

Fiducial Inference in Linear Mixed-Effects Models

Jie Yang, Xinmin Li , Hongwei Gao and Chenchen Zou * 

School of Mathematics and Statistics, Qingdao University, Qingdao 266071, China; xmli@qdu.edu.cn (X.L.); gaohongwei@qdu.edu.cn (H.G.)

* Correspondence: zouchenchen@hotmail.com

Abstract: We develop a novel framework for fiducial inference in linear mixed-effects (LME) models, with the standard deviation of random effects reformulated as coefficients. The exact fiducial density is derived as the equilibrium measure of a reversible Markov chain over the parameter space. The density is equivalent in form to a Bayesian LME with noninformative prior, while the underlying fiducial structure adds new benefits to unify the inference of random effects and all other parameters in a neat and simultaneous way. Our fiducial LME needs no additional tests or statistics for zero variance and is more suitable for small sample sizes. In simulation and empirical analysis, our confidence intervals (CIs) are comparable to those based on Bayesian and likelihood profiling methods. And our inference for the variance of random effects has competitive power with the likelihood ratio test.

Keywords: confidence interval; fiducial inference; LME; MCMC; zero-variance inference

MSC: 62F99



Academic Editors: Ali Mohammad-Djafari and Geert Verdoolaege

Received: 30 September 2024

Revised: 20 January 2025

Accepted: 28 January 2025

Published: 3 February 2025

Citation: Yang, J.; Li, X.; Gao, H.; Zou, C. Fiducial Inference in Linear Mixed-Effects Models. *Entropy* **2025**, *27*, 161. <https://doi.org/10.3390/e27020161>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The linear mixed-effects (LME) model is very popular in many fields like physics, biology and social sciences [1], where both within-subject and between-subject variations are studied. There have been a wealth of LME-based methodologies developed in the past decades, such as model selection [2], model averaging [3], semiparametric [4] and nonparametric [5] estimations, etc. Large-sample asymptotics of likelihood-based LME models are well studied [6]. Bayesian LME models also perform well under medium-to-large samples. For small samples, the inference based on the maximum likelihood needs finite-sample corrections or approximations [7], and Bayesian models rely much more on the prior.

The analysis of variance (ANOVA) and t-test summaries are two common ways of statistical inference in LME models. For fixed effects, F statistics in ANOVA and t statistics are available in most software. However, testing zero-variance components in random effects is challenging for small samples and has not been well resolved in the statistical literature. When some of the within-subject variances are zero or close-to-zero, Bayesian estimations may encounter issues [8]. And the inference of zero variance involves additionally constructed statistics like the profiling restricted maximum likelihood (REML) ratio test [9] or posterior evidence ratio with properly built prior [10]. Comparisons among multiple nested models are often unskippable.

From a frequentist viewpoint, aside from asymptotic approximations, the inference can be facilitated by nonparametric approaches [11,12], which are also known to be performant in LME models for small sample sizes. According to a comparative study [13] on various

bootstrap methods in LME models, a nonparametric bootstrap with resampling on subjects is recommended [11].

Fiducial inference, initiated by Fisher [14] and revitalized in the 21st century [15], has unique advantages in small-sample inference in parametric statistical models, despite its limits in nonparametric statistics and computing cost [16]. Unlike Bayesian posterior built upon prior, fiducial distribution derived from data generation equations thus relieves the need for prior. For LME, Li et al. [17] proposed a fiducial test for within-subject variances, but they did not consider other parameters. Hari et al. [18] studied a two-component LME with partially known structures on covariances. Hannig et al. [19] developed a generalized fiducial inference (GFImm) for interval data. Although GFImm can be adapted to non-interval cases by artificially adding the intervals, the estimations vary with the interval widths. And its numerical algorithm in practice is inefficient on models with more than nine parameters.

In this study, we propose a new framework for non-asymptotic inference in LME models via a fiducial approach, which reduces computational burden and does not require duplicates within each group. Moreover, our approach naturally incorporates the zero-variance inference, thus cutting the labor of extra tests and calculation.

The rest of this article is organized as follows: Sections 2 and 3 are about the methodology and algorithms. Section 4 is the simulation and section 5 is a real data illustration. Section 6 is the conclusion and discussion of our study.

2. Fiducial Distribution in LME

We assume the data is generated as

$$y_i = \mathbf{x}_i\beta + \mathbf{z}_i\gamma_i + e_i, i = 1, \dots, m, \quad (1)$$

where $\beta \in R^{p \times 1}$ is a vector of fixed coefficients; $\gamma_i \sim N_q(0, D)$ is a random coefficient vector; y_i is the $n_i \times 1$ response of the i th subject; \mathbf{x}_i and \mathbf{z}_i are the $n_i \times p$ and $n_i \times q$ covariate matrix of fixed and random effects, respectively; and $e_i \sim N(0, \sigma^2 I_{n_i})$ is independent of γ_i , $n_i \geq 1$, $i = 1, \dots, m$, $\sum_{i=1}^m n_i = n$. In this study, we assume a diagonal $D = \text{diag}(\sigma_1^2, \dots, \sigma_q^2)$.

2.1. Conditional Fiducial Distribution

Let $\delta = (\sigma_1, \dots, \sigma_q)'$, $U_i \stackrel{i.i.d.}{\sim} N(0, I_q)$ and $z_{U,i} = \mathbf{z}_i \text{diag}(U_i)$, then model (1) can be rewritten as

$$y_i = \mathbf{x}_i\beta + z_{U,i}\delta + e_i, i = 1, \dots, m.$$

It can be further specified in matrix form as

$$Y = X\beta + Z_U\delta + \epsilon, \quad (2)$$

with $Y = (y_1', \dots, y_m)'$, $X = (\mathbf{x}_1', \dots, \mathbf{x}_m)'$, $Z_U = (\mathbf{z}'_{U,1}, \dots, \mathbf{z}'_{U,m})'$, $U = (U_1', \dots, U_m)'$ and $\epsilon = (e_1', \dots, e_m)'$. The σ_j in random effects are formulated as a coefficient vector δ in Equation (2). This brings two benefits: 1. The inference of random effects can be realized by a general fiducial recipe [16] like all other parameters, avoiding extra tests and statistics like the LR tests. 2. We do not need $n_i \geq q$. The model can still be estimated even if $n_i = 1$. So, our method can handle inadequate within-subject measures.

Let $\theta = (\beta', \delta)'$, $X_U = (X, Z_U)_{n \times (p+q)}$ and $P_U = X_U(X_U'X_U)^{-1}X_U'$, then model (2) given $U = u$ can be reformed as

$$Y = X\beta + Z_u\delta + \epsilon = X_u\theta + \epsilon,$$

For observed $Y = y$, the conditional fiducial density of θ and σ^2 given u , denoted as $r_y(\theta|\sigma^2, u)$ and $r_y(\sigma^2|u)$, respectively, could be derived from the conditional fiducial quantities [17] below:

$$Q_y(\sigma^2|u) = y'(I - P_u)y/\zeta, \tag{3}$$

$$Q_y(\theta|u) = (X'_u X_u)^{-1} X'_u y + Q^{\frac{1}{2}}_y(\sigma^2|u)(X'_u X_u)^{-\frac{1}{2}} \zeta, \tag{4}$$

where $\xi \sim \chi^2_{n-p-q}$, $\zeta \sim N(0, I_{p+q})$ independently, irrespective of any parameters in an LME model. Let $\eta = (\theta', \sigma^2)'$, then the conditional fiducial density of η given u is

$$\begin{aligned} r_y(\eta|u) &= r_y(\sigma^2|u)r_y(\theta|\sigma^2, u) \\ &= \frac{\exp\left(-\frac{y'(I-P_u)y}{2\sigma^2}\right)\left(\frac{y'(I-P_u)y}{2\sigma^2}\right)^{\frac{n-p-q}{2}-1} \exp\left(-\frac{1}{2\sigma^2}(\theta - (X'_u X_u)^{-1} X'_u y)'(X'_u X_u)(\theta - (X'_u X_u)^{-1} X'_u y)\right)}{\sigma^4 \Gamma\left(\frac{n-p-q}{2}\right) (2\pi)^{\frac{p+q}{2}} \sigma^{p+q} |X'_u X_u|^{-\frac{1}{2}}} \\ &= \frac{(y'(I - P_u)y)^{\frac{n-p-q}{2}-1} \exp\left(-\frac{\|y - X_u \theta\|^2}{2\sigma^2}\right)}{\sigma^{n-p-q+2} \Gamma\left(\frac{n-p-q}{2}\right) 2^{\frac{n-p-q}{2}-1} (2\pi)^{\frac{p+q}{2}} \sigma^{p+q} |X'_u X_u|^{-\frac{1}{2}}} \\ &= \frac{2|X'_u X_u|^{\frac{1}{2}} (y'(I - P_u)y)^{\frac{n-p-q}{2}-1} \exp\left(-\frac{\|y - X_u \theta\|^2}{2\sigma^2}\right)}{\Gamma\left(\frac{n-p-q}{2}\right) \pi^{\frac{p+q-n}{2}}} \triangleq C(X_u, y) \frac{p(y|u, \eta)}{\sigma^2}. \end{aligned} \tag{5}$$

Here, $C(X_u, y) = \frac{2|X'_u X_u|^{\frac{1}{2}} (y'(I - P_u)y)^{\frac{n-p-q}{2}-1}}{\Gamma\left(\frac{n-p-q}{2}\right) \pi^{(p+q-n)/2}}$ and $p(y|u, \eta)$ is the likelihood function of (2) given $U = u$.

U is invisible in reality. But, it is easy to see that $U_i|\eta, y \sim N_q(E(U_i|\eta, y), Var(U_i|\eta, y))$ with

$$E(U_i|\eta, y) = D^{\frac{1}{2}} z'_i (\sigma^2 I_{n_i} + z_i D z'_i)^{-1} (y_i - x'_i \beta), \tag{6}$$

$$Var(U_i|\eta, y) = I_q - D^{\frac{1}{2}} z'_i (\sigma^2 I_{n_i} + z_i D z'_i)^{-1} z_i D^{\frac{1}{2}},$$

$i = 1, \dots, m$. We denote the conditional density of U as $p(u|\eta, y)$ hereinafter.

2.2. Gibbs Sampler and the Final Fiducial Distribution

With both $r_y(\eta|u)$ and $p(u|\eta, y)$ available, the fiducial distribution could be realized by a Gibbs sampler. Let $K(\eta, \tilde{\eta}) \triangleq \int r_y(\eta|u)p(u|\tilde{\eta}, y)du$. By Equations (5) and (6), we have

$$\begin{aligned} K(\eta, \tilde{\eta}) &= \int C(X_u, y) \frac{p(y|u, \eta)}{\sigma^2} p(u|y, \tilde{\eta}) du \\ &= \int C(X_u, y) \frac{p(y, u|\eta)}{\sigma^2 p(u)} p(u|y, \tilde{\eta}) du \\ &= \frac{p(y|\eta)}{\sigma^2} \int \frac{C(X_u, y)}{p(u)} p(u|y, \eta) p(u|y, \tilde{\eta}) du \\ &\triangleq \frac{p(y|\eta)}{\sigma^2} h(\eta, \tilde{\eta}), \end{aligned} \tag{7}$$

where $h(\eta, \tilde{\eta}) = \int \frac{C(X_u, y)}{p(u)} p(u|y, \eta) p(u|y, \tilde{\eta}) du$ and $p(y|\eta)$ is the likelihood function with U integrated out. Obviously, $h(\cdot, \cdot)$ is symmetric and $\frac{p(y|\eta)}{\sigma^2}$ satisfies the detailed balance below:

$$\frac{p(y|\tilde{\eta})}{\tilde{\sigma}^2} K(\eta, \tilde{\eta}) = \frac{p(y|\tilde{\eta})}{\tilde{\sigma}^2} \frac{p(y|\eta)}{\sigma^2} h(\eta, \tilde{\eta}) = \frac{p(y|\eta)}{\sigma^2} \frac{p(y|\tilde{\eta})}{\tilde{\sigma}^2} h(\tilde{\eta}, \eta) = \frac{p(y|\eta)}{\sigma^2} K(\tilde{\eta}, \eta). \tag{8}$$

The final fiducial density, as the stationary distribution of η , can be derived by the reversibility [20] of the Markov chain on η :

$$r_y(\eta) \propto \frac{p(y|\eta)}{\sigma^2}, \tag{9}$$

which is equivalent to a Bayesian LME with uniform prior on β and δ and prior $\frac{1}{\sigma}$ on σ .

In practice, once a u sampled by Equation (6) makes $X_u(X'_u X_u)^{-1} X'_u y$ close to y , the $Q_y(\sigma^2|u)$ in Equation (3) also becomes close to 0, making $Q_y(\theta|u)$ in Equation (4) nearly constant and the Monte Carlo Markov Chain (MCMC) degenerated. So, we restrict σ 's parameter space away from zero by a small $c_0 > 0$ so that $\sigma^2 \in [c_0, +\infty)$ and use the constrained fiducial quantity by a trimmed χ^2_{n-p-q} as follows:

$$Q_y^*(\sigma^2|u) = \frac{y'(I - X_u(X'_u X_u)^{-1} X'_u)y}{F_{\chi^2_{n-p-q}}^{-1}(F_{\chi^2_{n-p-q}}(\frac{y'(I - X_u(X'_u X_u)^{-1} X'_u)y}{c_0})\mathcal{U}(0, 1))}, \tag{10}$$

where $F(\cdot)$ is the cumulative distribution function, $F^{-1}(\cdot)$ is its inverse, and $\mathcal{U}(0, 1)$ is a uniform random variable on $(0, 1)$. In this paper, we set $c_0 = 0.01$. The Gibbs sampling algorithm is given in Algorithm 1.

Algorithm 1: Gibbs sampling for $r_y(\eta)$

1. Initialize $U^{(0)}$ by $U_i^{(0)} \overset{i.i.d}{\sim} N(0, I_q), i = 1, \dots, m$.
 2. Given current $U^{(l)}$, sample $\beta^{(l)}, \delta^{(l)}$ from Equation (4) and $\sigma^{(l)}$ from Equation (10), respectively.
 3. Update $U^{(l+1)}$ by Equation (6).
 4. For $l = 0, 1, 2, \dots$ iterate between step 2 to 3 until the MCMC chain gets well mixed.
-

Remark 1. We can also start with Equation (9) by setting corresponding prior in ready-made Bayesian packages that employ algorithms with better quality like HMC or NUTS [21]. However, HMC is time consuming and loses the advantages of direct inference for $\sigma_j = 0$. The histogram of a Gibbs sample on δ can directly reflect how much the distribution is concentrated around zero. So, we mainly use Gibbs in this study and leave HMC for future research.

Remark 2. When D is non-diagonal, $U_i \overset{i.i.d}{\sim} N(0, I_q)$ becomes $U_i \overset{i.i.d}{\sim} N_q(0, R)$, where R is the correlation matrix. And $p(u|\eta, y)$ becomes $p(u|\eta, R, y)$. Given $U = u$, the fiducial density of R , say $r_u(R)$, is also available [22]. The Gibbs sampler turns into iterations among three parts: $r_y(\eta|u)$, $p(u|\eta, R, y)$ and $r_u(R)$. The fiducial density here is no longer equivalent to Equation (9) and cannot be readily implemented by Bayesian packages. We also leave this for future research.

3. Fiducial Inference for LME

Given the observed y , the fiducial distribution of a parameter of interest $g = g(\eta)$ can be given as

$$F_y(g) = \int_{\{g(\eta) \leq g\}} r_y(\eta) d\eta.$$

3.1. Interval Estimation

A $1 - \alpha$ fiducial interval for parameter $g(\eta)$ is formed by the $\frac{\alpha}{2}$ and $1 - \frac{\alpha}{2}$ quantiles of $F_y(g)$, denoted as $g_{\frac{\alpha}{2}}$ and $g_{1-\frac{\alpha}{2}}$, respectively, which are the solutions of

$$F_y(g_{\frac{\alpha}{2}}) = \frac{\alpha}{2}, F_y(g_{1-\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2}.$$

Then, the $1 - \alpha$ equal-tailed fiducial interval is $[g_{\frac{\alpha}{2}}, g_{1-\frac{\alpha}{2}}]$.

Once we find a numerical sample of η as Section 2.2 described, the $1 - \alpha$ fiducial interval of $g = g(\eta)$ can be obtained from the empirical distribution of $F_y(g)$, which is given in Algorithm 2.

Algorithm 2: Interval estimation for $g(\eta)$

1. Generate η_1, \dots, η_N according to Algorithm 1.
 2. Compute $g_i = g(\eta_i), i = 1, \dots, N$, and sort g_i in ascending order to get $(g_{(1)}, \dots, g_{(N)})$.
 3. Find the $\frac{\alpha}{2}$ sample quantile $g_{([N \cdot \frac{\alpha}{2}])}$ and $1 - \frac{\alpha}{2}$ sample quantile $g_{([N \cdot (1 - \frac{\alpha}{2})])}$. Then $[g_{([N \cdot \frac{\alpha}{2}])}, g_{([N \cdot (1 - \frac{\alpha}{2})])}]$ is the $1 - \alpha$ fiducial interval estimation.
-

Noting that σ_j is conventionally nonnegative, we use its zero-truncated distribution to construct the CI such that the left end is $\max\{0, \sigma_{(N \cdot \frac{\alpha}{2}), j}\}$. According to remark 2.1 and theorem 2.2 in [17], a truncated fiducial quantity uniquely solved from the fiducial structural equations has the same theoretical property as the untruncated one for inference in a restricted parameter space.

3.2. Fiducial p-Value

For a hypothesis test $H_0: g(\eta) = g_0$ with a two-sided alternative, the fiducial p-value is

$$p = 2\min\{F_y(g_0), 1 - F_y(g_0)\}.$$

Given a significance level α , H_0 is rejected when $p < \alpha$. The computing procedure is given in Algorithm 3.

Algorithm 3: Fiducial p-value for $H_0: g(\eta) = g_0$

1. Generate η_1, \dots, η_N according to Algorithm 1.
 2. Compute $g_i = g(\eta_i), i = 1, \dots, N$.
 3. Calculate $p = 2\min\left\{\frac{\sum_{i=1}^N 1_{[g_i > g_0]}}{N}, \frac{\sum_{i=1}^N 1_{[g_i \leq g_0]}}{N}\right\}$.
 4. Reject H_0 if $p < \alpha$.
-

Remark 3. Algorithms 2 and 3 are routine constructions of the CI and p value in unimodal distributions. All the histograms of δ that we have plotted so far are unimodal in a Gibbs sampler, probably because the sign of δ and that of U are interdependent in model (2) while Gibbs updates δ largely based on U . But, this does not mean the bimodal shape will not occur for δ in other algorithms bypassing the sampling of $r(\theta|u, y)$ since δ in Equation (9) is symmetric at about 0. For a bimodal distribution, routine p value and CI are not applicable but a histogram or density curve can directly tell.

4. Simulation

4.1. Confidence Intervals

We compare the coverages and lengths of CIs constructed by our method (Fiducial), the profiling method (Profiling) in R package lme4 [9], the Bayesian LME (Bayesian) with default prior in R package brms (version 2.22.0) [10], and the nonparametric bootstrap (Bootstrap) in [13]. We generate a chain over 6000 in length, with the first 300 as a warmup, and select every 3 steps to gather a sample size $N = 2000$ for the Gibbs sampler in Algorithm 1. We also try the HMC in Remark 1 by setting a uniform prior for β and σ_j , and $\frac{1}{\sigma}$ is approximated by an inverse gamma prior $IG(0.001, 0.001)$ in brms (The density of $IG(a, b)$

is $\frac{b^a e^{-b/\sigma}}{\Gamma(a)\sigma^{a+1}}$, $\sigma > 0$. $\frac{1}{\sigma} = \lim_{b \rightarrow 0} \lim_{a \rightarrow 0} \frac{b^a e^{-b/\sigma}}{\Gamma(a)\sigma^{a+1}}$ thus can be approximated by IG(a,b) with a and b close to 0.) Bootstrap CIs are constructed by percentiles of 2000 replicates resampled among subjects, and each replicate is estimated by REML (lmer in package lme4). All MCMCs are set as a single chain with a sample size of 2000. Coverages (nominal level 95%) are annotated in figures and tables, averaged over 1000 runs.

Example 1. $y_{ij} = \beta_0 + z_{ij}\gamma_{i1} + e_{ij}$ with $z_{ij} \sim N(0,1)$, $\gamma_{i1} \sim N(0,\sigma_1^2)$, $e_{ij} \sim N(0,\sigma^2)$, $1 \leq j \leq 2$, $1 \leq i \leq 10$, $\beta_0 = 0.8$, $\sigma_1 = 0.75$, $\sigma = 0.7$. In general, the Profiling CIs are narrower but susceptible to low coverages (Figure 1). Fiducial CIs are comparable to the Bayesian except those on σ_1 are often wider. This can be improved by HMC yet it is much slower. We use Gibbs for our method afterward. Bayesian LME is time-consuming yet performs well, partly due to its advantages in small samples and partly due to the suitability of the default prior for this parameter setting. Bootstrap has the lowest coverages; perhaps the sample size here is too small for it to take effect. The relative CPU time for model fit plus CI construction in fiducial (Gibbs), profiling, Bayesian, and bootstrap is about 2:1:30:40. Given the underperformance of bootstrap in coverage and computation efficiency, we do not discuss it in later examples. The variations in CI length are similar across different methods. So, we summarize the median length in the table hereinafter.

Example 2. $y_{ij} = \beta_0 + \gamma_{i0} + z_{ij}\gamma_{i1} + e_{ij}$ with $\gamma_{i0} \sim N(0,\sigma_0^2)$, $\sigma_0 = 0$, and the others the same as example 1. That is, we add a false random effect $\gamma_{i0} \equiv 0$. Bayesian LME is unable to directly detect zero variance (coverage = 0 for σ_0 , Table 1). Profiling CIs can have zero end points to keep reasonable coverages, but other parameters are affected, especially σ . Also, 5% ~ 10% of lmer’s output contains NA or ∞ when $n_i = q = 2$. This is not a big issue when $n_i = 4 > q$, but σ ’s coverage is still lower. In contrast, fiducial CIs capture both zero and nonzero variances in random effects and have stable coverages (Table 1).

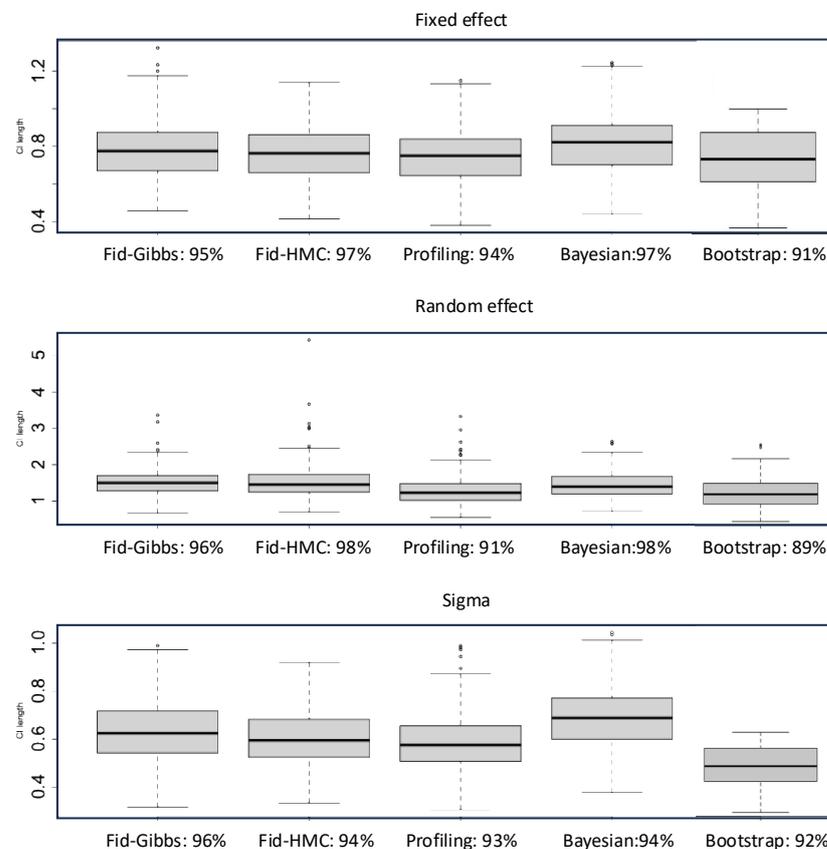


Figure 1. Confidence intervals in example 1.

Table 1. Coverage (length) of CIs in example 2.

$n_i = 2, m = 10$	$\beta_0 = 0.8$	$\sigma_0 = 0$	$\sigma_1 = 0.75$	$\sigma = 0.7$
Fiducial	0.99 (0.99)	0.95 (1.01)	0.97 (1.63)	0.95 (0.71)
Profiling	0.93 (0.80)	0.97 (0.70)	0.89 (1.38)	0.87 (0.64)
Bayesian	0.97 (1.03)	0.00 (0.98)	0.97 (1.49)	0.96 (0.77)
$n_i = 4, m = 5$	$\beta_0 = 0.8$	$\sigma_0 = 0$	$\sigma_1 = 0.75$	$\sigma = 0.7$
Fiducial	1.00 (1.30)	0.98 (1.50)	0.96 (2.21)	0.95 (0.60)
Profiling	0.95 (0.86)	0.99 (0.65)	0.93 (1.78)	0.92 (0.54)
Bayesian	1.00 (1.15)	0.00 (1.38)	0.97 (1.99)	0.95 (0.67)

Example 3. $y_i = \beta_0 + \gamma_{i0} + i \cdot \beta_1 + i \cdot \gamma_{i1} + e_i$ with $\gamma_{i0}, \gamma_{i1}, e_i$ all from $N(0, 1), i = 1, \dots, 10$, i.e., a longitudinal setting with $n_i = 1, q = 2$ so $n_i < q$. *lmer* does not work in this scenario, which makes profiling and bootstrap inapplicable. Bayesian CIs, overall, are wider with lower coverages than fiducial CIs (Table 2).

Table 2. Coverage (length) of CIs in example 3.

$n_i = 1, m = 10$	$\beta_0 = 10$	$\beta_1 = 0$	$\sigma_0 = 1$	$\sigma_1 = 1$	$\sigma = 1$
Fiducial	0.99 (16.31)	0.98 (3.36)	1.00 (9.89)	0.98 (2.06)	1.00 (6.26)
Profiling	NA	NA	NA	NA	NA
Bayesian	0.97 (20.87)	0.94 (4.36)	0.97 (9.01)	0.96 (2.79)	0.77 (8.58)

4.2. Zero-Variance Test for Random Effects

We plot the histograms of fiducial samples on zero and nonzero δ . They are indeed different (Figure 2). Negative δ generated by Equation (4) helps in measuring the significance of the random effect. We further compare p value distributions under null ($=0$) and alternative ($\neq 0$) hypotheses by treating example 2 as $y_{ij} = \beta_0 + \gamma_{i0} + z_{ij}\beta_1 + z_{ij}\gamma_{i1} + e_{ij}$ with $\beta_1 = 0$ and others unchanged. Frequentist p values are calculated by the package *lmerTest* [23] with fixed effects based on Satterthwaite-approximated t tests and random effects based on likelihood ratio (LR) tests. Fiducial p values are calculated simultaneously as described in Section 3. The type I error and power at a 5% significance level are annotated in Figure 3. The Satterthwaite p value has a higher type I error than the fiducial test (Figure 3). The type I errors of the fiducial and LR tests do not differ much, but the power of the fiducial test is higher. Moreover, the LR p value has a weird peak near 1 under alternative hypothesis (Figure 3), which is an obvious drawback in small data analysis. From this perspective, we think our fiducial approach can contribute to the inference in LME for small samples.

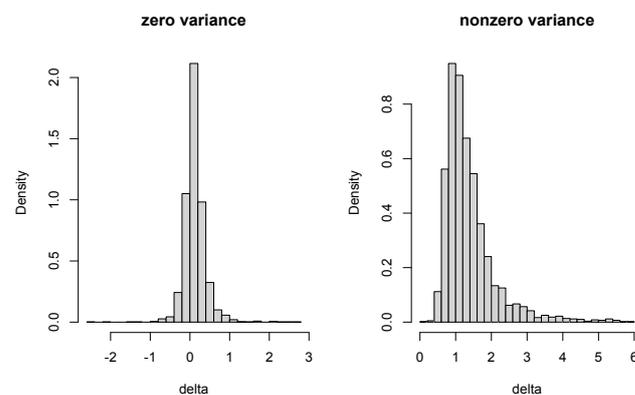


Figure 2. Fiducial density for $\delta = 0$ and $\delta = 1$.

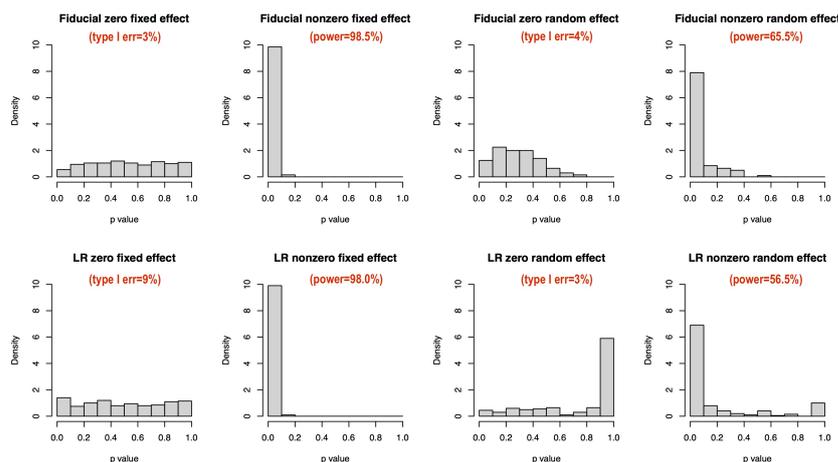


Figure 3. *p* value comparison for fixed and random effects, $n = 20$.

4.3. Comparison with GFllmm

For the GFllmm in [19] (simplified as GFI hereinafter), we set a simpler scenario that their R package gflmm (version 2.0.5) can run for comparison: $y_{ij} = \beta_0 + \gamma_i + \beta_1 x_{ij} + e_i$. Here, $\beta_0 = 0.8$, $\beta_1 = 0.5$, $\gamma_i \sim N(0, \delta^2)$, $x_{ij} \sim N(0, 1)$ and $e_i \sim N(0, 1)$, $j = 1, 2; i = 1, \dots, 10$. Intervals are added as $y_{ij} - 0.01 < y_{ij} < y_{ij} + 0.01$. Performances on both zero and nonzero δ are compared (Table 3). The CIs in the nonzero case are comparable, and GFI is better in the zero case. We also plot histograms of the two fiducial *p* values. At a 5% significance level, our method has lower type I error and higher power than GFI (Figure 4).

Table 3. Coverages (length) of CIs in comparison with GFI.

Parameter	$\beta_0 = 0.8$	$\beta_1 = 0.5$	$\delta = 0.75$	$\sigma = 1$
Fiducial	0.97 (1.56)	0.94 (0.83)	0.98 (1.55)	0.96 (0.82)
Profiling	0.94 (1.43)	0.92 (0.82)	0.94 (1.25)	0.96 (0.79)
GFI	0.96 (1.60)	0.96 (0.93)	1.0 (1.45)	0.95 (0.93)
Parameter	$\beta_0 = 0.8$	$\beta_1 = 0.5$	$\delta = 0$	$\sigma = 1$
Fiducial	0.98 (1.23)	0.95 (1.00)	0.98 (1.21)	0.95 (0.78)
Profiling	0.93 (0.96)	0.93 (0.75)	0.96 (0.88)	0.91 (0.66)
GFI	0.97 (1.15)	0.94 (0.82)	0.98 (1.09)	0.96 (0.77)

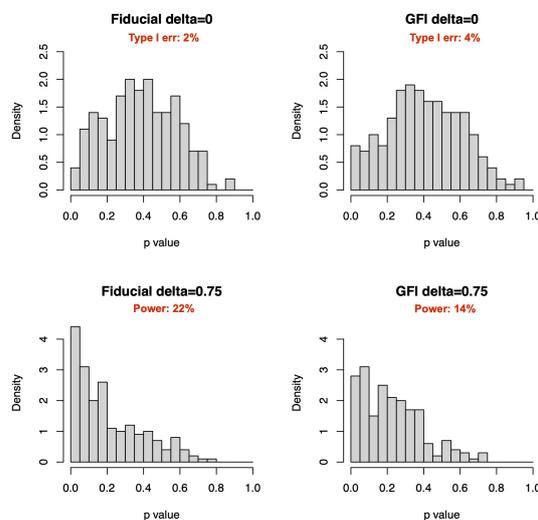


Figure 4. *p* value distributions on zero and nonzero random effects ($n_i = 2, m = 10$).

5. Empirical Analysis

We use the sleepstudy data in R package lme4 as an illustration, which contains a sleep-deprived group of 18 subjects for 10 days of the study in [24]. Day0 and Day1 were adaptation and training, so altogether $18 \times 8 = 144$ observations. The reaction time of the i th subject on Day_j after sleep deprivation is modeled as

$$reaction_{ij} = \beta_0 + \gamma_{i1} + \beta_1 Day_j + \gamma_{i2} Day_j + e_{ij},$$

with $\gamma_{i1} \sim N(0, \sigma_1^2)$, $\gamma_{i2} \sim N(0, \sigma_2^2)$, and $e_i \sim N(0, \sigma^2)$. Both the fixed and random effects are significant [9]. We resample 25% of the data for a subset with $n = 9 \times 4 = 36$ to see how much the inference based on a smaller sample agrees with that on the whole dataset. We conduct resampling 50 times and plot the CI lengths of β_0 , β_1 , σ_1 , and σ_2 in Figure 5. Different from the simulation, the Bayesian intervals are much wider this time since the default prior does not well suit the real data. We annotate the percentage of CIs excluding 0 in Figure 5 as an estimated power, on which the fiducial LME are generally higher. We further try a combination of $m = 18$ and $n_i = 2$, retaining $n = 36$, but neither lmer nor brms can fit all the cases successfully this time. Our fiducial LME still works. The percentages of significance are 100% on β_0 , 80% on β_1 , 21% on σ_1 , and 74% on σ_2 , respectively.

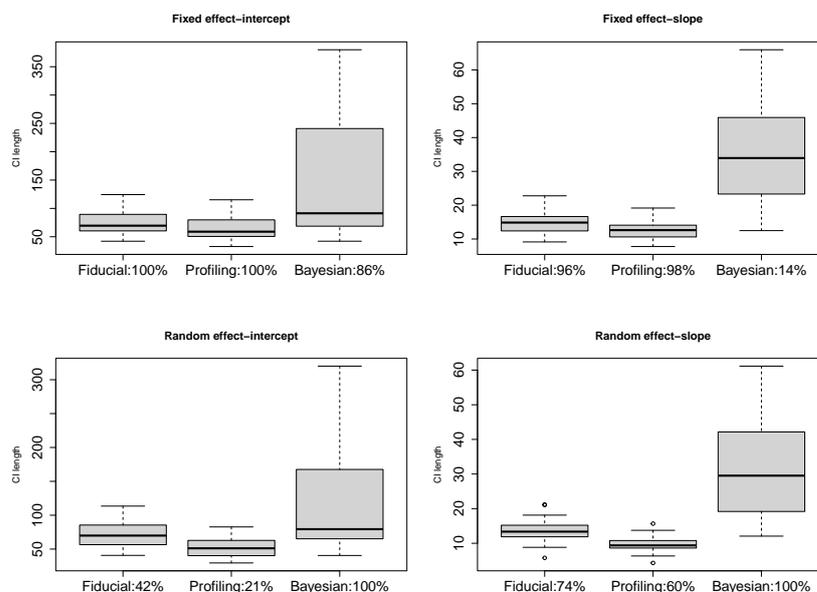


Figure 5. CI length in resampled sleep data.

6. Conclusions

We study the fiducial inference in LME with the group-level variation innovatively modeled as coefficients. We derive the fiducial density by the stationary theory of Markov chains and demonstrate its advantages in small sample sizes by using simulation and real data. Our study facilitates the inference of zero variance in random effects and reveals a deep relationship between fiducial and Bayesian inference. On one hand, it confirms the rationality of the commonly used noninformative prior in Bayesian LME. On the other, the equivalence to a Bayesian LME is only for independent random effects. It is difficult to make a direct inference of zero variance in random effects if we simply treat the final density as a posterior distribution with a flat prior. But, from the perspective of fiducial structural equations, it turns out to be feasible.

We compare four inferential methods on $n_i > q$, $n_i = q$, and $n_i < q$ respectively. Our approach outperforms others on $n_i \leq q$. This is helpful in prescreening random effects

since q can easily exceed n_i when we test if all the predictors have random effects. The minimal sample size on which our fiducial LME can work stably is much smaller than other methods. We also tried the parametric bootstrap in lme4. But, its improvement over nonparametric bootstrap is quite limited regarding coverage under the scenarios in Section 4.1.

There is also much space for improvement. The Gibbs algorithm is not fine enough. More advanced numerical algorithms can also be considered and will be studied in the future. The theoretical property of the truncated δ in the fiducial sample and the inference for correlated random effects (non-diagonal D) also need further investigation.

Author Contributions: C.Z. designed and conducted the research. X.L. provided theoretical support and edited the article. J.Y. participated in the simulation. H.G. provided resources, study materials and financial support. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Grants 11901328, 11871294 and 72171126 from the National Natural Science Foundation of China and Systems Science Plus Joint Research Program of Qingdao University (XT2024301).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Jiang, J. *Linear and Generalized Linear Mixed Models and Their Applications*; Springer: New York, NY, USA, 2007.
- Buscemi, S.; Plaia, A. Model selection in linear mixed-effect models. *Asta-Adv. Stat. Anal.* **2019**, *104*, 1–47. [[CrossRef](#)]
- Zhang, X.; Zou, G.; Liang, H. Model averaging and weight choice in linear mixed-effects models. *Biometrika* **2014**, *101*, 205–218. [[CrossRef](#)]
- Waterman, M.J.; Birch, J.B.; Abdel-Salam, A.-S.G. Several nonparametric and semiparametric approaches to linear mixed model regression. *J. Stat. Comput. Simul.* **2015**, *85*, 956–977. [[CrossRef](#)]
- Dion, C. New adaptive strategies for nonparametric estimation in linear mixed models. *J. Stat. Plan. Inference* **2014**, *150*, 30–48. [[CrossRef](#)]
- Jiang, J. *Asymptotic Analysis of Mixed Effects Models: Theory, Applications, and Open Problems*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2017.
- Breslow, N.E.; Clayton, D.G. Approximate inference in generalized linear mixed models. *J. Am. Stat. Assoc.* **1993**, *88*, 9–25. [[CrossRef](#)]
- Chung, Y.; Rabe-Hesketh, S.; Dorie, V.; Gelman, A.; Liu, J. A non-degenerate estimator for variance parameters in multilevel models via penalized likelihood estimation. *Psychometrika* **2013**, *78*, 685–709. [[CrossRef](#)] [[PubMed](#)]
- Bates, D. Fitting linear mixed-effects models using lme4. *arXiv* **2014**, arXiv:1406.5823.
- Burkner, P.-C. brms: An r package for bayesian multilevel models using stan. *J. Stat. Softw.* **2017**, *80*, 1–28. [[CrossRef](#)]
- Das, S.; Krishen, A. Some bootstrap methods in nonlinear mixed-effect models. *J. Stat. Plan. Inference* **1999**, *75*, 237–245. [[CrossRef](#)]
- Flores-Agreda, D.; Cantoni, E. Bootstrap estimation of uncertainty in prediction for generalized linear mixed models. *Comput. Stat. Data Anal.* **2019**, *130*, 1–17. [[CrossRef](#)]
- Thai, H.-T.; Mentre, F.; Holford, N.; Veyrat-Follet, C.; Comets, E. A comparison of bootstrap approaches for estimating uncertainty of parameters in linear mixed-effects models. *Pharm. Stat.* **2013**, *12*, 129–140. [[CrossRef](#)]
- Fisher, R.A. On a point raised by m. s. bartlett on fiducial probability. *Ann. Hum. Genet.* **1937**, *7*, 370–375. [[CrossRef](#)]
- Hannig, J.; Iyer, H.; Lai, R.C.S.; Lee, T.C.M. Generalized fiducial inference: A review and new results. *J. Am. Stat. Assoc.* **2016**, *111*, 1346–1361. [[CrossRef](#)]
- Martin, R.; Liu, C. Inferential models: A framework for prior-free posterior probabilistic inference. *J. Am. Stat. Assoc.* **2013**, *108*, 301–313. [[CrossRef](#)]
- Li, X.; Su, H.; Liang, H. Fiducial generalized p-values for testing zero-variance components in linear mixed-effects models. *Sci. China Math.* **2018**, *61*, 1–16. [[CrossRef](#)]
- Lidong, E.; Hannig, J.; Iyer, H.K. Fiducial intervals for variance components in an unbalanced two-component normal mixed linear model. *J. Am. Stat. Assoc.* **2008**, *103*, 854–865.
- Cisewski, J.; Hannig, J. Generalized fiducial inference for normal linear mixed models. *Ann. Stat.* **2012**, *40*, 2102–2127. [[CrossRef](#)]

20. Grimmett, G.R.; Stirzaker, D.R. *Probability and Random Processes*; Oxford University Press: Oxford, UK, 2001.
21. Betancourt, M. A Conceptual Introduction to Hamiltonian Monte Carlo. *arXiv* **2017**, arXiv:1701.02434. [[CrossRef](#)]
22. Shi, W.J.; Hannig, J.; Lai, R.C.S.; Lee, T.C.M. Covariance estimation via fiducial inference. *Stat. Theory Relat. Fields* **2021**, *5*, 316–331. [[PubMed](#)]
23. Kuznetsova, A.; Brockhoff, P.B.; Christensen, R.H.B. lmerTest package: Tests in linear mixed effects models. *J. Stat. Softw.* **2017**, *82*, 1–26. [[CrossRef](#)]
24. Belenky, G.; Wesensten, N.; Thorne, D.; Thomas, M.; Sing, H.; Redmond, D.; Russo, M.; Balkin, T. Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: A sleep dose-response study. *J. Sleep Res.* **2003**, *12*, 1–12. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.