Interval Estimations for variance Components: A Review and Implementations

Jorge Quiroz[1] and Richard Baumgartner [1]

[1]Merck Research Laboratories
Kenilworth, NJ

NCB 2017
Rutgers University
June, 2017

1. Introduction
   - Background
   - Motivation
   - The balanced one-way random model
2. Confidence Intervals
   - Frequentist Approach
     - Howe's Approach
     - Satterthwaite Approximation
     - Saddlepoint Confidence Intervals
     - Numerical Inversion of Characteristic Functions
   - Fiducial Inference: An Introduction
     - Generalized Confidence Intervals
   - Bayesian Methods
3. Simulation study
4. Numerical Examples
5. Conclusion

Background

- In general, confidence interval estimates are essential to account for the uncertainty associated with the data.
- Confidence ntervals of variance components are ubiquitous in drug development applications in and beyond non-clinical areas.
- More specifically, confidence intervals of variance components are important statistical tool to determine whether a level of precision is consistent with specified validation limits.

Non-clinical and Beyond

- Method transfer studies
- Method validation studies
- Gauge *R&R* studies
- Design of experiment
- Preclinical: Imaging assays that enable repeated measurements
- Clinical: Proof of concepts studies of novel biomarkers

Motivation

We survey and compare, via a comprehensive simulation study, new and established methods for constructing intervals for variance components in a balanced normal-based random effects design. Two immediate results of our discussion are:

- To facilitate the implementation of innovative new methods not currently available for practitioners
- Automation (e.g. R-package) will allows practitioners to have more opportunities to experience the benefit of these methods.

The balanced one-way random model

The balanced normal-based mixed design are commonly used to properly undestand and quantify measures of precision.

The balanced one-way random model is

$$\begin{aligned} Y_{ij} &= \mu + A_i + E_{ij} \\ i &= 1, \ldots, g, \\ j &= 1, \ldots, n, \end{aligned}$$

where μ is an unknown constant, A_i and E_{ij} are mutually independent normal random variables with means of zero and variances σ_A^2 and σ_E^2, respectively.

The variances σ_A^2 and σ_E^2 are denoted the *variance components*.

ANOVA Table

Table: ANOVA for the one-way random model

Source of variation	DF	MS	EMS
Between groups (A)	$d_a = g - 1$	S_A^2	$\theta_A = \sigma_E^2 + n\sigma_A^2$
Within groups (E)	$d_e = g(n-1)$	S_E^2	$\theta_E = \sigma_E^2$

Distributional Properties

$$\frac{(g-1)S_A^2}{\theta_A} \sim \chi^2_{g-1}$$

$$\frac{g(n-1)S_E^2}{\theta_E} \sim \chi^2_{g(n-1)}$$

Linear combination of variance components

Many parameters of interest are expressed as linear combination of variance components.

$$\theta = \sum_{k=1}^{K} c_k \theta_k$$

where c_k are selected constants that could be positive or negative and θ_k are the expected means of squares.

Parameters of interest

$$\begin{aligned} \sigma_A^2 &= \frac{1}{n}\theta_A - \frac{1}{n}\theta_E, \\ \sigma_A^2 + \sigma_E^2 &= \frac{1}{n}\theta_A + \left(1 - \frac{1}{n}\right)\theta_E. \end{aligned}$$

Howe's Approach

The basic steps for the confidence intervals construction are as follows:

1. $E_k = c_k S_k^2$ for $c_k \theta_k$, for $k = 1, \ldots, K$.

2. The upper and lower bounds for $c_k \theta_k$ are exact :

$$U_k = \begin{cases} \frac{c_k d_k S_k^2}{\chi^2_{\alpha/2, d_k}} & \text{for } c_k > 0 \\ \frac{c_k d_k S_k^2}{\chi^2_{1-\alpha/2, d_k}} & \text{for } c_k < 0 \end{cases} \qquad L_k = \begin{cases} \frac{c_k d_k S_k^2}{\chi^2_{1-\alpha/2, d_k}} & \text{for } c_k > 0 \\ \frac{c_k d_k S_k^2}{\chi^2_{\alpha/2, d_k}} & \text{for } c_k < 0 \end{cases}$$

   where $\chi^2_{p,df}$ is the *pth* quantile of a central chi-squared random variable with df degrees of freedom.

3. The respective upper and lower bounds of an approximate $100(1-\alpha)\%$ two-sided confidence interval on θ based on Howe's method are

$$\mathcal{U} = \sum_{k=1}^{K} E_k + \sqrt{\sum_{k=1}^{K} (U_k - E_k)^2} \qquad \mathcal{L} = \sum_{k=1}^{K} E_k - \sqrt{\sum_{k=1}^{K} (L_k - E_k)^2}.$$

   Negative bounds are set to zero.

Satterthwaite's confidence interval

Satterthwaite approach for constructing confidence interval on θ is based on approximating the random variable $\widehat{\theta} = \sum_{k=1}^{K} c_k S_k^2$ with a chi-squared distribution with corresponding degrees of freedom, d^*.

$$\frac{d^* \widehat{\theta}}{\theta} \sim \chi^2_{d^*}$$

where

$$d^* = \frac{\left(\sum_{k=1}^{K} c_k S_k^2\right)^2}{\sum_{k=1}^{K} \frac{(c_k S_k^2)^2}{d_k}}.$$

A $100(1-\alpha)\%$ two-sided Sattertwhwaite confidence intervals on θ is

$$\left[\frac{d^* \widehat{\theta}}{\chi^2_{1-\alpha/2;\ d^*}}; \quad \frac{d^* \widehat{\theta}}{\chi^2_{\alpha/2;\ d^*}}\right]$$

Saddlepoint Confidence Intervals

Consider the pivotal quantity $Q = \widehat{\theta}/\theta$.

$$Q = \frac{\sum_{k=1}^{K} c_k S_k^2}{\theta} = \sum_{k=1}^{K} w_k S_k^2 \tag{1}$$

where $w_k = c_k/\theta$.
The cumulant generating function (CGF) is

$$K(s) = -\frac{1}{2}\sum_{k=1}^{K} d_k \ln(1 - 2\lambda_k s), \tag{2}$$

where $\lambda_k = \frac{c_k \theta_k}{d_k \theta}$, d_k is the corresponding degrees of freedom for S_k^2, and $(1 - 2\lambda_k s) > 0$ for all $k = 1, \ldots, K$.
The saddlepoint approximation to the cumulative density function of Q, denoted by F^{sp} is

$$F^{sp}(x) = Prob[Q \leq x] \approx \Phi\left(w + \frac{1}{w}\ln\left(\frac{v}{w}\right)\right) \tag{3}$$

where

$$w = sign(\tilde{u})\sqrt{2\left[\tilde{u}x - K'(\tilde{u})\right]}, \qquad v = \tilde{u}\sqrt{K''(\tilde{u})},$$

and $\tilde{u} = \tilde{u}(x)$, denoted the saddlepoint at the value x, is the unique value of u that satisfies $K'(u) = x$.

Saddlepoint Confidence Intervals

A two-sided $100(1-\alpha)\%$ confidence interval for θ based on the saddlepoint approximation is

$$\left[\frac{\widehat{\theta}}{Q_{1-\alpha/2}}; \quad \frac{\widehat{\theta}}{Q_{\alpha/2}}\right]$$

where $\widehat{\theta} = \sum_{k=1}^{K} c_k S_k^2$, P_q is defined as the value that satisfies $F^{sp}(Q_q) = q^*$ and q^* is calculated accordingly to ensure that it is positive.

Numerical Inversion of CF

Once again consider the pivotal quantity

$$Q = \frac{\sum_{k=1}^{K} c_k S_k^2}{\theta} = \sum_{k=1}^{K} w_k S_k^2$$

where $w_k = c_k/\theta$.
The characteristic function (CF) is CF of Q is

$$\phi_Q(s) = \prod_{k=1}^{K} (1 - 2i\lambda_k s)^{-d_k/2}$$

where $i = \sqrt{-1}$, d_k represents the degrees of freedom, and $\lambda_k = \frac{c_k \theta_k}{d_k \theta}$.
Using the numerical inversion of developed by Gil-Pelaez of the characteristics function, the cumulative density function for the pivotal quantity Q is

$$F^{ni}(x) = Prob[Q \leq x] = \frac{1}{2} - \frac{1}{2\pi} \int_{-\infty}^{\infty} Im\left(\frac{\phi_P(s) e^{-isx}}{s}\right) ds.$$

The built-in functions `imhof` from the add-on R package `CompQuadForm` in R can be used to compute this integral.

Inversion confidence interval

A two-sided $100(1-\alpha)\%$ confidence interval for θ based on the numerical inversion formula is

$$\left[\frac{\widehat{\theta}}{Q_{1-\alpha/2}}; \quad \frac{\widehat{\theta}}{Q_{\alpha/2}}\right]$$

where $\widehat{\theta} = \sum_{k=1}^{K} c_k S_k^2$, Q_q is defined as the value that satisfies $F^{ni}(Q_q) = q^*$ and q^* is calculated accordingly to ensure that it is positive.

In simulation studies, we have the advantage of knowing the true values of λ_k. In these cases, the confidence intervals are exact and it performance can be used to benchmark the other intervals.

What is Fiducial?

Fiducial: Based on or relating to faith or trust. Here is a simple example to see the connection between "fiducial" with Fisher idea[1].

- $X \sim \mathsf{N}(\theta, 1)$.
- $X = \theta + E, \quad \text{where} \quad E \sim \mathsf{N}(0, 1)$.
- If we trust that a realized value $X = x$ does not change our belief that $E \sim \mathsf{N}(0, 1)$. We can write $\theta = x - E$.
- $\mathsf{N}(x, 1)$ can be used to summarize uncertainty about θ given the realized value x.

[1] Teaching notes by Martin Ryan, University of Illinois at Chicago

Generalized Pivotal Quantity

Generalized pivotal quantities are a fundamental underpinning for constructing generalized confidence intervals (GCIs), which can be considered a special case of fiducial inference.

A *Generalized pivotal quantity* (GPQ), $\mathcal{R}(\theta_A)$, for θ_A

$$W_A = \frac{(g-1)S_A^2}{\theta_A} \quad \Rightarrow \quad \mathcal{R}(\theta_A) = \frac{(g-1)s_A^2}{W_A},$$

where W_A is a chi-squared rv with $g-1$ DF, and s_A^2 is a realized value of S_A^2. $\mathcal{R}(\theta_A)$ can be used to construct confidence interval on θ_A given the realized value s_A^2.

Generalized Confidence Intervals

A *generalized pivotal quantity* for $\theta = \sum_{k=1}^{K} c_k\theta_k$

$$\mathcal{R}(\theta) = \sum_{i=1}^{K} c_k \frac{d_k s_k^2}{W_k}$$

where W_k is a chi-squared random variables with degrees of freedom d_k, and s_k^2 is a realized value of the random variable S_k^2.

Generalized confidence intervals on θ can be constructed via simulation of values from $\mathcal{R}(\theta)$ or by calculating exact values for the CDF using numerical inversion. Negative values are set to zero.

Bayesian Intervals

- Bayesian approaches require calculation of marginal posterior distributions
  1. Gibbs sampler techniques enable calculation of intervals by sampling from of the posterior distribution in several useful scenarios in straightforward ways
     - Inverse-gamma priors
     - Half-Cauchy priors
  2. Non-conjugate priors-modeled using the non-informative Jeffrey's reference priors. Posterior is numerically equivalent to GPQ.

Simulation Study

Model: One-way random effects model

Scenarios

σ_A^2	0.11	0.25	0.43	0.67	1	
σ_E^2	1	1	1	1	1	
ρ	0.1	0.2	0.3	0.4	0.5	
σ_A^2	1.5	2.33	4	10	20	50
σ_E^2	1	1	1	1	1	1
ρ	0.6	0.7	0.8	0.9	0.95	0.99

Sample sizes Groups = 5, 15, 15, replicates = 5, 10.

Condition To compare all the methods, only samples with $\widehat{\theta} > 0$ were considered.

Parameters σ_A^2 and $\sigma_T^2 = \sigma_A^2 + \sigma_E^2$.

Iterations 2000.

Criteria Empirical level and average length.

Results for σ_A^2

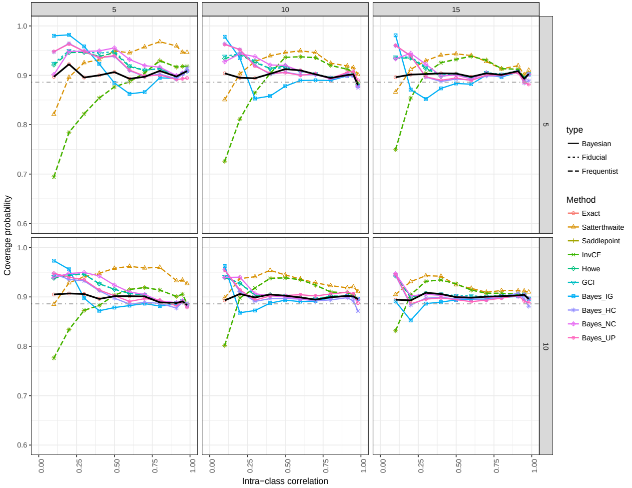

Figure: Coverage probability for σ_A^2

Results for σ_A^2

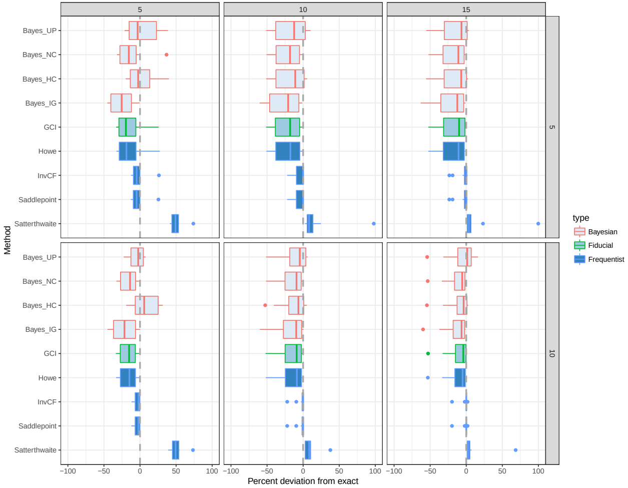

Figure: Average length for σ_A^2

Results for $\sigma_T^2 = \sigma_A^2 + \sigma_E^2$

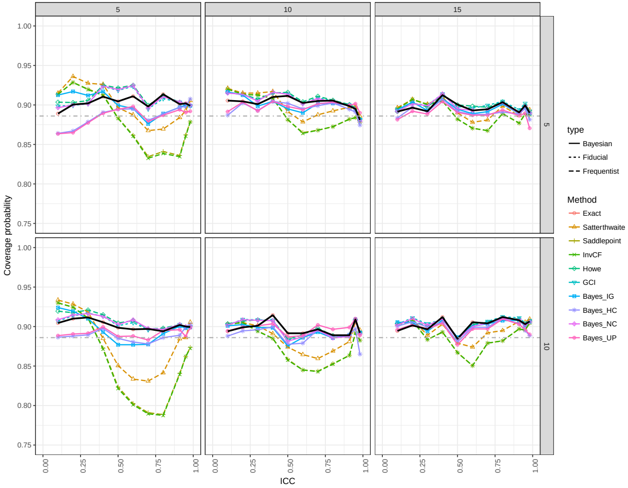

Figure: Coverage probability for σ_T^2

Results for $\sigma_T^2 = \sigma_A^2 + \sigma_E^2$

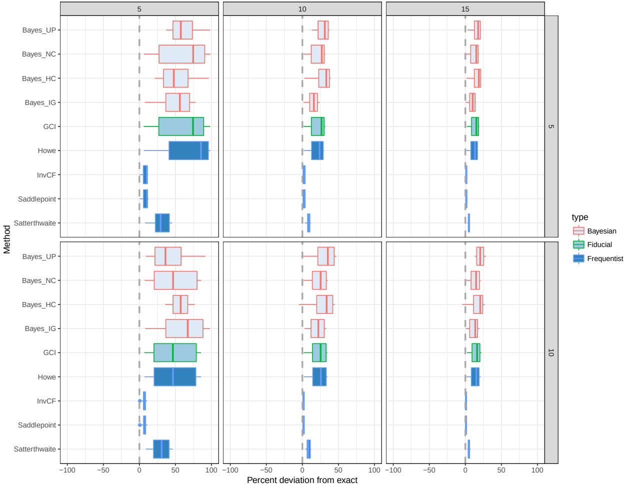

Figure: Averge length for σ_T^2

Recommendations from simulation study

Method	Maintain the nominal level?	Recommendation
Satterthwaite	No	Large sample
Saddlepoint	No	Large sample
Inv CF	No	Large sample
Howe	**Yes**	All scenarios
GCI	**Yes**	All scenarios
Bayes (IG)	No	Large sample
Bayes (HC)	**Yes**	All scenarios
Bayes (NC)	**Yes**	All scenrarios

Numerical Examples

Example 1 The data represent measures the thermal impedance of semiconductor power modules from a from a gauge $R\&R$ study. Ten randomly selected parts were measures 3 times by the same operators.

Example 2 The data represent a property related to stickness of samples of blood. Six subjects were selected at random from a large population, and a property related to stickness of samples of blood was measured 7 times on each subject.

Numerical Examples

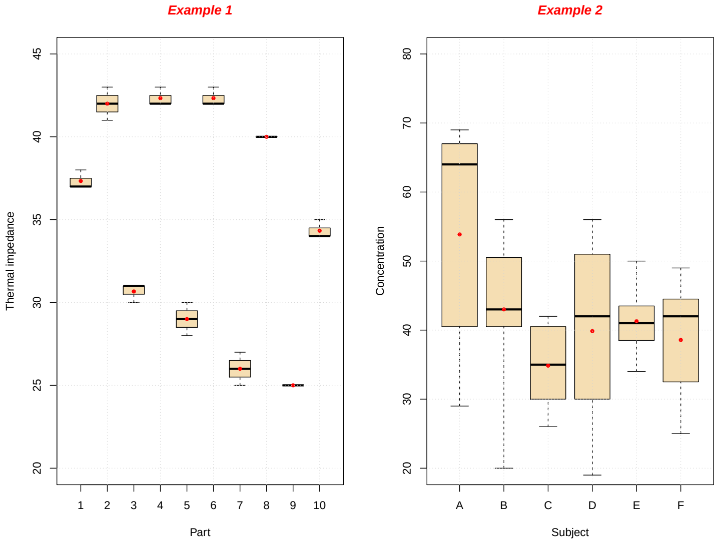

Figure: Coverage probability for σ^2

Results for Example 1

Table: 95% two-sided intervals for example 1

	σ_P		$\sigma_T = \sqrt{\sigma_P^2 + \sigma_E^2}$	
Method	LB	UB	LB	UB
Satterthwaite	4.62	13.11	4.73	12.56
Saddlepoint	4.70	12.54	4.74	12.46
Inv CF	4.70	12.55	4.74	12.46
Howe	4.70	12.51	4.75	12.53
GCI (S)	4.69	12.50	4.77	12.60
GCI (E)	4.70	12.51	4.75	12.53
Bayes (IG)	4.71	12.59	4.77	12.61
Bayes (HC)	4.81	11.16	4.87	11.18
Bayes (NC)	4.72	12.55	4.77	12.57
Bayes (AP)	4.79	13.73	4.85	13.75

Results for Example 2

Table: 95% two-sided intervals for example 2

	σ_S		$\sqrt{\sigma_S^2 + \sigma_E^2}$	
Method	LB	UB	LB	UB
Satterthwaite	2.18	155.77	9.89	16.13
Saddlepoint	2.46	21.45	9.85	15.93
Inv CF	2.46	21.49	9.85	15.93
Howe	0	15.30	10.27	19.32
GCI (S)	0	15.19	10.22	19.28
GCI (E)	0	15.26	10.22	19.32
Bayes (IG)	0.04	11.53	10.07	16.87
Bayes (HC)	0.36	13.95	10.29	18.66
Bayes (NC)	1.43	15.51	10.24	19.37
Bayes (AP)	0.36	15.97	10.35	19.98

Conclusion

- We survey and compare, via a comprehensive simulation study, methods for constructing intervals for linear combination of variance components in a balanced normal-based random effects design.
- The simulation study indicated that Howe, GCI, and Bayesian methods maintained the stated confidence level. They are the recommended methods.
- The simulation study indicated that Saddlepoint and numerical inversion methods did not generally maintained the stated confidence levels. These methods can be potentially improved by using more robust point estimates for variance components.
- All the method discussed here can be made readily accessible for multiple applications in relatively simple computer codes. For example, an `R` package.