Journal of Multivariate Analysis 100 (2009) 1384–1397

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Coverage of generalized confidence intervals

Anindya Roy [a,*], Arup Bose [b]

[a] University of Maryland Baltimore County, Baltimore, MD 21250, USA
[b] Indian Statistical Institute, 203 B.T. Road, Kolkata, 700108, India

ARTICLE INFO

Article history:
Received 1 November 2007
Available online 10 December 2008

AMS subject classifications:
62F03
62F05
62F40

Keywords:
Average bioequivalence
Behrens–Fisher problem
Bootstrap
Cornish–Fisher expansion
Coverage probability
Edgeworth expansion

ABSTRACT

Generalized confidence intervals provide confidence intervals for complicated parametric functions in many common practical problems. They do not have exact frequentist coverage in general, but often provide coverage close to the nominal value and have the correct asymptotic coverage. However, in many applications generalized confidence intervals do not have satisfactory finite sample performance. We derive expansions of coverage probabilities of one-sided generalized confidence intervals and use the expansions to explain the nonuniform performance of the generalized intervals. We then show how to use these expansions to obtain improved coverage by suitable calibration. The benefits of the proposed modification are illustrated via several examples.

1. Introduction

Generalized confidence intervals and generalized test statistics have been introduced by Tsui and Weerahandi [1] and Weerahandi [2–5] as easy tractable means of constructing confidence intervals for, and testing hypothesis about, complicated parametric functions. There have been several articles demonstrating the merits and drawbacks of generalized confidence intervals in routinely used applications. The advantages and disadvantages of the generalized intervals over other resampling methods have been established in numerous articles.

The primary focus of this paper is to provide theoretical explanation of the observed empirical behavior of the generalized intervals and to suggest ways of improving the finite sample performance of the generalized intervals. We establish that in general the generalized confidence intervals are not first order accurate, i.e., accurate only up to the $n^{-1/2}$ term. We provide a necessary and sufficient condition for the generalized intervals to be first order accurate. We suggest a modification to make the intervals first order accurate and hence, improve the coverage of the intervals significantly in many applications. Although we only discuss generalized confidence intervals in this paper, the results can be easily extended to improve the power properties of generalized tests as well.

Many articles over the last decade have shown that for messy parametric problems with certain pivotal structure, the generalized intervals perform adequately in the repeated sampling setup as well (even though the generalized intervals are not motivated from a repeated sampling argument). Hence they are appealing also to practitioners who are comfortable with the classical approach of frequentist confidence intervals. Generalized procedures have been successfully applied to several problems of practical importance. The areas of applications include comparison of means, testing and estimation

* Corresponding author.
E-mail addresses: anindya@math.umbc.edu (A. Roy), abose@isical.ac.in (A. Bose).

0047-259X/$ – see front matter
doi:10.1016/j.jmva.2008.12.001

of functions of parameters of normal and related distributions (Weerahandi, [2–5], Krishnamoorthy and Mathew [6], Johnson and Weerahandi [7], Gamage, Mathew and Weerahandi [8]); testing fixed effects and variance components in repeated measures and mixed effects ANOVA models (Zhou and Mathew [9], Gamage and Weerahandi [10], Chiang [11], Krishnamoorthy and Mathew [6], Weerahandi [5], Mathew and Webb [12], Arendacka [13]); interlaboratory testing (Iyer, Wang and Mathew [14]); bioequivalence (McNally, Iyer and Mathew [15]); growth curve modeling (Weerahandi and Berger [16], Lin and Lee [17]); reliability and system engineering (Roy and Mathew [18], Tian and Cappelleri [19], Mathew, Kurian and Sebastian [20]); process control (Burdick, Borror and Montgomery [21], Mathew, Kurian and Sebastian [22]); environmental health (Krishnamoorthy, Mathew and Ramachandran [23]) and many others. The simulation studies in Johnson and Weerahandi [7], Weerahandi [4,5] Zhou and Mathew [9], Gamage and Weerahandi [10], among others have demonstrated the success of the generalized procedure in many problems where the classical approach fails to yield adequate confidence intervals.

There has been some theoretical investigation of the success of generalized intervals in the frequentist sense. Hannig, Iyer and Patterson [24] have shown that asymptotically the generalized intervals maintain the target coverage level for a large class of problems. Hannig [25] has also investigated the connection between the generalized procedures and fiducial inference.

In Section 3 we give our main results. In Section 4, we investigate the magnitude of the first order term in the coverage probability and illustrate our methodology in the context of several examples. All technical proofs are given in an appendix. In the next section, we describe the assumptions and the notations.

2. Preliminaries and assumptions

Let $\mathbf{x} \in \mathbb{R}^d$ denote a d-dimensional statistic (and by abuse of notation, also the observed value of the statistic) whose distribution is indexed by a parameter $\theta \in \Theta \subseteq \mathbb{R}^q$. The parameter space Θ is assumed to be an open subset of $\mathbb{R}^q$. Let $1/2 < \alpha < 1$. Our interest is in constructing a $100\alpha\%$ one-sided confidence interval for a one-dimensional parametric function $\pi(\theta)$ based on the observed value $\mathbf{x}$.

In classical statistics, confidence intervals for the parameter $\pi(\theta)$ would be constructed by inverting the distribution of a pivotal quantity. However, depending on the nature of the parametric problem, such pivotal quantities may not be available. Weerahandi [2] suggested constructing the confidence interval by inverting the distribution of a generalized pivotal quantity.

Our definition of a generalized pivotal quantity is that due to Hannig et al. [24] which is an adaptation of the original definition of Weerahandi [2]. Let $\mathbf{X}$ denote an independent and identical copy (but unobserved) of the observable random vector $\mathbf{x}$. Then $T_\theta(\mathbf{x}, \mathbf{X})$ is a *Generalized Pivotal Quantity* for $\pi(\theta)$ if it satisfies:

(i) The distribution of $T_\theta(\mathbf{x}, \mathbf{X})$ conditional on $\mathbf{x}$ is free from θ.
(ii) $T_\theta(\mathbf{x}, \mathbf{x}) = \pi(\theta)$ for all $\mathbf{x} \in \mathbb{R}^d$.

Let $w_{n,\mathbf{x}}(\alpha)$ be the upper α percentile of the distribution of $T_\theta(\mathbf{x}, \mathbf{X})$ conditional on $\mathbf{x}$, i.e.,

$$P_{\mathbf{X}}(T_\theta(\mathbf{x}, \mathbf{X}) < w_{n,\mathbf{x}}(\alpha)|\mathbf{x}) = \alpha.$$

Then we define the $100\alpha\%$ upper generalized confidence interval for $\pi(\theta)$ as

$$\mathcal{I}_{n,\mathbf{x}}(\alpha) \equiv (-\infty, w_{n,\mathbf{x}}(\alpha)). \tag{1}$$

Similarly, we can define a lower confidence interval. Let a $100\alpha\%$ lower generalized confidence interval be defined as

$$\mathcal{I}'_{n,\mathbf{x}}(\alpha) \equiv (w'_{n,\mathbf{x}}(\alpha), \infty) \tag{2}$$

where $w'_{n,\mathbf{x}}(\alpha)$ is the lower α percentile of the distribution of $T_\theta(\mathbf{x}, \mathbf{X})$ conditional on $\mathbf{x}$.

The main objective of this paper is to investigate the frequentist coverage of this interval, i.e., $P_{\mathbf{x}}(\pi(\theta) \in \mathcal{I}_{n,\mathbf{x}}(\alpha))$. We will state and prove our results for $100\alpha\%$ upper confidence intervals and state the corresponding results for lower confidence intervals without proofs.

In order to derive expansions of the coverage probability of confidence intervals, we will assume the following general model for the random vector $\mathbf{x}$. The model is same as that for the regular case of bootstrap where the quantities involved all have absolutely continuous Lebesgue density (which is typically the setup for problems where the generalized inference methodology is useful) and hence Cramer's condition for validity of Edgeworth expansion is automatically satisfied. The statistic of interest is a vector of smooth functions of sample moments. Let $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_n$ be independent and identically distributed random column d-vectors and let $\mathbf{x} = n^{-1}\sum_{i=1}^n \mathbf{Y}_i$. In the context of a random sample $Z_1, Z_2, \ldots, Z_n$ from the distribution indexed by the parameter θ, the components of the d-vectors Y could be $Y_{ij} = g_j(Z_i)$; $j = 1, \ldots, d$, where $g_j(\cdot)$ are known functions with nonzero derivative at the expected value of Z, and Y_{ij} is the jth component of $\mathbf{Y}_i$. We make the following assumptions about the structure of the parametric problem.

(A1) $E(\|\mathbf{x}\|^4) < \infty$, where $\|\cdot\|$ denotes the usual Euclidean norm.
(A2) For each θ, the generalized pivot $T_\theta(\mathbf{x}, \mathbf{X})$ has continuous mixed partial derivatives up to order 4 in a neighborhood of $(\boldsymbol{\mu}, \boldsymbol{\mu})$.

Assumption (A2) is about the smoothness of the function $T_\theta(\mathbf{x}, \mathbf{X})$ and it is needed for a valid Edgeworth expansion of the distribution of $T_\theta(\mathbf{x}, \mathbf{X})$. Such expansions hold even in more general setup. Expansions for coverage probability of other type of nonsmooth generalized pivots is a topic of future investigation. The generalized procedure is typically applied in common parametric setups where Assumption (A1) usually hold.

We shall also need the following notation to state our results. Let $S(\mathbf{x}, \mathbf{y}) : \mathbb{R}^d \times \mathbb{R}^d \longrightarrow \mathbb{R}$ be any real valued function of two d-dimensional arguments $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$. For $\omega, \lambda \in \mathbb{N}^d$ let

$$D^{\omega,\lambda} S(\mathbf{x}, \mathbf{y}) = \frac{\partial_{(1)}^{|\omega|}}{\partial_{(1)}^{\omega_1}\mathbf{x}_1 \cdots \partial_{(1)}^{\omega_d}\mathbf{x}_d} \; \frac{\partial_{(2)}^{|\lambda|}}{\partial_{(2)}^{\lambda_1}\mathbf{y}_1 \cdots \partial_{(2)}^{\lambda_d}\mathbf{y}_d} S(\mathbf{x}, \mathbf{y}),$$

where $|\omega| = \sum_{i=1}^d \omega_i$ and $\partial_{(i)}, i = 1, 2$, denotes the partial with respect to the coordinates of the ith d-dimensional component of $S(\cdot, \cdot)$. Let $D^{\omega,0} S(\mathbf{x}, \mathbf{y}) = \frac{\partial_{(1)}^{|\omega|}}{\partial_{(1)}^{\omega_1}\mathbf{x}_1 \cdots \partial_{(1)}^{\omega_d}\mathbf{x}_d} S(\mathbf{x}, \mathbf{y})$, $D^{0,\lambda} S(\mathbf{x}, \mathbf{y}) = \frac{\partial_{(2)}^{|\lambda|}}{\partial_{(2)}^{\lambda_1}\mathbf{y}_1 \cdots \partial_{(2)}^{\lambda_d}\mathbf{y}_d} S(\mathbf{x}, \mathbf{y})$ and let $e_{i_1 \cdots i_k}$ denote a d-vector with ones at the $i_1, \ldots, i_k$th places and rest zeros. Define

$$\begin{aligned} a^{(1)}_{i_1 \ldots i_k}(\mathbf{x}, \mathbf{y}) &= D^{e_{i_1 \ldots i_k}, 0} T_\theta(\mathbf{x}, \mathbf{y}) \\ a^{(2)}_{j_1 \ldots j_l}(\mathbf{x}, \mathbf{y}) &= D^{0, e_{j_1 \ldots j_l}} T_\theta(\mathbf{x}, \mathbf{y}) \\ a^{(12)}_{i_1 \ldots i_k, j_1 \ldots j_l}(\mathbf{x}, \mathbf{y}) &= D^{e_{i_1 \ldots i_k}, e_{j_1 \ldots j_l}} T_\theta(\mathbf{x}, \mathbf{y}). \end{aligned} \tag{3}$$

Also $a^{(\cdot)}_{\cdots}(\mathbf{x}, \boldsymbol{\mu})$, $a^{(\cdot)}_{\cdots}(\boldsymbol{\mu}, \mathbf{y})$ and $a^{(\cdot)}_{\cdots}(\boldsymbol{\mu}, \boldsymbol{\mu})$ will be the values of the function $a^{(\cdot)}_{\cdots}(\mathbf{x}, \mathbf{y})$ evaluated at $\mathbf{y} = \boldsymbol{\mu}, \mathbf{x} = \boldsymbol{\mu}$ and $(\mathbf{x}, \mathbf{y}) = (\boldsymbol{\mu}, \boldsymbol{\mu})$, respectively. When there is no confusion we will write function values evaluated at $(\boldsymbol{\mu}, \boldsymbol{\mu})$ such as $a^{(\cdot)}_{\cdots}(\boldsymbol{\mu}, \boldsymbol{\mu})$ as simply $a^{(\cdot)}_{\cdots}$.

Let $\mathbf{x}_i$ denote the ith element of $\mathbf{x}$, $\beta_{i_1 \ldots i_j, n} = E\{(\mathbf{x} - \boldsymbol{\mu})_{i_1} \cdots (\mathbf{x} - \boldsymbol{\mu})_{i_j}\}$ and let

$$\beta_{i_1 \ldots i_j, n} = \mu_{i_1 \ldots i_j} + O(n^{-1}).$$

The leading terms $\mu_{i_1, \ldots, i_j}$ are computed using delta method (e.g. see the calculation in the examples given in this paper). Define the standardized (with respect to the $P_{\mathbf{X}}$ probability) quantity:

$$Z_{n,\mathbf{X}}(\mathbf{x}) = [n/A_0(\mathbf{x}, \boldsymbol{\mu})]^{1/2} (T_\theta(\mathbf{x}, \mathbf{X}) - T_\theta(\mathbf{x}, \boldsymbol{\mu})) \tag{4}$$

where

$$A_0(\mathbf{x}, \boldsymbol{\mu}) = \sum_{i=1}^d \sum_{j=1}^d a^{(2)}_i(\mathbf{x}, \boldsymbol{\mu}) a^{(2)}_j(\mathbf{x}, \boldsymbol{\mu}) \mu_{ij}. \tag{5}$$

Let

$$p_{1,\mathbf{x}}(z) = -\left\{ A_0(\mathbf{x}, \boldsymbol{\mu})^{-1/2} A_1(\mathbf{x}, \boldsymbol{\mu}) + \frac{1}{6} A_0(\mathbf{x}, \boldsymbol{\mu})^{-3/2} A_2(\mathbf{x}, \boldsymbol{\mu}) [z^2 - 1] \right\} \tag{6}$$

be the second degree even polynomial where

$$\begin{aligned} A_1(\mathbf{x}, \boldsymbol{\mu}) &= \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d a^{(2)}_{ij}(\mathbf{x}, \boldsymbol{\mu}) \mu_{ij}, \\ A_2(\mathbf{x}, \boldsymbol{\mu}) &= A_{21}(\mathbf{x}, \boldsymbol{\mu}) + 3\, A_{22}(\mathbf{x}, \boldsymbol{\mu}), \\ A_{21}(\mathbf{x}, \boldsymbol{\mu}) &= \sum_{i=1}^d \sum_{j=1}^d \sum_{k=1}^d A^{(2)}_i(\mathbf{x}, \boldsymbol{\mu}) a^{(2)}_j(\mathbf{x}, \boldsymbol{\mu}) a^{(2)}_k(\mathbf{x}, \boldsymbol{\mu}) \mu_{ijk}, \\ A_{22}(\mathbf{x}, \boldsymbol{\mu}) &= \sum_{i=1}^d \sum_{j=1}^d \sum_{k=1}^d \sum_{l=1}^d a^{(2)}_i(\mathbf{x}, \boldsymbol{\mu}) a^{(2)}_j(\mathbf{x}, \boldsymbol{\mu}) a^{(2)}_{kl}(\mathbf{x}, \boldsymbol{\mu}) \mu_{ik} \mu_{jl}. \end{aligned} \tag{7}$$

Define

$$\Delta = \frac{1}{A_0^{3/2}} \sum_{i=1}^d \sum_{j=1}^d \sum_{k=1}^d \sum_{l=1}^d \{ a^{(12)}_{i,j} a^{(2)}_k a^{(2)}_l - a^{(2)}_i a^{(12)}_{j,k} a^{(2)}_l \} \mu_{ij} \mu_{kl}. \tag{8}$$

Let $\phi(\cdot)$ and $\Phi(\cdot)$ denote the standard normal probability density function and the standard normal probability distribution function, respectively. The 100βth percentile of the standard normal distribution will be denoted by z_β.

3. Improved generalized intervals

This section contains the main results. All proofs are relegated to the Appendix. Consider the quantity

$$Z_n(\mathbf{x}) = [n/A_0(\mathbf{x}, \boldsymbol{\mu})]^{1/2} (T_\theta(\mathbf{x}, \mathbf{x}) - T_\theta(\mathbf{x}, \boldsymbol{\mu})). \tag{9}$$

To derive an expansion for the frequentist coverage of the generalized interval we will need an expansion of the distribution of $Z_n(\mathbf{x})$. Recall that we have assumed that $\mathbf{X}$ has Lebesgue density. The following theorem gives the expansion of the distribution of $Z_n(\mathbf{x})$.

Theorem 1. *Let Assumptions* (A1)–(A2) *hold. Let $Z_n(\mathbf{x})$ be the studentized statistic defined in (9). Then the distribution of $Z_n(\mathbf{x})$ has a valid Edgeworth expansion given by (up to one term)*

$$P_{\mathbf{x}}(Z_n(\mathbf{x}) < z) = \Phi(z) + n^{-1/2} q_{1,\boldsymbol{\mu}}(z)\phi(z) + O(n^{-1}), \tag{10}$$

where $q_{1,\boldsymbol{\mu}}(z) = p_{1,\boldsymbol{\mu}}(z) - \Delta$, with $p_{1,\boldsymbol{\mu}}(\cdot)$ defined in (6).

Remark 1. If the components of $\mathbf{x}$ are asymptotically independent, i.e., $\mu_{ij} = 0$ if $i \neq j$, then the expression for Δ reduces to

$$\Delta = A_0^{-3/2} \sum_{i=1}^{d} \sum_{k=1}^{d} \{a_{ii}a_k^2 - a_i a_k a_{ik}\}\mu_{ii}\mu_{kk}. \tag{11}$$

The following result is needed in the proof of the subsequent theorems and is reminiscent of expansion results in bootstrap, with added claims on the rates in (13) and (15). These error rates are obtained by keeping track of the relevant terms in the expansion.

Theorem 2. *Let Assumptions* (A1)–(A2) *hold. Then for each fixed $\mathbf{x} \in \mathbb{R}^d$ we have $n^{1/2}(T_\theta(\mathbf{x}, \mathbf{X}) - T_\theta(\mathbf{x}, \boldsymbol{\mu})) \xrightarrow{\mathcal{L}} N(0, A_0(\mathbf{x}, \boldsymbol{\mu}))$ where $A_0(\mathbf{x}, \boldsymbol{\mu})$ is defined in (5). Let $F_{n,\mathbf{X},\mathbf{x}}(\cdot)$ be the distribution function of $Z_{n,\mathbf{X}}(\mathbf{x})$. Then $F_{n,\mathbf{X},\mathbf{x}}(\cdot)$ has a valid Edgeworth expansion for every $\mathbf{x} \in \mathbb{R}^d$ given by (up to order $n^{-1/2}$ term)*

$$F_{n,\mathbf{X},\mathbf{x}}(z) = P_{\mathbf{X}}(Z_{n,\mathbf{X}}(\mathbf{x}) < z) = \Phi(z) + n^{-1/2} p_{1,\mathbf{x}}(z)\phi(z) + \epsilon_n(\mathbf{x}; z) \tag{12}$$

where $p_{1,\mathbf{x}}(\cdot)$ is defined in (6). For $\lambda > 1/2$, there exists a constant C_1 such that

$$P_{\mathbf{x}}\Big(\sup_{-\infty<z<\infty} |\epsilon_n(\mathbf{x}; z)| > C_1 n^{-1/2}\Big) = O(n^{-\lambda}). \tag{13}$$

Also, for any $\delta \in (0, 1/2)$, there exist constants $\epsilon > 0$ and $C_2 > 0$ such that the percentiles of the distribution of $Z_{n,\mathbf{X}}(\mathbf{x})$ admit Cornish–Fisher expansions (up to order $n^{-1/2}$ term) of the form

$$F^{-1}_{n,\mathbf{X},\mathbf{x}}(\beta) = z_\beta - n^{-1/2} p_{1,\mathbf{x}}(z_\beta) + C_n(\mathbf{x}, \beta), \tag{14}$$

uniformly in β for $n^{-\epsilon} \leq \beta \leq 1 - n^{-\epsilon}$ and the remainder term satisfies

$$P_{\mathbf{x}}\Big(\sup_{n^{-\epsilon}<\beta<1-n^{-\epsilon}} |C_n(\mathbf{x}; \beta)| > C_2 n^{-1+\delta}\Big) = O(n^{-\lambda}). \tag{15}$$

The results provide an asymptotic expansion of the coverage probability of the generalized confidence interval.

Theorem 3. *Let $1/2 < \alpha < 1$ be fixed. Suppose the assumptions of Theorem 1 hold. Then the coverage probability of the $100\alpha\%$ upper generalized confidence interval $\mathcal{I}_{n,\mathbf{x}}(\alpha)$ defined in (1) is given by*

$$P_{\mathbf{x}}[\pi(\theta) \in \mathcal{I}_{n,\mathbf{x}}(\alpha)] = \alpha - n^{-1/2}\Delta\phi(z_\alpha) + o(n^{-1/2}), \tag{16}$$

where Δ is defined in (8).

Analogous results can be obtained for $100\alpha\%$ lower generalized confidence intervals.

Corollary 1. *Let $1/2 < \alpha < 1$ be fixed. Suppose assumptions of Theorem 1 hold. Let Δ be as defined in (8). Then the coverage probability of the $100\alpha\%$ lower generalized confidence interval $\mathcal{I}'_{n,\mathbf{x}}(\alpha)$ defined in (2) is given by*

$$P_{\mathbf{x}}[\pi(\theta) \in \mathcal{I}'_{n,\mathbf{x}}(\alpha)] = \alpha + n^{-1/2}\Delta\phi(z_\alpha) + o(n^{-1/2}). \tag{17}$$

Remark 2. Note that as $n \to \infty$, $w_{n,\mathbf{x}}(\alpha)$ becomes larger than $w'_{n,\mathbf{x}}(\alpha)$. Thus, using Eqs. (16) and (17) we have

$$\begin{aligned} P_{\mathbf{x}}[\pi(\theta) \in [w'_{n,\mathbf{x}}(\alpha), w_{n,\mathbf{x}}(\alpha)]) &= 1 - (1 - \alpha + n^{-1/2}\Delta\phi(z_\alpha) + o(n^{-1/2}) + 1 - \alpha - n^{-1/2}\Delta\phi(z_\alpha) - o(n^{-1/2})) \\ &= 2\alpha - 1 + o(n^{-1/2}). \end{aligned} \tag{18}$$

Therefore, the two-sided interval is first order accurate.

One may improve the accuracy further as follows. We only sketch the main steps involved. The details will be similar to the arguments given in Hall [26], Chapter 3. Eq. (17) is sharpened to

$$P_{\mathbf{x}}[\pi(\theta) \in \mathcal{I}_{n,\mathbf{x}}(\alpha)] = \alpha - n^{-1/2}\Delta\phi(z_\alpha) + n^{-1}S(\theta, z_\alpha) + o(n^{-1}).$$

Similarly, Eq. (16) is sharpened. Using the standard methods of modifying confidence intervals based on estimates for Δ and $S(\theta, z_\alpha)$ [see for example Hall [26], Chapter 3] one may define modified confidence intervals $\mathcal{J}_{n,\mathbf{x}}(\alpha)$ and $\mathcal{J}'_{n,\mathbf{x}}(\alpha)$ such that

$$P_{\mathbf{x}}[\pi(\theta) \in \mathcal{J}_{n,\mathbf{x}}(\alpha)] = \alpha - n^{-1}S_2(\theta, z_\alpha) + o(n^{-1}),$$
$$P_{\mathbf{x}}[\pi(\theta) \in \mathcal{J}'_{n,\mathbf{x}}(\alpha)] = \alpha + n^{-1}S_2(\theta, z_\alpha) + o(n^{-1}),$$

This leads to a two-sided interval $[J', J]$ of accuracy $o(n^{-1})$. The $o(n^{-1})$ can be shown to be $O(n^{-3/2})$ under mild additional conditions, but the algebra is tedious and the gains are negligible in moderate samples.

Generally Δ is unknown. But a suitable estimate of Δ may be used to define modified intervals with improved accuracy. Let $\hat{\Delta}$ be an estimator of Δ. Let $\alpha_0 > 0$ be given and $\hat{\alpha}_n$ be a solution to

$$\alpha - n^{-1/2}\hat{\Delta}\phi(z_\alpha) = \alpha_0. \tag{19}$$

Theorem 4. *Let $\hat{\Delta}$ be a $\sqrt{n}$-consistent estimator of Δ. Let $\hat{\alpha}_n$ be a solution to Eq. (19) and let $\tilde{\alpha}_n = \max\{n^{-\epsilon}, \min(\hat{\alpha}_n, 1 - n^{-\epsilon})\}$ where ϵ is defined in Theorem 2. Let the assumptions of Theorem 3 hold. Then*

$$P_{\mathbf{x}}[\pi(\theta) \in \mathcal{I}_{n,\mathbf{x}}(\tilde{\alpha}_n)] = \alpha_0 + o(n^{-1/2}).$$

Remark 3. Note that $\hat{\alpha}_n = \alpha_0 + O_p(n^{-1/2})$. Hence, for all practical problems with moderate sample size $\tilde{\alpha}_n = \hat{\alpha}_n$ and thus the exact value of ϵ is irrelevant in applications.

4. The role of Δ

In this section we explore the role of Δ and its relation to coverage and the nature of the pivots. We show that in some cases Δ is free of parameters and may even equal zero. In others, it may be estimated and be used to improve the coverage. The gains can be significant even in small samples.

Let $A = ((a_{ij}^{(12)}))$, $\Sigma = ((\mu_{ij}))$ and $a = (a_1^{(2)}, \ldots, a_d^{(2)})'$. For notational simplicity we will omit the superscripts (1), (2) and (12) in the notation of the derivatives of the pivot. Let $B = 0.5\Sigma^{1/2}(A + A')\Sigma^{1/2}$ be the symmetric version of $\Sigma^{1/2}A\Sigma^{1/2}$ and let $b = (a'\Sigma a)^{-1/2}\Sigma^{1/2}a$. Then

$$\Delta = A_0^{-1/2}\Big[\mathrm{tr}(B) - b'Bb\Big]. \tag{20}$$

Note that $A_0 = (a'\Sigma a)$ and $\|b\| = 1$. The quantities B and b are parametric functions and we can write them as $B(\theta)$ and $b(\theta)$, respectively. Eq. (20) is essentially a way of rewriting the numerator of Δ in the matrix form. However, the expression is compact and insightful. We may now state the necessary and sufficient condition for the generalized intervals to be first order accurate as

Proposition 1. *A necessary and sufficient condition for $\Delta = 0$ is*

$$\mathrm{tr}\{B(\theta)\} = b(\theta)'B(\theta)b(\theta), \quad \text{for the true value } \theta. \tag{21}$$

Note that the above criterion depends on the unknown true parameter θ. In practice, to get an idea whether the generalized intervals are going to be accurate in a specific problem, one may consider devising a suitable test for testing $\mathrm{tr}\{B(\theta)\} - b(\theta)'B(\theta)b(\theta)$ equal to zero.

4.1. $\Delta = 0$ and exact frequentist coverage

When do generalized confidence intervals have exact frequentist coverage? A sufficient condition is the following.

Proposition 2. *Let $T_\theta(\mathbf{x}, \mathbf{X})$ be a generalized pivot for the parametric function $\pi(\theta)$. Then the generalized confidence interval constructed for $\pi(\theta)$ based on $T_\theta(\mathbf{x}, \mathbf{X})$ has exact frequentist coverage and the corresponding Δ is equal to zero if the following hold.*

(i) *There exists a function $\psi(\mathbf{x}, y, z) : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ such that for every $\mathbf{x} \in \mathbb{R}^d$, $\psi(\mathbf{x}, y, z) < 0$ iff $y < z$.*

(ii) *There exists a function $\tau_\theta(\mathbf{x}) : \mathbb{R}^d \to \mathbb{R}$ such that $\tau_\theta(\mathbf{x})$ has continuous distribution and $T_\theta(\mathbf{x}, \mathbf{X}) - T_\theta(\mathbf{x}, \mathbf{x}) = \psi(\mathbf{x}, \tau_\theta(\mathbf{x}), \tau_\theta(\mathbf{X}))$.*

Proof. That the generalized confidence interval based on $T_\theta(\mathbf{x}, \mathbf{X})$ has exact coverage in this case follows easily from the fact that $P_{\mathbf{X}}(T_\theta(\mathbf{x}, \mathbf{X}) < \pi(\theta)) = P_{\mathbf{X}}(T_\theta(\mathbf{x}, \mathbf{X}) < T_\theta(\mathbf{x}, \mathbf{x})) = P_{\mathbf{X}}(\psi(\mathbf{x}, \tau_\theta(\mathbf{x}), \tau_\theta(\mathbf{X})) < 0) = P_{\mathbf{X}}(\tau_\theta(\mathbf{x}) < \tau_\theta(\mathbf{X}))$ is uniformly distributed on $[0, 1]$ under $P_{\mathbf{x}}$ probability. For exact coverage the first order term is necessarily zero. However, it is insightful to show how the structural assumptions about $T_\theta(\mathbf{x}, \mathbf{X})$ lead to $\Delta = 0$. Let b_i denote the derivative $\frac{\partial \tau_\theta(\mathbf{x})}{\partial \mathbf{x}_i}$ evaluated at $\mathbf{x} = \boldsymbol{\mu}$. Then the derivatives of the pivot are given by $a_i^{(2)} = -\psi_1^{(2)} b_i$; $a_{i,j}^{(2)} = -\psi_{i,1}^{(12)} b_j - \psi_{11}^{(2)} b_i b_j$. Hence the numerator of Δ is

$$\begin{aligned}\sum_{i,j,k,l} [a_{i,j} a_k a_l - a_i a_{j,k} a_l] \mu_{ij} \mu_{kl} &= \sum_{i,j,k,l} \Big[(\psi_{j,1}^{(12)} b_k + \psi_{11}^{(2)} b_j b_k) b_i - (\psi_{i,1}^{(12)} b_j + \psi_{11}^{(2)} b_i b_j) b_k \Big] b_l \mu_{ij} \mu_{kl} \\ &= \sum_{i,j,k,l} (\psi_{j,1}^{(12)} b_i - \psi_{i,1}^{(12)} b_j) b_k b_l \mu_{ij} \mu_{kl} = 0. \quad \square\end{aligned}$$

The proposition supplements Remark 7 in Hannig et al. [24]. For the t-statistic the function $\psi(\mathbf{x}, y, z) = x_2(y - z)$ where $\mathbf{x} = (x_1, x_2)$ and $\tau_\theta(\mathbf{x})$ is the t-statistic, $(x_1 - \mu)/x_2$. Thus, for the t-statistic case the conventional pivot naturally leads to a generalized pivot. In general, if $g_{\mathbf{x}}(\pi(\theta))$ is a conventional pivot for $\pi(\theta)$ which is invertible, then $T_\theta(\mathbf{x}, \mathbf{X}) = g_{\mathbf{x}}(g_{\mathbf{X}}^{-1}(\pi(\theta)))$ is a generalized pivot for $\pi(\theta)$.

4.2. $\Delta = 0$, *but no exact coverage*

Of course, $\Delta = 0$ does not guarantee exact frequentist coverage. However, $\Delta = 0$ leads to very good performance of the generalized intervals. We illustrate this using the following example.

Example 1 (*Behrens–Fisher Problem*). This is a well-analyzed problem in statistics where the objective is to construct a confidence interval for the difference of means of two normal populations and the variances of the populations are not known. No exact frequentist confidence intervals are available.

Let $Z_{ij} \sim N(\tau_i, \sigma_i^2)$ for $i = 1, 2$ for the test group and the reference group, respectively, and $j = 1, \ldots, n$. The parameter vector is $\theta = (\tau_1, \tau_2, \sigma_1^2, \sigma_2^2)$ and the parameter of interest is $\pi_1(\theta) = \tau_1 - \tau_2$. The statistics used in the construction of the interval are $\mathbf{X} = (X_1, X_2, X_3, X_4)' := (\bar{Z}_1, \bar{Z}_2, S_1^2, S_2^2)'$ where $\bar{Z}_i = n^{-1} \sum_{j=1}^n Z_{ij}$ and $S_i^2 = (n-1)^{-1} \sum_{j=1}^n (Z_{ij} - \bar{Z}_i)^2$. Let $\mathbf{x}$ be the observed value of $\mathbf{X}$. The asymptotic distributional result in this case is $\sqrt{n}(\mathbf{X} - \boldsymbol{\mu}) \xrightarrow{\mathcal{L}} N_4(0, \Sigma)$, where $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3, \mu_4)' = \theta$ and $\Sigma = \big((\mu_{ij})\big) = \mathrm{Diag}(\mu_3, \mu_4, 2\mu_3^2, 2\mu_4^2)$. The obvious generalized pivot is

$$T_\theta(\mathbf{x}, \mathbf{X}) = (x_1 - x_2) - \left[\frac{(X_1 - \mu_1)\sqrt{x_3}}{\sqrt{X_3}} - \frac{(X_2 - \mu_2)\sqrt{x_4}}{\sqrt{X_4}} \right]. \tag{22}$$

The derivatives of the generalized pivot (algebraic details are omitted) are given by $(a_1^{(2)}, a_2^{(2)}, a_3^{(2)}, a_4^{(2)}) = (-1, 1, 0, 0)$ and

$$\left((a_{i,j}^{(12)}) \right) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\dfrac{1}{2\mu_3} & 0 & 0 & 0 \\ 0 & \dfrac{1}{2\mu_4} & 0 & 0 \end{pmatrix}. \tag{23}$$

Substituting the values of the derivatives and the asymptotic variances we find that $\Delta = 0$. A generalized interval yields good results (Tsui and Weerahandi [1]). The interval does not have exact frequentist coverage. However, from the above calculations, the intervals are first order accurate. This explains the near exact performance of the generalized intervals for this example.

4.3. *Magnitude of* Δ

Generalized confidence intervals often perform quite well in small samples even when they are not first order correct (that is $\Delta \neq 0$). What then could be the reason for their good performance? Actually in many common problems (e.g. Examples 2 and 3 in this article), the magnitude of the first order term, $|n^{-1/2} \Delta \phi(z_\alpha)|$, is bounded by a small quantity and thus the coverage error is insignificant even for moderate sample sizes. From (20) we see that Δ is contained in the interval $A_0^{-1/2} \Big[\sum_{i=1}^{d-1} \lambda_i(B), \sum_{i=2}^{d} \lambda_i(B) \Big]$ where $\lambda_1(B) < \cdots < \lambda_d(B)$ are the eigenvalues of B. Thus, a very conservative upper bound for $|\Delta|$ would be

$$|\Delta| < d(d-1) M_1 M_2 \tag{24}$$

where $M_1 = \max_{i,j} (\mu_{ii}\mu_{jj}/A_0)^{1/2}$ and $M_2 = \max_{i,j} 0.5[a_{ij} + a_{ji}]$. In practice, the term may be much smaller than the bound, and hence the first order term may be negligible.

We illustrate this by two examples. In order to do the necessary derivations for the examples, we need the following facts which are obvious consequences of the Edgeworth expansions for smooth functions of sample means.

In many applications, the function of interest, $A(\bar{\mathbf{x}})$, where $\bar{\mathbf{x}} = (\bar{x}_1, \ldots, \bar{x}_d)$ is the mean of a d-variate random vector, can be alternatively written as a function, $\tilde{A}(\bar{\mathbf{y}})$ where $\bar{\mathbf{y}} = (\bar{y}_1, \ldots, \bar{y}_r)$ and each $\bar{y}_i$ is of the form $\bar{y}_i = n_i^{-1} \sum_{j=1}^{n_i} y_{i,j}$ and $n_i = c_i n[1+O(n^{-1})]$ for some positive integers c_i. The y variables may be latent and not directly observable. In many of these problems, it may be easier to derive the coefficients of the polynomials in the Edgeworth expansion for $P(n^{1/2}\tilde{A}(\bar{\mathbf{y}}) < w)$ than those in the Edgeworth expansion of $P(n^{1/2}A(\bar{\mathbf{x}}) < w)$ because of the simpler moment structure of the $y_{i,j}$ compared to those of $x_{i,j}$. For example, the $\bar{y}_i$ could be independent for different i.

However, the different component of $\bar{\mathbf{y}}$ may be means of possibly unequal number of observations and the number of observations may be potentially of different order. This does not pose a problem while deriving an Edgeworth expansion for $n^{1/2}\tilde{A}(\bar{\mathbf{y}})$ as long as one is interested in an expansion up to the order $n^{-1/2}$. This is evident from the fact that the cumulant expressions of $\bar{\mathbf{y}}$ agree up to order n^{-1} with those of $\bar{\mathbf{y}}_{(n)}$ where $\bar{\mathbf{y}}_{(n)} = (\bar{y}_{1,(n)}, \ldots, \bar{y}_{r,(n)})$, $\bar{y}_{i,(n)} = (c_i n)^{-1} \sum_{j=1}^{c_i n} y_{i,j}$. Further, the fact that the order of the means, $c_i n$, are different for different i can be taken care of by differentially pooling the observations for the different means, $\bar{y}_{i,(n)}$ and adjusting the moment expressions for the means while deriving the coefficients of the polynomials in the expansion.

For example, let $z_{i,j} = c_i^{-1} \sum_{k=1}^{c_i} y_{i,(j-1)c_i+k}$ and let $\bar{z}_i = n^{-1} \sum_{j=1}^{n} z_{i,j}$. Define $\bar{\mathbf{z}} = (\bar{z}_1, \ldots, \bar{z}_r)$. Then the components of the mean vector $\bar{\mathbf{z}}$ are means of equal number of observations and the function $\tilde{A}(\bar{\mathbf{y}}_{(n)})$ can be written as $\tilde{A}(\bar{\mathbf{z}})$. Furthermore, $\mu_i^y = E(y_{i,j}) = E(z_{i,j}) = \mu_i^z$. Also $\mu_{ij}^z = E[(z_{i,k} - \mu_i^z)(z_{j,l} - \mu_j^z)] = \delta_{ij}\mu_{ij}^y$ where $\delta_{ij} = (c_i)^{-1}$ if $i = j$ and 1 if $i \neq j$. Higher order moments should be adjusted in a similar way.

For illustration consider the following example of a studentized statistic for the two sample mean problem in a normal model. Let $u_1, \ldots, u_n$ be iid $N(\mu_1, \sigma^2)$ and $v_1, \ldots, v_n$ be iid $N(\mu_2, \sigma^2)$ and the samples are mutually independent. Then the studentized statistic for testing $\mu_1 = \mu_2$ is $A(\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4) = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{[n\{(\bar{x}_3 + \bar{x}_4) - (\bar{x}_1^2 + \bar{x}_2^2)\}/(2n-2)]^{1/2}}$ where $\bar{x}_1 = n^{-1} \sum_{i=1}^{n} u_i$, $\bar{x}_2 = n^{-1} \sum_{i=1}^{n} v_i$, $\bar{x}_3 = n^{-1} \sum_{i=1}^{n} u_i^2$, $\bar{x}_4 = n^{-1} \sum_{i=1}^{n} v_i^2$. However note that $\bar{x}_1 - \bar{x}_2 = \bar{y}_1$ where $\bar{y}_1 = n_1^{-1} \sum_{j=1}^{n} y_{1,j}$ and $y_{1,j}$ are iid $N(\mu_1 - \mu_2, 2\sigma^2)$. Also, $n\{(\bar{x}_3 + \bar{x}_4) - (\bar{x}_1^2 + \bar{x}_2^2)\}/(2n-2)$ is the mean of $n_2 = 2(n-1)$ iid random variables each of which is distributed as σ^2 times a χ^2 random variable with one degree of freedom. Thus, in this example $c_1 = 1$ and $c_2 = 2$ with $n_i = 2n(1 + O(n^{-1}))$. In order to reduce means to averages of equal (or almost equal) number of random variable we can use $\bar{z}_1 = \bar{y}_1$ and $\bar{z}_2$ which is a mean of $(n-1)$ iid random variables each of which are distributed as $\sigma^2/2$ times a χ^2 random variable with two degrees of freedom. The function can be now written as $\tilde{A}(\bar{z}_1, \bar{z}_2) = \frac{\bar{z}_1 - (\mu_1 - \mu_2)}{\sqrt{\bar{z}_2}}$. The biggest advantage is that now the variables $\bar{z}_1$ and $\bar{z}_2$ are independent and the coefficients of the polynomial $p_{1,\boldsymbol{\mu}}$ in the Edgeworth expansion reduce to simpler expressions.

Example 2 (*One Way Random Model*). Consider the one-way random effect model $Y_{ij} = \beta_0 + \beta_i + \epsilon_{ij}, i = 1, \ldots, k; j = 1, \ldots, n$, where $\beta_i \sim N(0, \sigma_\beta^2)$, $\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$ and $\{\beta_i\}$ and $\{\epsilon_{ij}\}$ are mutually independent. Note that the asymptotics in this problem are with respect to the number of groups, k. Define the between groups mean square (S_B^2) and within group mean square (S_W^2) as

$$S_B^2 = \frac{n \sum_{i=1}^{k} (\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2}{k-1}, \qquad S_W^2 = \frac{\sum_{i=1}^{k} \sum_{j=1}^{n} (\bar{Y}_{ij} - \bar{Y}_{i\cdot})^2}{k(n-1)},$$

where $\bar{Y}_{i\cdot} = n^{-1} \sum_{j=1}^{n} Y_{ij}$ and $\bar{Y}_{\cdot\cdot} = k^{-1} \sum_{i=1}^{k} \bar{Y}_{i\cdot}$. Let $\mu_1 = (\sigma_\epsilon^2 + n\sigma_\beta^2)$ and $\mu_2 = \sigma_\epsilon^2$. Suppose the parameter of interest is the variance component $\pi(\theta) = \sigma_\beta^2 = n^{-1}(\mu_1 - \mu_2)$. There are no lower confidence interval for $\pi(\theta)$ that has exact frequentist coverage but there are many approximate intervals available in the literature (e.g. see Weerahandi [5] page 90). Note that $(k-1)S_B^2 \sim \mu_1 \chi_{k-1}^2$ and $k(n-1)S_W^2 \sim \mu_2 \chi_{k(n-1)}^2$. Based on these distributions, Weerahandi ([4] pp 152) proposed a generalized pivot

$$T_\theta(\mathbf{x}, \mathbf{X}) = n^{-1} \left[\frac{\mu_1 x_1}{X_1} - \frac{\mu_2 x_2}{X_2} \right], \tag{25}$$

where $\mathbf{x} = (x_1, x_2) = (s_B^2, s_W^2)$ and $\mathbf{X} = (X_1, X_2) = (S_B^2, S_W^2)$ are the observed values and a random copy of the mean squares, respectively. Hannig, et al. [24] have shown that the generalized confidence interval based on (25) has exact asymptotic coverage. However, the actual coverage could be different from the nominal level in small samples and depends typically on the quantity $0 < \lambda = \mu_2/\mu_1 < 1$. From the discussion in Section 4.3, we can derive the Edgeworth expansion in terms of (x_1, x_2, X_1, X_2) as long as we recognize that x_1 and X_1 are means of $(k-1) = k[1 + O(k^{-1})]$ iid random variables and (x_2, X_2) are means of k iid random variables each of which are means of $(n-1)$ iid random variables distributed as σ_ϵ^2 times χ^2 random variables with one degree of freedom. We have

$$\sqrt{k} \begin{pmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \end{pmatrix} \sim N_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \mu_{11} & \mu_{12} \\ \mu_{21} & \mu_{22} \end{pmatrix} \right)$$

where $\mu_{11} = 2\mu_1^2$; $\mu_{22} = 2\mu_2^2/(n-1)$ and $\mu_{12} = \mu_{21} = 0$. By computing and plugging in the derivatives, the expression for Δ reduces to

$$\Delta = \frac{n^{-3}[\mu_2^{-1} - \mu_1^{-1}]\mu_{11}\mu_{22}}{n^{-3}(\mu_{11}+\mu_{22})^{3/2}} = \frac{\sqrt{2}\lambda(1-\lambda)}{(n-1)(1+\lambda^2/(n-1))^{3/2}}.$$

For any pair (k, n) the first order term in the coverage error is bounded above by $\sqrt{2}\phi(z_\alpha)/[4(n-1)\sqrt{k}]$ which is much smaller than the conservative bound given in (24). The term is decreasing in α. However, even for α as small as 0.8 (which corresponds to a 80% confidence interval) the first order term is bounded above by $[10(n-1)\sqrt{k}]^{-1}$. As long as $(n-1)\sqrt{k} > 10$ we have $[10(n-1)\sqrt{k}]^{-1} < 0.1$. Thus, even for small sample sizes (e.g. $(n, k) = (6, 4)$) the coverage error is expected to be around 1%.

Example 3 (*Average Bioequivalence*). One of the criteria approved by the U.S. Food and Drug Administration (FDA) is that of average bioequivalence. In average bioequivalence, the parameter of interest is $\pi(\theta) = \mu_T/\mu_R$ where μ_T and μ_R are the mean responses of the test drug and the reference drug, respectively. As per FDA guidelines, the logarithm of the responses are analyzed and they are generally believed to be normally distributed. Generalized inference for some of the problems related to bioequivalence has been discussed in McNally et al. [15].

This specific statistical model is $Z_{ij} = \log Y_{ij} \sim N(\tau_i, \sigma_i^2)$ where $i = 1, 2$ for the test group and the reference group, respectively, and $j = 1, \ldots, n$. The parameter of interest is $\pi_1(\theta) = \exp\{\tau_1 - \tau_2 + \frac{1}{2}(\sigma_1^2 - \sigma_2^2)\}$, the ratio of the two lognormal means. Since the intervals are equivariant under monotone transformation, we will take $\pi(\theta) = [\tau_1 - \tau_2 + \frac{1}{2}(\sigma_1^2 - \sigma_2^2)]$ as the parameter of interest. The statistics involved are described in the Behrens–Fisher example. The asymptotic mean vector for the statistics is $(\mu_1, \mu_2, \mu_3, \mu_4) = (\tau_1, \tau_2, \sigma_1^2, \sigma_2^2)$. The obvious generalized pivot is

$$T_\theta(\mathbf{x}, \mathbf{X}) = (x_1 - x_2) - \left[\frac{(X_1-\mu_1)\sqrt{x_3}}{\sqrt{X_3}} - \frac{(X_2-\mu_2)\sqrt{x_4}}{\sqrt{X_4}}\right] + \frac{1}{2}\left(\frac{\mu_3 x_3}{X_3} - \frac{\mu_4 x_4}{X_4}\right).$$

After some algebra we have,

$$\Delta = \left[\frac{\frac{1}{2}(\mu_3-\mu_4)(\mu_3\mu_4 - \mu_3 - \mu_4)}{\{\mu_3 + \mu_4 + \frac{1}{2}(\mu_3^2+\mu_4^2)\}^{3/2}}\right]. \tag{26}$$

Therefore, Δ can be both negative and positive. From Proposition 3 in Appendix, we have $|\Delta| < 0.5$. The magnitude of the first order term at $\alpha = 0.9$ is bounded by $1/\sqrt{n}$. Therefore for sample sizes $n \geq 10$, the contribution of the first order term is no more than 1% toward the coverage error of the generalized intervals.

4.4. $\Delta \neq 0$

In this section, we discuss two scenarios. If Δ is known, then one can modify the interval in a straightforward way. If Δ is unknown, it needs to be first estimated and then the modification carried out. We give examples of both situations.

Case I: $\Delta \neq 0$ but known.

In this case, if the intended nominal level is α_0 then from Theorem 4 any percentile α_n of the generalized pivot distribution, $F_{n,\mathbf{X},\mathbf{x}}(\cdot)$, that satisfies the equation

$$\alpha - n^{-1/2}\Delta\phi(z_\alpha) = \alpha_0, \tag{27}$$

will yield upper confidence intervals with coverage probabilities that are first order accurate. Here is an example.

Example 4. Consider a system consisting of components connected in series and each component has an exponentially distributed survival time and an exponentially distributed repair time. Klion [27] and Jobe [28] have proposed a performance measure called mean corrective maintenance time per average unit operating time (MTUT) for such systems. The measure is defined as follows. Suppose there are d components connected in series and the mean failure time and the mean repair time for the ith component are λ_i and ν_i, respectively. Then the MTUT is defined as

$$M = \sum_{i=1}^{d}(\lambda_i/\nu_i).$$

Practitioners are interested in estimating M and obtaining upper confidence bounds for it. The measure M is closely related to the summary measure, called Availability, for system reliability and maintainability; see Knezevic [29]. Ananda [30] applied generalized procedure for computing confidence interval for availability and the results reported in [30] show that the generalized procedure works very well and the generalized confidence intervals for availability have near exact coverage probability for the parameter values investigated in that paper. We found that the frequentist coverage of generalized intervals for availability to be not as good for other parameter values. For instance, if there are many components ($d > 4$) and if the ratio of the mean repair time to the mean survival time is high (>0.3) for each component, then the difference of the nominal coverage and the intended coverage for the generalized confidence interval for availability could be as high as 6%. In this section, however, we only give results for the measure M. There are no exact confidence bounds available

Table 1
Performance of the corrected generalized intervals for Example 4.

α_0	$n=5$			$n=10$		
	MLE	Uncorrected	Corrected	MLE	Uncorrected	Corrected
0.800	0.782	0.887	0.788	0.800	0.876	0.801
0.850	0.822	0.921	0.836	0.833	0.913	0.857
0.900	0.866	0.954	0.892	0.877	0.944	0.903
0.950	0.915	0.984	0.942	0.919	0.972	0.946

in the literature for M. Jobe and David [31] proposed Buehler confidence bounds for M. Generalized confidence intervals provide an easy solution. However, the generalized bounds are typically very conservative. We will now investigate the performance of the modified generalized bounds by evaluating Δ and making the modification for a first order correction. Suppose the number of failure and repair time observations available for the ith component is n_i. We work out the details for the case when the number of observations for failure time is equal to the number of observations for repair time for each observation; however the derivation can be easily extended to the general case when there are possibly unequal number of observations for failure time and repair time for each component. We also assume that there are positive integers $c_1, \ldots, c_d$ whose greatest common factor is one and $n_i = nc_i$ for some positive integer n. For deriving the large sample correction we will assume $n \to \infty$. Following the arguments in Section 4.2.1, the general formula for Edgeworth expansion holds up to the $n^{-1/2}$ term provided the samples from the ith components are blocked into n blocks of c_i observations and the block averages are used as the modified observations. Thus the failure time observations associated with the ith component have mean λ_i and variance λ_i^2/c_i and the corresponding repair times have mean ν_i and variance ν_i^2/c_i. Let $x_{\lambda,1}, \ldots, x_{\lambda,d}$ be the sample mean of the failure times of the d components and $x_{\nu,1}, \ldots, x_{\nu,d}$ be those for the repair times. Let $X_{\lambda,1}, \ldots, X_{\lambda,d}$ and $X_{\nu,1}, \ldots, X_{\nu,d}$ be the corresponding copies. Here the dimension of the statistic is $D = 2d$ and the statistic is $\mathbf{x} = (x_{\lambda,1}, x_{\nu,1}, \ldots, x_{\lambda,d}, x_{\nu,d})$. Then the generalized pivot is defined as $T_\theta(\mathbf{x}, \mathbf{X}) = \sum_{i=1}^{d} \frac{X_{\lambda,i}\, x_{\nu_i}}{X_{\nu,i}\, x_{\lambda,i}}$. Next we identify the quantities needed for the computation of Δ. Since the samples are independent, we have $\mu_{ij} = 0$ if $i \neq j$, and $\mu_{ii} = \lambda_k^2/c_k$ if $i = 2k-1$ and $\mu_{ii} = \nu_k^2/c_k$ if $i = 2k$ for $k = 1, 2, \ldots, d$. Also, $a_i^{(2)} = -\lambda_i^{-1}$ if i is odd and $a_i^{(2)} = \nu_i^{-1}$ if i is even. Moreover,

$$a_{i,j}^{(12)} = \begin{cases} -\lambda_i^{-2}, & \text{if } i = j;\ i \text{ is odd}, \\ -\nu_i^{-2}, & \text{if } i = j;\ i \text{ is even}, \\ \lambda_i^{-1}\nu_i^{-1}, & \text{if } j = i+1;\ i \text{ is odd or } j = i-1;\ i \text{ is even}, \\ 0, & \text{otherwise}. \end{cases} \tag{28}$$

The asymptotic variance of the pivot is

$$A_0 = \sum_{i=1}^{d} [a_{2i-1}^{(2)} a_{2i-1}^{(2)} \mu_{2i-1,2i-1} + a_{2i}^{(2)} a_{2i}^{(2)} \mu_{2i,2i}] = \sum_{i=1}^{d} \frac{2}{c_i}.$$

Similarly, the numerator of Δ can be shown to be $-\sum\sum_{i\neq j} \frac{4}{c_i c_j}$. Therefore, $\Delta = -[\sum\sum_{i\neq j} \frac{4}{c_i c_j}]/[\sum_{i=1}^{d} \frac{2}{c_i}]^{-3/2}$ which does not depend on the unknown parameters but is nonzero and is a function of the sample size. Let us consider a specific case when $d = 2$ and let $n_1 = n$ and $n_2 = 2n$. For this example, $c_1 = 1$ and $c_2 = 2$. Then, $\Delta = -\frac{4}{3\sqrt{3}}$. Thus, for a system with two components if the number of observations for one component is nearly double that of the other component, the generalized confidence interval should be constructed using the α percentile of the generalized pivot distribution where α is a solution to

$$\alpha + \frac{4}{3\sqrt{3n}} \phi(z_\alpha) = \alpha_0.$$

Table 1 gives the results of a simulation experiment with $d = 2, n_1 = 5, 10, n_2 = 2n_1$ and the parameters $[\lambda_1, \nu_1, \lambda_2, \nu_2] = [2, 0.1, 5, 0.05]$. The simulation are based on 1000 samples and the percentiles of the generalized pivot distribution are computed based on 10,000 replications. For comparison, the coverage probability constructed based on the asymptotic normality of the MLE is also given under the column label 'MLE'. The corrected generalized intervals have better coverage probability than both, the intervals based on MLE and the uncorrected generalized intervals.

Case II: Δ unknown

Typically the value of Δ is a function of the parameters and hence unknown. Correction of generalized intervals based on an estimate of Δ is useful particularly for examples where Δ is large and an efficient estimator of Δ is available. Given that Δ is estimated from the sample, for small samples the estimation error may outweigh the possible gain from correcting the first order term in the coverage error. However, there are a number of examples where one would benefit from modifying the usual generalized intervals to make them first order correct.

Example 5. In economics and health research we find examples of data arising that are skewed and generally modeled as normally distributed quantity after log transformation. Even though the median is the more natural measure to analyze for skewed data, often practitioners are interested in the mean for such data as well. If multiple groups are involved then ratios

Table 2
Modified generalized intervals vs the usual generalized intervals in Example 5.

α_0	$n = 25$		$n = 50$	
	Uncorrected	Corrected	Uncorrected	Corrected
0.800	0.901	0.867	0.856	0.820
0.850	0.928	0.894	0.888	0.862
0.900	0.963	0.955	0.928	0.908
0.950	0.998	0.989	0.962	0.951

of lognormal means can be used for relative comparisons. However, in certain situations, the difference of the lognormal means may be the quantity of interest. Krishnamoorthy and Mathew [6] have proposed a generalized pivotal approach for the difference of two lognormal means. The generalized intervals perform adequately for the parameter values reported. However, if the population variances are large compared to the means then the generalized intervals are very conservative.

Let $Y_{1,1}, \ldots, Y_{1,n}$ be a random sample from lognormal distribution with parameters τ_1 and σ_1^2 and let $Y_{2,1}, \ldots, Y_{2,n}$ be a random sample from lognormal distribution with parameters τ_2 and σ_2^2. We will assume the samples are independent and that variances are known to be equal, i.e., $\sigma_1^2 = \sigma_2^2 = \sigma^2$. The parameter of interest is the difference of the two lognormal means, $\theta = \theta_1 - \theta_2 = \exp\{\tau_1 + 0.5\sigma^2\} - \exp\{\tau_2 + 0.5\sigma^2\}$. Let $X_i = n^{-1}\sum_{j=1}^{n} \log(Y_{i,j}),\ i = 1, 2$, and X_3 is the combined sample variance defined by $X_3 = (2n-2)^{-1}\sum_{i=1}^{2}\sum_{j=1}^{n}(\log(Y_{i,j}) - X_i)^2$. The asymptotic mean vector for the statistic is $(\mu_1,\ \mu_2,\ \mu_3) = (\tau_1,\ \tau_2,\ \sigma^2)$. Then the generalized pivot for θ is

$$T_\theta(\mathbf{x}, \mathbf{X}) = \exp\{x_1 - (X_1 - \mu_1)\sqrt{x_3/X_3} + 0.5\sigma^2 x_3/X_3\} - \exp\{x_2 - (X_2 - \mu_2)\sqrt{x_3/X_3} + 0.5\sigma^2 x_3/X_3\}.$$

After some algebra we have,

$$\Delta = \sigma\theta[\theta_1\theta_2(0.5\sigma^2 - 1) - 0.25\theta^2]/[\theta^2(1 + 0.25\sigma^2) + 2\theta_1\theta_2]^{3/2}.$$

For parameter configuration with σ^2 large compared to θ, Δ is potentially large. We investigated the performance of the generalized intervals and modified generalized intervals in a limited simulation study with parameters $(\mu_1, \mu_2, \sigma^2) = (0, 1, 9)$ and the results are given in Table 2.

The reduction in coverage error is not as significant as in the example where Δ is known. Even for sample size $n = 25$, the error in estimation of Δ is large enough to mitigate any gain from correcting the first order term in the expansion of coverage probability. For larger sample sizes, such as $n = 50$, the estimation error in Δ is small and the correction does produce generalized intervals which are less conservative. Improved estimation of Δ may improve the situation somewhat.

5. Conclusion

In this paper we have proposed a methodology for improving the finite sample coverage properties of generalized confidence intervals. The methodology works well when the finite sample properties of the traditional generalized intervals are poor. We have derived our results under the assumptions of smoothness for the generalized pivots. Analogous results for nondifferentiable parametric functions is a topic of future research. The modification suggested in this paper depends on evaluation of the derivatives of the generalized pivot at the estimated moment of the random quantities involved in the pivot. For applications where the pivot is based on matrix valued functions of the random quantities, the evaluation of the derivative can be challenging. However, since we are interested in the derivative value only at a certain point, numerical differentiation methods can be employed.

Acknowledgments

We are grateful to the referees for their constructive and insightful comments.

Appendix. Proofs

Proof of Theorem 1. Let $\tilde{A}(\mathbf{x}) = H(\mathbf{x})G(\mathbf{x})$ where $H(\mathbf{x}) = [T_\theta(\mathbf{x}, \mathbf{x}) - T_\theta(\mathbf{x}, \boldsymbol{\mu})]$ and $G(\mathbf{x}) = A_0^{-1/2}(\mathbf{x}, \boldsymbol{\mu})$. Note that $\tilde{A}(\mu) = 0$. We will denote the functions corresponding to A_0, A_1, A_2, A_{21} and A_{22} for $\tilde{A}(\mathbf{x})$ as $\tilde{A}_0, \tilde{A}_1, \tilde{A}_2, \tilde{A}_{21}$ and $\tilde{A}_{22}$. Then

$$\begin{aligned}\frac{\partial H(\mathbf{x})}{\partial \mathbf{x}_i} &= a_i^{(1)}(\mathbf{x}, \mathbf{x}) + a_i^{(2)}(\mathbf{x}, \mathbf{x}) - a_i^{(1)}(\mathbf{x}, \boldsymbol{\mu}) \\ &= -a_i^{(1)}(\mathbf{x}, \boldsymbol{\mu}) \quad [\text{because } T_\theta(\mathbf{x}, \mathbf{x}) = \pi(\theta)],\end{aligned}$$

$$\begin{aligned}\frac{\partial G(\mathbf{x})}{\partial \mathbf{x}_i} &= \frac{-1}{2A_0^{3/2}(\mathbf{x},\boldsymbol{\mu})}\sum_{k,l}[a_{i,k}^{(12)}(\mathbf{x},\boldsymbol{\mu})a_l^{(2)}(\mathbf{x},\boldsymbol{\mu})+a_{i,l}^{(12)}(\mathbf{x},\boldsymbol{\mu})a_k^{(2)}(\mathbf{x},\boldsymbol{\mu})]\mu_{kl}\\ &= \frac{-1}{A_0^{3/2}(\mathbf{x},\boldsymbol{\mu})}\sum_{k,l}[a_{i,k}^{(12)}(\mathbf{x},\boldsymbol{\mu})a_l^{(2)}(\mathbf{x},\boldsymbol{\mu})]\mu_{kl},\\ \frac{\partial^2 H(\mathbf{x})}{\partial \mathbf{x}_j\partial \mathbf{x}_i} &= a_{ij}^{(1)}(\mathbf{x},\mathbf{x})+a_{i,j}^{(12)}(\mathbf{x},\mathbf{x})+a_{j,i}^{(12)}(\mathbf{x},\mathbf{x})+a_{ij}^{(2)}(\mathbf{x},\mathbf{x})-a_{ij}^{(1)}(\mathbf{x},\boldsymbol{\mu})\\ &= -a_{ij}^{(11)}(\mathbf{x},\boldsymbol{\mu})\quad [\text{because } T_\theta(\mathbf{x},\mathbf{x})=\pi(\theta)].\end{aligned}$$

Therefore

$$\begin{aligned}&\left.\frac{\partial H(\mathbf{x})}{\partial \mathbf{x}_i}\right|_{\mathbf{x}=\boldsymbol{\mu}} = a_i^{(2)},\qquad \left.\frac{\partial G(\mathbf{x})}{\partial \mathbf{x}_i}\right|_{\mathbf{x}=\boldsymbol{\mu}} = A_0^{-3/2}\sum_{k,l}a_{i,k}^{(12)}a_l^{(2)}\mu_{kl},\\ &\left.\frac{\partial^2 \tilde{A}(\mathbf{x})}{\partial \mathbf{x}_j\partial \mathbf{x}_i}\right|_{\mathbf{x}=\boldsymbol{\mu}} = A_0^{-1/2}\left\{\left.\frac{\partial^2 H(\mathbf{x})}{\partial \mathbf{x}_j\partial \mathbf{x}_i}\right|_{\mathbf{x}=\boldsymbol{\mu}}\right\}+\left\{\frac{\partial H(\mathbf{x})}{\partial \mathbf{x}_j}\frac{\partial G(\mathbf{x})}{\partial \mathbf{x}_i}+\frac{\partial H(\mathbf{x})}{\partial \mathbf{x}_i}\frac{\partial G(\mathbf{x})}{\partial \mathbf{x}_j}\right\}_{\mathbf{x}=\boldsymbol{\mu}}\\ &\qquad = A_0^{-1/2}(a_{ij}^{(2)}+a_{i,j}^{(12)}+a_{j,i}^{(12)})-A_0^{-3/2}\sum_{k,l}[a_j^{(2)}a_{i,k}^{(12)}a_l^{(2)}+a_i^{(2)}a_{j,l}^{(12)}a_k^{(2)}]\mu_{kl}\\ &\qquad = A_0^{-1/2}a_{ij}^{(2)}+A_0^{-3/2}\sum_{k,l}[(a_{i,j}^{(12)}+a_{j,i}^{(12)})a_k^{(2)}a_l^{(2)}-a_j^{(2)}a_{i,k}^{(12)}a_l^{(2)}-a_i^{(2)}a_{j,l}^{(12)}a_k^{(2)}]\mu_{kl}.\end{aligned}$$

Thus,

$$\tilde{A}_0 = A_0^{-1}\sum_{ij}a_i^{(2)}a_j^{(2)}\mu_{ij} = A_0^{-1}A_0 = 1, \tag{A.1}$$

$$\begin{aligned}\tilde{A}_1 &= A_0^{-1/2}A_1+2^{-1}A_0^{-3/2}\sum_{i,j,k,l}\Big[(a_{i,j}^{(12)}+a_{j,i}^{(12)})a_k^{(2)}a_l^{(2)}-a_j^{(2)}a_{i,k}^{(12)}a_l^{(2)}-a_i^{(2)}a_{j,l}^{(12)}a_k^{(2)}\Big]\mu_{ij}\mu_{kl}\\ &= A_0^{-1/2}A_1+A_0^{-3/2}\sum_{i,j,k,l}\Big[a_{i,j}^{(12)}a_k^{(2)}a_l^{(2)}-a_i^{(2)}a_{j,l}^{(12)}a_k^{(2)}\Big]\mu_{ij}\mu_{kl}\\ &= A_0^{-1/2}A_1+\Delta.\end{aligned} \tag{A.2}$$

Also

$$\begin{aligned}\tilde{A}_{21} &= A_0^{-3/2}\sum_{ijk}a_i^{(2)}a_j^{(2)}a_k^{(2)}\mu_{ijk} = A_0^{-3/2}A_{21},\\ \tilde{A}_{22} &= A_0^{-3/2}A_{22}+A_0^{-5/2}\sum_{ijkl}\Big[a_i^{(2)}a_j^{(2)}(a_{k,l}^{(12)}+a_{l,k}^{(12)})a_m^{(2)}a_n^{(2)}-a_i^{(2)}a_j^{(2)}(a_{k,m}^{(12)}a_l^{(2)}a_n^{(2)}+a_k^{(2)}a_{l,n}^{(12)}a_m^{(2)})\Big]\mu_{ik}\mu_{jl}\mu_{mn}\\ &:= A_0^{-3/2}A_{22}+\Delta_2.\end{aligned} \tag{A.3}$$

The second part in the last equality in (A.3) is

$$\begin{aligned}\Delta_2 &= A_0^{-5/2}\sum_{ijklmn}a_i^{(2)}a_j^{(2)}\Big[(a_{k,l}^{(12)}+\ axylk)a_m^{(2)}a_n^{(2)}-a_{k,m}^{(12)}a_l^{(2)}a_n^{(2)}-a_k^{(2)}a_{l,n}^{(12)}a_m^{(2)}\Big]\mu_{ik}\mu_{jl}\mu_{mn}\\ &= A_0^{-5/2}\sum_{ijklmn}[a_i^{(2)}a_j^{(2)}a_{k,l}^{(12)}a_m^{(2)}a_n^{(2)}-a_i^{(2)}a_j^{(2)}a_{k,m}^{(12)}a_l^{(2)}a_n^{(2)}]\mu_{ik}\mu_{jl}\mu_{mn} = 0.\end{aligned}$$

Hence the result. □

Proof of Theorem 2. The results of Theorem 2 are similar to those for the distribution of analogous bootstrap pivotal quantities. We only give a sketch of the proof.

The claim about Edgeworth expansion (10) follows directly along the lines of Theorem 5.1 of Hall ([26]; pp. 239). Also see Bhattacharya and Ranga Rao [32]. The bootstrap probability measure (conditional on the sample) in Theorem 5.1 has to be replaced by the measure associated with the pseudorandomization due to $\mathbf{X}$ in the generalized pivot $T_\theta(\mathbf{x},\mathbf{X})$. We do not work out all the tedious details of the Edgeworth expansion of the pivot distribution under the $P_{\mathbf{X}}$ measure conditional on the sample $\mathbf{x}$. We merely point out that the bulk of the arguments rests on the behavior of the cumulants for the bootstrap probability measure conditional on the sample. The behavior of the cumulants under the pseudorandomization of the generalized pivot in our setup is almost identical to those for the bootstrap measure.

To prove the claim about the generalized pivot Cornish–Fisher expansion (14), we use the generalized pivot Edgeworth expansion (12) and follow the proof of Theorem 5.2 of Hall ([26]; pp. 241). The main idea is to establish that the coefficients of the polynomial $p_{1,\mathbf{x}}(z_\alpha)$ are bounded by a fixed constant C_3 with probability approaching one at the rate $1-O(n^{-\lambda})$.

However, by assumption (A2), the coefficients of the polynomial $A_1(\mathbf{x},\boldsymbol{\mu})/A_0(\mathbf{x},\boldsymbol{\mu})^{1/2}$ and $A_2(\mathbf{x},\boldsymbol{\mu})/A_0(\mathbf{x},\boldsymbol{\mu})^{3/2}$ are smooth functions of $\mathbf{x}$ with bounded derivatives and hence the claim follows. □

Proof of Theorem 3.

$$\begin{aligned} P_{\mathbf{x}}[\pi(\theta) \in \mathcal{I}_{n,\mathbf{x}}(\alpha)] &= P_{\mathbf{x}}[P_{\mathbf{X}}\{T_\theta(\mathbf{x},\mathbf{X}) < T_\theta(\mathbf{x},\mathbf{x})\} < \alpha] \\ &= P_{\mathbf{x}}[P_{\mathbf{X}}\{Z_{n,\mathbf{X}}(\mathbf{x}) < Z_n(\mathbf{x})\} < \alpha] = P_{\mathbf{x}}(Z_n(\mathbf{x}) < F^{-1}_{n,\mathbf{x},\mathbf{x}}(\alpha)) \\ &= P_{\mathbf{x}}(Z_n(\mathbf{x}) - C_n(\mathbf{x},\alpha) < z_\alpha - n^{-1/2}p_{1,\mathbf{x}}(z_\alpha)). \end{aligned} \tag{A.4}$$

Let

$$g_n(\mathbf{x},\alpha) = z_\alpha - n^{-1/2}p_{1,\mathbf{x}}(z_\alpha). \tag{A.5}$$

and let $D_n(\mathbf{x},\alpha) = g_n(\mathbf{x},\alpha) - g_n(\boldsymbol{\mu},\alpha) = n^{-1/2}(p_{1,\mathbf{x}}(z_\alpha) - p_{1,\mu}(z_\alpha))$. Therefore the required probability is

$$\begin{aligned} P_{\mathbf{x}}[\pi(\theta) \in \mathcal{I}_{n,\mathbf{x}}(\alpha)] &= P_{\mathbf{x}}(Z_n(\mathbf{x}) - C_n(\mathbf{x},\alpha) < g_n(\mathbf{x},\alpha)) \\ &= P_{\mathbf{x}}(Z_n(\mathbf{x}) < g_n(\boldsymbol{\mu},\alpha) + C_n(\mathbf{x},\alpha) + D_n(\mathbf{x},\alpha)). \end{aligned}$$

We will show that for some $\epsilon_n = o(n^{-1/2})$,

$$P_{\mathbf{x}}[\pi(\theta) \in \mathcal{I}_{n,\mathbf{x}}(\alpha)] = P_{\mathbf{x}}(Z_n(\mathbf{x}) < \tilde{g}_n(\boldsymbol{\mu},\alpha)) + o(n^{-1/2}),$$

where $\tilde{g}_n(\boldsymbol{\mu},\alpha) = g_n(\boldsymbol{\mu},\alpha) + 2C_n\epsilon_n$. It is enough to prove that

$$P_{\mathbf{x}}\{|C_n(\mathbf{x},\alpha)| > C_3\epsilon_n\} = o(n^{-1/2}), \tag{A.6}$$

and

$$P_{\mathbf{x}}\{|D_n(\mathbf{x},\alpha)| > C_3\epsilon_n\} = o(n^{-1/2}), \tag{A.7}$$

for some constant $C_3 > 0$. Choose $\epsilon_n = n^{-\beta}$ where $\beta > 1/2$ but close to 1/2. The relation (A.6) follows from (15). Now,

$$P_{\mathbf{x}}\{|g_n(\mathbf{x},\alpha) - g_n(\boldsymbol{\mu},\alpha)| > C_3\epsilon_n\} = P_{\mathbf{x}}\{|p_{1,\mathbf{x}}(z_\alpha) - p_{1,\mu}(z_\alpha)| > C_3 n^{1/2}\epsilon_n\}.$$

By assumption (A2), p_1, as a function of $\mathbf{x}$ is twice differentiable in a neighborhood of $\boldsymbol{\mu}$ and has bounded derivatives. Therefore, there exists a constant $C_4 > 0$, such that $|p_{1,\mathbf{x}}(z_\alpha) - p_{1,\boldsymbol{\mu}}(z_\alpha)| \le C_4\|\mathbf{x} - \boldsymbol{\mu}\|$ for all $\mathbf{x}$ in a δ_n neighborhood of $\boldsymbol{\mu}$ where $\delta_n = o(n^{1/2}\epsilon_n)$. By Assumption (A1), we have

$$P_{\mathbf{x}}\{\|\mathbf{x} - \boldsymbol{\mu}\| \ge \delta_n\} \le C_4[\delta_n n^{1/2}]^{-3}. \tag{A.8}$$

Therefore,

$$P_{\mathbf{x}}\{|D_n(\mathbf{x},\alpha)| > C_3\epsilon_n\} = O(n^{-3}\epsilon_n^{-3}). \tag{A.9}$$

It is easy to see that one can choose δ_n and β such that the $O(n^{-3}\epsilon_n^{-3})$ term in (A.9) is indeed $o(n^{-1/2})$. Thus,

$$P_{\mathbf{x}}[\pi(\theta) \in \mathcal{I}_{n,\mathbf{x}}(\alpha)] = P_{\mathbf{x}}(Z_n(\mathbf{x}) < \tilde{g}_n(\boldsymbol{\mu},\alpha)) + o(n^{-1/2}). \tag{A.10}$$

Using the one term Edgeworth expansion (10) for the studentized statistic $Z_n(\mathbf{x})$, and Taylor expansion of $\Phi(\tilde{g}_n(\boldsymbol{\mu},\alpha))$ up to order $n^{-1/2}$ term and noting that the leading term in $g_n(\boldsymbol{\mu},\alpha)$ is z_α we have

$$\begin{aligned} P_{\mathbf{x}}[\pi(\theta) \in \mathcal{I}_{n,\mathbf{x}}(\alpha)] &= \Phi(\tilde{g}_n(\boldsymbol{\mu},\alpha)) + n^{-1/2}q_{1,\boldsymbol{\mu}}(\tilde{g}_n(\boldsymbol{\mu},\alpha))\phi(\tilde{g}_n(\boldsymbol{\mu},\alpha)) + o(n^{-1/2}) \\ &= \alpha - n^{-1/2}[p_{1,\boldsymbol{\mu}}(z_\alpha) - q_{1,\boldsymbol{\mu}}(z_\alpha)]\phi(z_\alpha) + o(n^{-1/2}) \\ &= \alpha - n^{-1/2}\Delta\phi(z_\alpha) + o(n^{-1/2}). \end{aligned} \tag{A.11}$$

This completes the proof of Theorem 3. □

The error term in (A.11) can be actually shown to be $O(n^{-1})$ in many examples. For example, if the quantities are such that we can apply *Delta Method* as in Hall ([26], pp. 76), then the error term can be shown to be $O(n^{-1})$.

Before we prove Theorem 4 we establish the following lemma.

Lemma 1. *Let the assumptions of Theorem 3 hold and let $\hat{\Delta}$ be a $\sqrt{n}$-consistent estimator of Δ which satisfies*

$$P_{\mathbf{x}}\{|\hat{\Delta} - \Delta| > C_5 n^{\delta-1/2}\} = o(n^{-1/2}), \tag{A.12}$$

for some $0 < \delta < 1/2$ and some constant $C_5 < 0$. Let $r_n(\mathbf{x},\alpha) = g_n(\mathbf{x},\hat{\alpha}_n) - g_n(\mathbf{x},\alpha_n)$ where $g_n(\mathbf{x},\alpha)$ is defined in (A.5). Then there exists a constant $C_6 > 0$ such that $P_{\mathbf{x}}\{|r_n(\mathbf{x},\alpha)| > C_6 n^{-1+\delta}\} = o(n^{-1/2})$.

Proof. The case when $\Delta = 0$ can be easily dealt with as in that case $\alpha_n = \alpha_0$ and $\hat{\alpha}_n - \alpha_0$ is $n^{-1/2}\phi(z_{\hat{\alpha}_n})$ and ϕ is a bounded quantity. Thus, we will consider only the case when $\Delta \neq 0$. From definition $\hat{\alpha}_n - n^{-1/2}\hat{\Delta}\phi(z_{\hat{\alpha}_n}) = \alpha_0$ and $\alpha_n - n^{-1/2}\Delta\phi(z_{\alpha_n}) = \alpha_0$. Subtracting the second equation from the first and rearranging the term and expanding $\phi(z_{\hat{\alpha}_n})$ around α_n we have

$$(\hat{\alpha}_n - \alpha_n) = \frac{-n^{-1/2}(\hat{\Delta} - \Delta)\phi(z_{\hat{\alpha}_n})}{1 + n^{-1/2}\Delta z_{\alpha_n^*}}, \tag{A.13}$$

where α_n^* is between α_n and $\hat{\alpha}_n$. By the mean value theorem, $z_{\hat{\alpha}_n} - z_{\alpha_n} = \frac{(\hat{\alpha}_n - \alpha_n)}{\phi(z_{\alpha_n^{**}})}$, where α_n^{**} is between $\hat{\alpha}_n$ and α_n. Hence from (A.13) we have

$$r_n(\mathbf{x}, \alpha) = n^{-1}[\sqrt{n}(\hat{\Delta} - \Delta)]v_n(\mathbf{x}, \alpha),$$

where

$$v_n(\mathbf{x}, \alpha) = \frac{\phi(z_{\hat{\alpha}_n})[1 - n^{-1/2}\frac{A_2(\mathbf{x},\boldsymbol{\mu})}{A_0^{3/2}(\mathbf{x},\boldsymbol{\mu})}(z_{\hat{\alpha}_n} + z_{\alpha_n})]}{\phi(z_{\alpha_n^{**}})[1 + n^{-1/2}\Delta z_{\alpha_n^*}]}. \tag{A.14}$$

Because $\alpha_n^* \in [n^{-\epsilon}, 1 - n^{-\epsilon}]$ we have $n^{-1/2}\Delta z_{\alpha_n^*} < C_7\sqrt{\log n/n}$ for some constant $C_7 < 0$. Hence for large enough n, $[1 + n^{-1/2}\Delta z_{\alpha_n^*}]^{-1}$ is less than 2. Also, note that $\frac{\phi(z_{\hat{\alpha}_n})}{\phi(z_{\alpha_n^{**}})} \leq \max\{1, \frac{\phi(z_{\hat{\alpha}_n})}{\phi(z_{\alpha_n})} I(|\alpha_n - 1/2| > |\hat{\alpha}_n - 1/2|)\}$. Now from definition, $\frac{\phi(z_{\hat{\alpha}_n})}{\phi(z_{\alpha_n^{**}})} = \frac{\Delta}{\hat{\Delta}}\frac{\hat{\alpha}_n - \alpha_0}{\alpha_n - \alpha_0}$. Therefore, if $|\alpha_n - \alpha_0| > \eta$ for some constant $\eta > 0$ then $\frac{\phi(z_{\hat{\alpha}_n})}{\phi(z_{\alpha_n^{**}})} \leq \frac{C_8}{\hat{\Delta}} = \frac{C_8}{\hat{\Delta} - \Delta + \Delta}$ for some constant $C_8 > 0$. For $|\alpha_n - \alpha_0| \leq \eta$ we have $\frac{\phi(z_{\hat{\alpha}_n})}{\phi(z_{\alpha_n^{**}})} \leq e^{\max\{z^2_{\alpha_0+\eta} + z^2_{\alpha_0-\eta}\}}$. Therefore by assumption (A.12) we have

$$P_{\mathbf{x}}\left\{\frac{\phi(z_{\hat{\alpha}_n})}{\phi(z_{\alpha_n^{**}})} \geq C_{10}\right\} = o(n^{-1/2}), \tag{A.15}$$

for some constant $C_{10} > 0$. Again, since $\hat{\alpha}_n$ and α_n are in $[n^{-\epsilon}, 1 - n^{-\epsilon}]$ we have $n^{-1/2}(z_{\hat{\alpha}_n} + z_{\alpha_n}) < C_{11}\sqrt{\log n/n}$ for some constant $C_{11} > 0$. By the moment assumptions on $\mathbf{x}$ and the smoothness assumption on the derivatives $a_i(\mathbf{x}, \boldsymbol{\mu})$ and $a_{ij}(\mathbf{x}, \boldsymbol{\mu})$, for some constant $C_{12} < 0$, we have $P_{\mathbf{x}}\{|\frac{A_2(\mathbf{x},\boldsymbol{\mu})}{A_0(\mathbf{x},\boldsymbol{\mu})^{3/2}}| > C_{12}\} = o(n^{-1/2})$. Thus, for large enough n,

$$P_{\mathbf{x}}\left\{1 - n^{-1/2}(z_{\hat{\alpha}_n} + z_{\alpha_n})\frac{A_2(\mathbf{x}, \boldsymbol{\mu})}{A_0(\mathbf{x}, \boldsymbol{\mu})^{3/2}} > 1/2\right\} = o(n^{-1/2}). \tag{A.16}$$

Therefore, from (A.16), (A.15) and the fact that for large n, $[1 + n^{-1/2}\Delta z_{\alpha_n^*}]^{-1}$ is less than 2, we have

$$P_{\mathbf{x}}\{|v_n\mathbf{x}, \alpha| > C_{10}\} = o(n^{-1/2}). \tag{A.17}$$

Then by assumption (A.12) we have the result. □

Proof of Theorem 4. By (A.4), we have

$$P_{\mathbf{x}}[\pi(\theta) \in \mathcal{I}_{n,\mathbf{x}}(\tilde{\alpha}_n)] = P_{\mathbf{x}}(Z_n(\mathbf{x}) - C_n(\mathbf{x}, \tilde{\alpha}_n) < g_n(\mathbf{x}, \tilde{\alpha}_n)).$$

By definition $\tilde{\alpha}_n \in [n^{-\epsilon}, 1 - n^{-\epsilon}]$. Thus, as in proof of Theorem 3, we can bring the remainder term, $C_n(\mathbf{x}, \tilde{\alpha}_n)$, outside the probability as a $o(n^{-1/2})$ term. Therefore,

$$\begin{aligned} P_{\mathbf{x}}[\pi(\theta) \in \mathcal{I}_{n,\mathbf{x}}(\tilde{\alpha}_n)] &= P_{\mathbf{x}}(Z_n(\mathbf{x}) < g_n(\mathbf{x}, \tilde{\alpha}_n)) + o(n^{-1/2}) \\ &= P_{\mathbf{x}}(Z_n(\mathbf{x}) < g_n(\mathbf{x}, \alpha_n) + \tilde{r}_n(\mathbf{x}, \alpha)) + o(n^{-1/2}) \end{aligned} \tag{A.18}$$

where $\tilde{r}_n(\mathbf{x}, \alpha) = g_n(\mathbf{x}, \tilde{\alpha}_n) - g_n(\mathbf{x}, \alpha_n) = (z_{\tilde{\alpha}_n} - z_{\alpha_n})[1 - n^{-1/2}\frac{A_2(\mathbf{x},\boldsymbol{\mu})}{A_0^{3/2}(\mathbf{x},\boldsymbol{\mu})}(z_{\tilde{\alpha}_n} + z_{\alpha_n})]$. Similarly define

$$r_n(\mathbf{x}, \alpha) = g_n(\mathbf{x}, \hat{\alpha}_n) - g_n(\mathbf{x}, \alpha_n).$$

We can replace $\tilde{r}_n(\mathbf{x}, \alpha)$ with $r_n(\mathbf{x}, \alpha)$ in (A.18) as $P_{\mathbf{x}}(\tilde{\alpha}_n \neq \hat{\alpha}_n)$ is $o(n^{-1/2})$. By Lemma 1, we have $P_{\mathbf{x}}\{|r_n(\mathbf{x}, \alpha)| > C_6 n^{-1+\delta}\} = o(n^{-1/2})$. Then, by arguments similar to those in the proof of Theorem 3, we have

$$\begin{aligned} P_{\mathbf{x}}[\pi(\theta) \in \mathcal{I}_{n,\mathbf{x}}(\tilde{\alpha}_n)] &= P_{\mathbf{x}}(Z_n(\mathbf{x}) < g_n(\boldsymbol{\mu}, \alpha_n)) + o(n^{-1/2}) \\ &= \alpha_n - n^{-1/2}\Delta\phi(z_{\alpha_n}) + o(n^{-1/2}) \\ &= \alpha_0 + o(n^{-1/2}). \quad \square \end{aligned}$$

Finally we prove the Proposition which was used in Example 3.

Proposition 3. *Let $x, y > 0$ be positive real numbers. Then*

$$f(x, y) = \{x + y + 0.5(x^2 + y^2)\}^3 - (x - y)^2(x + y - xy)^2 > 0.$$

Proof. We prove the result using a combination of sum of squares optimization methods (Parrilo and Sturmfels [33]) and first principles. Rearranging the terms we have $f(x, y) = (1/8)[\mathcal{P}_3(x, y) + \mathcal{P}_4(x, y) + \mathcal{P}_5(x, y) + \mathcal{P}_6(x, y)]$, where $\mathcal{P}_r(x, y)$ is an rth degree polynomial in (x, y) and the polynomials are given by

$$\mathcal{P}_3(x, y) = 8(x + y)^3,$$
$$\mathcal{P}_4(x, y) = 12(x + y)^2(x^2 + y^2) - 8(x - y)^2(x + y)^2,$$
$$\mathcal{P}_5(x, y) = 6(x + y)(x^2 + y^2)^2 + 16xy(x - y)^2(x + y),$$
$$\mathcal{P}_6(x, y) = (x^2 + y^2)^3 - 8(x - y)^2x^2y^2.$$

The polynomial $\mathcal{P}_4(x, y)$ is positive because $3(x^2 + y^2) - 2(x - y)^2 > 0$. The polynomial $\mathcal{P}_5(x, y)$ can be written as $(x + y)[6(x^2 - y^2)^2 + 8xy(x^2 + y^2 - xy)]$ and hence positive. Since $f(x, y)$ is symmetric about $x = y$, we can assume $0 < y < x$ without loss of generality. Let $r = (y/x)^2$. Then $\mathcal{P}_6(x, y) > 0$ is equivalent to showing $1 - 5r + 11r^2 + r^3 > 0$ for $r \in (0, 1)$. A simple one variable analysis establishes the claim. □

References

[1] K.W. Tsui, S. Weerahandi, Generalized *p*-values in significance testing of hypotheses in the presence of nuisance parameters, J. Amer. Statist. Assoc. 84 (1989) 602–607.
[2] S. Weerahandi, Testing variance components in mixed models with generalized *p* values, J. Amer. Statist. Assoc. 86 (1991) 151–153.
[3] S. Weerahandi, Generalized confidence intervals, J. Amer. Statist. Assoc. 88 (1993) 899–905.
[4] S. Weerahandi, Exact Statistical Methods for Data Analysis, Springer, New York, 1995.
[5] S. Weerahandi, Generalized Inference in Repeated Measures: Exact Methods in MANOVA and Mixed Models, John Wiley and Sons, Inc., New York, 2004.
[6] K. Krishnamoorthy, T. Mathew, Inferences on the means of lognormal distributions using generalized *p*-values and generalized confidence intervals, J. Statist. Plann. Inference 115 (2003) 103–120.
[7] R.A. Johnson, S. Weerahandi, A Bayesian Solution to the multivariate Behrens–Fisher problem, J. Amer. Statist. Assoc. 83 (1998) 145–149.
[8] J. Gamage, T. Mathew, S. Weerahandi, Generalized *p*-values and generalized confidence regions for the multivariate Behrens–Fisher problem, J. Multivariate Anal. 88 (2004) 177–189.
[9] L. Zhou, T. Mathew, Some tests for variance components using generalized *p*-values, Technometrics 36 (1994) 394–421.
[10] J. Gamage, S. Weerahandi, Size performance of some tests in one-way ANOVA, Comm. Statist. Theory Methods 27 (1998) 625–640.
[11] A.K.L. Chiang, Simple general method for constructing confidence intervals for functions of variance components, Technometrics 43 (2001) 356–367.
[12] T. Mathew, D.W. Webb, Generalized *p* values and confidence intervals for variance components: Applications to army test and evaluation, Technometrics 47 (2005) 3120–3322.
[13] B. Arendacka, Generalized confidence intervals on the variance component in mixed linear models with two variance components, Statistics 39 (2005) 275–286.
[14] H.K. Iyer, C.M. Wang, T. Mathew, Models and intervals for true values in interlaboratory trials, J. Amer. Statist. Assoc. 99 (2004) 1060–1071.
[15] R.J. McNally, H. Iyer, T. Mathew, Tests for individual and population bioequivalence based on generalized *p*-values, Stat. Med. 22 (2003) 31–53.
[16] S. Weerahandi, V.W. Berger, Exact inference for growth curves with intraclass correlation structure, Biometrics 55 (1999) 921–924.
[17] S.H. Lin, J.C. Lee, Exact tests in simple growth curve models and one-way ANOVA with equicorrelation error structure, J. Multivariate Anal. 84 (2003) 351–368.
[18] A. Roy, T. Mathew, A generalized confidence limit for the reliability function of a two-parameter exponential distribution, J. Statist. Plann. Inference 128 (2005) 509–517.
[19] L. Tian, J.C. Cappelleri, A new approach for interval estimation and hypothesis testing of a certain intraclass correlation coefficient: The generalized variable method, Stat. Med. 23 (2004) 2125–2135.
[20] T. Mathew, K.M. Kurian, G. Sebastian, Generalized confidence intervals for process capability indices, Qual. Reliab. Eng. Int. 23 (2007a) 471–481.
[21] R.K. Burdick, C.M. Borror, D.C. Montgomery, Design and analysis of gauge R & R studies: Making decisions with confidence intervals in random and mixed ANOVA models, in: ASA-SIAM Series on Statistics and Applied Probability, 2005.
[22] T. Mathew, K.M. Kurian, G. Sebastian, Generalized confidence intervals for process capability indices in the one-way random model, Metrika 67 (2007b) 83–92.
[23] K. Krishnamoorthy, T. Mathew, G. Ramachandran, Generalized *p*-values and confidence intervals: A novel approach for analyzing lognormally distributed exposure data, J. Occup. Environ. Health 3 (2006) 642–650.
[24] J. Hannig, H. Iyer, P. Patterson, Fiducial generalized confidence intervals, J. Amer. Statist. Assoc. 101 (2006) 254–269.
[25] J. Hannig, On fiducial inference, the good, the bad, and the ugly, Technical Report 2006/3, Colorado State University, Dept. of Statistics, 2006.
[26] P. Hall, The Bootstrap and Edgeworth Expansion, Springer-Verlag, 1992.
[27] J. Klion, Specifying maintainability — A new approach, in: Proceedings of the Annual Reliability and Maintainability Symposium, IEEE, Las Vegas, 1986, pp. 338–343.
[28] J.M. Jobe, A new R & M measure and its estimation, IEEE Trans. Reliab. 37 (1988) 413–420.
[29] J. Knezevic, Reliability, Maintainability and Supportability: A Probabilistic Approach, McGraw-Hill International (UK) Limited, 1993.
[30] M.M.A. Ananda, Estimation and testing of availability of a parallel system with exponential failure and repair times, J. Statist. Plann. Inference 77 (1999) 237–246.
[31] J.M. Jobe, H.T. David, Buehler confidence bounds for a reliability–maintainability measure, Technometrics 34 (1992) 214–222.
[32] R.N. Bhattacharya, R. Ranga Rao, Normal Approximation and Asymptotic Expansions, John Wiley & Sons Inc., New York, 1976.
[33] P.A. Parrilo, B. Sturmfels, Minimizing polynomial functions. Algorithmic and quantitative real algebraic geometry, DIMACS Ser. Discrete Math. Theor. Comput. Sci. 60 (2003) 83–99.