Behav Genet (2012) 42:886–898
DOI 10.1007/s10519-012-9560-z

ORIGINAL RESEARCH

Adjusted Confidence Intervals for a Bounded Parameter

Hao Wu · Michael C. Neale

Received: 4 November 2011 / Accepted: 13 August 2012 / Published online: 13 September 2012

Abstract It is well known that the regular likelihood ratio test of a bounded parameter is not valid if the boundary value is being tested. This is the case for testing the null value of a scalar variance component. Although an adjusted test of variance component has been suggested to account for the effect of its lower bound of zero, no adjustment of its interval estimate has ever been proposed. If left unadjusted, the confidence interval of the variance may still contain zero when the adjusted test rejects the null hypothesis of a zero variance, leading to conflicting conclusions. In this research, we propose two ways to adjust the confidence interval of a parameter subject to a lower bound, one based on the Wald test and the other on the likelihood ratio test. Both are compatible to the adjusted test and parametrization-invariant. A simulation study and two examples are given in the framework of ACDE models in twin studies.

Keywords Confidence interval · Variance component · Likelihood ratio test · Wald test · ACDE models

Edited by Gitta Lubke.

H. Wu (✉)
Department of Psychology, Boston College, 300 McGuinn, 140 Commonwealth Avenue, Chestnut Hil, MA 02467, USA
e-mail: hao.wu.5@bc.edu

M. C. Neale
Department of Psychiatry and Virginia Institute for Psychiatric and Behavior Genetics, Virginia Commonwealth University, E. 800 Leigh St, Richmond, VA 23219, USA

Introduction

Confidence intervals

Statistical models, such as structural equation models (SEM), have been widely used in the analysis of twin and family data. Consequently, confidence intervals (CIs) are routinely presented along with point estimates of model parameters to quantify the uncertainty of parameter estimates due to sampling errors. Let $\boldsymbol{\xi} = (\theta, \boldsymbol{\zeta}')'$ be the parameter vector in a model in which θ is the parameter of interest and $\boldsymbol{\zeta}$ is the vector of other parameters. A $100(1 - \alpha)$ % CI of θ is defined as the set of all test values θ_0 such that the following statistical testing problem does not reject the null hypothesis at level α (Lehmann 1986, section 5.6):

$$\mathrm{H}_0: \quad \theta = \theta_0 \leftrightarrow \mathrm{H}_1: \quad \theta \neq \theta_0 \tag{1}$$

By definition, the construction of a CI is usually performed by inverting a test, or finding the true values that cannot be rejected by this test. If test (1) is performed using the Wald statistic

$$W = (\hat{\theta} - \theta_0)^2 / \sigma^2(\hat{\boldsymbol{\xi}}), \tag{2}$$

the CI can be constructed as $(\hat{\theta}_L, \hat{\theta}_U)$ with $\hat{\theta}_L = \hat{\theta} - \sigma(\hat{\boldsymbol{\xi}}) z_{\alpha/2}$ and $\hat{\theta}_U = \hat{\theta} + \sigma(\hat{\boldsymbol{\xi}}) z_{\alpha/2}$, where $\sigma(\hat{\boldsymbol{\xi}})$ is a consistent estimate of the standard error (SE) of $\hat{\theta}$ and z_α is the $1 - \alpha$ quantile of a standard normal distribution. The subscripts $_L$ and $_U$ are used to denote lower and upper limits throughout this article. Usually, $\sigma^2(\hat{\boldsymbol{\xi}})$ is taken as the corresponding diagonal element of the inverse Fisher information matrix. Note that a 95 % CI would require 2.5 % of the distribution at either side, which is why α is divided by two to find the upper and lower limits.

Although the SE-based CIs are widely used and easy to compute, they inherit some problems from the Wald test.

Because the Wald test is not parametrization-invariant, the SE based CIs of different monotonic functions of the same parameter are usually incompatible with each other. For example, the upper and lower limits of a CI constructed for a variance are usually not the squares of those of a CI for the corresponding standard deviation (Neale et al. 1989). In addition, the performance of a SE-based CI depends on whether the Wald statistic is close to normally distributed for the given parametrization and sample size. For example, Fisher's z transformation of a correlation usually requires smaller sample size to achieve asymptotic normality than does the original correlation parameter, and the coverage of a CI based on the z transformation is usually closer to the nominal level.

To avoid the problems above, Neale and Miller (1997) proposed a likelihood-based CI by inverting the likelihood ratio test (LRT). Because the LRT is parametrization-invariant, the resultant CIs have similar behavior. To obtain the likelihood-based CI, the function

$$(F(\hat{\xi}) - F(\xi) + z_{\alpha/2}^2)^2 \pm \theta \quad (3)$$

is minimized with respect to parameter vector $\xi = (\theta, \zeta')'$, where F is the negative twice log likelihood function and $\hat{\xi}$ is the maximum likelihood estimate (MLE). The resultant values of θ at the minima are the upper and lower limits of the CI. It was noted that this procedure is not strictly equivalent to inverting a LRT because the value of the likelihood function at the minima differs from their desired values by a (usually slight) bias.

Confidence intervals close to a boundary

Because CIs are formed by inverting tests, they are no longer valid when the tests are not valid. This situation is encountered when the true parameter of interest θ_0 is on or close to its boundary. When θ_0 is on its boundary, test (1) becomes a one sided test because the alternative hypothesis now lies only on one side of the null hypothesis, and the χ^2 asymptotic distribution in a regular LRT is no longer valid. LRTs under such situations have been the subject of a series of studies in general multivariate statistics and in ACE models (Self and Liang 1987; Shapiro 1988; Dominicus et al. 2006; Visscher 2006; Wu and Neale in press). In general, when the parameter of interest is the only parameter on a boundary and the Fisher information matrix has full rank, the asymptotic sampling distribution of the LRT statistic for test (1) is an equally weighted mixture of a χ_1^2 distribution and a point mass at 0 (Self and Liang 1987), and the p value is half of that of a regular LRT (Dominicus et al. 2006).

Under such situation, the CI constructed by inverting the LRT test also needs to be corrected, or it will no longer be compatible with the test. It may happen that a (corrected) LRT test rejects the null hypothesis because the corrected p value is smaller than the specified α level, but the uncorrected CI still contains the boundary because the p value of a regular LRT is still larger than α. Nevertheless, modifying the CI involves more than modifying a boundary LRT because the corrected LRT can only suggest whether the boundary value is inside the CI, while the modified CI needs to determine all true values that would not be reject by the LRT, which concerns LRTs with true values *close* to the boundary.

When the true value θ_0 is not on the boundary, the chance that the MLE $\hat{\theta}$ falls beyond the boundary can be made arbitrarily small by increasing sample size (which decreases the variance of $\hat{\theta}$). This implies that the χ^2 distribution is still valid as an *asymptotic* distribution. Unfortunately, this asymptotic result is not helpful for the current purpose of CI construction because (1) the actual sample size required for a valid regular LRT increases unboundedly as θ_0 gets closer and closer to its boundary, and (2) it cannot give a CI that is compatible with the corrected boundary LRT. The construction of a CI of θ, which concerns a range of true values θ_0 in the neighborhood of the boundary, must take into account the effect of a finite sample to bridge the gap between the regular and corrected asymptotic distributions.

In this paper, we propose two methods that yield a CI compatible with the corrected LRT, one based on the inversion of a Wald test and the other based on the inversion of a LRT. For both methods, we begin with a simplified example of a bounded mean parameter of a normal distribution and then discuss models with bounded parameters in general. Algorithms for computation will be given and simulation studies and examples will also be presented.

Correction based on a Wald test

Testing a normal mean close to a bound

We begin with a brief review of the hypothesis testing procedure and CI construction of the mean parameter μ in a normal distribution $N(\mu,1)$. If a random sample $\{X_1, X_2, \ldots, X_n\}$ of size n is obtained, we have $\hat{\mu} = \bar{X}$. Remember that $\hat{\mu}$ is a random variable and has sampling distribution $N(\mu_0, \sigma_0^2)$, where μ_0 is the true value of μ in the population and $\sigma_0^2 = 1/n$ is the variance of the sampling distribution. For a given sample and observed value $\hat{\mu}_{obs}$ and a given test value μ_0, the p value is defined as the probability to obtain a $\hat{\mu}$ that is more extreme than the current observation $\hat{\mu}_{obs}$, or

$$p = \Pr(|\hat{\mu} - \mu_0| \geq |\hat{\mu}_{obs} - \mu_0|) = 2\Phi(-|\hat{\mu}_{obs} - \mu_0|/\sigma_0), \quad (4)$$

where $\Phi()$ is the cumulative distribution function (cdf) of a standard normal distribution. If this p value is smaller than or equal to α, the specified significance level, the null hypothesis of test (1) is rejected. The $100(1-\alpha)$ % CI is obtained by solving for the μ_0 from the inequality $p > \alpha$ and is given by $\hat{\mu}_{\text{obs}} \pm \sigma_0 z_{\alpha/2}$.

When the boundary condition $\mu \geq 0$ is imposed, the MLE becomes $\hat{\mu}^+ = \max\{\bar{X}, 0\}$. It coincides with $\hat{\mu}$ if $\hat{\mu} \geq 0$ but remains at zero if $\hat{\mu} < 0$. As a result, it has a similar distribution to $\hat{\mu}$ except that a point mass on the value zero replaces the part of the distribution below zero. The size of this point mass is given by

$$q_0 = \Pr(\hat{\mu}^+ = 0) = \Pr(\bar{X} \leq 0) = \Phi(-\mu_0/\sigma_0).$$

Because this distribution has both discrete and continuous parts, neither a probability mass function (pmf) nor a probability density function (pdf) can describe it appropriately. Its cdf is shown in Fig. 1.

For a given observed value of the restricted MLE, $\hat{\mu}^+_{\text{obs}}$, the p value can be calculated as

$$\begin{aligned} p &= \Pr(|\hat{\mu}^+ - \mu_0| \geq |\hat{\mu}^+_{\text{obs}} - \mu_0|) \\ &= \Pr\left(\hat{\mu}^+ - \mu_0 \geq |\hat{\mu}^+_{\text{obs}} - \mu_0|\right) + \Pr\left(\hat{\mu}^+ - \mu_0 \leq -|\hat{\mu}^+_{\text{obs}} - \mu_0|\right) \\ &= \Phi\left(-|\hat{\mu}^+_{\text{obs}} - \mu_0|/\sigma_0\right) \\ &\quad + \begin{cases} \Phi\left(-|\hat{\mu}^+_{\text{obs}} - \mu_0|/\sigma_0\right) & \text{if} \quad \mu_0 - |\hat{\mu}^+_{\text{obs}} - \mu_0| \geq 0 \\ 0 & \text{if} \quad \mu_0 - |\hat{\mu}^+_{\text{obs}} - \mu_0| < 0 \end{cases} \qquad (5) \\ &= \begin{cases} 2\Phi\left(-|\hat{\mu}^+_{\text{obs}} - \mu_0|/\sigma_0\right) & \text{if} \quad 0 \leq \hat{\mu}^+_{\text{obs}} \leq 2\mu_0 \\ \Phi\left(-(\hat{\mu}^+_{\text{obs}} - \mu_0)/\sigma_0\right) & \text{if} \quad \hat{\mu}^+_{\text{obs}} > 2\mu_0 \end{cases}, \qquad (6) \end{aligned}$$

To understand this result, we first note that it is different from Eq. (4) because the sampling distribution of $\hat{\mu}^+$ is different from that of $\hat{\mu}$. In particular, when

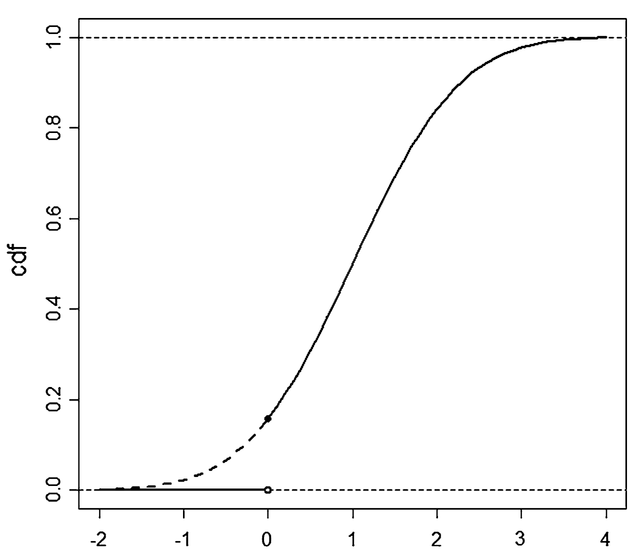

Fig. 1 The cumulative distribution function (cdf) of the sampling distribution of $\hat{\mu}^+ = \max\{\bar{X}, 0\}$ for a normal model N(μ,1) with a bounded mean $\mu \geq 0$. Note the jump indicates the value 0 has a positive probability. A regular normal cdf would follow the *dashed curve* for negative values and no single value would have positive probability

$\mu_0 - |\hat{\mu}^+_{\text{obs}} - \mu_0| < 0$, the second term in (5) is 0 because $\hat{\mu}^+$ cannot take negative values. As a result, the p value in this case is just a half of what it would be when no boundary condition were imposed on parameter μ. Note the condition $\mu_0 - |\hat{\mu}^+_{\text{obs}} - \mu_0| < 0$ simplifies to $\hat{\mu}^+_{\text{obs}} > 2\mu_0$ in Eq. (6). When $\mu_0 - |\hat{\mu}^+_{\text{obs}} - \mu_0| \geq 0$, or equivalently $0 \leq \hat{\mu}^+_{\text{obs}} \leq 2\mu_0$, the p value remains the same as in a non-bounded normal mean problem.

The p value given by Eq. 6 is plotted as a function of the observed MLE $\hat{\mu}^+$ for $\mu_0 = 1.5$ and $\sigma_0 = 1$ in Fig. 2. We note that in the region of $[0, 2\mu_0]$, the curve follows a two-sided p value and is symmetric around the test value μ_0. However, a dent occurs at $2\mu_0$ where it switches from a two-sided p value to a one-sided p value. If a horizontal line of $p = \alpha$ intersects with the symmetric part of the curve, the α level test is the regular two sided test; if the line intersects with the trailing part of the curve, the α level test is the regular one-sided test; if the line goes through the jump, the α level test is one-sided with rejection region $\hat{\mu}^+_{\text{obs}} > 2\mu_0$ and its actual type I error rate is smaller than α. The last situation occurs when α lies between $2\Phi(-\mu_0/\sigma_0)$ and $\Phi(-\mu_0/\sigma_0)$, or equivalently, when μ_0 lies between $z_{\alpha/2}\sigma_0$ and $z_\alpha \sigma_0$.

A confidence interval for a bounded normal mean

The Wald test adjusted CI of μ can be obtained by collecting all true values μ_0 that would yield a p value greater than α. When the estimate is on the boundary, or $\hat{\mu}^+_{\text{obs}} = 0$, the CI is given by $[0, \sigma_0 z_{\alpha/2}]$. When the estimate is not on the boundary, solving for μ_0 from $p \geq \alpha$ using Eq. (6) can be quite cumbersome due to the complicated form of the p value. Instead, we plot the p value given by Eq. (6) as a function of test value μ_0 for a fixed $\hat{\mu}^+_{\text{obs}} > 0$. This plot is shown in Fig. 3. We can see that the upper limit of the CI is always calculated from the solid curve of a regular two-sided p value and is always the same as the upper limit of a regular CI for a normal mean, or $\hat{\mu}^+_{\text{obs}} + \sigma_0 z_{\alpha/2}$. The lower limit is more complicated, depending on which part of the p value curve the horizontal line of $p = \alpha$ intersects. If the line $p = \alpha$ intersects the solid part of the p value curve, the lower limit of CI is given by $\hat{\mu}_L^{\alpha/2} = \hat{\mu}^+_{\text{obs}} - \sigma_0 z_{\alpha/2}$, the lower limit of a regular CI, and it must be greater than or equal to $\hat{\mu}^+_{\text{obs}}/2$, where the p value curve jumps. If the line $p = \alpha$ intersects the dashed part of the p value curve, the lower limit of CI is given by $\max\{\hat{\mu}_L^\alpha, 0\}$, where $\hat{\mu}_L^\alpha = \hat{\mu}^+_{\text{obs}} - \sigma_0 z_\alpha$, the lower limit of a $100(1-2\alpha)$ % CI, and it must be smaller than $\hat{\mu}^+_{\text{obs}}/2$. If the line $p = \alpha$ goes through the jump between the two branches, the lower limit of CI is given by $\hat{\mu}^+_{\text{obs}}/2$, the location of the jump. In practice, one may calculate $\hat{\mu}_L^{\alpha/2}$, $\hat{\mu}_L^\alpha$ and $\hat{\mu}^+_{\text{obs}}/2$, the three potential candidates of the lower limit, and determine

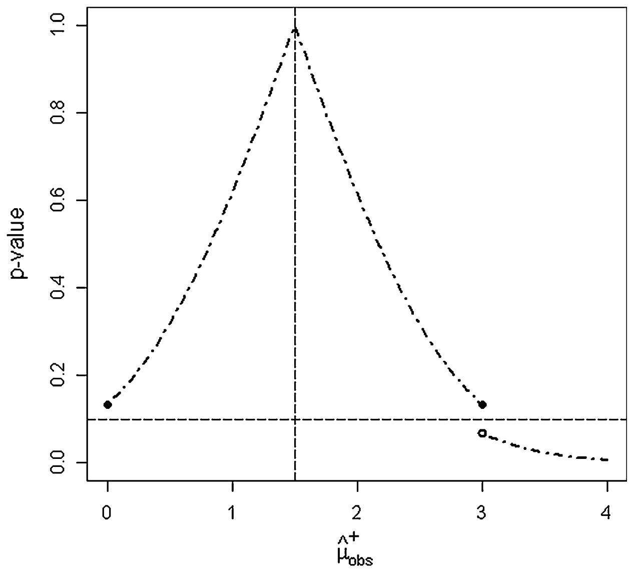

Fig. 2 The p value of a Wald test as a function of observed MLE $\hat{\mu}^+_{obs}$ for $\mu_0 = 1.5$ and $\sigma_0 = 1$ in a bounded normal mean problem (see Eq. 6). The *horizontal line* marks $p = 0.1$

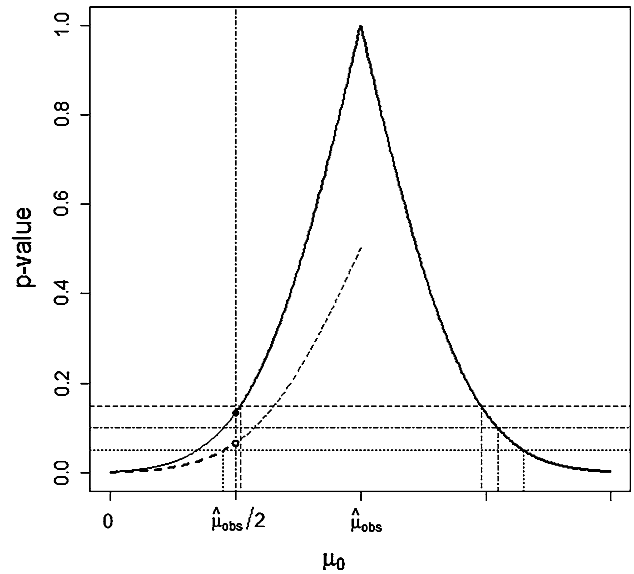

Fig. 3 Plot of p value in a Wald test as a function of test value μ_0 given $\hat{\mu}_{obs} > 0$ in a bounded normal mean problem (see Eq. 6). The (*bold* and *regular*) *solid curve* shows the two sided p value when μ is not constrained in H_1. The (*bold* and *regular*) *dashed curve* gives the one sided p value. The discontinuous *bold* (*solid* and *dashed*) *curve* shows the p value of a Wald test as given by Eq. 6. *Three horizontal lines* denotes three different α-levels and their corresponding CIs are also marked. They correspond to the three different situations when calculating $\hat{\mu}_L$. One should note that these different situations can happen for a single α level if $\hat{\mu}_{obs}$ varies. In this plot $\hat{\mu}_{obs} = 3\sigma_0$

which one to use according to their relative locations. The above result is summarized in the following expression for the adjusted CI $(\hat{\mu}_L, \hat{\mu}_U)$:[1]

$$\hat{\mu}_U = \hat{\mu}_U^{\alpha/2} \quad \text{and}$$
$$\hat{\mu}_L = \begin{cases} \hat{\mu}_L^{\alpha/2} & \text{if} \quad \hat{\mu}^+_{obs}/2 \le \hat{\mu}_L^{\alpha/2} \\ \hat{\mu}^+_{obs}/2 & \text{if} \quad \hat{\mu}_L^{\alpha/2} < \hat{\mu}^+_{obs}/2 \le \hat{\mu}_L^{\alpha} \\ \max\{\hat{\mu}_L^{\alpha}, 0\} & \text{if} \quad \hat{\mu}_L^{\alpha} < \hat{\mu}^+_{obs}/2 \end{cases} \tag{7}$$

where $\hat{\mu}_L^{\alpha} = \hat{\mu}^+_{obs} - \sigma_0 z_\alpha$ and $\hat{\mu}_U^{\alpha} = \hat{\mu}^+_{obs} + \sigma_0 z_\alpha$. It should be noted that when the true value lies between $z_{\alpha/2}\sigma_0$ and $z_\alpha\sigma_0$, the Wald test adjusted CI is conservative, or has more confidence than its nominal level $1 - \alpha$, because the rejection rate of the corresponding Wald test is smaller than α as discussed above.

The general case

In light of the above analysis of the simple problem of a bounded normal mean, the general problem for a bounded parameter in a parametric model can be solved. Let $f(\boldsymbol{x}|\boldsymbol{\xi})$ be a statistical model with parameter vector $\boldsymbol{\xi} = (\theta, \boldsymbol{\zeta}')'$, where θ is the parameter of interest. Let $\hat{\boldsymbol{\xi}}$ and $\hat{\theta}$ be the non-constrained maximum likelihood estimates (MLE) of $\boldsymbol{\xi}$ and θ. When a lower bound of 0 is imposed on θ, let $\hat{\boldsymbol{\xi}}^+$ and $\hat{\theta}^+$ be the constrained MLEs. We assume the central limit theorem (CLT) $\sqrt{n}(\hat{\theta} - \theta_0) \rightarrow N(0, \sigma^2(\boldsymbol{\xi}_0))$ holds for both boundary and nonboundary true values $\boldsymbol{\xi}_0$'s.[2] Derivations in the last section lead to a CI of the same expression as in Eq. 7 with the μ's replaced by θ's and σ_0 replaced by $\sigma(\hat{\boldsymbol{\xi}}^+_{obs})$ as its estimate. However, the CI in the last section was derived under the assumption that $\hat{\theta}$ follows a normal distribution with constant variance σ_0^2. In practice, normality only holds approximately through CLT with its validity depending on the parametrization, and the variance of $\hat{\theta}$ usually depends on the true value $\boldsymbol{\xi}_0$. To make the adjusted CI parametrization-invariant, the four critical quantities that appear in Eq. 7 must be based on the likelihood function. These quantities are $\hat{\theta}_U^{\alpha/2}, \hat{\theta}_L^{\alpha/2}, \hat{\theta}_L^{\alpha}$ and $\hat{\theta}^+_{obs}/2$.

When $\hat{\theta}^+_{obs} > 0$, or $\hat{\boldsymbol{\xi}}^+_{obs} = \hat{\boldsymbol{\xi}}_{obs}$ is not on the boundary, $\hat{\theta}_U^{\alpha/2}, \hat{\theta}_L^{\alpha/2}$ and $\hat{\theta}_L^{\alpha}$ can be calculated using a likelihood-based approach as explained by Neale and Miller (1997). For example, $\hat{\theta}_U^{\alpha/2}$ is the upper limit of a regular $100(1 - \alpha)$ % CI and it must satisfy the condition that the LRT statistic is

[1] All upper and lower limits of CIs in this article are non-inclusive except for two cases: (1) a zero lower limit in both adjusted CIs and (2) a lower limit in Wald test adjusted CI that takes the value of the middle point between the observed and the boundary values.

[2] This assumption is satisfied for most models as long as θ_0 is not an natural boundary and no nuisance parameter is on or close to their boundaries. See "Discussion" for the issues of a natural boundary and boundary nuisance parameters.

exactly $\chi^2_{1,\alpha} = z^2_{\alpha/2}$ when the test (1) is tested using the LRT, or

$$\min_{\theta=\hat{\theta}_U^{\alpha/2}} F(\xi) = F(\hat{\xi}^+_{\text{obs}}) + z^2_{\alpha/2} \tag{8}$$

where F is the negative twice log-likelihood function for the given sample. Unfortunately in Neale and Miller (1997), the limits of CI were computed using a suboptimal algorithm, which minimizes Eq. 3. Minimizing Eq. 3 is not equivalent to solving Eq. 8 and a bias is present in $\min_{\theta=\hat{\theta}_U^{\alpha/2}} F(\xi)$. Exact solution of $\hat{\theta}_U^{\alpha/2}$ can be found by maximizing θ with respect to the whole parameter vector ξ subject to the nonlinear constraint.[3]

$$F(\xi) = F(\hat{\xi}^+_{\text{obs}}) + z^2_{\alpha/2}, \tag{9}$$

which can be easily implemented in both the classic- and OpenMx (Neale 2004; Boker et al. 2011) for covariance structure models (e.g. the ACDE models) and in R for a general problem of a boundary parameter. Similarly, $\hat{\theta}_L^{\alpha/2}$ can be found by minimizing θ with respect to ξ subject to the constraint above. The third quantity $\hat{\theta}_L^{\alpha}$, or the lower limit of a regular $100(1 - 2\alpha)$ % CI can also be obtained.

The fourth quantity present in Eq. 7 is $\hat{\theta}^+_{\text{obs}}/2$, the middle point between the observed estimate $\hat{\theta}^+_{\text{obs}}$ and the boundary 0. In the case of a bounded normal mean, this middle point plays a significant role because when $\theta_0 = \hat{\theta}^+_{\text{obs}}/2, \Pr(\hat{\theta} \geq \hat{\theta}^+_{\text{obs}}) = \Pr(\hat{\theta} \leq 0)$ due to the symmetry of the normal sampling distribution of $\hat{\theta}$. For a general parametric model, it need not be a half of $\hat{\theta}^+_{\text{obs}}$ because the sampling distribution of $\hat{\theta}$ may not be symmetric. We replace it by $\hat{\theta}_m$ defined below through the likelihood function. This desired midpoint $\hat{\theta}_m$ should satisfy

$$F(\hat{\xi}_m) - F(\hat{\xi}_{\text{obs}}) = \left\{F(\hat{\xi}_b) - F(\hat{\xi}_{\text{obs}})\right\}/4 \tag{10}$$

where $\hat{\xi}_b$ is the MLE of ξ restricted to be on the boundary $\theta = 0$. The above equation specifies that the "distance" between the middle point $\hat{\xi}_m$ and the observed value $\hat{\xi}_{\text{obs}}$ is just half of the distance from the observed value to the boundary. Remember that difference in F measures the squared distance. $\hat{\xi}_m$ can be found by minimizing θ subject to the constraint $F(\xi) = F(\hat{\xi}_{\text{obs}}) + (F(\hat{\xi}_b) - F(\hat{\xi}_{\text{obs}}))/4$.

When $\hat{\theta}^+_{\text{obs}} = 0$, the lower limit of the CI must be $\hat{\theta}_L = 0$. For the bounded normal mean problem, the upper limit is given by $\hat{\theta}_U = \sigma_0 z^2_{\alpha/2}$. This is the value, if taken as the true value, that would yield a sampling distribution whose point mass on 0 is exactly $q_0 = \alpha/2$. In terms of the likelihood function, the LRT value should be $z^2_{\alpha/2}$ for a data set whose non-constrained MLE $\hat{\xi}$ happens to be on the boundary, or the distance between $\hat{\theta}_U$ and the boundary is $z_{\alpha/2}$. If the non-constrained MLE of the current observed data $\hat{\xi}_{\text{obs}}$ with $\hat{\theta}_{\text{obs}} < 0$ is available, the squared distance between $\hat{\theta}_{\text{obs}}$ and the boundary $\hat{\theta}^+_{\text{obs}} = 0$ can be measured by $F(\hat{\xi}^+_{\text{obs}}) - F(\hat{\xi}_{\text{obs}})$. Because the desired distance from $\hat{\theta}_U$ to the boundary is $z_{\alpha/2}$, the desired distance between $\hat{\theta}_{\text{obs}}$ and $\hat{\theta}_U$ should be $\sqrt{F(\hat{\xi}^+_{\text{obs}}) - F(\hat{\xi}_{\text{obs}})} + z_{\alpha/2}$. Thus, $\hat{\theta}_U$ can be obtained by maximizing θ with respect to ξ subject to the constraint

$$F(\xi) - F(\hat{\xi}_{\text{obs}}) = \left\{\sqrt{F(\hat{\xi}^+_{\text{obs}}) - F(\hat{\xi}_{\text{obs}})} + z_{\alpha/2}\right\}^2 \tag{11}$$

The general procedure of obtaining an adjusted CI based on Wald test is illustrated in Fig. 4.

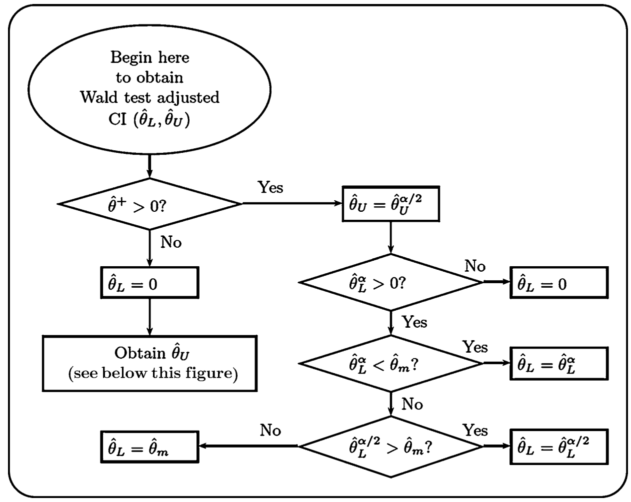

Fig. 4 Flowchart of the procedure to obtain a Wald test adjusted CI

Correction based on the likelihood ratio test

Testing a normal mean close to a bound

It should be noted that the CI proposed above is based on a Wald test corrected for the boundary condition. It is in general not equivalent to inverting a (corrected) LRT unless the test value is on the boundary. This non-equivalence can be observed by revisiting the case of a bounded normal mean. The LRT for the hypothesis testing problem in (1) would use the statistic

[3] The "=" sign in the constraint can be replaced by "$\leq$" if the constrained optimization algorithm requires that the starting value satisfy the constraint. This is also the case for constraints in Eqs. 10 and 11

$$T = \frac{(\hat{\mu} - \mu_0)^2 - (\hat{\mu} - \hat{\mu}^+)^2}{\sigma_0^2} = \begin{cases} (\hat{\mu} - \mu_0)^2/\sigma_0^2 & \text{if} \quad \hat{\mu} = \hat{\mu}^+ \geq 0 \\ \mu_0(\mu_0 - 2\hat{\mu})/\sigma_0^2 & \text{if} \quad \hat{\mu} < \hat{\mu}^+ = 0 \end{cases} \tag{12}$$

which is plotted in Fig. 5. When $\mu_0 = 0$, or the true value is on the boundary, $T = W$ and the two tests are equivalent. When $\mu_0 > 0$, this test statistic coincides with the Wald test statistic in Eq. 2 only if $\hat{\mu} = \hat{\mu}^+ \geq 0$. If the unrestricted MLE $\hat{\mu}$ violates the boundary condition, we have $\hat{\mu} < \hat{\mu}^+ = 0$ and T increases linearly in $|\hat{\mu}|$. As a result, as long as $\mu_0 > 0$, the LRT may reject the null hypothesis if the unrestricted MLE $\hat{\mu}$ is negative with a large absolute value, resulting in a rejection region for both positive and negative $\hat{\mu}$'s, as illustrated in Fig. 5. This is in contrast to the Wald test, which uses only the information of $\hat{\mu}^+$ and only rejects positive $\hat{\mu}$'s when $\mu_0 < z_{\alpha/2}\sigma_0$. Figure 5 shows when the lower rejection region is beyond the boundary, it is farther away from the true value than the upper rejection region, resulting in a smaller rejection probability on the left tail than on the right tail.

For $\mu_0 > 0$, given an observed LRT statistic t, from Eq. 12, the p value can be calculated as

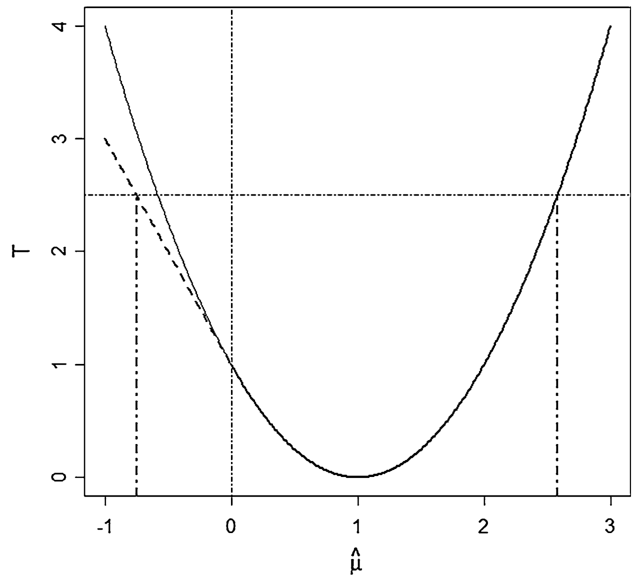

Fig. 5 Plot of LRT statistic T as a function of $\hat{\mu}$ for a normal distribution with $\sigma_0 = \mu_0 = 1$. If the alternative model has no boundary constraint on μ, T follows the (*bold* and *regular*) *solid curve*. If the alternative model imposes $\mu \geq 0$, T follows the *bold* (*dashed* and *solid*) *curve*. The rejection region $\{T > 2.5\}$ is also marked for the second case

$$p = \Pr(T \geq t) = \begin{cases} \Pr\left(\hat{\mu} \geq \mu_0 + \sigma_0 t^{\frac{1}{2}}\right) + \Pr\left(\hat{\mu} \leq \mu_0 - \sigma_0 t^{\frac{1}{2}}\right) & \text{if} \quad \mu_0 - \sigma_0 t^{\frac{1}{2}} \geq 0 \\ \Pr\left(\hat{\mu} \geq \mu_0 + \sigma_0 t^{\frac{1}{2}}\right) + \Pr\left(\hat{\mu} \leq (\mu_0^2 - \sigma_0^2 t)/2\mu_0\right) & \text{if} \quad \mu_0 - \sigma_0 t^{\frac{1}{2}} < 0 \end{cases}$$
$$= \begin{cases} 2\Phi(-t^{\frac{1}{2}}) & \text{if} \quad \mu_0 - \sigma_0 t^{\frac{1}{2}} \geq 0 \\ \Phi(-t^{\frac{1}{2}}) + \Phi\left(-\frac{1}{2}\left(\frac{\mu_0}{\sigma_0} + t\frac{\sigma_0}{\mu_0}\right)\right) & \text{if} \quad \mu_0 - \sigma_0 t^{\frac{1}{2}} < 0 \end{cases} \tag{13}$$

In terms of the observed MLE $\hat{\mu}_{obs}$, when $\hat{\mu}_{obs} = \hat{\mu}^+_{obs} \geq 0$, we have $t = (\hat{\mu}_{obs} - \mu_0)^2/\sigma_0^2$ and Eq. 13 becomes

$$p = \begin{cases} 2\Phi(-|\hat{\mu}_{obs} - \mu_0|/\sigma_0) & \text{if} \quad |\mu_0 - \hat{\mu}_{obs}| \leq \mu_0 \\ \Phi(-|\hat{\mu}_{obs} - \mu_0|/\sigma_0) + \Phi\left(-\mu_0/2\sigma_0 - (\hat{\mu}_{obs} - \mu_0)^2/2\sigma_0\mu_0\right) & \text{if} \quad |\mu_0 - \hat{\mu}_{obs}| > \mu_0 \end{cases} \tag{14}$$

Note in this case $\sigma_0 t^{\frac{1}{2}} = \sqrt{\mu_0(\mu_0 - 2\hat{\mu}_{obs})} > \mu_0$, so only the second line in Eq. 13 is relevant to produce Eq. 15.

When $\hat{\mu}_{obs} < 0 = \hat{\mu}^+_{obs}$, we have $t = \mu_0(\mu_0 - 2\hat{\mu}_{obs})/\sigma_0^2$ and Eq. 13 becomes

$$p = \Phi\left(-\sqrt{\mu_0(\mu_0 - 2\hat{\mu}_{obs})}/\sigma_0\right) + \Phi(-(\mu_0 - \hat{\mu}_{obs})/\sigma_0) \tag{15}$$

The p value is plotted in Fig. 6 as a function of the unrestricted MLE $\hat{\mu}_{obs}$ for $\mu_0 = 1.5$ and $\sigma_0 = 1$. Because this curve does not have jumps, a test with exact type I error rate of α exists for all α. In addition, the rejection region lies on both sides for all α.

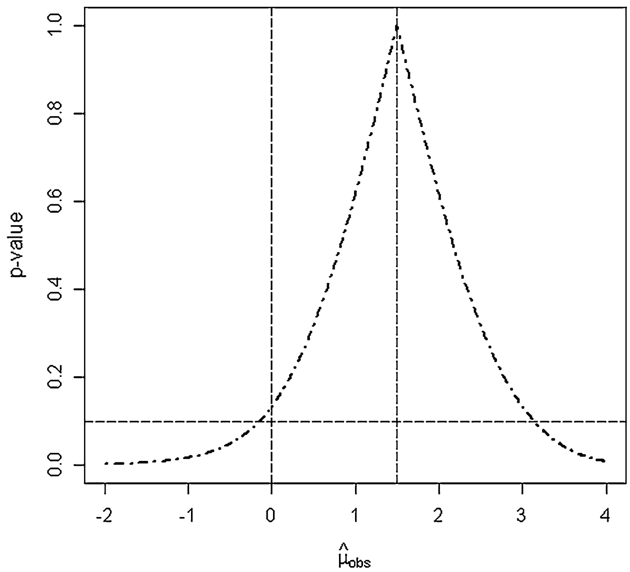

Fig. 6 The p value of a LRT as a function of observed MLE $\hat{\mu}_{obs}$ for $\mu_0 = 1.5$ and $\sigma_0 = 1$ in a bounded normal mean problem (see Eqs. 14 and 15). The *horizontal line* marks $p = 0.1$

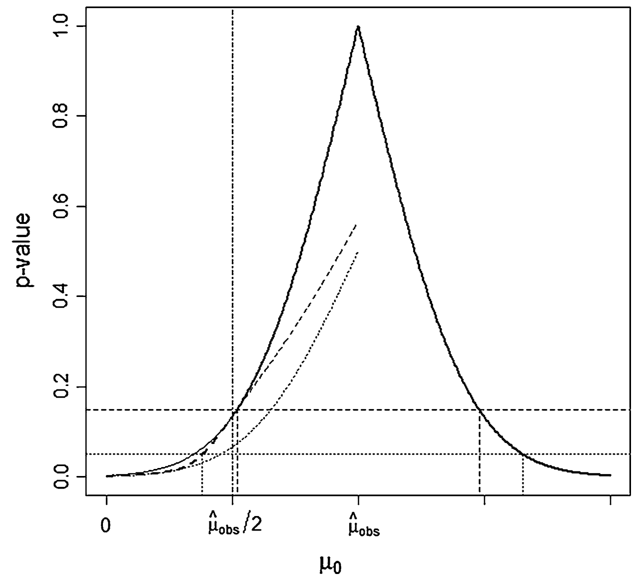

Fig. 7 Plot of p value in a LRT as a function of test value μ_0 given $\hat{\mu}_{obs} > 0$ in a bounded normal mean problem (see Eq. 14). The (*bold* and *regular*) *solid curve* shows the two sided p value when μ is not constrained in H_1. The *bold* (*solid* and *dashed*) curve shows the p value of a LRT as given by Eq. 13. The one sided p value is also included as a *dotted curve*. Note the *dotted curve* and the *bold p* value *curve* only intersect at $\mu_0 = 0$ but are very close to each other over a range of positive μ_0. *Two horizontal lines* denotes two different α-levels and their corresponding CIs are also marked. They correspond to the two different situations when calculating $\hat{\mu}_L$. One should note that these different situations can happen for a single α level but for different $\hat{\mu}_{obs}$'s. In this plot $\hat{\mu}_{obs} = 3\sigma_0$

A confidence interval for a bounded normal mean

To obtain the CI, we only need to find the true value μ_0 for which the p value is greater than α. If $\hat{\mu}_{obs} = \hat{\mu}^+_{obs} > 0$ the p value in Eq. 14 is plotted in Fig. 7 as a function of the test value μ_0. We can see the upper limit of the CI, similar to that of the Wald test adjusted CI, is always given by $\hat{\mu}_U^{\alpha/2} = \hat{\mu}_{obs} + \sigma_0 z_{\alpha/2}$, the upper limit of a regular $100(1-\alpha)$ % CI. The lower limit of the CI is more complicated, depending on whether the horizontal line $p = \alpha$ crosses the (bold) dashed part of the curve or the (bold) solid part of the curve. The two parts of the p value curves join each other at $\hat{\mu}^+/2$. If the horizontal line crosses the (bold) solid part of the p value curve, the resulting lower limit is given by $\hat{\mu}_L^{\alpha/2} = \hat{\mu}_{obs} - \sigma_0 z_{\alpha/2}$, the lower limit of a regular $100(1-\alpha)$ % CI, and it satisfies $\hat{\mu}_L^{\alpha/2} > \hat{\mu}_{obs}/2$. If the horizontal line crosses the (bold) dashed part of the p value curve, the resulting lower limit $\hat{\mu}^*_L$ satisfies the equations

$$\begin{aligned}&\Phi(-|\hat{\mu}_{obs} - \hat{\mu}^*_L|/\sigma_0)\\&\quad+\Phi\Big(-\hat{\mu}^*_L/2\sigma_0 - (\hat{\mu}_{obs} - \hat{\mu}^*_L)^2/2\sigma_0\hat{\mu}^*_L\Big)\\&\quad=\alpha,\end{aligned} \tag{16}$$

and $\hat{\mu}^*_L < \hat{\mu}_{obs}/2$. The above result is summarized below as

$$\hat{\mu}_U = \hat{\mu}_U^{\alpha/2} \quad \text{and} \quad \hat{\mu}_L = \begin{cases} \hat{\mu}_L^{\alpha/2} & \text{if} \quad \hat{\mu}_L^{\alpha/2} \geq \hat{\mu}_{obs}/2 \\ \hat{\mu}^*_L & \text{if} \quad \hat{\mu}_L^{\alpha/2} < \hat{\mu}_{obs}/2 \end{cases}$$

where $(\hat{\mu}_L^{\alpha/2}, \hat{\mu}_U^{\alpha/2})$ is a regular $100(1-\alpha)$ % CI, and $\hat{\mu}^*_L$ satisfies Eq. 16. If $\hat{\mu}_{obs} < \hat{\mu}^+_{obs} = 0$, immediately we have $\hat{\mu}_L = 0$ and $\hat{\mu}_U$ is given by the value $\hat{\mu}^*_U$ that satisfies

$$\Phi\left(-\sqrt{\hat{\mu}^*_U(\hat{\mu}^*_U - 2\hat{\mu}_{obs})}/\sigma_0\right) + \Phi\big(-(\hat{\mu}^*_U - \hat{\mu}_{obs})/\sigma_0\big) = \alpha \tag{17}$$

Comparing the LRT adjusted and the Wald test adjusted CIs, we can see that when $\hat{\theta}_{obs} = \hat{\theta}^+_{obs} \geq 0$, both methods produce the regular CI $(\hat{\theta}_L^{\alpha/2}, \hat{\theta}_U^{\alpha/2})$ if the lower limit is larger than the middle point between $\hat{\theta}_{obs}$ and the boundary 0. If this is not true, the two methods differ in how to adjust the lower limit. As can be observed from Fig. 7, the Wald test adjusted CI gives a higher lower limit than the LRT adjusted CI, though these two lower limits may be numerically the same over a region above the boundary. When the observed data violate the boundary condition, the lower limit of the CI must be 0, and the two methods differ on how the upper limit be calculated. The Wald test adjusted CI has the same upper limit irrespective of the unrestricted MLE $\hat{\theta}_{obs}$, while LRT adjusted CI considers the information from $\hat{\theta}_{obs} < 0$ and has a smaller upper limit.

The general case

To extend the above arguments to the general case of a bounded parameter and to achieve invariance under reparametrization, quantities in the CI must be likelihood-based. In the end of Section "Correction based on a Wald test" we provided the algorithms for finding $\hat{\theta}_L^{\alpha/2}, \hat{\theta}_U^{\alpha/2}$ and the middle point $\hat{\theta}_m$. Below we give the procedures to calculate likelihood based $\hat{\theta}_L^*$ and $\hat{\theta}_U^*$, the counterparts of $\hat{\mu}_L^*$ and $\hat{\mu}_U^*$ in Eqs 16 and 17.

When $\hat{\theta}_{obs} = \hat{\theta}_{obs}^+ > 0, \hat{\theta}_L^*$ needs to be calculated. Its counterpart in the normal model, $\hat{\mu}_L^*$, satisfies Eq. (16), where $|\hat{\mu}_{obs} - \hat{\mu}_L^*|/\sigma_0$ gives the distance between $\hat{\mu}_{obs}$ and $\hat{\mu}_L^*$, and $\hat{\mu}_L^*/\sigma_0$ is the distance from $\hat{\mu}_L^*$ to the boundary 0. For a general parametric model with bounded parameter $\theta, |\hat{\mu}_{obs} - \hat{\mu}_L^*|/\sigma_0$ can be replaced by $r^* = \sqrt{\min_{\theta=\theta_L^*} F(\xi) - F(\hat{\xi}_{obs})}$, an approximation of the distance from $\hat{\theta}_L^*$ to $\hat{\theta}_{obs}$, and $\hat{\mu}_L^*/\sigma_0$ by $r_b - r^*$, where $r_b = \sqrt{F(\hat{\xi}_b) - F(\hat{\xi})}$ is an approximation of the distance from $\hat{\theta}$ to the boundary. Note that $\hat{\xi}_b$ is the MLE constrained to be on the boundary (i.e., $\hat{\theta}_b = 0$). Thus, $\hat{\theta}_L^*$ satisfies

$$\Phi(-r^*) + \Phi\left(-\frac{1}{2}\left(r_b - r^* + \frac{r^{*2}}{r_b - r^*}\right)\right) = \alpha \qquad (18)$$

where r^* as explained above is a function of $\hat{\theta}_L^*$. Following from this equation, one may obtain $\hat{\theta}_L^*$ by minimizing θ with respect to ξ subject to

$$\Phi(-r) + \Phi\left(-\frac{1}{2}\left(r_b - r + \frac{r^2}{r_b - r}\right)\right) \geq \alpha \qquad (19)$$

where $r^2 = F(\xi) - F(\hat{\xi})$. Note r_b can be obtained before the optimization.

When $\hat{\theta}_{obs} < 0 = \hat{\theta}_{obs}^+$, the upper limit of the CI is $\hat{\theta}_U^*$. For $\hat{\theta}_U^*$, its counterpart $\hat{\mu}_U^*$ in the normal model satisfies Eq. 17, where $\sqrt{\hat{\mu}_U^*(\hat{\mu}_U^* - 2\hat{\mu}_{obs})}/\sigma_0$ is the square root of the LRT statistic t if $\mu_0 = \mu_U^*$ in the test (see Eq. 12), and $(\hat{\mu}_U^* - \hat{\mu}_{obs})/\sigma_0$ is the distance between $\hat{\mu}_U^*$ and $\hat{\mu}_{obs}$. As a result, $\hat{\theta}_U^*$ should satisfy

$$\Phi(-r_1^*) + \Phi(-r_2^*) = \alpha \qquad (20)$$

where $r_1^* = \sqrt{\min_{\theta=\hat{\theta}_U^*} F(\xi) - F(\hat{\xi}_{obs}^+)}$ and $r_2^* = \sqrt{\min_{\theta=\hat{\theta}_U^*} F(\xi) - F(\hat{\xi}_{obs})}$ are both functions of $\hat{\theta}_U^*.\hat{\theta}_U^*$ can be obtained by maximizing θ with respect to ξ under the constraint

$$\Phi(-r_1) + \Phi(-r_2) \geq \alpha, \qquad (21)$$

where $r_1^2 = F(\xi) - F(\hat{\xi}_{obs})$ and $r_2^2 = F(\xi) - F(\hat{\xi}_{obs}^+)$.

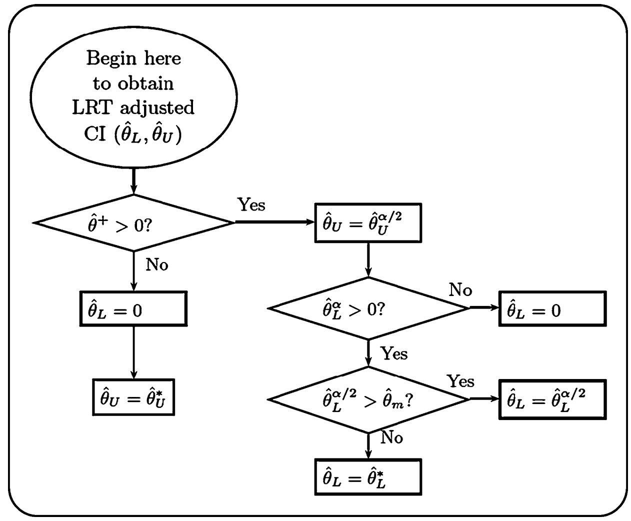

Fig. 8 Flowchart of the procedure to obtain a LRT adjusted CI

Both the two optimization problems above for obtaining $\hat{\theta}_U^*$ and $\hat{\theta}_L^*$ can be solved numerically. However, they are less stable than those for obtaining the Wald test based CI. One reason is that the nonlinear constraints in these problems are expressed in terms of the p values, which, when small, become very insensitive to parameters. The general procedure of obtaining an adjusted CI based on LRT is illustrated in Fig. 8.

Simulation study

Study I

We first present a simulation study for the simple case of a bounded normal mean $\mu \geq 0$. In this study, we set $\sigma_0 = 1$ and vary the true value μ_0, which takes four values 0, 1, 1.9 and 2.5. 10,000 replications are used for each true value to simulate the missing probabilities[4] of the two adjusted CIs and an unadjusted CI. The unadjusted CI is an naively computed CI whose upper and lower limits are determined to give an increase of $z_{\alpha/2}^2$ in negative twice log-likelihood. The result is summarized in Table 1.

We can see that the unadjusted CI has correct missing probabilities from below the lower limit in all cases because the upper tail of the sampling distribution is not affected by the boundary. However, the missing probability from above its upper limit is smaller than the nominal value. Especially, when the true value is on the boundary, because the upper limit of an unadjusted CI cannot be below the boundary, the missing probability from above

[4] i.e. the probability that the true value is not covered by the CI.

Table 1 Missing probabilities of the unadjusted and two adjusted CIs in simulation study I

		$\mu = 0$	$\mu = 1$	$\mu = 1.9$	$\mu = 2.5$
Unadjusted	Lower	2.69	2.18	2.36	2.36
	Upper	0	0.66	2.31	2.42
	Total	2.69	2.84	4.67	4.78
Wald test adjusted	Lower	5.07	4.74	2.71	2.36
	Upper	0	0	0	2.42
	Total	5.07	4.74	2.71	4.78
LRT adjusted	Lower	5.07	3.24	2.36	2.36
	Upper	0	1.48	2.31	2.42
	Total	5.07	4.72	4.67	4.78

The lower and upper missing probabilities are probabilities that the true value falls below the lower limit or above the upper limit of the CI. Entries are percentages based on 10,000 replications. Monte Carlo error is about ±0.43 % for missing probabilities that are supposedly 5 % and ± 0.31 % for those supposedly 2.5 %

the upper limit is 0, making the total missing rate only half of its nominal value.

The LRT adjusted CI have the correct overall missing probabilities. In particular, when $\mu_0 = 0$, the missing probability come from below its lower limit, which corresponds to samples on the upper tail of the sampling distribution. This is not surprising because in this case the test of $\mu = \mu_0$ is one-sided and only samples on the upper tail of the sampling distribution should be rejected. As μ_0 moves away from the boundary, the CI becomes gradually balanced. When $\mu_0 > z_{\alpha/2}$, the rejection region of the LRT test coincide with the traditional two sided z-test and the CI has balanced missing probabilities.

For the Wald test adjusted CI, the missing probability from above its upper limit stays at zero for $\mu_0 < z_{\alpha/2}$ and becomes half the nominal missing rate for $\mu_0 > z_{\alpha/2}$. This is consistent with the fact that a Wald test is one-sided in the former case and becomes two-sided in the latter case as explained by Fig. 2. The missing probability from below its lower limit is 5 % for $\mu_0 < z_\alpha$ and is 2.5 % for $\mu_0 > z_{\alpha/2}$, as expected for a one-sided test and a two-sided test, respectively. When $\mu_0 = 1.9$, which is between z_α and $z_{\alpha/2}$, the CI has smaller missing probability than its nominal level.

Study II

The second simulation study is conducted with a univariate ACE model, in which the variance components a^2, c^2 and e^2 have lower bounds of 0. In this study, we are interested in the estimation of the standardized component of common environment c^2. Four different conditions are chosen with $c^2 = 0$, 0.15, 0.2 and 0.25. In all four conditions, e^2 is set at $e^2 = 0.1$ and a^2 is chosen such that the total variance is 1. The sample sizes are $n_{\mathrm{MZ}} = n_{\mathrm{DZ}} = 150$ for the MZ and DZ twins. 10,000 replications are used and all replications are convergent for all procedures.[5] The missing probabilities from below the lower limits and above the upper limits of different types of CIs are shown in Table 2.

Table 2 Missing probabilities of CIs in simulation study II

		$c^2 = 0$	$c^2 = 0.15$	$c^2 = 0.2$	$c^2 = 0.25$
Unconstrained	Lower	2.62	2.67	2.72	2.86
	Upper	2.55	2.48	2.21	2.28
	Total	5.17	5.15	4.93	5.14
Unadjusted	Lower	2.62	2.67	2.72	2.86
	Upper	0	1.76	2.15	2.28
	Total	2.62	4.43	4.87	5.14
Wald test adjusted	Lower	5.16	5.21	5.03	2.86
	Upper	0	0	0	2.28
	Total	5.16	5.21	5.03	5.14
LRT adjusted	Lower	5.16	3.25	2.80	2.86
	Upper	0	2.16	2.18	2.28
	Total	5.16	5.41	4.98	5.14

The lower and upper missing probabilities are probabilities that the true value falls below the lower limit or above the upper limit of the CI. Entries are percentages based on 10,000 replications. Monte Carlo error is about ±0.43 % for missing probabilities that are supposedly 5 % and ±0.31 % for those supposedly 2.5 %

The unconstrained CIs are CIs calculated for an ACE model with c^2 allowed to become negative. The unadjusted CIs are CIs obtained from the ACE model in the traditional way without adjustment. They differ from the unconstrained CIs in two ways: first, their lower limit cannot be negative; second, when a boundary MLE is obtained, their upper limit produces an increase of $z^2_{\alpha/2}$ in negative twice log-likelihood computed at the boundary MLE instead of at a non-constrained MLE and is therefore greater than that of the unconstrained CIs.

The unconstrained CI has correct total missing probabilities under all conditions. This is not surprising as no boundary issue arise in this case. This CI is included to show that the χ^2 approximation works well for the given sample sizes when no boundary is present. The missing probabilities for the remaining three types of CIs in the table exhibit a similar pattern to simulation study I, except that a missing probability smaller than the nominal value is not observed for the Wald test adjusted CI in this study.

Table 3 compares the sizes of the adjusted CIs. When boundary MLE is obtained, the lower limits of both CIs must be zero and the LRT based CI has smaller upper limit and is therefore shorter. When the MLE is not on the

[5] Non-convergent replications were handled by manually adjusting the starting values until convergence was reached.

Table 3 Comparison of adjusted CIs in simulation study II

		$c^2 = 0$	$c^2 = 0.15$	$c^2 = 0.2$	$c^2 = 0.25$
$\hat{c} = 0$	$U_W > U_{LRT}$	49.77	9.95	3.79	1.28
$\hat{c} > 0$	$0 = L_{LRT} = L_W$	45.07	55.20	42.53	27.23
	$0 < L_{LRT} = L_W$	2.63	11.79	14.92	15.98
	$0 < L_{LRT} < L_W$	2.53	23.06	38.76	55.61
	Subtotal	50.23	90.05	96.21	98.82
Overall	$CI_W > CI_{LRT}$	49.77	9.95	3.79	1.28
Length	$CI_W = CI_{LRT}$	47.70	66.99	57.45	43.11
Comparison	$CI_W < CI_{LRT}$	2.53	23.06	38.76	55.61

Entries are percentages based on 10,000 replications

Table 4 Comparison of expected lengths of the two adjusted CIs in simulation study II

		$c^2 = 0$	$c^2 = 0.15$	$c^2 = 0.2$	$c^2 = 0.25$
$\hat{c} = 0$	CI_W	0.220	0.221	0.222	0.224
	CI_{LRT}	0.152	0.178	0.185	0.188
$\hat{c} > 0$	CI_W	0.296	0.331	0.341	0.344
	CI_{LRT}	0.296	0.333	0.346	0.351
Overall	CI_W	0.258	0.320	0.336	0.342
	CI_{LRT}	0.225	0.318	0.339	0.349

Entries are based on 10,000 replications

boundary, both CIs have the same upper limit, which is also the upper limit of an unconstrained CI. In this case, the lower limit of the LRT adjusted CI is smaller than or equal to that of the Wald test adjusted CI. The two types of CIs have the same lower limits if both limits are close or equal to 0 or are above the middle point $\hat{c}_m^2$. The overall frequencies of relative sizes depends on the frequency of occurrence of boundary MLEs. However, the expected lengths of the two kinds of CIs are very close to each other, as suggested by Table 4.

Figure 9 plots the calculated middle point $\hat{c}_m^2$ against the MLE $\hat{c}^{2+}$. A reference line with slope 0.5 is also plotted. It can be observed that as the estimate is away from the boundary, the calculated middle point deviates from and is greater than half of the standardized estimate.

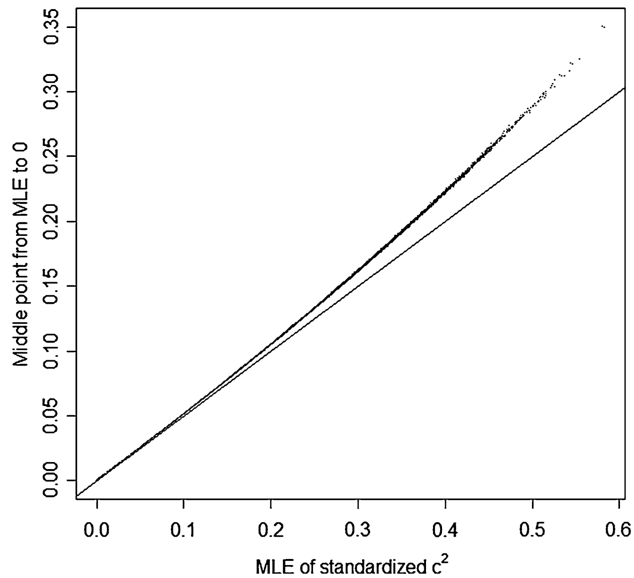

Fig. 9 Plot of middle point $\hat{c}_m^2$ against $\hat{c}^2/2$ in simulation study II

Table 5 Contingency tables of twin pair diagnosis of lifetime major depressive illness

Twin 2	MZ		DZ	
	Twin 1		Twin 1	
	Normal	Depressed	Normal	Depressed
Normal	329	83	201	94
Depressed	95	83	82	63

Examples

Depression data

For illustration, we use the data published by (1992, Chapter 6) as reproduced in Table 5. These data come from a sample of adult female twin pairs drawn from birth records in the Commonwealth of Virginia, USA. An ADE model is fitted with different thresholds for the two twin types and the parameter estimates and unadjusted 95 % CIs are summarized in Table 6. As the total variance is constrained to unity for the purpose of identification, all variance components are standardized components and are bounded between zero and one. For the A and D components the lower bound of zero is attainable while for the E component the upper bound of one is attainable.[6] As the MLE is not on any boundary, we only need to consider the

[6] See "Discussion" section for the distinction between natural and attainable boundaries.

Table 6 Parameter estimates and CIs in the ADE model fitted to the depression data

Parameter	MLE	Unadjusted and Wald/LRT adjusted CI	
a^2	0.303	0.000	0.530
d^2	0.131	0.000	0.542
e^2	0.566	0.452	0.689
t_{MZ}	0.549	0.464	0.634
t_{DZ}	0.404	0.314	0.494
$t_{MZ} - t_{DZ}$	0.145	0.021	0.269

The negative twice ML is 2509.632. The two adjusted CIs both coincide with the unadjusted CI

Table 7 Covariance matrices of MZ and DZ male twin pairs of age under 30 for body mass index: 1981 Australian Survey

MZ twins ($n_{MZ} = 251$)		DZ twins ($n_{MZ} = 184$)	
0.597	0.448	0.719	0.245
0.448	0.569	0.245	0.818

The raw data was transformed as $7\ln(BMI) - 21$

lower limits of the CIs of A and D components and the upper limit of the CI of the E component. Fortunately the 90 % CIs of the A and D components both have lower limit of 0, so their adjusted 95 % CIs must have lower limits zero. For the E parameter, the 90 % CI has an upper limit of 0.672. When $e^2 = 1$, the negative twice ML is 2556.185, about 46.553 larger than that of the ADE model. To find the middle point between $\hat{e}^2$ and its boundary, we set the negative twice ML at $(46.553/4 + 2509.632 = 2521.270)$ and maximize the e^2 parameter, yielding a value of 0.780. This middle point is closer to the boundary than both the upper limits of the 95 % and 90 % CIs, so the CI need not be adjusted as well. It should be noted that because more than one parameter in this example is close to its boundary, all CIs displayed in Table 6 may be conservative.

Body mass index data

Another example includes body mass index (BMI) data from a survey of volunteers from the Australian NH & MRC twin register (Neale and Cardon 1992, Chapter 6). Table 7 contains the covariance matrices of log-transformed BMI of male MZ and DZ twins with age less than 30. An ADE model is fitted to the data and the estimates are summarized in Table 8. For the original parameters, both the A and D components have an attainable lower bound of zero and their estimates are positive, so the lower limits of their CIs may need to be adjusted. For the A component, the 90 % CI also has a lower limit of zero, so the adjusted CI must have a lower limit of zero. For the D component, the 90 % CI has a lower limit of 0.038 and the

Table 8 Parameter estimates and CIs in the ADE model fitted to the BMI data

Parameter	MLE	Unadjusted CI		Wald/LRT adjusted CIs	
Original					
a^2	0.248	0.000	0.571	0.000	0.571
d^2	0.295	0.000	0.592	0.038	0.592
e^2	0.137	0.116	0.165	–	–
Standardized					
a^2	0.365	0.000	0.805	0.000	0.805
d^2	0.433	0.000	0.824	0.055	0.824
e^2	0.202	0.165	0.248	0.165	0.248

The two adjusted CIs coincide to the accuracy presented

middle point between $\hat{d}^2$ and zero is 0.144, so the adjusted lower limit should be 0.038 if the Wald type adjustment is used. If the LRT type adjustment is used, the lower limit should be separately calculated and is also 0.038.

The standardized components are proportions of the A, D and E components in the total variance and are more interesting to researchers. The standardized a^2, d^2 parameters have an attainable lower bound of zero and the standardized e^2 have an attainable upper bound of one. Again, for the standardized A component, the 90 % CI has a zero lower limit and therefore the adjusted CI should also have lower limit zero. For the standardized C component, the lower limit of a 90 % CI is 0.055 and the middle point between the estimate and the boundary is 0.210, so the adjusted CI should have a lower limit of 0.055 if the Wald type adjustment is employed. If the LRT type adjustment is used, the lower limit is still 0.055. The standardized E component is estimated far away from its upper bound of one and its CI is unlikely affected. In fact, its 90 % CI has an upper limit of 0.240 and the middle point between its MLE 0.202 and its upper boundary $e^2 = 1$ is at 0.471, much larger than the upper limits of both the upper limits of the regular 90 and 95 % CIs.

Discussion

In this article we proposed two methods of CI adjustement when the true parameter is close to its boundary. These methods originate from Wald test and LRT of the a bounded normal mean parameter and are generalized and made parametrization-invariant. The proposed CIs do not touch the boundary if the (corrected) LRT rejects the boundary null hypothesis. Several points need to be considered before concluding the article.

First, Carey (2005) suggests the removal of the lower bounds on (scalar) variance components (and the more general inequality constraints of positive-semidefiniteness on a matrix variance component). This approach does remove the technical issue in hypothesis testing and CI, but on the other hand renders the model difficult to interpret because a variance cannot be negative. In particular, neglecting the lower bounds may yield a CI of a variance that includes negative values. This approach was motivated in a time when no adjusted LRT or CI was available. With the advance in statistical methodology (e.g. Dominicus et al. 2006), this approach may no longer be the best strategy for testing a variance component. The research presented in this article is to further provide CIs that give consistent results with the adjusted test.

Second, CI adjustment is only needed when the boundary of concern is not the natural boundary of the statistical model, which we now define. A parameter on or beyond its natural boundary leads to an invalid or degenerate distribution of the data. For example, a binomial parameter p beyond the boundary of (0,1) does not define a valid binomial distribution, and similarly a population correlation between observed variables cannot be beyond $(-1,1)$. However, a correlation between latent variables may still define a valid multivariate normal distribution even if it is beyond this range, as a large residual variance may guarantee the positive-definiteness of the covariance matrix of the observed variables. In this case, ± 1 are no longer natural boundaries of the parameter. In terms of an univariate ACE model, the lower bound of zero is the natural boundary for parameter e^2. Because the MLE of e^2 never violates this natural boundary, the regular LRT and the likelihood-based CI need not be modified in this case. In contrast, a^2, c^2 or d^2 can be estimated at zero, so these boundaries are attainable and their LRTs and CIs need to be corrected.

Third, both the two proposed adjusted CIs are unbalanced in the sense that the missing rates beyond the two limits of a CI are not the same. Although this seems a drawback of the current procedure, it is a necessary feature for adjusted CIs near a boundary. To see this, we note that when the true value θ_0 is on the boundary, the test in Eq. 1 is necessarily a one-sided test on the upper tail of the sampling distribution and θ_0 can only be rejected from below the CI. As a result, the missing rate from above the upper limit of the CI must be 0. On the other hand, if θ_0 is far from the boundary, the test is not affected by the boundary condition and the CI is balanced. To connect the above two extreme cases, the adjusted CI must become gradually unbalanced when the true value approaches the boundary, with the missing rate from above the upper limit decreasing from $\alpha/2$ to 0 and that from below the lower limit increasing from $\alpha/2$ to α.

Fourth, as we have discussed and illustrated with simulation, between the two proposed methods, the Wald-test adjusted CI is shorter when the MLE is not on the boundary and the LRT adjusted CI is shorter when the MLE is on the boundary. It seems appealing to choose the method of adjustment to produce the shorter CI. Unfortunately this is not a valid procedure as it would result in under-coverage of the CI. To see this, remember as a frequentist concept, coverage concerns not only the situation of the current observation but also observations that would have been observed. If under those hypothetical situations a different method would be chosen to produce a shorter CI, the coverage probability would be smaller than using one method consistently.

Last, the development of the two adjusted CIs assumes that the parameter of interest is only close to one boundary and is the only parameter that is close to a boundary. Because the sampling distributions of the Wald and LRT statistics are affected by boundary conditions of all parameters, the proposed CIs may not produce desired coverage if a parameter not being considered is close to its boundary or the parameter of interest is subject to additional boundary conditions.[7] However, in such cases, it is still advisable to use the proposed adjustments instead of a CI that neglects all boundary conditions. This is because the current approach still includes partial boundary information of the parameter space in such cases and is expected to produce CIs whose coverage probabilities are closer to the nominal ones. For example, the adjustment of CI of the standardized e^2 when e^2 is close to 1 in an ACE model is invalid, because the boundary condition of $e^2 = 1$ implies $a^2 = c^2 = 0$, which involves two parameters. Geometrically, the parameter space is given by $\{a^2 + c^2 + e^2 = 1, a^2 > 0, c^2 > 0 \text{ and } e^2 > 0\}$, which defines an angular region near the point $e^2 = 1$. The proposed adjustments assume the parameter space is a half plane $\{a^2 + c^2 + e^2 = 1 \text{ and } e^2 < 1\}$, which is larger than the true parameter space but still more restricted than the plane $\{a^2 + c^2 + e^2 = 1\}$ assumed in the regular CI.

Summary and conclusion

In this article we present two adjustments to traditional CIs when the parameter of interest is close to its boundary. The Wald type adjustment inverts a Wald test based on the inequality-constrained MLE. The LRT type adjustment inverts a LRT. Both adjustment methods are developed to be invariant under reparametrization and are consistent with the corrected LRT for a boundary null hypothesis. Both methods also give the regular CI when the MLE is far away from the boundary. The Wald type adjustment is relatively simpler but may give CIs with higher confidence than its nominal level in some cases. The LRT type adjustment is more complicated but does not suffer from

[7] Similarly, a regular CI for an unbounded parameter may not be valid when another parameter is close to its boundary.

this problem of being conservative. The Wald type adjustment tends to give longer CIs when the MLE is on boundary, while the LRT type adjustment tends to give longer CIs when the MLE is not on the boundary, but the expected lengths of the CIs do not differ much.

Acknowledgments This research is supported by the National Institute of Drug Abuse training grant R25DA026119 to the second author. We thank Gregory Carey, Sophie van der Sluis and Conor Dolan for their helpful comments that have led to the improvement of the style of this article.

References

Boker S, Neale M, Maes H, Wilde M, Spiegel M, Brick T, Spies J, Estabrook R, Kenny S, Bates T, Mehta P, Fox J (2011) OpenMx: an open source extended structural equation modeling framework. Psychometrika 76:306–317

Carey G (2005) Cholesky problems. Behav Genet 35:653–665

Dominicus A, Skrondal A, Gjessing HK, Pedersen NL, Palmgren J (2006) Likelihood ratio tests in behavioral genetics: problems and solutions. Behav Genet 36:331–340

Lehmann EL (1986) Testing statistical hypotheses. Wiley, New York

Neale MC (2004) Mx Manual. Available at http://www.vcu.edu/mx/documentation.html

Neale MC, Cardon LR (1992) Methodology for genetic studies of twins and families. Kluwer, Dordrecht

Neale MC, Heath AC, Hewitt JK, Eaves LJ, Fulker DW (1989) Fitting genetic models with LISREL: hypothesis testing. Behav Genet 19:37–49

Neale MC, Miller MB (1997) The use of likelihood based confidence intervals in genetic models. Behav Genet 27:113–120

Self SG, Liang K-Y (1987) Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under non-standard conditions. J Am Stat Assoc 82:605–610

Shapiro A (1988) Towards a unified theory of inequality constrained testing in multivariate analysis. Int Stat Rev 56:49–62

Visscher PM (2006) A note on the asymptotic distribution of likelihood ratio tests to test variance components. Twin Res Hum Genet 9:490–495

Wu H, Neale MC (in press) On the likelihood ratio tests in bivariate ACDE models. Psychometrika