



An adaptation of pseudo-score confidence interval method for linear mixed models

Hatice Tul Kubra Akdur*, Deniz Ozonur, Hulya Bayrak

Department of Statistics, Faculty of Science, Gazi University, Ankara, Turkey

Abstract

We adapt a confidence interval method based on a generalized Chi-Square test for fixed effect parameters of linear mixed models. Under different correlation structure of a response variable of a linear mixed model, the performances of the adaptation method pseudo-score and some of the existing confidence interval methods are investigated by carrying out a Monte Carlo simulation study. The simulation study suggests that pseudo-score method provides better results for small to moderate sample size cases with dependent observations in terms of coverage probability rates and average expected lengths. A depression study is analyzed for demonstrating the adaptation method.

Mathematics Subject Classification (2010). 62J05, 62F25, 62F05

Keywords. linear mixed models, pseudo-score, profile likelihood

1. Introduction

Linear mixed models (LMMs) including a very broad class of models are frequently utilized to fit and analyze longitudinal, clustered, or repeated measures datasets which are arisen from many areas of application such as agriculture, medical sciences, economics and engineering. An important issue in the analysis of longitudinal or repeated measures data is that the observations usually show dependency structure within-subject. In order to obtain valid inferences of model parameters of LMMs, the correlation structure of observations within-subject should be identified accurately in the first step of an analysis. An efficient way is to use extended LMMs which allow modelling within-subject correlation and variation [18]. In biopharmaceutical clinical trials, LMM is quiet pervasive method. For instance, placebo-controlled schizophrenia study was carried out by fitting three-level linear mixed model to the data in order to determine sample size for future studies in psychiatric research area [6]. In order to compare anti-depressant effects in an elderly depressed group of patients a linear mixed model including random intercept and slope was fitted to the dataset containing repeated measures of depression severity over time [12]. By allowing both fixed effects and random effects, LMMs provide very usable tools for modeling within-subject and between subject variations. A covariance structure of repeated measures can be explained as a function of time by including random effects to a linear model. In this study, main attention is the LMMs related to longitudinal

*Corresponding Author.

Email addresses: haticesenol@gazi.edu.tr (H. T. K. Akdur), denizozonur@gazi.edu.tr (D. Ozonur), hbayrak@gazi.edu.tr (H. Bayrak)

Received: 08.01.2019; Accepted: 11.03.2020

studies and repeated measures designs which include correlated measurements. Statistical inferences in LMMs usually focus on fixed effect parameters and they are implemented by using hypothesis tests or confidence intervals methods. In both cases, researchers need to find a good fixed effect estimator which has small mean square error (MSE). Constructing confidence intervals for fixed effects require estimators of fixed effect, variance estimations of random effects and error terms in LMMs. Most common variance estimations methods are ANOVA, maximum likelihood (ML), restricted (or residual) maximum likelihood (REML). ML based methods for model parameters in LMMs can be carried out by using numerical optimization algorithms such as Newton-Raphson, EM, Fisher-Scoring due to the model likelihood complexity. MSE estimator for REML of fixed effect in LMMs was suggested by expanding [13] approximation to reduce bias in small samples [14]. It has been recently pointed out that variance-covariance structures of LMMs affect power of approximate F test, Kenward-Roger or Satterthwaite approximation of degrees of freedom for hypothesis testing of fixed effect parameters based on simulation-based sample size study in LMMs [3]. Most widely used methods for constructing confidence intervals are Wald, t-naive [8], profile likelihood and Bayesian methods. Recent studies have provided that parametric bootstrap approach can be better alternative for making inference about fixed effect parameter of LMMs. Parametric bootstrap confidence interval methods (PBCI) have been investigated and reported that bootstrap-t is the best alternative among bootstrap methods [21, 22]. Also, some of existing confidence interval methods has been recently investigated for nested error regression model and it is reported that PBCI can be a good alternative for weak correlation under small samples [2]. Computational time and coverage rate accuracy of some confidence interval methods including parametric bootstrap (PB) were examined in repeated measures degradation model [16]. Even though PB method seems a good alternative, its computational time is quiet unbearable for a model with too many parameters. Furthermore, other drawback of PB method is that rely on distributional assumptions which may not be held. Many contributions to the statistics literature provide estimation theory of LMMs parameters and applications for LMMs. However, very few of these studies addresses the issue of performances of confidence interval (CI) methods under different correlation structures of the response variable of LMMs. Balancing the coverage accuracy and the narrowness of average length of a CI method is desired against the relative computational cost. Since the existing CI methods have their own drawbacks, our aim is to search for a new confidence interval method in LMMs. In this study, we have adapted pseudo-score confidence interval method for fixed effect parameters of LMMs under different correlation settings for small to moderate sample sizes by using Kenward-Roger method for obtaining standard errors of fixed effect estimators. Pseudo-score confidence interval was firstly introduced for discrete model parameters [1]. The generalized form of Pearson-Chi square statistic provided by [17] was used for constructing confidence interval for fixed effect parameters of LMMs in this study. In the simulation study, well-known covariance structures for longitudinal studies and repeated measures data-sets were exploited to model covariance structures of error terms within-subjects in order to investigate effect of covariance structures on the performances of the confidence interval methods.

In Section 2, general form of LMMs and the assumptions are provided briefly. In Section 3, some of existing confidence interval methods are summarized and pseudo-score confidence interval method for fixed effects of LMM are introduced. Section 4 includes a simulation study for scenarios of variance-covariance matrix for the response of a random slope model which is special case of LMMs. In Section 5, previously analyzed longitudinal study is used with a random slope model and the confidence intervals of fixed effects are obtained for each method.

2. Extended Gaussian linear mixed models

Extended Gaussian linear mixed model gives flexibility to basic LMM with independent errors within-subject by allowing heteroscedastic correlated errors within-subject. Extended Gaussian linear mixed model was given by [18] as follow,

$$\begin{cases} Y_i = X_i\beta + Z_i b_i + \varepsilon_i & i = 1, \dots, n \\ b_i \sim N(0, G), \quad \varepsilon_i \sim N(0, R_i), \quad \text{cov}(b_i, \varepsilon_i) = 0 \\ V(Y_i) = Z_i G Z_i' + R_i. \end{cases} \quad (2.1)$$

Equation 2.1 represents the marginal formulation of linear mixed model and n represents the number of independent experimental units (subjects or clusters). Y_i denotes a $(m \times 1)$ vector of observations within unit (for balanced sample size within each unit) i ; X_i is a $(m \times p)$ known covariates design matrix for unit i ; β is a $(p \times 1)$ unknown vector of fixed effects parameters. Z_i is a $(m \times q)$ known design matrix for unit i associated with a $(q \times 1)$ unknown vector of random effects b_i . b_i is a $(q \times 1)$ unknown random effect vector for unit i and it is distributed normally with zero mean and covariance matrix G . ε_i is an unobserved vector of random errors distributed normally with zero mean and covariance matrix R_i . The columns of Z_i can be subset of columns of X_i . This usage allows population parameters β vary randomly among independent experimental units. In this study, we focus on performances of confidence interval methods described in Section 3 when the covariance matrix of $V(Y_i)$ differs based on the covariance matrix R_i of random errors in the LMMs setting of complete and balanced sample size within-subject.

The likelihood function for the linear mixed model in Equation 2.1 is

$$L(\alpha, \beta) = \prod_{i=1}^n (2\pi)^{-m/2} |V_i(\alpha)|^{-1/2} \exp\left(-\frac{1}{2}(Y_i - X_i\beta)' V_i^{-1}(\alpha) (Y_i - X_i\beta)\right) \quad (2.2)$$

where α denote the vector of all variance and covariance parameters found in $V(Y_i) = Z_i G Z_i' + R_i$.

3. The confidence interval methods

This section describes some of the existing methods of confidence interval and the proposed method, namely pseudo-score confidence interval method, for fixed effect parameters of Gaussian linear mixed models.

3.1. Approximate Wald confidence interval method

In parameter space of fixed effect parameters β , $k = 1, \dots, p$, an approximate Wald test (z-test) and related confidence interval is obtained based on asymptotic normality of maximum likelihood estimators of fixed effect parameters. The Wald statistic is written for $k = 1, \dots, p$

$$\frac{\widehat{\beta}_k - \beta_k}{SE(\widehat{\beta}_k)} \quad (3.1)$$

where $\widehat{\beta}_k$ is a maximum likelihood estimator or a restricted maximum likelihood estimator of (β_k) and $SE(\widehat{\beta}_k)$ is a standard error of $\widehat{\beta}_k$. The standard error of $\widehat{\beta}_k$ is obtained by using negative expectation of Hessian matrix (Fisher information matrix) or observed information matrix. The Wald statistic in Equation 3.1 can be seen as a pivot in confidence interval framework. The pivot distribution is assumed to have standard normal distribution for sufficient sample sizes. Hence z-distribution percentiles are used for constructing approximate Wald confidence interval. $\widehat{\beta}_k \pm z_{1-\frac{\alpha}{2}} SE(\widehat{\beta}_k)$ denotes asymptotic

100(1 - α)% Wald confidence interval for β_k where $z_{1-\frac{\alpha}{2}}$ is the $1 - \frac{\alpha}{2}$ quantile of standard normal distribution.

3.2. Approximate t-naive confidence interval method

Under small samples when the variance of the fixed effect estimator is unknown, maximum likelihood estimator of a fixed effect may not be distributed normally. The pivot distribution in Equation 2.2 may not have standard normal. Also the estimated standard errors which underestimate the true variability in fixed effect estimators. Because they do not take into account the variability arisen from estimating random effects [23]. Student-t distribution cut-off is suggested [8] to use instead of z-distribution cut-off. Given the degrees of freedom $n * m - p$ for the pivot statistic, asymptotic 100(1 - α)% t-naive confidence interval for β_k is $\hat{\beta}_k \pm t_{n * m - p; 1 - \frac{\alpha}{2}} SE(\hat{\beta}_k)$ where $t_{n * m - p; 1 - \frac{\alpha}{2}}$ is the $1 - \frac{\alpha}{2}$ with $n * m - p$ (for balanced sample size cases) degree of freedom quantile of student-t distribution [8].

3.3. Approximate profile likelihood confidence interval method

Due to skewed shape of the log-likelihood function or excessive number of nuisance parameter in the model, Wald-type methods may fail to provide accurate interval estimation. Profile likelihood (PL) method is more preferable in this situation because it does not require asymptotic normality of maximum likelihood estimators. Assume that unknown parameter vector θ is partitioned as $\theta = (\beta_k, \delta)$ where β_k is the parameter of interest, fixed effect parameters of LMMs in this paper, and δ is the additional parameters of the model. Let $L(\beta_k, \delta)$ represents the likelihood function of the model. For β_k , profile likelihood function is $L_1(\beta_k) = \max L(\beta_k, \delta)$. $L_1(\beta_k)$ is the maximum likelihood function over the remaining parameters for each fixed value of $\beta_k = \beta_{k(0)}$. In order to obtain PL confidence interval for parameter of interest, the likelihood ratio test of a two-sided hypothesis should be inverted. Therefore, profile likelihood method still depends on asymptotic property of likelihood ratio test which is assumed to be distributed with chi-square distribution with one degree of freedom. Likelihood ratio test is used for comparing nested models with different mean structures. For a two-sided test of null hypothesis $H_0 : \beta_k = \beta_{k(0)}$, the likelihood ratio test statistic is the difference between log-likelihood of full model and reduced model:

$$2 \left[\log L(\hat{\beta}_k, \hat{\delta}) - \log L(\beta_{k(0)}, \hat{\delta}_0) \right] \quad (3.2)$$

where $\hat{\beta}_k$ and $\hat{\delta}$ are the MLEs for the full model and $\hat{\delta}_0$ is the MLE of δ for the reduced model with $\beta_k = \beta_{k(0)}$. Hence, a 100(1 - α)% CI for β_k includes the values of $\beta_{k(0)}$ where the test is non-significant at the α level as shown below in the inequality:

$$2 \left[\log L(\hat{\beta}_k, \hat{\delta}) - \log L(\beta_{k(0)}, \hat{\delta}_0) \right] < \chi_{1, \alpha}^2. \quad (3.3)$$

SAS Proc Mixed, Proc Glimmix procedures and R lme4 package provide a general framework for profile confidence interval for linear mixed model and generalized linear mixed model [4, 24]. However, in this paper, we used a simpler algorithm which in many situations, such as random slope model which is used in the simulation study, could be relatively easy to implement. Consider lower bound of confidence interval and assume that profile likelihood function is an increasing function. Firstly, ML estimates of (β_k, δ) should be obtained for full model. And then implemented by the following steps:

- (1) For computing a lower confidence limit, a lower bound can be obtained by $\beta'_k = \hat{\beta}_k - 5xSE(\hat{\beta}_k)$.
- (2) From β'_k to $\hat{\beta}_k$ define a grid of values in the range (300 points).
- (3) For each point $\beta_{k(i)}$ in the range, profile likelihood value is computed by maximizing $\log L(\beta_{k(i)}, \hat{\delta}(i))$ over δ values.

- (4) Lower bound of 95% confidence interval is taken as the smallest $\beta_{k(i)}$ value for which its profile likelihood value, $\log L_1(\beta_{k(i)})$, satisfies this inequality, $\log L_1(\beta_{k(i)}) > \log L(\hat{\beta}_k, \hat{\delta}) - 1.92$.

The upper confidence limit is similarly estimated using the algorithm. When necessary, the grid of values around lower bound or similarly for upper bound can be expanded for a greater accuracy. Step 4 and Step 5 are implemented using uniroot function of R which utilizes bi-section algorithm over the interval. More points may be required when the number of random effects and fixed effect parameters in the model are increased.

3.4. Approximate Pseudo-score confidence interval method

Assume that \hat{Y}_i is maximum likelihood fitted values of Y_i for a particular linear mixed model. Let \hat{Y}_{i0} be maximum likelihood fitted values of a reduced model or null model in which fixed effect parameter, β_k , takes a fixed value, β_{k0} , in a reasonable interval of its parameter space. This study implements generalized form of Pearson statistics suggested by [17] to construct a new confidence interval for a fixed effect for a LMM. The generalized form of Pearson statistic is given below:

$$X^2 = \sum_i \frac{(\hat{Y}_i - \hat{Y}_{i0})^2}{\text{var}(\hat{Y}_{i0})} = (\hat{Y} - \hat{Y}_0)^T \hat{V}_0^{-1} (\hat{Y} - \hat{Y}_0) \quad (3.4)$$

where $\text{var}(\hat{Y}_{i0})$ is the estimated variance of Y_i under the reduced/null model and \hat{V}_0 is the variance-covariance matrix of these values [17]. Although the Pearson statistic is used for constructing confidence interval for discrete model parameter, we adapted this generalized form for continuous response model by using (adjusted) estimated covariance matrix of linear mixed model [14] to construct an asymptotic confidence interval for a fixed effect parameter. Because Kenward and Roger stated that Wald-type test procedures and corresponding confidence intervals that are based on a conventional estimate of asymptotic covariance matrix ignore the variability in the estimate of $V(Y_i)$ and finally block-diagonal matrix of LMM, V . \hat{Y}_{i0} is obtained by fitting the reduced model under the constraint, $\beta = \beta_{k0}$. $\chi_{1,\alpha}^2$ is the $1 - \alpha$ of chi-squared distribution with one degree of freedom. An asymptotic $100(1 - \alpha)\%$ CI for any fixed effect parameter, β_k , is obtained by inverting the generalized Pearson statistic. The asymptotic CI for β_k includes such values of β_{k0} , that satisfies the following inequality,

$$X^2(\beta_{k0}) = \sum_i \frac{(\hat{Y}_i - \hat{Y}_{i0})^2}{\text{var}(\hat{Y}_{i0})} < \chi_{1,\alpha}^2. \quad (3.5)$$

4. Simulation study

In order to investigate performances of four confidence interval methods presented in Section 3 in terms of coverage probability rate and average expected length, a broad simulation study is carried out with different simulation settings. In the context of longitudinal study and repeated measures design, our simulation study is formed by differentiating within-subject covariance matrix, R_i . For all simulation settings, the following LMM with random intercept and random slope that vary randomly among subjects for the i^{th} experimental unit at the j^{th} measurement time point is taken into the simulation study,

$$Y_{ij} = \beta_0 + \beta_1 x_{ij} + \beta_2 t_{ij} + b_{1i} + b_{2i} t_{ij} + \varepsilon_{ij} \quad i = 1, \dots, n; \quad j = 1, \dots, m. \quad (4.1)$$

In the random slope model in the equation 4.1, the fixed effect parameters are β_0 , β_1 , β_2 and random effect parameters are b_{1i} , b_{2i} for $i = 1, \dots, n$. The continuous covariates were generated from the normal distribution, $x_{ij} \sim \text{Uniform}(0, 4)$. It is known that two

sample size concepts are involved in LMMs. One of them represents number of experimental unit/ cluster, n , and the other represents number of observations/ measurements within each experimental unit/ subject, m . In the simulation study, $\beta_0=0$, $\beta_1=3$, $\beta_2=0.87$ are taken as values of fixed effect parameters. The variance-covariance matrix for random effects in the model is taken as unstructured matrix, $G = Var(b_i) = \begin{bmatrix} \sigma_{b_1}^2 & \sigma_{b_1, b_2} \\ \sigma_{b_2, b_1} & \sigma_{b_2}^2 \end{bmatrix} = \begin{bmatrix} 5.55 & 0.8183 \\ 0.8183 & 8.0238 \end{bmatrix}$ for all simulation cases. We were inspired by real applications when choosing G and R_i matrices for random slope model in the simulation cases [10]. Multivariate normal distribution was used to generate random effects with zero mean and covariance matrix, G , using `mvtnorm` R package [10]. Alternatively, Gaussian copula with copula package can be also used to jointly generate b_{1i} and b_{2i} with a given correlation coefficient using the normal distribution margins instead of multivariate normal distribution because Gaussian copula describes dependence in the same way that multivariate normal distribution does. [7, 26]. Using the same mean model and random effect covariance structure, different covariance structures for random error terms, $Var(\varepsilon_i) = R_i$, and $\sigma_{\varepsilon_i}^2 = 4.41$ were taken into account in the simulation study. MLE-based estimations of fixed effect parameters, Wald CIs for β_1 , and conventional variance-covariance matrices were obtained by performing `lme4` R package (version 1.1-15) [4]. The covariance estimations based on Kenward-Roger approximation for pseudo-score method were obtained by using `pbrkrtest` R package [11]. Two-sided 95% confidence intervals were constructed for each case. The simulation cases were summarized below:

Case 1. R_i is chosen as AR(1), auto-regressive correlation is assumed between error terms within each unit. AR(1) structure provides that the correlations decline over time as the time interval between observations of repeated measures increases. Repeated measurements should be measured at equal points of time in order to use this correlation structure.

$$t_{ij} = (8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46); m = 20.$$

$$t_{ij} = (8, 10, 12, 14, 16, 18, 20, 22, 24, 26); m = 10.$$

$$t_{ij} = (8, 10, 12, 14, 16); m = 5.$$

Case 2. R_i is chosen as compound symmetry structure, same correlation is assumed between error terms within each unit. Same vectors with AR(1) covariance model for time covariates t_{ij} are used for this covariance model.

Case 3. R_i is chosen as an exponential structure, it is suitable for the situation that the measurements are not made equally spaced over time. $\{t_{i1}, \dots, t_{im}\}$ represents the time points for the i^{th} experimental unit under the constant variance assumption for random error terms within each experimental unit,

$$\begin{aligned} corr(e_{ij}, e_{ik}) &= \rho^{|t_{ij} - t_{ik}|}, \\ cov(e_{ij}, e_{ik}) &= \sigma^2 \rho^{|t_{ij} - t_{ik}|} = \sigma^2 \exp(-\theta |t_{ij} - t_{ik}|), \end{aligned} \quad (4.2)$$

where $\theta = -\log(\rho)$ for $\theta > 0$. The structure allows the correlation between error terms or any pair of measurements within each experimental unit decreases exponentially as the time intervals increase between any pair of measurements [9].

$$t_{ij} = (2, 3, 5, 8, 9, 10, 13, 15, 19, 20, 24, 26, 29, 30, 33, 35, 38, 40, 41, 44); m = 20.$$

$$t_{ij} = (2, 3, 5, 8, 9, 10, 13, 15, 19, 20); m = 10.$$

$$t_{ij} = (2, 3, 5, 8, 9); m = 5.$$

Case 4. R_i is chosen as a banded unstructured. The banded covariance structure assumes that correlation is zero after some specified time point. A banded pattern can be applied to any covariance structure. Same vectors with AR(1) covariance model for time covariate t_{ij} is used for this covariance model. For the time points $m = 5$ and banding parameter $c = 1$, unstructured banded covariance matrix example which is used in the simulation study is given below:

$$R_i = \begin{bmatrix} 2.8 & 0.1 & 0 & 0 & 0 \\ 0.1 & 0.3 & 0.2 & 0 & 0 \\ 0 & 0.2 & 0.4 & 0.1 & 0 \\ 0 & 0 & 0.1 & 0.5 & 0.23 \\ 0 & 0 & 0 & 0.23 & 0.6 \end{bmatrix} \tag{4.3}$$

The simulation results of the confidence interval methods were summarized in Table 1 to Table 4 for all cases.

Table 1. Coverage probability rates and average expected lengths of the confidence interval methods for the Case 1 for $\rho = 0.25, 0.5, 0.75$ at the 95% confidence level.

Sample Sizes			Estimated Coverage Probability Rates				Estimated Average Lengths			
ρ	n	m	Wald	Naive	Pseudo-Score	Profile	Wald	Naive	Pseudo-Score	Profile
0.25	5	5	0.8897	0.9083	0.9205	0.9936	3.1325	3.3145	3.5581	5.9136
	5	10	0.9318	0.9380	0.9626	0.9944	2.1795	2.2370	2.5840	3.4851
	5	20	0.946	0.9482	0.9538	0.9924	1.5281	1.5474	1.5932	2.2557
	10	5	0.9237	0.9309	0.9360	0.9911	2.2827	2.3430	2.4472	3.6309
	10	10	0.9404	0.9428	0.9474	0.9918	1.5685	1.5883	1.6418	2.2568
	10	20	0.9454	0.9464	0.9528	0.9911	1.0883	1.0950	1.1230	1.4784
0.5	5	5	0.9032	0.9182	0.9370	0.9921	3.0426	3.2193	3.7994	5.7295
	5	10	0.9296	0.9366	0.9398	0.9944	2.1428	2.1994	2.2912	3.4321
	5	20	0.9386	0.9414	0.9482	0.9924	1.5100	1.5291	1.5706	2.22394
	10	5	0.9272	0.9328	0.9394	0.9950	2.2279	2.2867	2.3996	3.5443
	10	10	0.9416	0.9456	0.9516	0.9926	1.5373	1.5569	1.6081	2.2141
	10	20	0.9466	0.9488	0.9534	0.9901	1.073	1.0802	1.1070	1.4650
0.75	5	5	0.8837	0.9007	0.9371	0.9931	2.9496	3.1210	3.7253	5.5352
	5	10	0.9313	0.9397	0.9459	0.9935	2.0723	2.1270	2.2128	3.3158
	5	20	0.9492	0.9518	0.9554	0.9954	1.4659	1.4843	1.5240	2.1820
	10	5	0.9256	0.9330	0.9438	0.9916	2.1742	2.2316	2.3553	3.4567
	10	10	0.9406	0.9442	0.9504	0.9915	1.4912	1.5101	1.5602	2.1558
	10	20	0.9496	0.9514	0.9544	0.9894	1.0461	1.0525	1.0779	1.4400

Table 2. Coverage probability rates and average expected lengths of the confidence interval methods for the Case 2 for $\rho = 0.25, 0.5, 0.75$ at the 95% confidence level.

Sample Sizes			Estimated Coverage Probability Rates				Estimated Average Lengths			
ρ	n	m	Wald	Naive	Pseudo-Score	Profile	Wald	Naive	Pseudo-Score	Profile
0.25	5	5	0.8869	0.9040	0.9632	0.9887	1.3844	1.4648	1.8342	2.3499
	5	10	0.9240	0.9310	0.9672	0.9876	0.9352	0.9599	1.1236	1.3820
	5	20	0.9372	0.9400	0.9488	0.9888	0.6432	0.6513	0.6665	0.9041
	10	5	0.9212	0.9280	0.9424	0.9884	1.0221	1.0491	1.1166	1.4826
	10	10	0.9430	0.9464	0.9534	0.9894	0.6730	0.6815	0.6997	0.9073
	10	20	0.9456	0.9464	0.9520	0.9874	0.4570	0.4598	0.4690	0.6003
0.5	5	5	0.8896	0.9116	0.9344	0.9858	1.1531	1.2201	1.3434	1.9381
	5	10	0.9353	0.9413	0.9705	0.9858	0.7750	0.7955	0.9238	1.1414
	5	20	0.9428	0.9450	0.9501	0.9928	0.5275	0.5342	0.5446	0.7451
	10	5	0.9252	0.9326	0.9470	0.9868	0.8564	0.8790	0.9255	1.2329
	10	10	0.9430	0.9468	0.9522	0.9904	0.5545	0.5615	0.5740	0.7509
	10	20	0.9406	0.9430	0.9478	0.9880	0.3742	0.3765	0.3836	0.4966
0.75	5	5	0.9012	0.9172	0.9644	0.9882	0.8506	0.9000	1.0810	1.4078
	5	10	0.9390	0.9442	0.9496	0.9908	0.5524	0.5669	0.5793	0.8121
	5	20	0.9472	0.9504	0.9540	0.9936	0.3729	0.3776	0.3837	0.5293
	10	5	0.9282	0.9346	0.9424	0.9862	0.6221	0.6385	0.6614	0.8886
	10	10	0.9482	0.9514	0.9554	0.9906	0.3953	0.4003	0.4075	0.5356
	10	20	0.9482	0.9514	0.9560	0.9914	0.2648	0.2664	0.2712	0.3533

Table 3. Coverage probability rates and average expected lengths of the confidence interval methods for the Case 3 for $\rho = 0.25, 0.5, 0.75$ at the 95% confidence level.

Sample Sizes			Estimated Coverage Probability Rates				Estimated Average Lengths			
ρ	n	m	Wald	Naive	Pseudo-Score	Profile	Wald	Naive	Pseudo-Score	Profile
0.25	5	5	0.8889	0.9050	0.9560	0.9860	1.4478	1.5637	1.9592	2.5233
	5	10	0.9312	0.9364	0.9688	0.9894	1.0454	1.0730	1.2596	1.5450
	5	20	0.9432	0.9452	0.9504	0.9878	0.7330	0.7423	0.7611	1.0240
	10	5	0.9288	0.9364	0.9496	0.9900	1.0944	1.1233	1.1908	1.5919
	10	10	0.9450	0.9468	0.9530	0.9902	0.7530	0.7625	0.7836	1.0130
	10	20	0.9458	0.9469	0.9580	0.9859	0.5215	0.5247	0.5356	0.6792
0.5	5	5	0.8958	0.9126	0.9655	0.9895	1.3210	1.3978	1.7132	2.2502
	5	10	0.9352	0.9418	0.9720	0.9908	0.9930	1.0192	1.1912	1.4707
	5	20	0.9420	0.9452	0.9510	0.9896	0.7129	0.7218	0.7391	1.0014
	10	5	0.9238	0.9306	0.9488	0.9880	1.0909	1.1197	1.1874	1.5863
	10	10	0.9419	0.9452	0.9524	0.9896	0.7125	0.7214	0.7400	0.9629
	10	20	0.9496	0.9504	0.9544	0.9874	0.5208	0.5240	0.5348	0.6789
0.75	5	5	0.9012	0.9159	0.9588	0.9866	1.0028	1.0611	1.2701	1.6921
	5	10	0.9374	0.9442	0.9742	0.9910	0.8527	0.8752	1.0142	1.2701
	5	20	0.9516	0.9538	0.9590	0.9942	0.6547	0.6629	0.6774	0.9265
	10	5	0.9360	0.9436	0.9506	0.9896	0.7333	0.7526	0.7781	1.0675
	10	10	0.9400	0.9426	0.9490	0.9880	0.6123	0.6200	0.6334	0.8362
	10	20	0.9510	0.9520	0.9578	0.9876	0.4668	0.4697	0.4788	0.6182

As shown in Table 1, the coverage rates of Wald and t-naive methods often approximately were around 0.90 for small sample size cases for $n = 5$ and $m = 5, 10$ for AR(1), compound symmetry and exponential covariance structures. Approximate pseudo-score method provided higher coverage probability rates than Wald and t-naive methods. Profile likelihood method provided the highest coverage probability rates due to the widest

average expected lengths. As the correlation terms of AR(1), compound symmetry and exponential structures got stronger, coverage probability rates of Wald and t-naive methods are generally slightly increased.

The performances of confidence interval methods for the compound symmetry covariance structure of random error terms showed similar behavior to the autoregressive case above. Coverage probability rates of pseudo-score method in compound symmetry structure slightly were better than those in AR(1) covariance structure for especially in small samples.

Table 4. Coverage probability rates and average expected lengths of the confidence interval methods for the Case 4 at the 95% confidence level.

Sample Sizes		Estimated Coverage Probability Rates				Estimated Average Lengths			
n	m	Wald	Naive	Pseudo-Score	Profile	Wald	Naive	Pseudo-Score	Profile
5	5	0.8909	0.9084	0.9647	0.9899	1.2590	1.3331	1.6539	2.1068
5	10	0.9300	0.9376	0.9488	0.9886	1.1920	1.2235	1.2838	1.7707
5	20	0.9444	0.9466	0.9496	0.9911	1.1355	1.1499	1.1893	1.6170
10	5	0.9280	0.9356	0.9478	0.9854	0.9839	1.0099	1.0722	1.4037
10	10	0.9410	0.9440	0.9510	0.9876	0.8587	0.8695	0.8967	1.1499
10	20	0.9377	0.9403	0.9470	0.9887	0.8112	0.8162	0.8366	1.069

For banded covariance structure of random errors, approximate pseudo-score method produced the closest coverage probability rates to the nominal confidence level even for small sample sizes.

5. Application

A psychological longitudinal study is used as an example dataset and a sketch of it is given in Table 5 to demonstrate the confidence interval methods [15]. Since we consider small-sample cases in our simulation study, we only include a subset of the dataset which contains observations of ten subjects for our analysis. This longitudinal study focuses on new treatment on the depression patients aged 18 to 75 who are not treated with any treatment method [20]. The patients are randomly allocated to two treatment groups: computer based treatment (Beating the Blues) and usual treatment group (Tau). The response variable, depression score of the patients is measured by Beck Depression Inventory [5] on five time points: prior to treatment (baseline), 2 months later, and at 1, 3, and 6 months follow-up. The effect of duration of depression is also included in the study by categorizing the duration time: if a patient had been ill for longer than six months coded duration=1 or for less than six months coded duration=0. In this paper, the patients with missing values are excluded from the dataset only the patients with complete measurements in all time points are included to the analysis. The random slope and random intercept model is fitted to the dataset. In the model, depression scores as a response variable, baseline as a continuous covariate, treatment coded 1 and 0, respectively shows Beating the Blues and Tau, as categorical covariate and duration as an another categorical covariate are taken. Here the following linear mixed model for the response is given on the patient i at time t_j for $i = 1, \dots, 10; j = 1, \dots, 4, ,$

$$Y_{ij} = \beta_0 + \beta_1 Treatment_i + \beta_2 Duration_i + \beta_3 Baseline_i + \beta_4 t_j + b_{1i} + b_{2i} t_j + \varepsilon_{ij}. \quad (5.1)$$

In order to choose an appropriate covariance structure, compound symmetry, exponential, unstructured and AR(1) covariance models were considered and fitted to the dataset by using nlme R package [19]. Since time intervals are not equal in the application as given in the sketch of the dataset, exponential covariance structure would be more preferable

Table 5. A subset of a dataset of a psychological longitudinal study

Subject	Duration	Treatment	Baseline	Month	Depr-Score
2	1	1	32	2	16
2	1	1	32	3	24
2	1	1	32	5	17
2	1	1	32	8	20
4	1	1	21	2	17
4	1	1	21	3	16
4	1	1	21	5	10
4	1	1	21	8	9
7	0	0	17	2	7
7	0	0	17	3	7
7	0	0	17	5	3
7	0	0	17	8	7
8	1	0	20	2	20
8	1	0	20	3	21
8	1	0	20	5	19
8	1	0	20	8	13

for this dataset according to Table 6. Parameter estimations, standard errors, t-values and confidence interval estimations of fixed effects under this covariance structure are summarized in Table 7.

Table 6. Likelihood-based values of Several Covariance Models

Covariance Model	AIC	BIC	log-likelihood
AR(1)	254.71	270.26	-117.35
Unstructured	259.87	283.20	-114.93
Compound Symmetry	256.57	272.12	-118.28
Exponential	254.71	270.27	-117.35

Table 7. Parameter estimations, standard errors, t-values and confidence interval estimations of fixed effects at the 95% confidence level

Parameter	β_0	β_1	β_2	β_3	β_4
Estimation(REML)	-1.704	-0.800	-2.405	2.342	0.846
Standard Error	8.123	0.581	3.424	5.033	0.404
K-R Stn. Error	9.230	0.581	3.954	5.812	0.466
Wald CI	(-17.626; 14.218)	(-1.938; 0.338)	(-9.118; 4.307)	(-7.522; 12.208)	(0.054; 1.639)
Profile CI	(-15.496; 12.338)	(-1.992; 0.392)	(-9.757; 4.946)	(-6.105; 10.790)	(0.131; 1.561)
t-naive CI	(-18.196; 14.788)	(-1.979; 0.379)	(-9.358; 4.547)	(-7.876; 12.561)	(0.026; 1.667)
Pseudo-score CI	(-16.476; 13.318)	(-1.962; 0.362)	(-9.314; 4.504)	(-7.105; 11.790)	(0.031; 1.661)

Based on the simulation study results, it was concluded that pseudo-score method was more preferable than the other methods. Therefore, pseudo-score method confidence interval results for the fixed effect parameters would be more appropriate for making inference in this dataset.

6. Discussion

This paper contributed the idea of using pseudo-score confidence interval method with Kenward-Roger variance estimation method for obtaining confidence intervals for fixed

effect parameters in linear mixed models. In this study, the covariance structure of response of a LMM was composed from two sources of variation: first variation was derived from random effects and its covariance structure was called G (unstructured) and second variation was derived from random error terms and covariance structure was modeled by R_i . Hence a hybrid covariance structure of V_i was produced for each covariance case in the simulation study. Approximate pseudo-score method was adapted for random slope model when the dependent observations within-subject was existed.

The simulation study suggested that pseudo-score confidence interval method usually provided close coverage rates to nominal 95% confidence level for small to moderate sample sizes. Wald method provides liberal results whereas profile likelihood method provides very conservative results. Pseudo-Score method generally produces better coverage rates than Wald and t-naive methods for especially small sample size cases. In cases where the covariance structure of random error terms within subject are autoregressive, compound symmetry, exponential and banded it can be said that more preferable confidence interval method is approximate pseudo-score method in terms of coverage probability rates. Pseudo-score method not only regulates the disadvantage of wider lengths of confidence intervals provided by profile likelihood but it also regulates poor coverage rates of Wald and t-naive methods. Therefore, pseudo-score confidence interval method usually balances these two disadvantages of the existing methods. There can be limitation of the simulation in this study. The proposed approach can be investigated with other covariance structures of linear mixed models such as ARMA or MA as a future study. Recently, Wu and de Leon (2014) have proposed Gaussian copula mixed models for correlated mixed response [25]. Gaussian copula mixed models are flexible in terms of modeling multiple correlated discrete or continuous responses together in a joint analysis. The proposed confidence interval method can be extended for fixed effect model parameters of Gaussian copula mixed models as an another future study.

Acknowledgment. The authors are very grateful to Dr. Alan Agresti for his valuable suggestions. The authors thank to Dr. Alex De Leon for contributing the paper with his valuable comments and suggestions. Furthermore, the part of the simulation results for the manuscript are from the author, H.T.K. AKDUR, PhD thesis that was submitted to Gazi University in July 2017. The author was supported by Scientific Council of Turkey (TUBITAK) and Gazi University by a visiting PhD student grant (2214) for studying abroad in University of Calgary under supervision of Dr. Alex De Leon during the part of this research in 2015.

References

- [1] A. Agresti and E. Ryu, *Pseudo-score confidence intervals for parameters in discrete statistical models*, *Biometrika* **97** (1), 215-222, 2010.
- [2] H.T.K. Akdur, D. Ozonur and H. Bayrak, *A Comparison of Confidence Interval Methods of Fixed Effect in Nested Error Regression Model*, *SDU J Nat Appl Sci* **20** (2), 2016.
- [3] M.K. Bahçecitapar, *Some factors affecting statistical power of approximate tests in the linear mixed model for longitudinal data*, *Comm. Statist. Simulation Comput.* **47** (1), 2018.
- [4] D. Bates, M. Machler, B. Bolker and S. Walker, *Fitting linear mixed-effects models using lme4*, arXiv preprint **1406.5823**, 2014.
- [5] A.T. Beck, R.A. Steer and G.K. Brown, *Beck depression inventory-II*, *Psychological Corporation* **78** (2), 490-498, San Antonio, TX, 1996.

- [6] D.K. Bhaumik, A. Roy, S. Aryal, K. Hur, N. Duan, S.L.T. Normand, C.H. Brown, and R.D. Gibbons, *Sample size determination for studies with repeated continuous outcomes*, *Psychiatr. Ann* **38** (12), 2008.
- [7] A. R. De Leon and K. C. Chough, *Analysis of mixed data: methods and applications*, CRC Press, 2013.
- [8] D. Eugene, *Mixed Models: Theory and Applications*, Hoboken: Wiley, 2004.
- [9] G.M. Fitzmaurice, N.M. Laird and J.H. Ware *Applied longitudinal analysis (Vol. 998)*, John Wiley and Sons, 2012.
- [10] A. Genz, F. Bretz, T. Miwa, ... , and M.T. Hothorn, *Package mvtnorm*, *J. Stat. Softw.* **11**, 950-971, 2020.
- [11] U. Halekoh and S. Højsgaard, *A kenward-roger approximation and parametric bootstrap methods for tests in linear mixed modelsthe R package pbkrtest*, *J. Stat. Softw.* **59** (9), 1-30, 2014.
- [12] P.R. Houck, S. Mazumdar, T. Koru-Sengul, G. Tang, B.H. Mulsant, B.G. Pollock and C.F. Reynolds III, *Estimating treatment effects from longitudinal clinical trial data with missing values: comparative analyses using different methods*, *Psychiatry Research* **129** (2), 209-215, 2004.
- [13] R.N. Kacker and D.A. Harville, *Approximations for standard errors of estimators of fixed and random effects in mixed linear models*, *J. Am. Stat. Assoc* **79** (388), 853-862, 1984.
- [14] M. G. Kenward and J. H. Roger, *Small sample inference for fixed effects from restricted maximum likelihood*, *Biometrics* **53** (3), 983-997, 1997.
- [15] S. Landau, *A handbook of statistical analyses using SPSS*, CRC, 2004.
- [16] J. Liu, *Statistical inference for functions of the parameters of a linear mixed model*, Unpublished doctoral dissertation, Iowa State University, 2013.
- [17] G. Lovison, *On rao score and pearson χ^2 statistics in generalized linear models*, *Statist. Papers* **46** (4), 555-574, 2005.
- [18] J.C. Pinheiro and D.M. Bates, *Mixed-effects models in S and S-Plus*, Springer. New York. US, 2000.
- [19] J.C. Pinheiro, D.M. Bates, S. DebRoy, D. Sarkar, and R. C. Team, *nlme: Linear and nonlinear mixed effects models*. R package version, **3** (1), 111.
- [20] J. Proudfoot, D. Goldberg, A. Mann, B. Everitt, I. Marks and J. Gray, *Computerized, interactive, multimedia cognitive-behavioural program for anxiety and depression in general practice*, *Psychol. Med.* **33** (2), 217-227, 2003.
- [21] V.S. Staggs, *Parametric bootstrap interval approach to inference for fixed effects in the mixed linear model*, Unpublished doctoral dissertation, University of Kansas, 2009.
- [22] V.S. Staggs, *Comparison of naive, kenwardroger, and parametric bootstrap interval approaches to small-sample inference in linear mixed models*, *Comm. Statist. Simulation Comput.* **46** (3), 1933-1943, 2017.
- [23] G. Verbeke *Linear mixed models for longitudinal data. In Linear mixed models in practice*, 631-53, Springer, 1997.
- [24] W. W. Stroup, G. A. Milliken, E. A. Claassen, R. D. Wolfinger, *SAS for Mixed Models: Introduction and Basic Applications.*, SAS Institute, 2018.
- [25] B. Wu and A.R. de Leon *Gaussian copula mixed models for clustered mixed outcomes, with application in developmental toxicology*, *J. Agric. Biol. Environ. Stat.***19** (1), 39-56, 2014.
- [26] J. Yan, *Enjoy the joy of copulas: with a package copula*, *J. Stat. Softw.* **21** (4), 1-21, 2007.