Hindawi Publishing Corporation
Journal of Probability and Statistics
Volume 2011, Article ID 474826, 16 pages
doi:10.1155/2011/474826

Research Article

Equivariance and Generalized Inference in Two-Sample Location-Scale Families

Sévérien Nkurunziza and Fuqi Chen

Department of Mathematics and Statistics, University of Windsor, 401 Sunset Avenue, Windsor, ON, Canada N9B 3P4

Correspondence should be addressed to Sévérien Nkurunziza, severien@uwindsor.ca

Received 17 May 2011; Revised 15 August 2011; Accepted 15 August 2011

Academic Editor: Ricardas Zitikis

We are interested in-typical Behrens-Fisher problem in general location-scale families. We present a method of constructing generalized pivotal quantity (GPQ) and generalized P value (GPV) for the difference between two location parameters. The suggested method is based on the minimum risk equivariant estimators (MREs), and thus, it is an extension of the methods based on maximum likelihood estimators and conditional inference, which have been, so far, applied to some specific distributions. The efficiency of the procedure is illustrated by Monte Carlo simulation studies. Finally, we apply the proposed method to two real datasets.

1. Introduction

In statistical problems involving nuisance parameters, the small-sample optimal solution may not be available. For example, for the difference between means of two exponential distributions, or two normal distributions with different variances, small sample-optimal test and confidence intervals do not exist (see, [1]). To overcome this problem, Tsui and Weerahandi [2] introduced the concept of generalized P value (GPV) and generalized test variable (GTV). Further, Weerahandi [1] developed the concept of generalized pivotal quantity (GPQ) and generalized confidence interval (GCI). The GCI and GPV have been revealed to perform well for some small-sample problems where classical procedures are not optimal. For example, Weerahandi [1] applied the GCIs to the difference in two exponential means and two normal means. In addition, Bebu and Mathew [3] developed a generalized pivotal quantity for comparing the means and variances of a bivariate log-normal distribution.

In this paper, we present a method of constructing the GPQ and GTV in two-sample location-scale families. Also, we extend the method in Sprott [4] where the author applied conditional inference to some particular bivariate location-scale families. In the quoted book, the author uses the maximum likelihood estimator (MLE). However, it is well known that the MLE does not exist in some location and scale families. For more details, we refer to Pitman [5], and Gupta and Székely [6] among others.

Our proposed method is based on Pitman estimator that is the minimum risk equivariant estimator (MRE). It is noticed that, when MLE of a location parameter (or scale parameter) exists, it is an equivariant estimator. Indeed, the suggested method is more general, and our simulation studies show that it provides a high coverage probability, high power and preserves the nominal level of the test.

The rest of this paper is organized as follows. In Section 1, we present some background about generalized inference in location and scale family. We also establish in Section 1 the proposed generalized pivotal quantity and generalized test variable in two-sample location-scale family. Section 2 gives the main result of this paper. Namely, in this section, we present the algorithm of the proposed method. In Section 3, we discuss the application of the method in some specific location-scale families. Section 4 presents some simulation studies as well as analysis results of two real datasets. Finally, Section 5 gives discussion and concluding remarks. Details and technical results are outlined in two appendices.

2. Background and Preliminary Results

In this section, we present some concepts of generalized inference for the convenience of the reader. Also, we set up notation which is used in this paper. For more details about the concepts of GPQ, GTV, and GPV, the reader is referred to Tsui and Weerahandi [2], Weerahandi [1], and Krishnamoorthy et al. [7] among others. Let $X_1, \ldots, X_n$ be i.i.d. random variables from the population probability density function (pdf) $f_x(x \mid \boldsymbol{\eta}_1)$. Also, let $Y_1, \ldots, Y_m$ be iid random variables from the population pdf $f_y(y \mid \boldsymbol{\eta}_2)$. We assume that the two random samples $(X_1, \ldots, X_n)$ and $(Y_1, \ldots, Y_m)$ are independent. Also, let $\boldsymbol{\eta} = (\boldsymbol{\eta}_1', \boldsymbol{\eta}_2')'$ be a p-column vector of unknown parameters (with $p \geq 2$). Further, let $\tau(\boldsymbol{\eta})$ be a q-column vector function of $\boldsymbol{\eta}$ with $q \leq p$, and to simplify the notation, let $\tau(\boldsymbol{\eta}) = \boldsymbol{\theta} = (\theta_1, \boldsymbol{\theta}_2')'$ where θ_1 is the parameter of interest and $\boldsymbol{\theta}_2$ is a vector of nuisance parameters. Let $\mathcal{X}$ denote the sample space of possible values of $(X, Y)'$, where $X = X_1, \ldots, X_n$, $Y = Y_1, \ldots, Y_m$, and let Θ denote the parameter space of $\boldsymbol{\theta}$. In addition, we denote $(x, y)((x, y) \in \mathcal{X})$ as an observation from (X, Y). Given this statistical model, two statistical problems about θ_1 are considered.

First, we are interested in deriving confidence interval estimation of θ_1. Second, for a given θ_0, we consider the testing problem

$$H_0 : \theta_1 \geq \theta_0 \text{ versus } H_1 : \theta_1 < \theta_0. \tag{2.1}$$

Definition 2.1. Let $R = R(X, Y, x, y, \boldsymbol{\theta})$ be a function of X, Y, x, y, θ, where $\theta = (\theta_1, \boldsymbol{\theta}_2)$. Then, the function R is said to be a generalized pivotal quantity for θ_1 if

(1) given x, y, the distribution of R is free from unknown parameters;

(2) the observed value, defined as $R_{\text{obs}} = R(x, y, x, y, \boldsymbol{\theta})$, does not depend on the nuisance parameter $\boldsymbol{\theta}_2$.

Definition 2.2. The generalized test variable for θ_1, is defined as a function of $(X, Y, x, y, \boldsymbol{\theta})$, say $T(X, Y, x, y, \boldsymbol{\theta})$, which satisfies the following requirements.

(1) $t = T(x, y, x, y, \boldsymbol{\theta})$ is free of $\boldsymbol{\theta}_2$.

(2) For fixed x, y, $\boldsymbol{\theta}$, the distribution of $T(X, Y, x, y, \boldsymbol{\theta})$ is free of the nuisance parameter $\boldsymbol{\theta}_2$.

(3) For fixed x, y and $\boldsymbol{\theta}_2$, $P[T(X, Y, x, y, \boldsymbol{\theta}) \geq T(x, y, x, y, \boldsymbol{\theta})]$ is stochastically monotone in θ_1.

To make the connection between GPQ and GTV, it is noticed that the GTV can be derived from GPQ $R(X, Y, x, y, \boldsymbol{\theta})$. In fact, if $R(X, Y, x, y, \boldsymbol{\theta})$ is a GPQ for θ_1 then,

$$T_1(X, Y, x, y, \theta) = R(X, Y, x, y, \boldsymbol{\theta}) - R(x, y, x, y, \boldsymbol{\theta}) \tag{2.2}$$

is a GTV. For instance, if $R(x, y, x, y, \boldsymbol{\theta}) = \theta_1$, we have $T_1(X, Y, x, y, \boldsymbol{\theta}) = R(X, Y, x, y, \boldsymbol{\theta}) - \theta_1$. For more details, see Krishnamoorthy et al. [7]. Also, the generalized P value for the testing problem (2.1) is defined as $p = \sup_{H_0} P[T_1(X, Y, x, y, \boldsymbol{\theta}) \geq 0]$. More specifically, for the case where $T_1(X, Y, x, y, \boldsymbol{\theta}) = R(X, Y, x, y, \boldsymbol{\theta}) - \theta_1$, the GPV for the testing problem (2.1) becomes

$$p = \sup_{H_0} P\left[R(X, Y, x, y, \boldsymbol{\theta}) - \theta_0 \geq 0\right] = P\left(R(X, Y, x, y, \boldsymbol{\theta}) \geq \theta_0\right). \tag{2.3}$$

Thus, since the distribution of $R(X, Y, x, y, \boldsymbol{\theta})$ is free of any unknown parameters, the GPV for θ_1 can be obtained from (2.3) by either analytical method or Monte Carlo simulation. We consider the case where $\boldsymbol{\eta} = (\mu_1, \sigma_1, \mu_2, \sigma_2)'$, $\tau(\boldsymbol{\eta}) = \mu_1 - \mu_2$. Thus, we present the GPQ and GTV for the difference between two location parameters $\delta = \mu_1 - \mu_2$. On one hand, we are interested in deriving GCIs for δ. On the other hand, we consider solving the following testing problem:

$$H_0 : \delta \geq \delta_0 \text{ versus } H_1 : \delta < \delta_0. \tag{2.4}$$

Let R_δ denote the GPQ for δ. For the testing problem (2.4), the generalized P value is

$$p_\delta = P(R_\delta \geq \delta_0). \tag{2.5}$$

2.1. Equivariance and Minimum Risk Equivariant Estimators

In this subsection, we give a brief background about the concept of equivariance and minimum risk equivariant estimators in location-scale family. For more details about this concept, we refer to Lehmann and Casella [8, page 171–173], Schervish [9, chapter 6] among others. To set up some notation, let $\mathbf{X} = (X_1, X_2, \ldots, X_n)'$ be a random sample whose joint pdf can be written as $(1/\sigma^n)\prod_{i=1}^{n} g((x_i - \mu)/\sigma)$ where g is a pdf which does not depend on μ and σ. Then, $\mathbf{X}$ is said to be from the location-scale family with location parameter μ and scale parameter σ.

An estimator $\delta_1(\mathbf{X})$ for the scale parameter σ is said to be equivariant if it satisfies

$$\delta_1(b\mathbf{X}+a) = b\delta_1(\mathbf{X}), \quad \forall -\infty < a < \infty,\ b > 0. \tag{2.6}$$

An estimator $\delta_2(\mathbf{X})$ for the location parameter μ is said to be equivariant if it satisfies

$$\delta_2(b\mathbf{X}+a) = b\delta_2(\mathbf{X}) + a, \quad \forall -\infty < a < \infty,\ b > 0. \tag{2.7}$$

Also, let $\delta(\mathbf{X})$ be equivariant estimator for the scale (or location) parameter θ and let $R(\delta(\mathbf{X}),\theta)$ be its risk function, that is, the expected value of a certain loss function which is invariant under the scale (or location) transformation. Then, the estimator $\delta(\mathbf{X})$ is said to be minimum risk equivariant estimator (MRE) if for any other equivariant estimator for θ, $\phi(\mathbf{X})$, we have

$$R(\delta(\mathbf{X}),\theta) \le R(\phi(\mathbf{X}),\theta), \quad \forall\theta \in \Theta. \tag{2.8}$$

In this paper, the loss function under consideration is the quadratic error loss function, and in this case, the minimum risk equivariant estimator is also known as Pitman estimator (see Lehmann and Casella [8, pages 154–174]).

In particular, let $\widehat{\mu}_{lp}$ and $\widehat{\sigma}_{lp}$, $l = 1,2$ denote the minimum risk equivariant estimator for μ_l and σ_l, $l = 1,2$, respectively. In this notation, the subscript p refers to Pitman estimator. Further, let $\widehat{\mu}_{\text{lobs}}$, $\widehat{\sigma}_{\text{lobs}}$ denote the observed values of $\widehat{\mu}_{lp}$ and $\widehat{\sigma}_{lp}$, $l = 1,2$, respectively. We close this section by recalling the result which is used in computing $\widehat{\mu}_{lp}$, and $\widehat{\sigma}_{lp}$, $l = 1,2$.

Theorem 2.3. *Let $X_1, X_2, \dots, X_n$ be iid random sample from scale-location family with pdf $f(x \mid \mu,\sigma) = \sigma^{-1}\ g((x-\mu)/\sigma)$, where μ and σ are unknown. Also, under quadratic error loss function, suppose that there exists an equivariant estimator with finite risk. Then, under quadratic loss function the MRE of μ and σ are, respectively*

$$\begin{aligned}
\widehat{\mu}_p(\mathbf{x}) &= \int_0^\infty\int_{-\infty}^\infty uv^{n+1}\prod_{i=1}^n g((x_i-u)v)\,du\,dv \times \left(\int_0^\infty\int_{-\infty}^\infty v^{n+1}\prod_{i=1}^n g((x_i-u)v)\,du\,dv\right)^{-1}, \\
\widehat{\sigma}_p(\mathbf{x}) &= \int_0^\infty v^n\int_{-\infty}^\infty \prod_{i=1}^n g((x_i-u)v)\,du\,dv \times \left(\int_0^\infty v^{n+1}\int_{-\infty}^\infty \prod_{i=1}^n g((x_i-u)v)\,du\,dv\right)^{-1}.
\end{aligned} \tag{2.9}$$

For proof, the reader is referred to Lehmann and Casella [8, page 154], Schervish [9, chapter 6] and references therein.

2.2. GPQ and GTV in Two-Sample Location-Scale Families

Let $(X_1,\dots,X_n)$ and $(Y_1,\dots,Y_m)$ be two-sample iid from the population pdfs

$$f_x(x_i \mid \mu_1,\sigma_1) = \sigma_1^{-1}g_1\left(\frac{x_i-\mu_1}{\sigma_1}\right), \qquad f_y(y_j \mid \mu_2,\sigma_2) = \sigma_2^{-1}g_2\left(\frac{y_j-\mu_2}{\sigma_2}\right), \tag{2.10}$$

respectively with $i = 1, \ldots, n$; $j = 1, \ldots, m$; where μ_1, μ_2, σ_1, σ_2 unknown parameters, and g_1 and g_2 are pdfs. Then, the joint pdf of $X_1, \ldots, X_n$, $Y_1, \ldots, Y_m$ is given by

$$f_{xy}(x_1, \ldots, x_n, y_1, \ldots, y_m) = \sigma_1^{-n} \sigma_2^{-m} \prod_{i=1}^{n} g_1\left(\frac{x_i - \mu_1}{\sigma_1}\right) \prod_{j=1}^{m} g_2\left(\frac{y_j - \mu_2}{\sigma_2}\right). \qquad (2.11)$$

We are interested in inference problems concerning the difference between the location parameters $\delta = \mu_1 - \mu_2$, with $\rho = \sigma_2/\sigma_1$ unknown. To set up notation, let $\mathbf{Z}_3 = (Z_{31}, Z_{32})$, where

$$Z_{3l} = \frac{\widehat{\mu}_{lp} - \mu_l}{\widehat{\sigma}_{lp}}, \quad l = 1, 2. \qquad (2.12)$$

By using equivariance property of $\widehat{\mu}_l$ and $\widehat{\sigma}_l$, $l = 1, 2$, we derive the GPQ, GTV of δ. Indeed, let

$$R_{31} = \widehat{\mu}_{1p_{\text{obs}}} - \widehat{\sigma}_{1p_{\text{obs}}} Z_{31}, \qquad R_{32} = \widehat{\mu}_{2p_{\text{obs}}} - \widehat{\sigma}_{2p_{\text{obs}}} Z_{32}, \qquad (2.13)$$

where Z_{3l}, $l = 1, 2$ are defined by (2.12). Then, R_{3l} are GPQ for μ_l, $l = 1, 2$. Using R_{3l}, $l = 1, 2$, we derive the GPQ and GTV for δ as given by the following proposition.

Proposition 2.4. *If the two samples are from the pdf in* (2.10)*, the GPQ for δ is*

$$R_{\delta} = \widehat{\mu}_{1p_{obs}} - \widehat{\mu}_{2p_{obs}} - \widehat{\sigma}_{1p_{obs}} Z_{31} + \widehat{\sigma}_{2p_{obs}} Z_{32}. \qquad (2.14)$$

Furthermore, the GTV is $T_3 = R_{\delta} - \delta$.

Proof. Obviously, the observed value of R_{δ} is δ. Further, since $\widehat{\mu}_l$ and $\widehat{\sigma}_l$, $l = 1, 2$ are equivariant for μ_l and σ_l, respectively, by using Lemma A.3, we conclude that the distributions of R_{3l}, $l = 1, 2$ are not dependent on parameter. Therefore, the distribution of $R_{\delta} = R_{31} - R_{32}$ does not depend on parameter, and this completes the proof. □

In the following section, we present an algorithm which is used in computing the GCI and GPV. The proposed algorithm extensively uses Proposition 2.5 and Corollary 2.6 given below. To the best of our knowledge, these two results are not in the existing literature. To set up notation, let $\mathbf{a} = (a_1, \ldots, a_n)'$, and let $\mathbf{b} = (b_1, \ldots, b_m)'$, where

$$a_i = \frac{X_i - \widehat{\mu}_{1p}}{\widehat{\sigma}_{1p}}, \qquad b_j = \frac{Y_j - \widehat{\mu}_{2p}}{\widehat{\sigma}_{2p}}, \quad i = 1, \ldots, n, \ j = 1, \ldots, m. \qquad (2.15)$$

Further, let $\underline{a} = (a_1, a_2, \ldots, a_{n-2})'$, let $\underline{b} = (b_1, b_2, \ldots, b_{n-2})'$ and let $Z_{4l} = \widehat{\sigma}_{lp} / \sigma_l$, $l = 1, 2$.

Proposition 2.5. *Assume that the two random samples are from the pdfs in* (2.10)*. Then, conditionally to $\underline{a}$, $\underline{b}$, the joint pdf of $(Z_{31}, Z_{32}, Z_{41}, Z_{42})$ is*

$$f(x, y, t, s \mid \underline{a}, \underline{b}) = C^{-1} t^{n-1} s^{m-1} \prod_{i=1}^{n} g_1(t(x + a_i)) \prod_{j=1}^{m} g_2(s(y + b_j)), \qquad (2.16)$$

$t, s > 0$, $-\infty < x, y < +\infty$ where

$$C = \iint_{-\infty}^{\infty} \iint_{0}^{\infty} z_1^{n-1} z_2^{m-1} \prod_{i=1}^{n} g_1(z_1(u + a_i)) \prod_{j=1}^{m} g_2(z_2(w + b_j)) dz_1\, dz_2\, du\, dw. \tag{2.17}$$

Proposition 2.5 extends Corollary A.2 that is established in Appendix A. The proof follows from similar arguments as for Corollary A.2. Further, from Proposition 2.5, we establish Corollary 2.6 that gives the joint pdf of (Z_{31}, Z_{32}) conditionally to $\underline{a}$, $\underline{b}$.

Corollary 2.6. *If Proposition 2.5 holds then, conditionally to $\underline{a}$, $\underline{b}$, the joint pdf of (Z_{31}, Z_{32}) is*

$$f_3(x, y \mid \underline{a}, \underline{b}) = C^{-1} \iint_{0}^{\infty} t^{n-1} s^{m-1} \prod_{i=1}^{n} g_1(t(x + a_i)) \prod_{j=1}^{m} g_2(s(y + b_j)) ds\, dt, \tag{2.18}$$

where with C given in (2.17).

Proof. The proof follows directly from Proposition 2.5. □

3. Framework

In general, the distributions of the GPQ R_δ in (2.14) do not have a closed form. Accordingly, Monte Carlo simulations are needed in order to compute numerically the distributions of R_δ. In this section, we present an algorithm which is used in computing the GCI and GPV for δ. The proposed algorithm is applicable to all members of location-scale families, and in particular, it is applicable to the normal family that is the most commonly discussed in the literature. To the best of our knowledge, there does not exist a similar algorithm in the literature.

The proposed GCI and GPV are obtained by using the following algorithm.

(1) For a given dataset $\mathbf{x}$, using Theorem 2.3, compute $\widehat{\mu}_{l\text{obs}}(\mathbf{x})$, $\widehat{\sigma}_{l\text{obs}}(\mathbf{x})$, the observed values of $\widehat{\mu}_l(\mathbf{X})$, $\widehat{\sigma}_l(\mathbf{X})$, $l = 1, 2$, respectively.

(2) By using (2.15), compute $\{a_i\}$, $i = 1, 2, \ldots, n$, and $\{b_j\}$, $j = 1, 2, \ldots, m$.

(3) Generate $U_{1l} \sim U(0, 1)$ for $l = 1, 2$.

(4) From the pdf of (Z_{31}, Z_{32}), $f_3(x, y \mid \underline{a}, \underline{b})$ given in (2.18), determine Z_{31} and Z_{32} such that $\int_{-\infty}^{Z_{31}} \int_{-\infty}^{\infty} f_3(x, y \mid \underline{a}, \underline{b}) dy\, dx = U_{11}$, $\int_{-\infty}^{Z_{32}} \int_{-\infty}^{\infty} f_3(x, y \mid \underline{a}, \underline{b}) dx\, dy = U_{12}$.

(5) By using (2.14), compute R_δ.

(6) Repeat from step (3) to (5), M times (with M large), and set $R_{\delta,\, k}$ the value of R_δ obtained at the kth replicate, $k = 1, 2, \ldots, M$.

(7) Find $\mathcal{R}_{\delta,\alpha/2}(\mathbf{x})$ and $\mathcal{R}_{\delta,1-\alpha/2}(\mathbf{x})$ as, respectively, $100\alpha/2$ and $100(1 - \alpha/2)$ percentiles of $R_{\delta,1}, R_{\delta,2}, \ldots, R_{\delta,M}$.

(8) Let $\mathbb{I}_A$ denote the indicator function of the event A. Using (2.3), estimate the GPV for δ by $\widehat{p}_\delta = M^{-1} \sum_{k=1}^{M} \mathbb{I}_{\{R_{\delta,\, k} \geq \delta_0\}}$.

Remark 3.1. The equations in step (4) of the above algorithm do not generally give a closed-form solution. Thus, some numerical methods are needed in order to find the quantiles Z_{31} and Z_{32}. In this paper, we applied Newton's method.

Remark 3.2. For the normal sample case, the proposed algorithm produces the same solution as in Weerahandi [1]. Indeed, at normal case, as established in Section 3, Z_{31} and Z_{32} have Student's distributions with, respectively, $n-1$ and $m-1$ degrees of freedom.

4. Some Cases of Two-Sample Location-Scale Families

In this section, we discuss the application of the proposed method to some specific two-sample location-scale families. More precisely, we discuss the application of the proposed method to the two-sample location-scale families for which MLEs do not exist. Also, in order to illustrate the fact that the proposed approach generalizes the method designed at normal case, we discuss briefly the two-sample normal families case.

4.1. Two-Sample Normal Case

Let $\mathbf{x} = (x_1, \ldots, x_n)'$ and let $\mathbf{y} = (y_1, \ldots, y_m)'$. From (2.3), we have

$$f_{xy}(\mathbf{x}, \mathbf{y}) = \sigma_1^{-n} \sigma_2^{-m} (2\pi)^{-(n+m)/2} \exp\left[-\left(2\sigma_1^2\right)^{-1} \sum_{i=1}^{n} (x_i - \mu_1)^2 - \left(2\sigma_2^2\right)^{-1} \sum_{j=1}^{m} (y_j - \mu_2)^2 \right]. \tag{4.1}$$

Under the model in (4.1), we illustrate the computation of GCI and GPV, based on the proposed GPQ. To set up notation, let

$$\overline{a} = \frac{1}{n}\sum_{i=1}^{n} a_i, \qquad S_1^2 = \sum_{i=1}^{n} (a_i - \overline{a})^2, \qquad \overline{b} = \frac{1}{m}\sum_{j=1}^{m} b_i, \qquad S_2^2 = \sum_{j=1}^{m} \left(b_i - \overline{b}\right)^2. \tag{4.2}$$

If $X_i \sim \mathcal{N}(\mu_1, \sigma_1)$ and $Y_i \sim \mathcal{N}(\mu_2, \sigma_2)$, using Theorem 2.3, we have

$$\widehat{\mu}_{1p} = n^{-1} \sum_{i=1}^{n} X_i, \qquad \widehat{\mu}_{2p} = m^{-1} \sum_{j=1}^{m} Y_j, \qquad \widehat{\sigma}_{1p} = \left(\frac{\Gamma(n/2)}{\Gamma((n+1)/2)} \right) \sqrt{\frac{1}{2} \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2}, \tag{4.3}$$

and $\widehat{\sigma}_{2p} = (\Gamma(m/2)/\Gamma((m+1)/2))\sqrt{\sum_{j=1}^{m} (Y_j - \overline{Y})^2/2}$. Also, $\overline{a} = \overline{b} = 0$, $S_1^2 = (2\Gamma^2((n+1)/2))/\Gamma^2(n/2)$, $S_2^2 = (2\Gamma^2((m+1)/2))/\Gamma^2(m/2)$.

Then, using Proposition 2.4 and some computations, we have

$$R_{31} = \widehat{\mu}_{1p_{\text{obs}}} - \widehat{\sigma}_{1p_{\text{obs}}}\left(\frac{S_{1\text{obs}}}{\sqrt{n(n-1)}}\right) \times \left(\frac{\sqrt{n(n-1)}\left(\left(\widehat{\mu}_{1p}-\mu_1\right)/\widehat{\sigma}_{1p}\right)}{S_1}\right)$$
$$= \widehat{\mu}_{1p_{\text{obs}}} - \widehat{\sigma}_{1p_{\text{obs}}}\left(\frac{S_{1\text{obs}}\mathcal{T}_{n-1}}{\sqrt{n(n-1)}}\right), \tag{4.4}$$

where $\mathcal{T}_{n-1}$ stands for a Student't variate with $n-1$ degrees of freedom. Similarly

$$R_{32} = \widehat{\mu}_{2p_{\text{obs}}} - \widehat{\sigma}_{2p_{\text{obs}}}\left(\frac{S_{2\text{obs}}}{\sqrt{n(n-1)}}\right)\left(\frac{\sqrt{n(n-1)}\left(\left(\widehat{\mu}_{2p}-\mu_2\right)/\widehat{\sigma}_{2p}\right)}{S_2}\right)$$
$$= \widehat{\mu}_{2p_{\text{obs}}} - \widehat{\sigma}_{2p_{\text{obs}}}\left(\frac{S_{2\text{obs}}\mathcal{T}_{m-1}}{\sqrt{m(m-1)}}\right), \tag{4.5}$$

and taking $R_{31} - R_{32}$, we get,

$$R_{\delta} = \widehat{\mu}_{1p_{\text{obs}}} - \widehat{\mu}_{2p_{\text{obs}}} - \frac{\widehat{\sigma}_{1p_{\text{obs}}}S_{1\text{obs}}\mathcal{T}_{n-1}}{\sqrt{n(n-1)}} + \frac{\widehat{\sigma}_{2p_{\text{obs}}}S_{2\text{obs}}\mathcal{T}_{m-1}}{\sqrt{m(m-1)}}. \tag{4.6}$$

4.2. Two Location-Scale Families Case Where MLE Does Not Exist

Following Gupta and Székely [6], let the families $\sigma_l^{-1}g_l((x-\mu_l)/\sigma_l)$, $l = 1,2$ where

$$g_1(x) = c\left(x\log^2 x\right)^{-1}, \qquad g_2(y) = c\left(y\log^2 y\right)^{-1}, \tag{4.7}$$

$0 < x, y \le k < 1$, k is any constant that satisfies $0 < k < 1$ and $c = -1/\log(k)$ is a constant. Gupta and Székely [6] proved that MLEs for σ_l, μ_l, $l = 1,2$ do not exist.

The second illustrative example is based on the result in Pitman [5]. Namely, we consider families $\sigma_l^{-1}g_l((x-\mu_l)/\sigma_l)$, $l = 1,2$, where

$$g_l(x_l) = \left(2(1+|x_l|)(1+\log(1+|x_l|))^{\,2}\right)^{-1}, \quad -\infty < x_l < \infty, \ l = 1,2. \tag{4.8}$$

Pitman [5] proved that MLEs for σ_l, μ_l, $l = 1,2$ do not exist. For the families in (4.7) and (4.8), the pdf of R_{3l} and R_{4l}, $l = 1,2$ do not have a closed form and thus, the distribution of R_{δ} is obtained numerically by using the algorithm given in Section 2.

5. Simulation Study and Data Analysis

5.1. Simulation Study

In this section, we carry out intensive simulation studies in order to evaluate the performances of the suggested approach in small and moderate sample sizes. To this end,

Table 1: The coverage probabilities (CPR) of the 95% GCI for δ (family in (4.8)).

Sizes (n, m)	CPR: $(\mu_1, \mu_2, \sigma_1, \sigma_2) = (2,\ 2,\ 2,\ 2)$	CPR: $(\mu_1, \mu_2, \sigma_1, \sigma_2) = (2,\ 1,\ 2,\ 1)$
(5, 5)	0.847	0.839
(10, 10)	0.925	0.920
(20, 20)	0.942	0.938
(50, 50)	0.946	0.942
(60, 60)	0.946	0.945
(10, 20)	0.967	0.957
(10, 50)	0.970	0.972
(30, 50)	0.962	0.969

Table 2: The simulated powers for δ (the location-scale family in (4.8)).

Sizes (n, m)	$(\delta, \sigma_1, \sigma_2)$	Power	Sizes (n, m)	$(\delta, \sigma_1, \sigma_2)$	Power
	(−2, 2, 2)	0.194		(−2, 2, 2)	0.816
(5,5)	(−1, 2, 2)	0.102	(20,20)	(−1, 2, 2)	0.453
	(0, 2, 2)	0.037		(0, 2, 2)	0.046
	(−2, 2, 2)	0.527		(−2, 2, 2)	0.946
(10,10)	(−1, 2, 2)	0.251	(50,50)	(−1, 2, 2)	0.752
	(0, 2, 2)	0.044		(0, 2, 2)	0.054
	(−4, 2, 2)	0.831		(−4, 2, 2)	0.819
(10,20)	(−2, 2, 2)	0.510	(10,50)	(−2, 2, 2)	0.465
	(0, 2, 2)	0.042		(0, 2, 2)	0.038

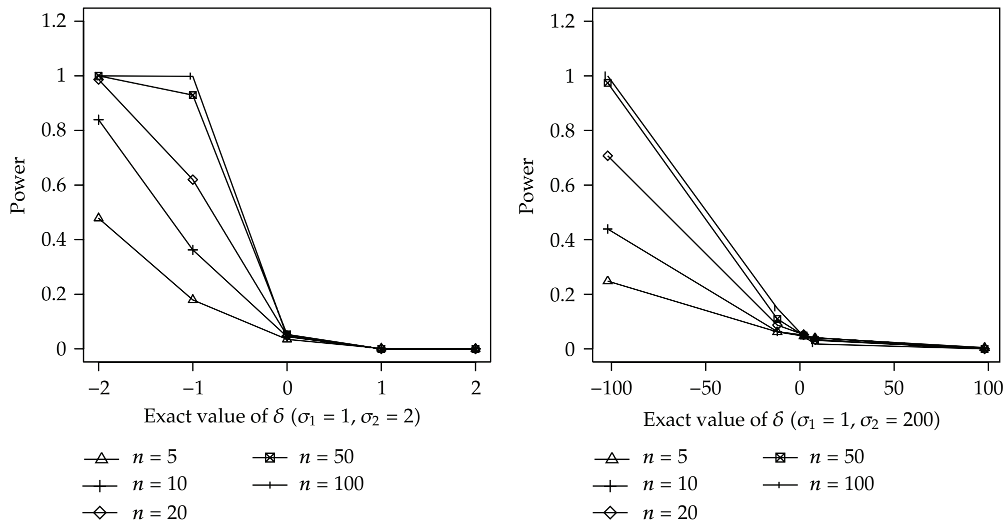
we generate 10000 two-samples from logistic distribution, from the distribution in (4.7), and from the distribution in (4.8). In order to save space, we report below the empirical coverage probability and the empirical power for the location-scale family given in (4.8). Namely, the simulated coverage probabilities of the 95% GCI are presented in Table 1, and the empirical powers of the proposed test are given in Table 2, at significance level $\alpha = .05$.

In particular, concerning the GCI of δ, Table 1 shows that, for $n \geq 20$, the coverage probabilities are also relatively close to the nominal confidence level of 95%. Interestingly, the case of equal scale parameters and that of unequal scale parameters seem to provide similar results. Further, it is noticed that as the sample size increases, the coverage probability gets closer to the nominal confidence level (95%). Concerning the performance of the solution to the testing problem (2.4), Table 2 shows that the power function varies with different values of m, n, μ_1, σ_1, σ_2, and $\delta = \mu_1 - \mu_2$. In fact, from Figure 1, it can be seen that when $\delta = \delta_0 = 0$, the powers are all approximately equal to 0.05. But on the left-hand side of 0, the power continually increases to 1 when the distance between δ and 0 increases. Also, in the right hand side, the power decreases to 0 as the distance increases. Furthermore, in the left hand side of 0, for each exact value of δ, the power increases as the sample size increases.

5.2. Illustrative Examples and Data Analysis

5.2.1. Normal Body Temperature Dataset

This dataset is found in Mackowiak et al. [10]. In this dataset, a total number of 130 patients have been assigned, with 65 males and 65 females. Their body temperatures have been tested



(a) If $\sigma_1 = 1, \sigma_2 = 2$ (the family in (4.8))

(b) If $\sigma_1 = 1, \sigma_2 = 200$ (the family in (4.8))

Figure 1: The simulated power versus δ (the family case in (4.8)).

Table 3: Numerical results for the normal body temperature dataset.

Parameters of interest	Point estimate	95% GCI
μ_1	98.1	(97.92873, 98.28112)
σ_1	.696	(0.5283083, 0.8186700)
μ_2	98.39	(98.20655, 98.57497)
σ_2	0.74	(0.5623908, 0.8679169)
δ	−.2892308	(−0.54288915, −0.03725161)

and recorded. Furthermore, it is already confirmed that the temperatures in these 2 gender groups are normally distributed. In particular, for the male group, one can consider $X \sim \mathcal{N}(\mu_1, \sigma_1)$, and for the female group, one can consider $Y \sim \mathcal{N}(\mu_2, \sigma_2)$. From Table 3, a 95% GCI for δ is $(-0.54288915, \ -0.03725161)$ and thus, since the interval does not contain 0, there is a significant difference between the two location parameters. By applying (2.5) to the testing problem $H_0 : \delta \geq 0$ versus $H_0 : \delta < 0$, the GPV is found to be 0.0133, and this result indicates that the null hypothesis should be rejected at 2% significant level, that is, this confirms that $\mu_1 < \mu_2$.

5.2.2. Cloud Seeding Dataset

The cloud seeding dataset consists of the amount of rainfall (in acre-feet) which have been recorded. The dataset is given in Krishnamoorthy, and Mathew [11]. For this dataset, 26 clouds were randomly seeded with silver nitrate, and 26 others were unseeded. In the above quoted paper, the authors showed that lognormal model fits the dataset very well. Thus, we assume unseeded cloud group $X_1 \sim \text{Lognormal}(\mu_1, \sigma_1)$ and seeded cloud group $Y_1 \sim \text{Lognormal}(\mu_2, \sigma_2)$. We set $X = \log(X_1) \sim \mathcal{N}(\mu_1, \sigma_1)$ and $Y = \log(Y_1) \sim \mathcal{N}(\mu_2, \sigma_2)$.

Table 4: Numerical results for the cloud seeding dataset.

Parameters of interest	Point estimate	95% GCI
μ_1	3.990406	(3.325968, 4.641512)
σ_1	1.625515	(1.071410, 2.148553)
μ_2	5.134187	(4.496115, 5.775415)
σ_2	1.583602	(1.019519, 2.084743)
δ	−1.143781	(−2.073196, −0.2211248)

From Table 4, the GCI for δ indicates that the difference between the two location parameters is statistically significant. Also, for the testing problem $H_0 : \delta \geq 0$ versus $H_1 : \delta < 0$, the GPV is 0.007 which indicates that $\mu_1 < \mu_2$. Note that this finding corroborates the result given in Krishnamoorthy and Mathew [11], where the authors concluded that μ_1 is statistically different from μ_2.

6. Conclusion

In this paper, we proposed a solution of typical Behrens-Fisher problem in the general setting where two independent samples are from location-scale families. We presented a general statistical method for constructing GPQ and GTV for the difference between two location parameters of location-scale families. The proposed method is based on the minimum risk equivariant estimators which are known to be more general and more efficient than the MLEs. The simulation studies show that the proposed methods provide CIs and tests with high coverage probability and power, and the resulting tests preserve the significance level.

The proposed method applies to all members of the location-scale families, as opposed to the methods given in the literature, as Welch's method, which are designed only for the normal case. In addition to this generality, our method is at least as good as Welch's method in the normal Behrens-Fisher problem (see simulation results in Appendix B).

Appendices

A. Technical Results and Proof of Proposition 2.5

In this subsection, we present some results which are useful in deriving Proposition 2.5. Recall that this last proposition is used in deriving Corollary 2.6 that plays a central role in the proposed algorithm as given in Section 2. For the sake of simplicity, the results are outlined for the case where $\sigma_1 = \sigma_2 = 1$, that is, when the two samples $(X_1, X_2, \ldots, X_n)$ and $(Y_1, Y_2, \ldots, Y_m)$ are independent from location families with location parameters μ_l, $l = 1, 2$.

Let $\widehat{\mu}_{lp}$ be the equivariant estimator of μ_l, $l = 1, 2$. Also, let

$$c_i = X_i - \widehat{\mu}_{1p}, \quad i = 1, \ldots, n; \quad d_j = Y_j - \widehat{\mu}_{2p}, \quad j = 1, \ldots, m, \tag{A.1}$$

where the two samples $(X_1, X_2, \ldots, X_n)$ and $(Y_1, Y_2, \ldots, Y_m)$ are independent from location families with location parameters μ_l, $l = 1, 2$.

Proposition A.1. *Assume two random samples are from two independent location families and assume that relation* (A.1) *holds. Then* $c_1, \ldots, c_{n-1}$, $d_1, \ldots, d_{m-1}$ *are ancillary statistics. Furthermore, the joint pdf of* $c_1, \ldots, c_{n-1}$, $d_1, \ldots, d_{m-1}$ *is*

$$f_{cd}(\mathbf{x}, \mathbf{y}) = \iint_{-\infty}^{\infty} \prod_{i=1}^{n} g_1(x_i + z_1) \prod_{j=1}^{m} g_2(y_j + z_2) dz_1\, dz_2. \tag{A.2}$$

Proof. From the fact that $\widehat{\mu}_{1p}$ and $\widehat{\mu}_{2p}$ are equivariant estimators for μ_1 and μ_2, respectively, we conclude that $c_1, \ldots, c_{n-1}$, $d_1, \ldots, d_{m-1}$ are ancillary statistics. Further, without loss of generality, assume that $\sigma_1 = \sigma_2 = 1$. Also, let us define c_n, and d_n by $X_n = c_n + \widehat{\mu}_{1p}$, $Y_m = d_m + \widehat{\mu}_{2p}$. Then, since $\widehat{\mu}_{1p}$ and $\widehat{\mu}_{2p}$ are equivariant, c_n and d_m can be expressed as a function of $c_1, \ldots, c_{n-1}$, $d_1, d_2, \ldots, d_{m-1}$ and thus, one can set $c_n = T_1(c_1, \ldots, c_{n-1})$, $d_m = T_2(d_1, \ldots, d_{n-1})$. Then

$$\begin{aligned} x_i &= c_i + \widehat{\mu}_{1p}, \quad i = 1, \ldots, n-1, \qquad x_n = c_n + \widehat{\mu}_{1p}, \\ y_j &= d_j + \widehat{\mu}_{2p}, \quad i = 1, \ldots, m-1, \qquad y_m = d_m + \widehat{\mu}_{2p}. \end{aligned} \tag{A.3}$$

Let $X = (X_1, \ldots, X_n)$, let $Y = (Y_1, \ldots, Y_m)$, let $x = (x_1, \ldots, x_n)$, and let $y = (y_1, \ldots, y_m)$. We have $f(x, y) = \prod_{i=1}^{n} g_1(x_i - \mu_1) \prod_{j=1}^{m} g_2(y_j - \mu_2)$. Also, let $\mathbf{c} = (c_1, \ldots, c_{n-1})$, $\mathbf{d} = (d_1, \ldots, d_{m-1})$. The joint pdf of $(\mathbf{c}, \mathbf{d}, \widehat{\mu}_{1p}, \widehat{\mu}_{2p})$ is

$$f(c, d, \widehat{\mu}_{1p}, \widehat{\mu}_{2p}) = |\mathbf{J}| \prod_{i=1}^{n} g_1(c_i + \widehat{\mu}_{1p} - \mu_1) \prod_{j=1}^{m} g_2(d_j + \widehat{\mu}_{2p} - \mu_2), \tag{A.4}$$

where $\mathbf{J}$ is the Jacobian matrix. We have

$$\mathbf{J} = \begin{pmatrix} \mathbf{J}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_2 \end{pmatrix} \quad \text{with } \mathbf{J}_1 = \begin{pmatrix} \dfrac{\partial x_1}{\partial c_1} & \cdots & \dfrac{\partial x_1}{\partial c_{n-1}} & \dfrac{\partial x_1}{\partial \widehat{\mu}_1} \\ \vdots & \cdots & \vdots & \vdots \\ \dfrac{\partial x_n}{\partial c_1} & \cdots & \dfrac{\partial x_n}{\partial c_{n-1}} & \dfrac{\partial x_n}{\partial \widehat{\mu}_1} \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 0 & 1 \\ \vdots & \cdots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix},$$
$$\mathbf{J}_2 = \begin{pmatrix} \dfrac{\partial y_1}{\partial d_1} & \cdots & \dfrac{\partial y_1}{\partial d_{m-1}} & \dfrac{\partial y_1}{\partial \widehat{\mu}_2} \\ \vdots & \cdots & \vdots & \vdots \\ \dfrac{\partial y_m}{\partial d_1} & \cdots & \dfrac{\partial y_m}{\partial d_{m-1}} & \dfrac{\partial y_m}{\partial \widehat{\mu}_2} \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 0 & 1 \\ \vdots & \cdots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix}. \tag{A.5}$$

Therefore, from (A.4) and letting $z_l = \widehat{\mu}_l - \mu_l$, $l = 1, 2$, we get $f_{cd}(\mathbf{x}, \mathbf{y})$ as stated in the proposition, and that completes the proof. □

Corollary A.2. *If two random samples are from the pdfs in* (2.10) *with* $\sigma_1 = \sigma_2 = 1$ *then, conditionally to* $\mathbf{c}$, $\mathbf{d}$, *the joint pdf of* $\widehat{\mu}_{1p} - \mu_1$, $\widehat{\mu}_{2p} - \mu_2$ *is*

$$f(x, y \mid \mathbf{c}, \mathbf{d}) = \frac{\left(\prod_{i=1}^{n} g_1(c_i + x) \prod_{j=1}^{m} g_2(d_j + y)\right)}{f_{cd}(\mathbf{c}, \mathbf{d})}, \quad -\infty < x,\ y < +\infty, \tag{A.6}$$

where $f_{cd}(\mathbf{c}, \mathbf{d})$ *is the pdf given in Proposition A.1.*

Proof. From (A.4), we directly get the conditional joint pdf of $(\widehat{\mu}_{1p}, \widehat{\mu}_{2p})$ given $\mathbf{c}$, $\mathbf{d}$. By algebraic computations, we verify that the conditional pdf of $(\widehat{\mu}_{1p} - \mu_1, \widehat{\mu}_{2p} - \mu_2)$ corresponds to that stated in the corollary. □

Lemma A.3. *Let* $\mathbf{X} = (X_1, X_2, \ldots, X_n)'$ *be a random sample from the location-scale family with location parameter* μ *and scale parameter* σ. *Also, let* $\widehat{\mu}_p$, $\widehat{\sigma}_p$ *be equivariant estimators for* μ *and* σ, *respectively. Then, the distributions of* $(\widehat{\mu}_p - \mu)/\widehat{\sigma}$ *and* $\widehat{\sigma}_p/\sigma$ *do not depend on the parameters* μ *and* σ.

Proof. Let $\delta_1(\mathbf{X}) = \widehat{\sigma}_p$, let $\delta_2(\mathbf{X}) = \widehat{\mu}_p$, let $\psi_1(\mathbf{X}, \mu, \sigma) = \widehat{\sigma}_p/\sigma$, and let $\psi_2(\mathbf{X}, \mu, \sigma) = (\widehat{\mu}_p - \mu)/\widehat{\sigma}$. Since $\delta_1(\mathbf{X})$ and $\delta_2(\mathbf{X})$ are equivariant for σ and μ, respectively, we have (see Lehmann and Casella [8, pages 171–173])

$$\delta_1(b\mathbf{X} + a) = b\delta_1(\mathbf{X}), \quad \delta_2(b\mathbf{X} + a) = b\delta_2(\mathbf{X}) + a, \quad \forall -\infty < a < \infty, \quad b > 0. \tag{A.7}$$

Hence, taking $a = -\mu/\sigma$, $b = 1/\sigma$, we obtain

$$\delta_1\left(\frac{\mathbf{X} - \mu}{\sigma}\right) = \frac{\delta_1(\mathbf{X})}{\sigma} = \psi_1(\mathbf{X}, \mu, \sigma), \tag{A.8}$$

$$\delta_2\left(\frac{\mathbf{X} - \mu}{\sigma}\right) = \frac{(\delta_2(\mathbf{X}) - \mu)}{\sigma} = \psi_2(\mathbf{X}, \mu, \sigma). \tag{A.9}$$

Further, since $\mathbf{X}$ is from a location-scale family with location and scale parameters μ and σ respectively, the distribution of $(\mathbf{X} - \mu)/\sigma$ does not depend on parameter. Indeed, the joint pdf of $\mathbf{X}$ can be written as $(1/\sigma^n)\prod_{i=1}^{n} g((x_i - \mu)/\sigma)$ where g is a pdf which does not depend on μ and σ. Then, the joint pdf of $(\mathbf{X} - \mu)/\sigma$ is $\prod_{i=1}^{n} g(y_i)$. Hence, the distributions of $\delta_1((\mathbf{X} - \mu)/\sigma)$ and $\delta_2((\mathbf{X} - \mu)/\sigma)$ do not depend on parameter. Therefore, from (A.8) and (A.9), we conclude that the distributions of $\psi_1(\mathbf{X}, \mu, \sigma)$ and $\psi_2(\mathbf{X}, \mu, \sigma)$ do not depend on parameter, and this completes the proof. □

B. Simulation Results in Normal Samples Case

In this section we present some numerical results for the normal samples case. Indeed, the proposed approach generalizes the existing methods used in solving the well-known Behrens-Fisher problem.

Table 5: Coverage probability for δ (normal case with unequal sample sizes).

Sizes (m, n)	$(\delta, \sigma_1, \sigma_2)$	Coverage proba.		Average width		St. error of width	
		Welch	GTV	Welch	GTV	Welch	GTV
$(5, 10)$	$(0, 2, 1)$	0.9404	0.9522	4.6383	4.8018	1.6342	1.6096
$(5, 100)$	$(0, 2, 1)$	0.9452	0.9478	4.2551	4.2896	1.6747	1.6605
$(50, 100)$	$(0, 2, 1)$	0.9490	0.9496	1.1938	1.1971	0.1085	0.1082
$(10, 5)$	$(0, 2, 1)$	0.9556	0.9554	3.3334	3.7019	0.8165	0.7930
$(100, 5)$	$(0, 2, 1)$	0.9436	0.9542	2.7044	2.3623	0.7050	0.6994
$(100, 50)$	$(0, 2, 1)$	0.9494	0.9506	0.9656	0.9722	0.0566	0.0560
$(5, 10)$	$(0, 200, 2)$	0.9484	0.9484	466.6274	466.4071	171.1893	171.1114
$(5, 100)$	$(0, 200, 2)$	0.9522	0.9516	470.7588	470.4365	172.2657	172.2866
$(50, 100)$	$(0, 200, 2)$	0.9510	0.9502	113.3475	113.2865	11.4323	11.4340
$(10, 5)$	$(0, 200, 2)$	0.9526	0.9520	278.2983	278.2280	65.7577	65.7696
$(100, 5)$	$(0, 200, 2)$	0.9514	0.9512	79.3017	79.2324	5.5889	5.6043
$(100, 50)$	$(0, 200, 2)$	0.9482	0.9496	79.2472	79.2421	5.5864	5.6010

Table 6: Comparison with the bootstrap method (normal case with unequal sample sizes).

Sizes (m, n)	$(\delta, \sigma_1, \sigma_2)$	Coverage proba.		Average width		St. error of width	
		Bootstr	GTV	Bootstr	GTV	Bootstr	GTV
$(5, 10)$	$(0, 2, 1)$	0.941	0.952	5.199	4.802	2.605	1.607
$(5, 100)$	$(0, 2, 1)$	0.946	0.948	6.284	4.290	3.499	1.661
$(50, 100)$	$(0, 2, 1)$	0.948	0.950	1.184	1.197	0.117	0.108
$(10, 5)$	$(0, 2, 1)$	0.953	0.955	3.424	3.702	0.726	0.793
$(100, 5)$	$(0, 2, 1)$	0.934	0.954	2.493	2.362	1.000	0.699
$(100, 50)$	$(0, 2, 1)$	0.944	0.951	0.963	0.972	0.067	0.056
$(5, 10)$	$(0, 200, 2)$	0.942	0.948	764.452	466.407	769.088	171.111
$(5, 100)$	$(0, 200, 2)$	0.946	0.952	698.5501	470.437	504.280	172.287
$(50, 100)$	$(0, 200, 2)$	0.953	0.950	111.983	113.287	12.110	11.434
$(10, 5)$	$(0, 200, 2)$	0.945	0.952	297.40	278.228	78.240	65.770
$(100, 5)$	$(0, 200, 2)$	0.942	0.951	78.867	79.232	6.198	5.604
$(100, 50)$	$(0, 200, 2)$	0.949	0.950	78.666	79.242	6.155	5.601

For comparison purposes, we compared the proposed method with bootstrap. In particular, Table 6 shows that the proposed method dominates the bootstrap in small sample cases, and it is at least as good as the bootstrap in large sample cases. Also, we present in Table 5 the coverage probability obtained by using the Welch's approximation method for the normal case. Also, we present in Tables 7 and 8 the empirical powers obtained by using the Welch approximation method for the normal case. In summary, for the Behrens-Fisher problem with unbalanced sample sizes, the proposed confidence interval is at least as accurate as that given by Welch method. Further, the proposed test is at least as powerful as the Welch approximation test. In addition, the proposed method has the advantage of being useful for the more general statistical model of two samples from location-scale family.

Table 7: Powers for δ (normal case with unequal sample sizes & scale parameters).

Sizes (m, n)	$(\delta, \sigma_1, \sigma_2)$	Power		$(\delta, \sigma_1, \sigma_2)$	Power	
		Welch	GTV		Welch	GTV
$(5, 10)$	$(-2,\ 2,\ 1)$	0.5472	0.550	$(-102,\ 200,\ 2)$	0.2462	0.2488
	$(-1,\ 2,\ 1)$	0.2274	0.2296	$(-12,\ 200,\ 2)$	0.0572	0.0608
	$(0,\ 2,\ 1)$	0.0578	0.0464	$(0,\ 200,\ 2)$	0.0542	0.0504
	$(1,\ 2,\ 1)$	0.0082	0.0066	$(8,\ 200,\ 2)$	0.046	0.037
$(10, 5)$	$(-2,\ 2,\ 1)$	0.7712	0.7791	$(-102,\ 200,\ 2)$	0.438	0.4402
	$(-1,\ 2,\ 1)$	0.3216	0.3304	$(-12,\ 200,\ 2)$	0.066	0.0678
	$(0,\ 2,\ 1)$	0.045	0.0458	$(0,\ 200,\ 2)$	0.0524	0.049
	$(1,\ 2,\ 1)$	0.0016	0.002	$(8,\ 200,\ 2)$	0.0448	0.042
$(5, 100)$	$(-2,\ 2,\ 1)$	0.5714	0.5804	$(-102,\ 200,\ 2)$	0.2414	0.2496
	$(-1,\ 2,\ 1)$	0.2322	0.2409	$(-12,\ 200,\ 2)$	0.0592	0.0688
	$(0,\ 2,\ 1)$	0.0542	0.0508	$(0,\ 200,\ 2)$	0.047	0.0486
	$(1,\ 2,\ 1)$	0.0044	0.004	$(8,\ 200,\ 2)$	0.0448	0.0486
$(50, 100)$	$(-2,\ 2,\ 1)$	1	1	$(-102,\ 200,\ 2)$	0.9712	0.9726
	$(-1,\ 2,\ 1)$	0.9522	0.9534	$(-12,\ 200,\ 2)$	0.110	0.116
	$(0,\ 2,\ 1)$	0.052	0.0488	$(0,\ 200,\ 2)$	0.047	0.049
	$(1,\ 2,\ 1)$	0	0	$(8,\ 200,\ 2)$	0.029	0.0252
$(100, 50)$	$(-2,\ 2,\ 1)$	1	1	$(-102,\ 200,\ 2)$	0.9984	0.999
	$(-1,\ 2,\ 1)$	0.992	0.9932	$(-12,\ 200,\ 2)$	0.1378	0.156
	$(0,\ 2,\ 1)$	0.0512	0.0492	$(0,\ 200,\ 2)$	0.051	0.0486

Table 8: Powers for δ for the normal case: other scenarios with unequal sample sizes.

Sizes (m, n)	$(\delta, \sigma_1, \sigma_2)$	Power		$(\delta, \sigma_1, \sigma_2)$	Power	
		Welch	GTV		Welch	GTV
$(10, 5)$	$(-2,\ 2,\ 1)$	0.7712	0.7791	$(-102,\ 200,\ 2)$	0.438	0.4402
	$(-1,\ 2,\ 1)$	0.3216	0.3304	$(-12,\ 200,\ 2)$	0.066	0.0678
	$(0,\ 2,\ 1)$	0.045	0.0458	$(0,\ 200,\ 2)$	0.0524	0.049
	$(1,\ 2,\ 1)$	0.0016	0.002	$(8,\ 200,\ 2)$	0.0448	0.042
	$(2,\ 2,\ 1)$	0	0	$(98,\ 200,\ 2)$	0.001	0
$(100, 5)$	$(-2,\ 2,\ 1)$	0.9616	0.9678	$(-102,\ 200,\ 2)$	0.9994	0.9996
	$(-1,\ 2,\ 1)$	0.555	0.558	$(-12,\ 200,\ 2)$	0.148	0.152
	$(0,\ 2,\ 1)$	0.0478	0.0544	$(0,\ 200,\ 2)$	0.0532	0.0476
	$(1,\ 2,\ 1)$	0	0	$(8,\ 200,\ 2)$	0.0214	0.0196
	$(2,\ 2,\ 1)$	0	0	$(98,\ 200,\ 2)$	0	0
$(100, 50)$	$(-2,\ 2,\ 1)$	1	1	$(-102,\ 200,\ 2)$	0.9984	0.999
	$(-1,\ 2,\ 1)$	0.992	0.9932	$(-12,\ 200,\ 2)$	0.1378	0.156
	$(0,\ 2,\ 1)$	0.0512	0.0492	$(0,\ 200,\ 2)$	0.051	0.0486
	$(1,\ 2,\ 1)$	0	0	$(8,\ 200,\ 2)$	0.0218	0.0214
	$(2,\ 2,\ 1)$	0	0	$(98,\ 200,\ 2)$	0	0

Acknowledgments

This paper received the financial support from the Natural Sciences and Engineering Research Council (NSERC) of Canada. The authors would like to thank Professor Mathew for providing interesting references. The authors would also like to thank anonymous referee for helpful comments.

References

[1] S. Weerahandi, "Generalized confidence intervals," *Journal of the American Statistical Association*, vol. 88, no. 423, pp. 899–905, 1993.
[2] K.-W. Tsui and S. Weerahandi, "Generalized *p*-values in significance testing of hypotheses in the presence of nuisance parameters," *Journal of the American Statistical Association*, vol. 84, no. 406, pp. 602–607, 1989.
[3] I. Bebu and T. Mathew, "Comparing the means and variances of a bivariate log-normal distribution," *Statistics in Medicine*, vol. 27, no. 14, pp. 2684–2696, 2008.
[4] D. A. Sprott, *Statistical Inference in Science*, Springer Series in Statistics, Springer, New York, NY, USA, 2000.
[5] E. J. G. Pitman, *Some Basic Theory for Statistical Inference*, Chapman and Hall, London, UK, 1979.
[6] A. K. Gupta and G. J. Székely, "On location and scale maximum likelihood estimators," *Proceedings of the American Mathematical Society*, vol. 120, no. 2, pp. 585–589, 1994.
[7] K. Krishnamoorthy, T. Mathew, and G Ramachandran, "Upper limits for exceedance probabilities under the one-way random effects model," *Annals of Occupational Hygiene*, vol. 51, no. 4, pp. 397–406, 2007.
[8] E. L. Lehmann and G. Casella, *Theory of Point Estimation*, Springer Texts in Statistics, Springer, New York, NY, USA, 2nd edition, 1998.
[9] M. J. Schervish, *Theory of Statistics*, Springer Series in Statistics, Springer, New York, NY, USA, 1997.
[10] P. A. Mackowiak, S. S. Wasserman, and M. M. Levine, "A Critical appraisal of 98.6 degrees F, the upper limit of the normal body temperature, and other legacies of Carl Reinhold August Wunderlich," *Journal of the American Medical Association*, vol. 268, pp. 1578–1580, 1992.
[11] K. Krishnamoorthy and T. Mathew, "Inferences on the means of lognormal distributions using generalized *p*-values and generalized confidence intervals," *Journal of Statistical Planning and Inference*, vol. 115, no. 1, pp. 103–121, 2003.