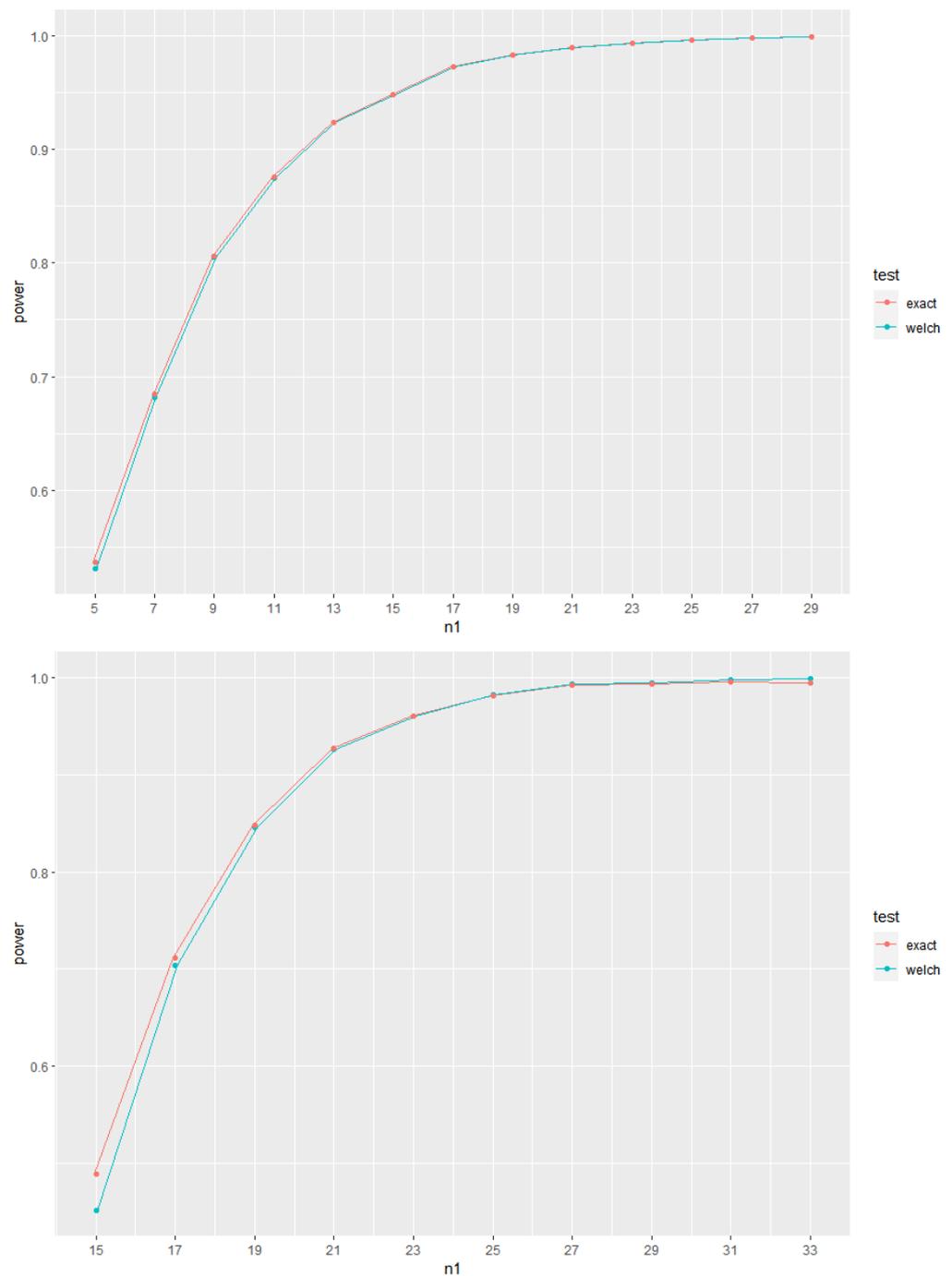


Figure 1. The left panel is based on $\sigma_1^2 = 5, \sigma_2^2 = 45$. $(n_1, n_2) = (5 + 2k, 5 + 2k)$ for $k = 0, 1, \dots, 12$ and $\delta_k = (\mu_1 - \mu_2) / \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{40} / \sqrt{\frac{5}{5+2k} + \frac{45}{5+2k}}$. The right panel is based on $\sigma_1^2 = 15, \sigma_2^2 = 27$ and $(n_1, n_2) = (15 + 2k, 3 + 2k)$ for $k = 0, 1, \dots, 9$ and $\delta_k = (\mu_1 - \mu_2) / \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{40} / \sqrt{\frac{15}{15+2k} + \frac{27}{3+2k}}$. Red and blue lines represent the power curves from exact distribution and Welch’s test, respectively.

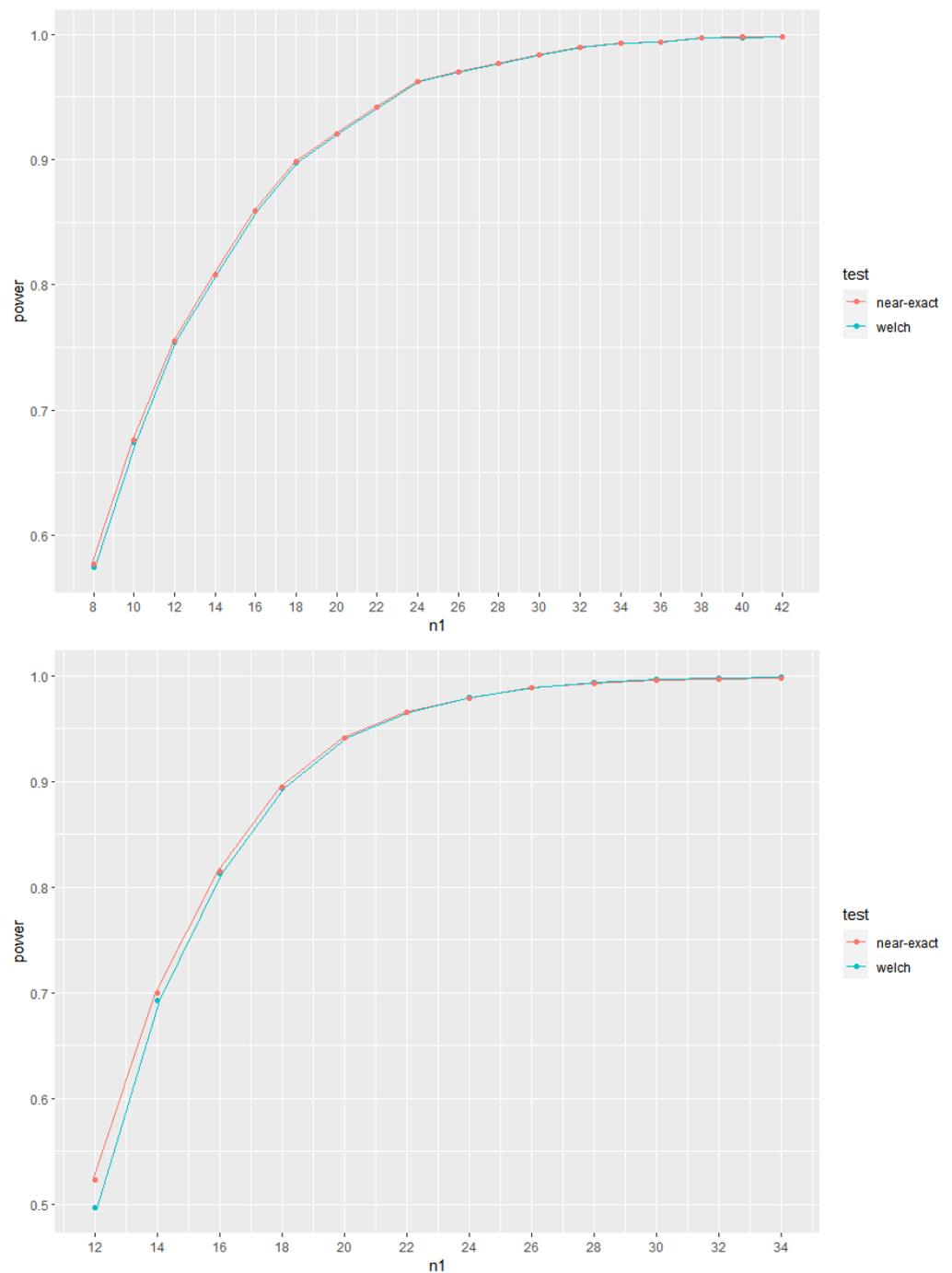


Figure 2. The left panel is based on $\sigma_1^2 = 8, \sigma_2^2 = 72$. $(n_1, n_2) = (8 + 2k, 8 + 2k)$ for $k = 0, 1, \dots, 17$ and $\delta_k = (\mu_1 - \mu_2) / \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{40} / \sqrt{\frac{8}{8+2k} + \frac{72}{8+2k}}$. The right panel is based on $\sigma_1^2 = 12, \sigma_2^2 = 36$ and $(n_1, n_2) = (12 + 2k, 4 + 2k)$ for $k = 0, 1, \dots, 11$ and $\delta_k = (\mu_1 - \mu_2) / \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{40} / \sqrt{\frac{12}{12+2k} + \frac{36}{4+2k}}$. Red and blue lines represent the power curves from the nearly-exact distribution and Welch’s test, respectively.

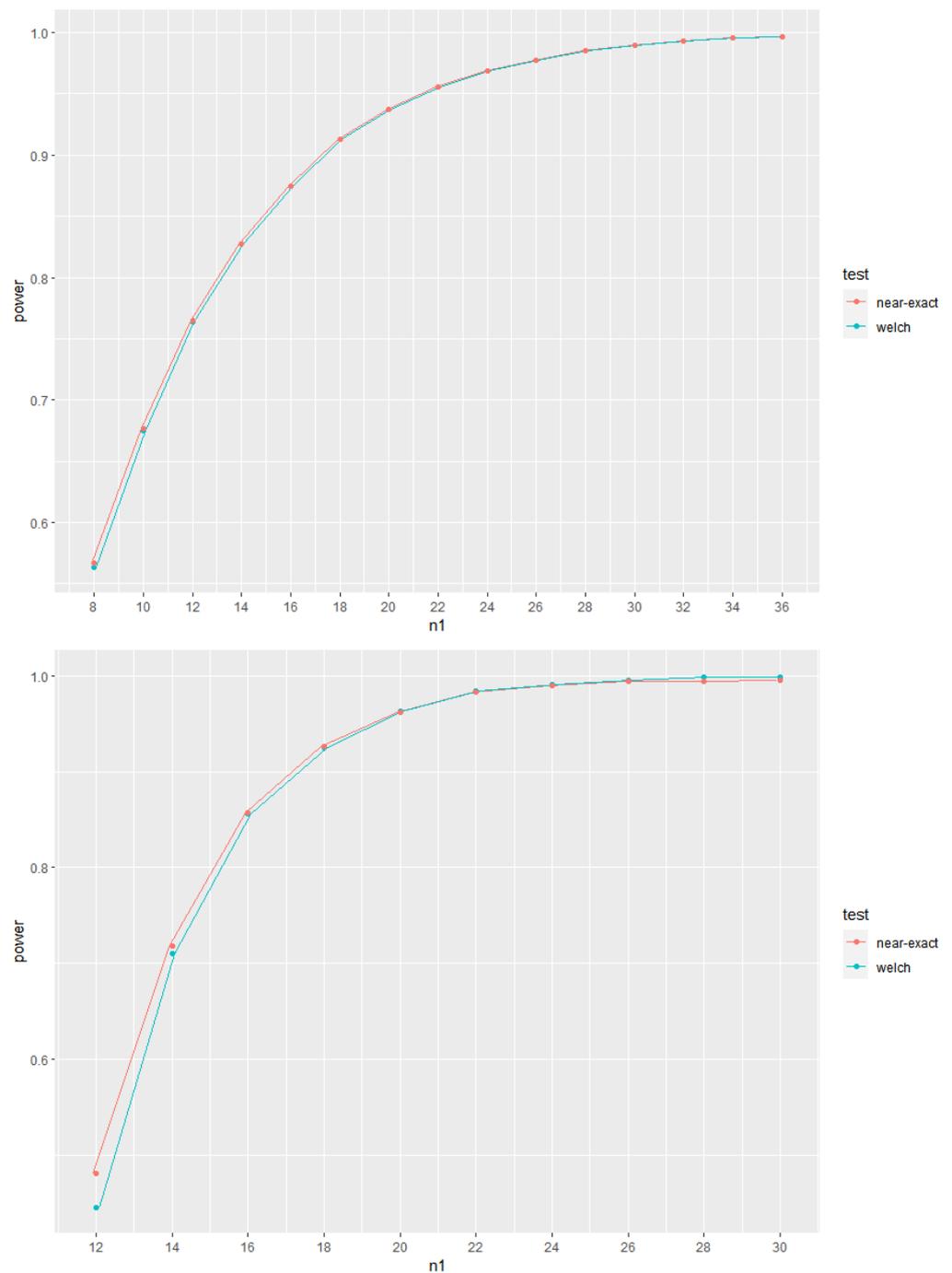


Figure 3. The left panel is based on $\sigma_1^2 = 8, \sigma_2^2 = 63$. $(n_1, n_2) = (8 + 2k, 7 + 2k)$ for $k = 0, 1, \dots, 14$ and $\delta_k = (\mu_1 - \mu_2) / \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{40} / \sqrt{\frac{8}{8+2k} + \frac{63}{7+2k}}$. The right panel is based on $\sigma_1^2 = 12, \sigma_2^2 = 27$ and $(n_1, n_2) = (12 + 2k, 3 + 2k)$ for $k = 0, 1, \dots, 9$ and $\delta_k = (\mu_1 - \mu_2) / \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{40} / \sqrt{\frac{12}{12+2k} + \frac{27}{3+2k}}$. Red and blue lines represent the power curves from the nearly exact distribution and Welch’s test, respectively.

6. Conclusions

Over the years since the Behrens–Fisher problem was first introduced, many different solutions have been presented for the problem. In this paper, we propose another approach for the Behrens–Fisher problem that is based on a version of the exact distribution and near-exact distributions for its statistic, which are based on GIG (generalized integer gamma) distributions. Overall, the differences between the sizes of the near-exact distribution

and Welch's t -test are negligible, while the use of the exact or near-exact distributions provide powers that are larger than those provided by Welch's t -test, mainly for the cases where the sample sizes and/or the variances are unbalanced, and namely when smaller sample sizes are associated with larger variances. The results thus show that, mainly for the cases of unbalanced sample sizes and/or unbalanced variances, the use of Welch's t -test leads to some loss in power compared with the use of the exact or near-exact distributions developed, thus advising towards the use of these latter ones.

The computation of the exact or near-exact distributions poses absolutely no problems, even for large sample sizes, with the computation times remaining in the hundredths of a second for sample sizes in the order of a few hundreds.

In order to decide which distribution to use, the user may want to test the hypothesis $\lambda_1 = \lambda_2$. We may note that, given the definition of λ_1 and λ_2 in (3), testing that hypothesis is indeed equivalent to testing the hypothesis

$$\frac{\sigma_2}{\sigma_1} \frac{n_1(n_1 - 1)}{n_2(n_2 - 1)} = 1,$$

which may be tested in as much the same way we run a test of equality of two variances based on two independent samples.

Author Contributions: Conceptualization, S.H., C.A.C. and J.P.; methodology, S.H., C.A.C. and J.P.; software, S.H.; validation, S.H., C.A.C. and J.P.; formal analysis, S.H.; investigation, C.A.C. and J.P.; writing—original draft preparation, S.H.; writing—review and editing, C.A.C. and J.P.; funding acquisition, C.A.C. and J.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2020R1A2C1A01100526) and also funded by Portuguese national funds through FCT - Fundação para a Ciência e a Tecnologia, I.P., under the scope of the projects UIDB/00297/2020 and UIDP/00297/2020 (Center for Mathematics and Applications-NOVAMath).

Conflicts of Interest: The authors declare no conflict of interest. The founding institutions had no role in the design of the study, in the collection, analysis, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

References

1. Linnik, J.V. *Statistical Problems with Nuisance Parameters (Scripta Technica, Trans.)*; (Original Work Published 1966); American Mathematical Society: Providence, RI, USA, 1968.
2. Fisher, R.A. The fiducial argument in statistical inference. *Ann. Eugen.* **1935**, *6*, 391–398. [[CrossRef](#)]
3. Fisher, R.A. The comparison of samples with possibly unequal variances. *Ann. Eugen.* **1939**, *9*, 174–180. [[CrossRef](#)]
4. Jeffreys, H. Note on the Behrens-Fisher formula. *Ann. Eugen.* **1940**, *10*, 48–51. [[CrossRef](#)]
5. Jeffreys, H. *Theory of Probability*, 3rd ed.; Oxford University Press: London, UK, 1961.
6. Welch, B.L. The significance of the difference between two means when the population variances are unequal. *Biometrika* **1938**, *29*, 350–362. [[CrossRef](#)]
7. Welch, B.L. The generalization of 'Student's' problem when several different population variances are involved. *Biometrika* **1947**, *34*, 28–35. [[CrossRef](#)] [[PubMed](#)]
8. Coelho, C.A. The Generalized Integer Gamma distribution—A basis for distributions in Multivariate Statistics. *J. Multivar. Anal.* **1998**, *64*, 86–102. [[CrossRef](#)]
9. Coelho, C.A. The wrapped Gamma distribution and wrapped sums and linear combinations of independent Gamma and Laplace distributions. *J. Stat. Theory Pract.* **2007**, *1*, 1–29. [[CrossRef](#)]
10. Abramowitz, M.; Stegun, I.A. *Handbook of Mathematical Functions*, 9th ed.; Dover: New York, NY, USA, 1974.