Edgeworth Corrected Generalized Confidence Intervals

Anindya Roy and Arup Bose
University of Maryland, Baltimore County and Indian Statistical Institute

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 asymptotic expansions for quantities associated with one-sided generalized confidence intervals, which helps to explain the above limited performance. We then show how to use these expansions to obtain improved coverage by suitable calibration. We implement these ideas in a few examples and demonstrate the improved finite sample performance of these corrected intervals.

AMS 2000 subject classifications: Primary 62F25; secondary 62F05, 62F40, 62J10.
Keywords and phrases: Average bioequivalence, Behrens-Fisher problem, bootstrap, Cornish-Fisher expansion, coverage probability, Edgeworth expansion, pivot..

1. Introduction

Generalized confidence intervals and generalized test statistics have been introduced by Tsui and Weerahandi [29] and Weerahandi [31, 32, 33] 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 generalized intervals are not motivated from a repeated sampling argument and confidence statements for these do not come from a long run frequency interpretation. Rather, as in the Bayesian approach, the sample at hand assumes paramount importance. Nevertheless, many authors quite naturally have asked questions about the frequentist properties of generalized intervals.

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 set up as well. 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 of functions of parameters of normal and related distributions (Weerahandi [31, 32, 33, 34], Chang and Huang [4], Krishnamoorthy and Mathew [19], Johnson and Weerahandi [16], Gamage, Mathew and Weerahandi [8]); testing fixed effects and variance components in repeated measures and mixed effects ANOVA models (Zhou and Mathew [38], Gamage and Weerahandi [7], Chiang [6], Krishnamoorthy and Mathew[20], Weerahandi [34], Mathew and Webb [23], Arendacka [2], Lin and Liao [22]); interlaboratory testing (Iyer, Wang and Mathew [13]); bioequivalence (McNally, Iyer and Mathew [24]); growth curve modeling (Weerahandi and Berger [36], Lin and Lee [21]); reliability and system engineering (Roy and Mathew [26], Tian and Cappelleri [27]); process control (Burdick, Borror and Montgomery [3]); health studies (Chen and Zhou [5]) and many others. The simulation studies in Griffiths and Judge [9], Thursby [28], Johnson and Weerahandi [16], Weerahandi [33], [34], Zhou and Mathew [38], Gamage and Weerahandi [7], Weerahandi and Amaratunga [35], 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. Recently Hannig *et al.* [12] have shown that asymptotically the generalized intervals maintain the target coverage level for a large class of parametric problems. These imply that the generalized confidence intervals are no worse asymptotically than the intervals based on asymptotic normality of the estimator of the parametric function. Hannig [11] has also investigated the connection between the generalized procedures and fiducial inference.

Our focus in this article is on one sided generalized confidence intervals. They are in the same spirit as the intervals constructed by the "other percentile method" [cf: Hall [10]] where the interval is obtained by inverting the resampling distribution of the estimator. Hall [10] notes that the "other percentile method" intervals are typically unappealing and have undesirable properties unless properly adjusted. Comparatively, the generalized intervals enjoy richer structural properties.

By deriving an expansion for the coverage probability, we show that in general the generalized confidence intervals are not first order accurate, i.e., accurate up to the $n^{-1/2}$ term. We also provide a necessary and sufficient condition for the generalized intervals to be first order accurate. When the $n^{-1/2}$ in the coverage error is not zero, we suggest a correction to make the intervals first order accurate. The modification works adequately both when the $n^{-1/2}$ term in the coverage error is known and also when it is a function of unknown parameters. See Section 3. Although we only discuss generalized confidence intervals in this paper, the results derived in Section 3 can be easily extended to improve the power properties of generalized tests as well.

In Section 4, we investigate the magnitude of the first order term in the coverage probability and illustrate our methodology in the context of some well known 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 [31] suggested constructing the confidence interval by inverting the distribution of a generalized pivotal quantity.

We first give the definition for a generalized pivotal quantity for the parameter of interest $\pi(\theta)$. Let $\mathbf{X}$ denote an independent and identical copy (but unobserved) of the random vector $\mathbf{x}$. Then $T_\theta(\mathbf{x}, \mathbf{X})$ is a *Generalized Pivotal Quantity* for $\pi(\theta)$ if it satisfies:
(i) $T_\theta(\mathbf{x}, \mathbf{x}) = \pi(\theta)$ for all $\mathbf{x} \in \mathbb{R}^d$.
(ii) The distribution of $T_\theta(\mathbf{x}, \mathbf{X})$ conditional on $\mathbf{x}$ is free from θ.

A $100\alpha\%$ upper generalized confidence interval for $\pi(\theta)$ is obtained by inverting the distribution of $T_\theta(\mathbf{x}, \mathbf{X})$ conditional on $\mathbf{x}$ and the interval is given by

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

where $w_{n,\mathbf{x}}(\alpha)$ is the upper α percentile of the distribution of $T_\theta(\mathbf{x}, \mathbf{X})$ conditional on $\mathbf{x}$. Similarly, we can defined a lower confidence interval. Let a $100(1-\alpha)\%$ 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 $(1-\alpha)$ 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. We will show that in most of the common examples where the frequentist coverage is not exact, the accuracy of coverage, $|P_{\mathbf{x}}\{\pi(\theta) \in \mathcal{I}_{n,\mathbf{x}}(\alpha)\} - \alpha|$ is $O(n^{-1/2})$. We will suggest modifications to the interval to improve the accuracy of coverage to $o(n^{-1/2})$.

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 statistic of interest is a vector of smooth functions of sample moments. Let $\boldsymbol{Y}_1, \boldsymbol{Y}_2, \ldots, \boldsymbol{Y}_n$ be independent and identically distributed random column d-vectors and let

$\mathbf{x} = n^{-1}\sum_{i=1}^{n} \boldsymbol{Y}_i$. 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 $\boldsymbol{Y}_i$. 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 set up. Expansions for coverage probability of other type of non-smooth generalized pivots is a topic of future investigation. The generalized procedure is typically applied in common parametric set ups where assumption (A1) usually hold. Hence, for these parametric problems, application of the central limit theorem also gives

(A3) There exists $\boldsymbol{\mu} := \boldsymbol{\mu}(\theta) = (\mu_1(\theta), \ldots, \mu_d(\theta))'$ and a positive definite matrix $\Sigma := \Sigma(\theta)$ such that $\sqrt{n}(\mathbf{x} - \boldsymbol{\mu}) \xrightarrow{\mathcal{L}} N_d(0, \Sigma)$ where $\xrightarrow{\mathcal{L}}$ denotes convergence in distribution and $N_d(0, \Sigma)$ denotes a d-dimensional normal distribution with covariance matrix equal to Σ.

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

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

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 $S(\cdot, \cdot)$. Also, let $D^{\omega,0} S(\mathbf{x}, \boldsymbol{y}) = \frac{\partial_{(1)}^{|\omega|}}{\partial_{(1)}^{\omega_1}\mathbf{x}_1 \cdots \partial_{(1)}^{\omega_d}\mathbf{x}_d} S(\mathbf{x}, \boldsymbol{y})$ and let $D^{0,\lambda} S(\mathbf{x}, \boldsymbol{y}) = \frac{\partial_{(2)}^{|\lambda|}}{\partial_{(2)}^{\lambda_1}\boldsymbol{y}_1 \cdots \partial_{(2)}^{\lambda_d}\boldsymbol{y}_d} S(\mathbf{x}, \boldsymbol{y})$. Let $e_{i_1 \cdots i_k}$ denote a d-vector with ones at the $i_1, \ldots, i_k$th places and rest zeroes. Define

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

Also define $a^{(\cdot)}_{\cdots}(\mathbf{x}, \boldsymbol{\mu}), a^{(\cdot)}_{\cdots}(\boldsymbol{\mu}, \boldsymbol{y})$ and $a^{(\cdot)}_{\cdots}(\boldsymbol{\mu}, \boldsymbol{\mu})$ will be the values of the function $a^{(\cdot)}_{\cdots}(\mathbf{x}, \boldsymbol{y})$ evaluated at $\boldsymbol{y} = \boldsymbol{\mu}, \mathbf{x} = \boldsymbol{\mu}$ and $(\mathbf{x}, \boldsymbol{y}) = (\boldsymbol{\mu}, \boldsymbol{\mu})$, 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 $\mathbf{x}$ and $\mu_{i_1 \ldots i_j} = E\{(\mathbf{x}-\boldsymbol{\mu})_{i_1} \cdots (\mathbf{x}-\boldsymbol{\mu})_{i_j}\}$.

Define (by abuse of the word "statistic"), the standardized statistic (with respect to the $P_{\mathbf{X}}$ probability) as

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

where

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

Let

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

be the second degree even polynomial where

$$\begin{aligned}
A_1(\mathbf{x},\boldsymbol{\mu}) &= \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d} a_{ij}^{(2)}(\mathbf{x},\boldsymbol{\mu})\mu_{ij}, \\
A_2(\mathbf{x},\boldsymbol{\mu}) &= A_{21}(\mathbf{x},\boldsymbol{\mu}) + 3\, A_{22}(\mathbf{x},\boldsymbol{\mu}), \\
A_{21}(\mathbf{x},\boldsymbol{\mu}) &= \sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{k=1}^{d} A_i^{(2)}(\mathbf{x},\boldsymbol{\mu}) a_j^{(2)}(\mathbf{x},\boldsymbol{\mu}) a_k^{(2)}(\mathbf{x},\boldsymbol{\mu})\mu_{ijk}, \\
A_{22}(\mathbf{x},\boldsymbol{\mu}) &= \sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{k=1}^{d}\sum_{l=1}^{d} a_i^{(2)}(\mathbf{x},\boldsymbol{\mu}) a_j^{(2)}(\mathbf{x},\boldsymbol{\mu}) a_{kl}^{(2)}(\mathbf{x},\boldsymbol{\mu})\mu_{ik}\mu_{jl}.
\end{aligned} \tag{7}$$

Define

$$\Delta = \frac{1}{A_0^{3/2}}\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{k=1}^{d}\sum_{l=1}^{d}\{a_{i,j}^{(12)} a_k^{(2)} a_l^{(2)} - a_i^{(2)} a_{j,k}^{(12)} a_l^{(2)}\}\mu_{ij}\mu_{kl}. \tag{8}$$

The 100βth percentile of the standard normal distribution will be denoted by z_β.

3. Main results

This section contains the main results. All proofs are relegated to the Appendix. The first result is reminiscent of expansion results in bootstrap, with added claims on the rates in (10) and (12). These error rates are obtained by keeping track of the relevant terms in the expansion.

Theorem 1. *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)*

$$F_{n,\mathbf{X},\mathbf{x}}(z) = P_{\mathbf{X}}(Z_{n,\mathbf{x}}(\mathbf{x}) < z) = \Phi(z) + n^{-1/2}p_{1,\mathbf{x}}(z)\phi(z) + \epsilon_n(\mathbf{x}; z) \tag{9}$$

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

$$P_{\mathbf{x}}(\sup_{-\infty<z<\infty} |\epsilon_n(\mathbf{x}; z)| > C_1 n^{-1/2}) = O(n^{-\lambda}). \tag{10}$$

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

$$F^{-1}_{n,\mathbf{X},\mathbf{x}}(\beta) = z_\beta - n^{-1/2} p_{1,\mathbf{x}}(z_\beta) + C_n(\mathbf{x}, \beta), \tag{11}$$

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

$$P_{\mathbf{x}}(\sup_{n^{-\epsilon}<\beta<1-n^{-\epsilon}} |C_n(\mathbf{x}; \beta)| > C_2 n^{-1+\delta}) = O(n^{-\lambda}). \tag{12}$$

Now consider (again by abuse of the word) the statistic

$$Z_n(\mathbf{x}) = [n/A_0(\mathbf{x}, \boldsymbol{\mu})]^{1/2} (T_\theta(\mathbf{x}, \mathbf{x}) - T_\theta(\mathbf{x}, \boldsymbol{\mu})). \tag{13}$$

To derive an expansion for the frequentist coverage of the generalized interval we will need an expansion of the distribution of $Z_n(\mathbf{x})$. The following theorem gives the expansion.

Theorem 2. *Let $Z_n(\mathbf{x})$ be the studentized statistic defined in (13). Let all assumptions of Theorem 1 hold. Then the distribution of $Z_n(\mathbf{x})$ has a valid Edgeworth expansion given by (up to one term)*

$$P_{\mathbf{x}}(Z_n(\mathbf{x}) < z) = \Phi(z) + n^{-1/2} q_{1,\boldsymbol{\mu}}(z)\phi(z) + O(n^{-1}), \tag{14}$$

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

Remark 1. *If the components of $\mathbf{x}$ are asymptotically independent, i.e., $\Sigma = \Big((\mu_{ij})\Big)$ is a diagonal matrix, then the expression for Δ reduces to*

$$\Delta = A_0^{-3/2} \sum_{i=1}^{d} \sum_{k=1}^{d} \{a_{ii} a_k^2 - a_i a_k a_{ik}\} \mu_{ii} \mu_{kk}. \tag{15}$$

Next, using the above Theorems, we obtain 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*

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

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. Then the coverage probability of the $100\alpha\%$ lower generalized confidence interval $\mathcal{I}^{'}_{n,\mathbf{x}}(\alpha)$ defined in* (2) *is given by*

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

where Δ is defined in (8).

Remark 2. *For the two-sided interval, using the expansions in Theorem 3 and Corollary 1, the coefficients of the $n^{-1/2}$ terms will cancel each other and thus it will be first order accurate. One may improve the accuracy further by first finding the $O(n^{-1})$ term and then adjusting the interval appropriately. 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 α_0 be given and $\hat{\alpha}_n$ be a solution to

$$\alpha - n^{-1/2}\hat{\Delta}\phi(z_\alpha) = \alpha_0. \tag{16}$$

Theorem 4. *Let $\hat{\Delta}$ be a $\sqrt{n}$-consistent estimator of Δ. Let $\hat{\alpha}_n$ be a solution to the equation (16) and let $\tilde{\alpha}_n = \max\{n^{-\epsilon}, \min(\hat{\alpha}_n, 1-n^{-\epsilon})\}$ where ϵ is defined in Theorem 1. Let the assumptions of Theorem 3 hold. Then*

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

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 $A = \left((a_{ij}^{(12)})\right)$ and $\boldsymbol{a} = (a_1^{(2)}, \ldots, a_d^{(2)})'$. 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 $\boldsymbol{b} = (\boldsymbol{a}'\Sigma\boldsymbol{a})^{-1/2}\Sigma^{1/2}\boldsymbol{a}$. Then

$$\Delta = A_0^{-1/2}\Big[\mathrm{tr}(B) - \boldsymbol{b}'B\boldsymbol{b}\Big]. \tag{17}$$

Note that $A_0 = (\boldsymbol{a}'\Sigma\boldsymbol{a})$ and $\|\boldsymbol{b}\| = 1$. The quantities B and $\boldsymbol{b}$ are parametric functions and we can write them as $B(\theta)$ and $\boldsymbol{b}(\theta)$, respectively. Equation (17) is essentially a way of rewriting the numerator of Δ in the matrix form. However, the expression is more compact and insightful. We may now restate the necessary and sufficient condition for the generalized intervals to be first order accurate as

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

$$\mathrm{tr}\{B(\theta)\} = \boldsymbol{b}(\theta)' B(\theta) \boldsymbol{b}(\theta), \quad \textit{for the true value } \theta. \tag{18}$$

4.1. $\boldsymbol{\Delta = 0}$ *and exact frequentist coverage*

When do generalized confidence intervals have exact frequentist coverage? A sufficient condition is the following. Note that in such cases Δ is necessarily zero.

Let $\psi(\mathbf{x}, t) : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ *be a function such that for every* $\mathbf{x} \in \mathbb{R}^d$, $\psi(\mathbf{x}, \cdot)$ *is a monotonically increasing function with* $\psi(\mathbf{x}, 0) = 0$. *Suppose there exists a function* $\tau_\theta : \mathbb{R}^d \to \mathbb{R}$ *such that* $T_\theta(\mathbf{x}, \mathbf{X}) = \psi(\mathbf{x}, \tau_\theta(\mathbf{x}) - \tau_\theta(\mathbf{X})) + T_\theta(\mathbf{x}, \mathbf{x})$. *Then the generalized confidence interval constructed for* $\pi(\theta)$ *based on* $T_\theta(\mathbf{x}, \mathbf{X})$ *has exact frequentist coverage and* $\Delta = 0$.

In this representation τ_θ is necessarily a pivotal quantity. However, it need not be a pivot for the parameter of interest. 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. The representation can be trivially obtained when a conventional pivot exists for the parameter of interest. An example where this representation is used to prove the exact coverage of the generalized pivot is give in Roy and Mathew [26]. Also see Hannig *et al.* [12].

Let us now verify that Δ is zero for this subclass of generalized pivots. 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

$$a_i^{(2)} = -\psi_1^{(2)} b_i; \quad a_{i,j}^{(2)} = -\psi_{i,1}^{(12)} b_j - \psi_{11}^{(2)} b_i b_j.$$

Hence the numerator of Δ is

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

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

Of course, $\Delta = 0$ does not guarantee exact frequentist coverage. Before we discuss an example where coverage is not exact in spite of Δ being zero, we point out the following facts which are obvious consequences of the Edgeworth expansions for smooth functions of sample means.

4.2.1. Edgeworth expansion using independent functions of 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, $\tilde{A}(\bar{\boldsymbol{y}})$ where $\bar{\boldsymbol{y}} = (\bar{y}_1, \ldots, \bar{y}_r)$ 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 maybe latent and not directly observable. In many of these problems, it maybe easier to derive the coefficients of the polynomials in the Edgeworth expansion for $P(n^{1/2}\tilde{A}(\bar{\boldsymbol{y}}) < w)$ 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 $\bar{\boldsymbol{y}}$ maybe means of possibly unequal number of observations and the number of observations maybe potentially of different order. This does not pose a problem while deriving an Edgeworth expansion for $n^{1/2}\tilde{A}(\bar{\boldsymbol{y}})$ 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 $\bar{\boldsymbol{y}}$ agree up to order n^{-1} with those of $\bar{\boldsymbol{y}}_{(n)}$ where $\bar{\boldsymbol{y}}_{(n)} = (\bar{y}_{1,(n)}, \ldots, \bar{y}_{r,(n)})$, $\bar{y}_{i,(n)} = (c_i n)^{-1} \sum_{j=1}^{c_i n} y_{i,j}$ and then proceed as in the proof of Theorem 2.1 in Hall [10]. Thus, if $P(n^{1/2}\tilde{A}(\bar{\boldsymbol{y}}) < w) = \Phi(w) + n^{-1/2} p_1(w)\phi(w) + O(n^{-1})$, we also have $P(n^{1/2}\tilde{A}(\bar{\boldsymbol{y}}_{(n)}) < w) = \Phi(w) + n^{-1/2} p_1(w)\phi(w) + O(n^{-1})$. 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 $\bar{\boldsymbol{z}} = (\bar{z}_1, \ldots, \bar{z}_r)$. Then the components of the mean vector $\bar{\boldsymbol{z}}$ are means of equal number of observations and the function $\tilde{A}(\bar{\boldsymbol{y}}_{(n)})$ can be written as $\tilde{A}(\bar{\boldsymbol{z}})$. 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 fashion.

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 are 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 take $\bar{z}_2$ as the 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 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. A generalized interval yields good results (Tsui and Weerahandi [29]). Though the pivot is not exact, we show that $\Delta = 0$. This explains the good performance of the generalized procedure.

Let $Z_{ij} \sim N(\tau_i, \sigma_i^2)$ for $i = 1, 2$ for the test group and the reference group, respectively, and $j = 1, \ldots, n$. Typically $n_1 \neq n_2$. But for deriving asymptotic properties we will take $n_1 = n_2 = n$. Otherwise one can work with $\min(n_1, n_2)$ as long as the sample sizes are of the same order. 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

$$\Sigma \;=\; ((\mu_{ij})) = \begin{pmatrix} \mu_3 & 0 & 0 & 0 \\ 0 & \mu_4 & 0 & 0 \\ 0 & 0 & 2\mu_3^2 & 0 \\ 0 & 0 & 0 & 2\mu_4^2 \end{pmatrix}. \tag{19}$$

The obvious generalized pivot is

$$T_\theta(\mathbf{x}, \mathbf{X}) = (x_1 - x_2) - \left[\frac{(X_1 - \mu_1)\sqrt{x_3}}{\sqrt{X_3}} - \frac{(X_2 - \mu_2)\sqrt{x_4}}{\sqrt{X_4}} \right]. \tag{20}$$

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

$$\left((a_{i,j}^{(12)}) \right) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\frac{1}{2\mu_3} & 0 & 0 & 0 \\ 0 & \frac{1}{2\mu_4} & 0 & 0 \end{pmatrix}. \tag{21}$$

Substituting the values of the derivatives and the asymptotic variances it can be shown that $\Delta = 0$.

4.3. Magnitude of $\boldsymbol{\Delta}$

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, 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 (17) 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

$$|\Delta| < d(d-1)M_1M_2 \tag{22}$$

where $M_1 = \max\limits_{i,j}(\mu_{ii}\mu_{jj}/A_0)^{1/2}$ and $M_2 = \max\limits_{i,j}0.5[a_{ij} + a_{ji}]$. In practice, the term maybe much smaller than the bound, and hence the first order term maybe negligible. We illustrate this by two examples.

Example 2: One way random model. Consider the one-way random effect model $Y_{ij} = \beta_0 + \beta_i + \epsilon_{ij}, \quad i = 1,\dots,k;\ j = 1,\dots,n$, 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

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

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 $\pi(\theta) = \sigma_\beta^2 = n^{-1}(\mu_1 - \mu_2)$. One must note that there are no lower confidence interval for $\pi(\theta)$ with exact frequentist coverage. Approximate confidence intervals such as Tukey-Williams confidence intervals (Tukey [30]; Williams [37]) have coverage that depend severely on the parameter values and can be very different from the nominal level depending on the sample size and the parameter configuration.

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 ([33] pp 152) proposed a generalized pivot

$$T_\theta(\mathbf{x},\mathbf{X}) = n^{-1}\Big[\frac{\mu_1x_1}{X_1} - \frac{\mu_2x_2}{X_2}\Big], \tag{23}$$

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.* (2006) have shown that the generalized confidence interval based on (23) has exact asymptotic coverage. However, the actual coverage could be different from the nominal level in small samples and depends typically on the quantity

$$0 < \lambda = \mu_2/\mu_1 < 1.$$

By the discussion in subsection 4.2.1, 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 $\mu_{12}=\mu_{21}=0$. The derivatives of the pivot are given by

$$\begin{aligned} \frac{\partial T_\theta(\mathbf{x},\mathbf{X})}{\partial X_1} &= \frac{-\mu_1 x_1}{nX_1^2};\ \frac{\partial T_\theta(\mathbf{x},\mathbf{X})}{\partial X_2} = \frac{\mu_2 x_2}{nX_2^2};\ \Rightarrow\ a_1^{(2)} = -\frac{1}{n};\ a_2^{(2)} = \frac{1}{n}; \\ \frac{\partial T_\theta(\mathbf{x},\mathbf{X})}{\partial x_1 \partial X_1} &= \frac{-\mu_1}{nX_1^2};\ \frac{\partial T_\theta(\mathbf{x},\mathbf{X})}{\partial x_2 \partial X_2} = \frac{\mu_2}{nX_2^2};\ \Rightarrow\ a_{1,1}^{(12)} = -\frac{1}{n\mu_1};\ a_{2,2}^{(12)} = \frac{1}{n\mu_2}. \end{aligned}$$

Also, $a_{i,j}^{(12)}=0$ if $i \neq j$. Thus, $A_0 = n^{-2}(\mu_{11}+\mu_{22})$.

Most of the cross terms are zero and the expression (8) reduces to

$$\begin{aligned} \Delta &= A_0^{-3/2}\sum_{i\neq j} a_{i,i}^{(2)}(a_j^{(2)})^2\mu_{ii}\mu_{jj} = \frac{n^{-3}[\mu_2^{-1}-\mu_1^{-1}]\mu_{11}\mu_{22}}{n^{-3}(\mu_{11}+\mu_{22})^{3/2}} \\ &= \frac{\sqrt{2}\lambda(1-\lambda)}{(n-1)(1+\lambda^2/(n-1))^{3/2}}. \end{aligned}$$

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 (22). 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. The responses are area under curve (AUC) for the plasma drug concentration curve over time since drug administration. As per FDA guidelines, the logarithm of AUC are analyzed and they are generally believed to be normally distributed, (thus making the responses lognormally distributed). Generalized inference for some of the problems related to bioequivalence has been discussed in McNally, Iyer and Mathew [24].

This specific statistical model is also discussed in Hannig *et al.* [12], where they show that the generalized confidence intervals have asymptotically correct coverage probability. Let $Z_{ij} = \log Y_{ij} \sim N(\tau_i, \sigma_i^2)$ where $i=1,2$ for the test group and the reference group, respectively, and $j=1,\ldots,n$. This is exactly the same set up as the Behrens-Fisher problem described in the previous section but 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 again 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

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

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

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

Substituting the values of the derivatives and the asymptotic variances in the expression for Δ, we have the numerator of Δ is equal to

$$\sum_{i,j,k,l} [a_{i,j}^{(12)} a_k^{(2)} a_l^{(2)} - a_i^{(2)} a_{j,k}^{(12)} a_l^{(2)}]\mu_{ij}\mu_{kl} \quad = \quad \frac{1}{2}(\mu_3 - \mu_4)(\mu_3\mu_4 - \mu_3 - \mu_4).$$

Similarly the asymptotic variance, A_0, appearing in the denominator of Δ is $\sum_{i,j} a_i^{(2)} a_j^{(2)} \mu_{ij} = \mu_3 + \mu_4 + \frac{1}{2}(\mu_3^2 + \mu_4^2)$. Thus,

$$\Delta = \left[\frac{\frac{1}{2}(\mu_3 - \mu_4)(\mu_3\mu_4 - \mu_3 - \mu_4)}{\{\mu_3 + \mu_4 + \frac{1}{2}(\mu_3^2 + \mu_4^2)\}^{3/2}}\right]. \tag{24}$$

Therefore, Δ can be both negative and positive. From Proposition 1 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.

4.4.1. $\Delta \neq 0$ *but known*

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

$$\alpha - n^{-1/2}\Delta\phi(z_\alpha) = \alpha_0, \tag{25}$$

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 [17] and Jobe [14] 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 [18]. Ananda [1] applied generalized procedure for computing confidence interval for availability and the results reported in [1] 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. There are however, no exact confidence bounds available in the literature for M. Jobe and David [15] 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 i th 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 the subsection 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,

$$a_{i,j}^{(12)} = \begin{cases} -\lambda_i^{-2}, & \text{if } i = j; i \text{ is odd}, \\ -\nu_i^{-2}, & \text{if } i = j; i \text{ is even}, \\ \lambda_i^{-1}\nu_i^{-1}, & \text{if } j = i+1; i \text{ is odd or } j = i-1; i \text{ is even}, \\ 0, & \text{otherwise}. \end{cases} \tag{26}$$

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 and 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.

TABLE 1
Performance of the corrected generalized intervals for Example 4

	$n = 5$		$n = 10$	
α_0	uncorrected	corrected	uncorrected	corrected
0.800	0.887	0.788	0.876	0.801
0.850	0.921	0.836	0.913	0.857
0.900	0.954	0.892	0.944	0.903
0.950	0.984	0.942	0.972	0.946

4.4.2. $\Delta \neq 0$ and is unknown

Typically the value of Δ is a function of the parameters and hence unknown. Let $\hat{\Delta}$ be a $\sqrt{n}$-consistent estimator of Δ. Such estimators are available in most common parametric problems where the generalized methods are used. Let $\hat{\alpha}_n$ be a solution to the equation (16). We may then use Theorem 4 to modify the generalized interval. Note that in order for the coverage probability expansion to hold the value of α needs to be in the interval $[n^{-\epsilon}, 1-n^{-\epsilon}]$ where ϵ was defined in Theorem 1. Thus, we use $\tilde{\alpha}_n$ instead of $\hat{\alpha}_n$ where $\tilde{\alpha}_n = \max\{n^{-\epsilon}, \min(\hat{\alpha}_n, 1-n^{-\epsilon})\}$. In practice, this would hardly make any difference as the target nominal level α_0 will be well within the interval (0,1) and the sequence $\hat{\alpha}_n$ will tend to α_0 in probability with $\hat{\alpha} - \alpha_0 = O_p(n^{-1})$ in the $P_{\mathbf{x}}$ probability. 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. The coefficient of n^{-1} in the expansion of the coverage probability maybe large and thus make a significant contribution to the coverage error when n is small. The first order term will be the dominating term in the coverage error for larger samples, but as pointed out earlier in many examples the regular generalized intervals perform adequately enough, specially for moderately large sample and thereby making any adjustment to the intervals quite unnecessary. However, we expect there are still a number of examples where one would benefit from modifying the usual generalized intervals to make them first order correct.

Example 5. There are ample examples of data arising in economics and health research 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 of lognormal means can be used for relative comparisons. However, in certain situations, the difference of the lognormal means maybe the quantity of interest. Krishnamoorthy and Mathew [19] have proposed a generalized pivotal approach for the difference of two lognormal means. Also see Chen and Zhou [5]. 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 parameter τ_1 and σ_1^2 and let $Y_{2,1}, \ldots, Y_{2,n}$ be a random sample from lognormal distribution with parameter 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 statistics is $(\mu_1,\ \mu_2,\ \mu_3) = (\tau_1,\ \tau_2,\ \sigma^2)$. Then the generalized pivot for θ is

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

The asymptotic covariance matrix for $(X_1,\ X_2,\ X_3)$ is a diagonal matrix with diagonal elements $(\sigma^2, \sigma^2, \sigma^4)$. The derivatives of the pivot are given by $(a_1^{(2)},\ a_2^{(2)},\ a_3^{(2)}) = (-\theta_1,\ \theta_2,\ \theta/2)$ and

$$\left((a_{ij}^{(12)})\right) = \begin{pmatrix} -\theta_1 & 0 & -0.5\theta_1 \\ 0 & \theta_2 & 0.5\theta_2 \\ -0.5\theta_1(1+\sigma^{-2}) & 0.5\theta_2(1+\sigma^{-2}) & -0.5\theta(0.5+\sigma^{-2}) \end{pmatrix}.$$

Thus, the relevant quantities in computation of Δ are:

$$A_0 = \theta_1^2\sigma^2 + \theta_2^2\sigma^2 + 0.25\theta^2\sigma^4 = \sigma^2[\theta^2(1+0.25\sigma^2) + 2\theta_1\theta_2],$$

$$\sum_{i,j=1}^{3}\sum_{k,l=1}^{3}[a_{ij}^{(12)}a_k^{(2)}a_l^{(2)} - a_i^{(2)}a_{jk}^{(12)}a_l^{(2)}]\mu_{ij}\mu_{kl} = \sigma^4\theta[\theta_1\theta_2(0.5\sigma^2 - 1) - 0.25\theta^2].$$

Hence,

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

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.

TABLE 2
Performance of the modified generalized intervals relative to the usual generalized intervals in Example 5

	$n = 25$		$n = 50$	
α_0	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

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.

Appendix

Proof of Theorem 1: The results of Theorem 1 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 (14) follows directly along the lines of Theorem 5.1 of Hall ([10]; pp-239). The bootstrap probability measure (conditional on the sample) in Theorem 5.1 has to be replaced by the measure associated with the pseudo randomization 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 rest on the behavior of the cumulants for the bootstrap probability measure conditional on the sample. The behavior of the cumulants under the pseudo randomization of the generalized pivot in our set up is almost identical to those for the bootstrap measure.

To prove the claim about the generalized pivot Cornish-Fisher expansion (11), we use the generalized pivot Edgeworth expansion (9) and follow the proof of Theorem 5.2 of Hall ([10]; 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 2: 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

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

Therefore

$$\begin{aligned}
\frac{\partial H(\mathbf{x})}{\partial \mathbf{x}_i}\Big|_{\mathbf{x}=\boldsymbol{\mu}} &= a_i^{(2)}, \\
\frac{\partial G(\mathbf{x})}{\partial \mathbf{x}_i}\Big|_{\mathbf{x}=\boldsymbol{\mu}} &= A_0^{-3/2} \sum_{k,l} a_{i,k}^{(12)} a_l^{(2)} \mu_{kl},
\end{aligned}$$

and

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

Thus,

$$\tilde{A}_0 = A_0^{-1} \sum_{ij} a_i^{(2)} a_j^{(2)} \mu_{ij} = A_0^{-1} A_0 = 1, \tag{27}$$

and

$$\begin{aligned}
\tilde{A}_1 &= A_0^{-1/2} A_1 + 2^{-1} A_0^{-3/2} \sum_{i,j,k,l} \Big[(a_{i,j}^{(12)} + a_{j,i}^{(12)}) a_k^{(2)} a_l^{(2)} \\
&\qquad - a_j^{(2)} a_{i,k}^{(12)} a_l^{(2)} - a_i^{(2)} a_{j,l}^{(12)} a_k^{(2)} \Big] \mu_{ij} \mu_{kl} \\
&= A_0^{-1/2} A_1 + A_0^{-3/2} \sum_{i,j,k,l} \Big[a_{i,j}^{(12)} a_k^{(2)} a_l^{(2)} - a_i^{(2)} a_{j,l}^{(12)} a_k^{(2)} \Big] \mu_{ij} \mu_{kl} \\
&= A_0^{-1/2} A_1 + \Delta.
\end{aligned} \tag{28}$$

Also

$$\begin{aligned}
\tilde{A}_{21} &= A_0^{-3/2} \sum_{ijk} a_i^{(2)} a_j^{(2)} a_k^{(2)} \mu_{ijk} = A_0^{-3/2} A_{21}, \\
\tilde{A}_{22} &= A_0^{-3/2} A_{22} + A_0^{-5/2} \sum_{ijkl} \Big[a_i^{(2)} a_j^{(2)} (a_{k,l}^{(12)} + a_{l,k}^{(12)}) a_m^{(2)} a_n^{(2)} - \\
&\qquad a_i^{(2)} a_j^{(2)} (a_{k,m}^{(12)} a_l^{(2)} a_n^{(2)} + a_k^{(2)} a_{l,n}^{(12)} a_m^{(2)}) \Big] \mu_{ik} \mu_{jl} \mu_{mn} \\
&:= A_0^{-3/2} A_{22} + \Delta_2.
\end{aligned} \tag{29}$$

The second part in the last equality in (29) is

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

Hence the result. □

Proof of Theorem 3:

$$\begin{aligned}P_{\mathbf{x}}[\pi(\theta) \in \mathcal{I}_{n,\mathbf{x}}(\alpha)] &= P_{\mathbf{x}}[P_{\mathbf{X}}\{T_\theta(\mathbf{x},\mathbf{X}) < T_\theta(\mathbf{x},\mathbf{x})\} < \alpha] \\ &= P_{\mathbf{x}}[P_{\mathbf{X}}\{Z_{n,\mathbf{X}}(\mathbf{x}) < Z_n(\mathbf{x})\} < \alpha] \\ &= P_{\mathbf{x}}(Z_n(\mathbf{x}) < F_{n,\mathbf{x},\mathbf{x}}^{-1}(\alpha)) \\ &= P_{\mathbf{x}}(Z_n(\mathbf{x}) - C_n(\mathbf{x},\alpha) < z_\alpha - n^{-1/2} p_{1,\mathbf{x}}(z_\alpha)) \quad (30)\end{aligned}$$

Let

$$g_n(\mathbf{x},\alpha) = z_\alpha - n^{-1/2} p_{1,\mathbf{x}}(z_\alpha). \quad (31)$$

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

$$P_{\mathbf{x}}\{|C_n(\mathbf{x},\alpha)| > C_3\epsilon_n\} = o(n^{-1/2}), \quad (32)$$

and

$$P_{\mathbf{x}}\{|D_n(\mathbf{x},\alpha)| > C_3\epsilon_n\} = o(n^{-1/2}), \quad (33)$$

for some constant $C_3 > 0$. Choose $\epsilon_n = n^{-\beta}$ where $\beta > 1/2$ but close to $1/2$. The relation (32) follows from (12). 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 neighbourhood of $\boldsymbol{\mu}$ where $\delta_n = o(n^{1/2}\epsilon_n)$. By assumption (A1), we have

$$P_{\mathbf{x}}\{\|\mathbf{x}-\boldsymbol{\mu}\| \ge \delta_n\} \le C_4[\delta_n n^{1/2}]^{-3}. \tag{34}$$

Therefore,

$$P_{\mathbf{x}}\{|D_n(\mathbf{x},\alpha)| > C_3\epsilon_n\} = O\Big(n^{-3}\epsilon_n^{-3}\Big). \tag{35}$$

It is easy to see that one can choose δ_n and β such that the $O\Big(n^{-3}\epsilon_n^{-3}\Big)$ term in (35) is indeed $o(n^{-1/2})$. Thus,

$$P_{\mathbf{x}}[\pi(\theta) \in \mathcal{I}_{n,\mathbf{x}}(\alpha)] = P_{\mathbf{x}}(Z_n(\mathbf{x}) < \tilde{g}_n(\boldsymbol{\mu},\alpha)) + o(n^{-1/2}). \tag{36}$$

Using the one term Edgeworth expansion (14) for the studentized statist ic $Z_n(\mathbf{x})$, 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

$$\begin{aligned}
P_{\mathbf{x}}[\pi(\theta) \in \mathcal{I}_{n,\mathbf{x}}(\alpha)] &= \Phi(\tilde{g}_n(\boldsymbol{\mu},\alpha)) \\
&\quad + n^{-1/2} q_{1,\boldsymbol{\mu}}(\tilde{g}_n(\boldsymbol{\mu},\alpha))\phi(\tilde{g}_n(\boldsymbol{\mu},\alpha)) + o(n^{-1/2}) \\
&= \alpha - n^{-1/2}[p_{1,\boldsymbol{\mu}}(z_\alpha) - q_{1,\boldsymbol{\mu}}(z_\alpha)]\phi(z_\alpha) + o(n^{-1/2}) \\
&= \alpha - n^{-1/2}\Delta\phi(z_\alpha) + o(n^{-1/2}). \qquad (37)
\end{aligned}$$

This completes the proof of Theorem 3. □

The error term in (37) 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 ([10], 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*

$$P_{\mathbf{x}}\{|\hat{\Delta} - \Delta| > C_5 n^{\delta - 1/2}\} = o(n^{-1/2}), \tag{38}$$

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 (31). 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 $\Delta \ne 0$. 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

$$(\hat{\alpha}_n - \alpha_n) = \frac{-n^{-1/2}(\hat{\Delta} - \Delta)\phi(z_{\hat{\alpha}_n})}{1 + n^{-1/2}\Delta z_{\alpha_n^*}}, \tag{39}$$

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 (39) we have

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

where

$$v_n(\mathbf{x}, \alpha) = \frac{\phi(z_{\hat{\alpha}_n})[1 - n^{-1/2}\frac{A_2(\mathbf{x},\boldsymbol{\mu})}{A_0^{3/2}(\mathbf{x},\boldsymbol{\mu})}(z_{\hat{\alpha}_n} + z_{\alpha_n})]}{\phi(z_{\alpha_n^{**}})[1 + n^{-1/2}\Delta z_{\alpha_n^*}]} \tag{40}$$

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 (38) we have

$$P_{\mathbf{x}}\{\frac{\phi(z_{\hat{\alpha}_n})}{\phi(z_{\alpha_n^{**}})} \geq C_{10}\} = o(n^{-1/2}), \tag{41}$$

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,

$$P_{\mathbf{x}}\{1 - n^{-1/2}(z_{\hat{\alpha}_n} + z_{\alpha_n})\frac{A_2(\mathbf{x}, \boldsymbol{\mu})}{A_0(\mathbf{x}, \boldsymbol{\mu})^{3/2}} > 1/2\} = o(n^{-1/2}). \tag{42}$$

Therefore, from (42), (41) and the fact that for large n, $[1 + n^{-1/2}\Delta z_{\alpha_n^*}]^{-1}$ is less than 2, we have

$$P_{\mathbf{x}}\{|v_n\mathbf{x}, \alpha| > C_{10}\} = o(n^{-1/2}). \tag{43}$$

Then by assumption (38) we have the result. □

Proof of Theorem 4: By (30), 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,

$$\begin{aligned} P_{\mathbf{x}}[\pi(\theta) \in \mathcal{I}_{n,\mathbf{x}}(\tilde{\alpha}_n)] &= P_{\mathbf{x}}(Z_n(\mathbf{x}) < g_n(\mathbf{x}, \tilde{\alpha}_n)) + o(n^{-1/2}) \\ &= P_{\mathbf{x}}(Z_n(\mathbf{x}) < g_n(\mathbf{x}, \alpha_n) + \tilde{r}_n(\mathbf{x}, \alpha)) + o(n^{-1/2}) \quad (44) \end{aligned}$$

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 (44) 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

$$\begin{aligned} P_{\mathbf{x}}[\pi(\theta) \in \mathcal{I}_{n,\mathbf{x}}(\tilde{\alpha}_n)] &= P_{\mathbf{x}}(Z_n(\mathbf{x}) < g_n(\boldsymbol{\mu},\alpha_n)) + o(n^{-1/2}) \\ &= \alpha_n - n^{-1/2}\Delta\phi(z_{\alpha_n}) + o(n^{-1/2}) \\ &= \alpha_0 + o(n^{-1/2}). \end{aligned} \tag{45}$$

Hence the result. □

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

PROPOSITION 1. *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 sums of squares optimization methods (Parrilo and Sturmfels [25]) 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

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

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] ANANDA, M. M. A. (1999). Estimation and testing of availability of a parallel system with exponential failure and repair times. *J. Statist. Plann. Inference* **77**, 237–246.

[2] ARENDACKA, B. (2005). Generalized confidence intervals on the variance component in mixed linear models with two variance components. *Statistics* **39**, 275-286.

[3] Burdick, R. K., Borror, C. M. and Montgomery, D. C. (2005). *Design and Analysis of Gauge R&R Studies: Making Decisions with Confidence Intervals in Random and Mixed ANOVA Models.* ASA-SIAM Series on Statistics and Applied Probability.

[4] Chang, Y. P. and Huang, W. T. (2000). Generalized confidence intervals for the largest value of some functions of parameters under normality, *Statist. Sinica* **10**, 1369–1383.

[5] Chen, Y. H. and Zhou, X. H. (2006). Generalized confidence intervals for the ratio or difference of two means for lognormal populations with zeros, *UW Biostatistics Working Paper Series. Working Paper 296.*

[6] Chiang, A. K. L. (2001). Simple General Method for Constructing Confidence Intervals for functions of variance components *Technometrics* **43**, 356–367.

[7] Gamage, J. and Weerahandi, S. (1998). Size performance of some tests in one-way ANOVA. *Comm. Statist. Theory Methods* **27**, 625–640.

[8] Gamage, J., Mathew, T. and Weerahandi, S. (2004). Generalized p-values and generalized confidence regions for the multivariate Behrens-Fisher problem. *J. Multivariate Anal.* **88**, 177–189.

[9] Griffiths, W. and Judge, G. (1992). Testing and estimating location vectors when the error covariance matrix is unknown. *J. Econometrics* **54**, 121–138.

[10] Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*, Springer-Verlag.

[11] Hannig, J. (2006). On fiducial inference, the good, the bad, and the ugly. Technical Report 2006/3, Colorado State University, Dept. of Statistics. Available at *http://www.stat.colostate.edu/research/tech_reports.html*

[12] Hannig, J., Iyer, H. and Patterson, P. (2006). Fiducial Generalized Confidence Intervals. *J. Amer. Statist. Assoc.* **101**, 254–269.

[13] Iyer, H. K. , Wang, C. M., and Mathew, T. (2004). Models and confidence intervals for true values in interlaboratory trials, *J. Amer. Statist. Assoc.* **99,** 1060-1071.

[14] Jobe, J. M. (1988). A new R&M measure and its estimation. *IEEE Trans. Reliability* **37**, 413-420.

[15] Jobe J. M. and David, H. T. (1992). Buehler confidence bounds for a reliability-maintainability measure. *Technometrics* **34,** 214–222.

[16] Johnson, R. A. and Weerahandi, S. (1998). A Bayesian Solution to the multivariate Behrens-Fisher problem. *J. Amer. Statist. Assoc.* **83**, 145–149.

[17] Klion, J. (1986). Specifying maintainability – a new approach. *Proceedings of the Annual Reliability and Maintainability Symposium,* Las Vegas: IEEE, 338–343.

[18] Knezevic, J. (1993). *Reliability, maintainability and supportability: a probabilistic approach.* McGraw-Hill International (UK) Limited.

[19] Krishnamoorthy K and Mathew T. (2003). Inferences on the means of lognormal distributions using generalized p-values and generalized confidence intervals. *J. Statist. Plann. Inference* **115**, 103–120.

[20] Krishnamoorthy, K. and Mathew, T. (2004). One-Sided tolerance lim-

its in balanced and unbalanced one-way random models based on generalized confidence limits, *Technometrics* **46**, 44–52.
[21] Lin, S. H. and Lee, J. C. (2003). Exact tests in simple growth curve models and one-way ANOVA with equicorrelation error structure, *J. Multivariate Anal.* **84**, 351 – 368.
[22] Lin, T. Y. and Liao, C. T. (2006). A β-expectation tolerance interval for general balanced mixed linear models, *Comput. Statist. Data Anal.* **50**, 911–925
[23] Mathew, T. and Webb, D. W. (2005). Generalized p values and confidence intervals for variance components: applications to army test and evaluation. *Technometrics* **47**, 3120–322.
[24] McNally, R.J., Iyer, H. and Mathew, T. (2003). Tests for individual and population bioequivalence based on generalized p-values. *Stat. Med.* **22**, 31–53.
[25] Parrilo, P. A. and Sturmfels, B. (2003). Minimizing polynomial functions. *Algorithmic and quantitative real algebraic geometry, DIMACS Series in Discrete Mathematics and Theoretical Computer Science* **60**, 83–99.
[26] Roy, A. and Mathew, T. (2005). A Generalized Confidence Limit for the Reliability Function of a Two-Parameter Exponential Distribution, *J. Statist. Plann. Inference* **128**, 509–517.
[27] Tian, L. and Cappelleri, J. C. (2004). A new approach for interval estimation and hypothesis testing of a certain intraclass correlation coefficient: the generalized variable method. *Stat. Med.* **23**, 2125–2135,
[28] Thursby, J.G. (1992). A comparison of several exact and approximate tests for structural shift under heteroscedasticity. *J. Econometrics* **53**, 363–386.
[29] Tsui, K. W. and Weerahandi, S. (1989). Generalized P-values in Significance Testing of Hypotheses in the Presence of Nuisance Parameters, *J. Amer. Statist. Assoc.,* **84**, 602–607.
[30] Tukey, J. W. (1951). Components in regression, *Biometrics* **7**, 33–69.
[31] Weerahandi, S. (1991). Testing Variance Components in Mixed Models with Generalized p values, *J. Amer. Statist. Assoc.* **86**, 151–153.
[32] Weerahandi, S. (1993). Generalized Confidence Intervals, *J. Amer. Statist. Assoc.* **88**, 899–905.
[33] Weerahandi, S. (1995). *Exact Statistical Methods for Data Analysis*, Springer-Verlag.
[34] Weerahandi, S. (2004). *Generalized Inference in Repeated Measures: Exact Methods in MANOVA and Mixed Models*, John Wiley and Sons, Inc.
[35] Weerhandi, S. and Amaratunga, D. (1999). A performance comparison of methods in mixed models. *Technical report, Bell Communications.*
[36] Weerahandi, S. and Berger, V. W. (1999). Exact inference for growth curves with intraclass correlation structure, *Biometrics* **55**, 921–924.
[37] Williams, J. S. 1962. A Confidence interval for variance components, *Biometrics* **49**, 278–281.
[38] Zhou, L. and Mathew, T. (1994). Some tests for variance components using generalized p-values. *Technometrics* **36**, 394–421.

Anindya Roy
Department of Mathematics and Statistics
University of Maryland
1000 Hilltop Cir.,
Baltimore, MD 21250, USA
anindya@math.umbc.edu

Arup Bose
Stat-Math Unit
Indian Statistical Institute
202 B. T. Road
Kolkata 700108, India
abose@isical.ac.in