Research Article

Empirical Likelihood Estimation for Population Pharmacokinetic Study Based on Generalized Linear Model

Fang-rong Yan,[1, 2, 3] Jin-guan Lin,[1] Yuan Huang,[2, 3] Jun-lin Liu,[2, 3] and Tao Lu[2, 3]

[1] *Department of Mathematics, Southeast University, Nanjing 210096, China*
[2] *Department of Mathematics, China Pharmaceutical University, Nanjing 210009, China*
[3] *State Key Laboratory of Natural Medicines, China Pharmaceutical University, Nanjing 210009, China*

Correspondence should be addressed to Jin-guan Lin, jglin@seu.edu.cn

To obtain efficient estimation of parameters is a major objective in population pharmacokinetic study. In this paper, we propose an empirical likelihood-based method to analyze the population pharmacokinetic data based on the generalized linear model. A nonparametric version of the Wilk's theorem for the limiting distributions of the empirical likelihood ratio is derived. Simulations are conducted to demonstrate the accuracy and efficiency of empirical likelihood method. An application illustrating our methods and supporting the simulation study results is presented. The results suggest that the proposed method is feasible for population pharmacokinetic data.

1. Introduction

The parameter estimation in population pharmacokinetics (PPK) is a significant problem in clinical research. It is a novel problem in pharmacokinetic (PK) study that combines classic PK models with group statistical models. The PPK parameters, including group typical values, fixed effect parameter, interindividual variation, and intraindividual variation, which are the determinant factors of drug concentration in patients, are taken into consideration. PPK is capable of quantitatively describing the effects of different factors in drug metabolism, such as pathology, physiology, and combined medication, then providing guidance on the adjustment of therapeutic regimen, thus increasing the efficacy and safety in a new drug evaluation.

Many statistical models have been proposed to fit PPK parameters. The most popular analytic statistical model for PPK data is the random effects model proposed by Larid

and Ware [1]. Besides, the classic compartmental model and nonlinear mixed effect model proposed by Sheiner et al. in 1977 [2] are the most commonly used statistical models for PPK. And recently, generalized linear model proposed by Salway and Wakefield in 2008 [3] has received wide attention.

In fact nonlinear models are used to estimate parameters of a chosen compartmental model. This model can provide generally good results. However, a disadvantage of the nonlinear models is that it is generally more difficult to fit. From a computational point of view, one also faces the usual challenges associated with nonlinear regression such as choosing starting values, problem with convergence, nonlinear regression diagnostics, and so forth. Other nonlinear models for the analysis of PPK data see [4–6]. Salway and Wakefield [3] proposed a generalized linear model with gamma distribution to deal with PPK data. The GLM is one of the most widely used regression models for statistical analysis. The relevant theoretical investigations are rather complete and the results of parameter estimation have been the standard results in the literature and textbooks. Many methods could solve the generalized linear models, including general least square (GLS) method, quasi-likelihood method, normal theory maximum likelihood, quadratic estimating equations, extended quasi-likelihood, and general estimation equation (GEE). McCulloch [7] proposed three kinds of effective algorithm which can estimate the parameter in generalized linear model accurately. They are Monte Carlo EM (MCEM) algorithm, Monte Carlo Newton-Raphson (MCNR) algorithm, and Simulated Maximum likelihood (SML) separately. In the progress of maximum, Newton-Raphson algorithms, Fisher algorithms, EM algorithms and other basic algorithms can also be used to get the approximate solution through iteration. However, referring to multidimensional integral, it is difficult to solve a likelihood equation, and sometimes the algorithm is not convergent.

In all these methods the covariance structure is an especially crucial problem. In general, we divided variance into four models following in normal distribution: (1) the constant error model, (2) the proportional error model, (3) a combined error model, (4) an alternative combined error model. Therefore, how to choose variance structure appropriately is really confusing. Wakefield et al. [8] suggested the use of integrated nested Laplace approximations for GLMs in this context. In this case, the empirical likelihood is an alternative good choice. Empirical likelihood as a nonparametric data-driven technique was first proposed by Owen [9]. It employs the likelihood function without specifically assuming a distribution for the data and incorporates side information through constraints or a prior distribution, which maximizes the efficiency of the method. What's more, if the calculation procedure of the algorithm involves iteration, the problem of convergence may exist. Sometimes we cannot obtain convergent solution. But in principle, empirical likelihood can give estimates and confidence regions with almost any method that yields a "reasonable" set of estimating function. By "reasonable" we mean that the estimating functions have zero mean, finite variance, and either based on independent and identically distributed sampling or have higher order moments which allow use of a triangular array argument. Empirical likelihood only requires that the estimating functions have expectation zero. Hence empirical likelihood will be robust to mis-specification of the variance function, as long as the equation for the mean is correct.

In the past two decades, unique and desirable properties of the empirical likelihood method have been studied by a number of researchers. These properties include, but are not limited to, range-respecting, transformation-preserving, asymmetric confidence intervals, Bartlett correctability, and better coverage probability or shorter confidence interval. Empirical likelihood has both the easy implementation of nonparametric inference

method and the effectiveness of likelihood inference method. Compared with parametric reference method, it just needs a mild assumption of the existence of the matrix. When the sample size is small, the confidence interval we get by this method is superior to that by the asymptotic normal method. The shape and direction of the confidence interval confidence depend completely on the data itself. The estimation of variance is generally not required. Besides, empirical likelihood has unique superiority in parameter estimation and test of goodness of fit. An updated comprehensive overview of the empirical likelihood is available in Owen [10]. In a large range of situations, empirical likelihood has been shown to have properties analogous to a real likelihood, see Qin and Lawless [11, 12], Li [13], Jing [14], and others. A general framework of empirical likelihood based on estimation function is discussed in Qin and Lawless [12]. Wang et al. [15] incorporated the within-subject correlation structure of the repeated measurements into the empirical likelihood method. They also performed empirical likelihood inference on individual components of regression function.

To the best of our knowledge, there is little literature on the empirical likelihood inference for PPK study with longitudinal data. In this paper, we develop an empirical likelihood method for inference of the parametric component in generalized linear models with population pharmacokinetic data through constructing auxiliary random vectors. The model is as follows:

$$E\big(y_{ij}\big(t_{ij}\big)\big) = \mu_{ij} = D_i \ \exp\left(\beta_{i0} + \beta_{i1}t_{ij} + \frac{\beta_{i2}}{t_{ij}}\right), \qquad i = 1\ldots n, \ \ j = 0,1,2, \qquad (1.1)$$
$$\beta_{ij} = \beta_i + \eta_{ij},$$

where y_{ij} is the response variable. A few conditions needed to be assumed for the transferred additive error model to be valid to use. Assumptions are listed as follows:

$$y_{ij} > 0, \qquad \beta_{i0} > 0, \qquad \beta_{i1} < 0, \qquad \beta_{i2} < 0. \qquad (1.2)$$

These assumptions are used to insure a ascending absorption phase and a descending elimination phase. The model is first proposed by Salway and Wakefield [3].

The rest of the paper is organized as follows. In Section 2, the empirical likelihood method for PPK models is proposed. The asymptotic normality of the proposed empirical estimator and asymptotic chi-square distribution of the proposed empirical likelihood ratio are derived under some regularity assumptions. In Section 3, the simulation is presented and the results also support the theory. A real data analysis is reported in Section 4. Finally, the paper concludes with some discussions in Section 5. Proofs are given in the appendix.

2. Methods

In recent years, numerous scholars have conducted a detailed study on PPK parameter estimation and a variety of methods appeared, for example GEE (Liang and Zeger [16]). Here, we use the method of quasi-likelihood to derive estimating functions for empirical likelihood. In the application part, we will compare the effectiveness of GEE and empirical likelihood method. Note that, since maximum likelihood and quasi-likelihood coincide for all linear exponential families, the class of estimating functions derived using the former method automatically falls within that of the latter. See Morris [17] and Nelder and Pregibon [18]

Standard empirical likelihood for generalized linear models has been considered by Kolaczyk [19], based on constraints derived from the score function of the quasi-likelihood. We employ empirical likelihood and deduce as follows.

Assume there are n subjects with m observations per subject in a study. From (1.1), the generalized error model can be rewritten as

$$Y_i = U_i + \varepsilon_i, \tag{2.1}$$

where $Y_i = [\log y_{i1}(t_{i1}), \log y_{i2}(t_{i2}), \ldots, \log y_{im}(t_{im})]$ and $U_i = [\log \mu_{i1}(t_{i1}), \log \mu_{i2}(t_{i2}), \ldots, \log \mu_{im}(t_{im})]$ with $\text{var}(\varepsilon_i) = \Sigma$.

Therefore, change both side of (1.1) to logarithmic, and following Xue and Zhu [20], we can define an empirical log-likelihood ratio function as follows. Suppose that β is a $3*1$ dimensional parameters vector to be estimated. From model (1.1), an auxiliary random vector can be defined as

$$Z_i(\beta) = X_i^T V_i^{-1}(\hat{\beta})(Y_i - U_i), \tag{2.2}$$

where $X_i = (1_{1\times m}, t_i, 1/t_i)$, $t_i = (t_{i1}, t_{i2}, \ldots, t_{im})$, and $V_i(\beta)$ is an invertible covariance matrix dependent on the parameter vector β, and will be equal to $\text{cov}(Y_i)$. $V_i(\beta)$ can be estimated by using the method of moments [16]. Note that $E(Z_i(\beta)) = 0$ if β is the true parameters. Using this, an empirical log-likelihood-ratio function is defined as

$$\tilde{l}_n(\beta) = -2 \sup \left\{ \sum_{i=1}^{n} \log(np_i) \mid \sum_{i=1}^{n} p_i Z_i(\beta) = 0, \sum_{i=1}^{n} p_i = 1, p_i \geq 0, i = 1, \ldots, n \right\}. \tag{2.3}$$

By the Lagrange multiplier method, the optimal value for p_i is given by

$$p_i = \frac{1}{n} \frac{1}{1 + t^T Z_i(\beta)}, \tag{2.4}$$

where t is Lagrange multiplier

By combining (2.3) and (2.4), $\tilde{l}_n(\beta)$ can be represented as

$$\tilde{l}_n(\beta) = -2 \sum_{i=1}^{n} \log\left(1 + t^T Z_i(\beta)\right), \tag{2.5}$$

where t is determined by

$$\frac{1}{n} \sum_{i=1}^{n} \frac{Z_i(\beta)}{1 + t^T Z_i(\beta)} = 0. \tag{2.6}$$

Define $\tilde{\beta}$ as the unique maximizer of $\tilde{l}_n(\beta)$. We state the asymptotic properties of the empirical likelihood ratio. The following theorems indicate that $\tilde{\beta}$ is asymptotically normal and

the empirical log-likelihood ratio for β is a standard chi-square distribution. Some assumptions and proofs of our results are given in the appendix.

Theorem 2.1. *In addition to Lemma A.1 mentioned in the appendix, one assumes that $\partial^2 Z_i(\beta)/\partial\beta\partial\beta^T$ is continuous in a neighborhood of the true value β_0. Then if $\|\partial^2 Z_i(\beta)/\partial\beta\partial\beta^T\|$ can bounded by some integrals function $Z(x)$ in the neighborhood, then one has*

$$\sqrt{n}\left(\widetilde{\beta} - \beta_0\right) \longrightarrow N(0, V), \tag{2.7}$$

where $V = [E(\partial Z/\partial\beta)^T (EZZ^T)^{-1} E(\partial Z/\partial\beta)]^{-1}$.

Theorem 2.2. *Using the assumptions of Theorem 2.1, if β is the true value of the parameter and the last component of β is a positive number, then one has*

$$l_n(\beta) \longrightarrow \chi_p^2, \tag{2.8}$$

where $\rightarrow$ stands for convergence in distribution, and χ_p^2 is a chi-squared distribution with p degrees of freedom.

As a consequence of Theorem 2.1, confidence regions for the parameter β can be constructed by (2.3). More precisely, for any $0 \le \alpha < 1$, let c_α be the fraction quintile such that $P(\chi_p^2 \le \alpha) = 1 - \alpha$. Then $l_\alpha(\widetilde{\beta}) = \{\widetilde{\beta} : l_\alpha(\widetilde{\beta}) \le c_\alpha\}$ constitutes a confidence region for β with asymptotically correct coverage probability $1 - \alpha$.

3. Simulation

In this section some simulation results are reported to show the performance of the proposed empirical likelihood method. A two-stage generalized linear random effect model that incorporates a finite mixture model is considered in the simulation.

At the first stage

$$Y_i = D \exp\left(T^* \beta_i\right) + \varepsilon_i, \tag{3.1}$$

where $\beta_i = (\beta_{i0}, \beta_{i1}, \beta_{i2})'$.

At the second stage

$$\beta_i = \beta + \eta_i, \tag{3.2}$$

where $\beta = (\beta_0, \beta_1, \beta_2)'$ is the associated fixed effect.

To perform the simulations, 500 datasets are generated, each consisting of $n = 10, 50, 100$ subjects. The time schedule $\{t_{ij}\}$ is set as $[1\ 2\ 4\ 8\ 10\ 12\ 16\ 32]/h$. Dosage is set to $5\,mg$. The fixed effect β is set as $[0.8 - 0.04 - 0.2]$. The random effect is generated from a normal distribution as follows: $\varepsilon_i \sim_{i.i.d} N(0, \Sigma)$, and $\eta_i \sim_{i.i.d} N(0, \Omega)$. For the variance matrix Σ and Ω, we choose a diagonal matrix with elements of 0.05^2 and 0.005^2, respectively, for the reason that neither within-individual nor between-individual variation has been considered here.

Table 1: Means of parameter estimates for β with its standard derivation (SD) and mean length (ML) of the confidence region, when the estimation method are empirical likelihood (EL) and general estimation equation (GEE), respectively.

n	EL		GEE	
	Mean	CI (95% level)	Mean	CI (95% level)
	0.8001	(0.7851, 0.8151)	0.8001	(0.7834, 0.8177)
10	−0.0398	(−0.0414, −0.0382)	−0.0398	(−0.0420, −0.0378)
	−0.2003	(−0.2049, −0.1956)	−0.2003	(−0.2059, −0.1946)
	0.8001	(0.7951, 0.8051)	0.8001	(0.7841, 0.8160)
50	−0.0401	(−0.0405, −0.0395)	−0.0401	(−0.0410, −0.0389)
	−0.2000	(−0.2016, −0.1984)	−0.2000	(−0.2029, −0.1973)
	0.8001	(0.7966, 0.8036)	0.8001	(0.7908, 0.8094)
100	−0.0400	(−0.0403, −0.0397)	−0.0400	(−0.0404, −0.0395)
	−0.2000	(−0.2011, −0.1989)	−0.2000	(−0.2023, −0.1984)

The conjugate gradient method used to search for the optimal empirical likelihood parameter estimates by modifying a routing is given in Press et al. [21]. For each dataset, we compute the MELE beta for fixed effects, and the 95% intervals. To illustration the efficacy of empirical likelihood, we also analysis each dataset by GEE method. The simulation results for different sample size are summarized in Table 1, based on 500 simulations.

The results are listed in Table 1. The "Mean" is constructed according to the 500 simulated point estimators, which was computed as the mean value of 500 simulated results of $\widetilde{\beta}$. The confidence intervals are listed in the columns "CI". The lower and upper bounds of "CI" are the average of 500 simulated lower bounds, and upper bounds, respectively. From Table 1, it is fairly that, as the sample size increases, the confidence intervals become smaller, and the length of EL confidence interval shows advantage over GEE method. For the estimator $\widetilde{\beta}$, these show that the empirical likelihood method for the model can obtain accurate estimator and can construct the effective confidence intervals of β.

4. Application

We apply the empirical likelihood in the proposed generalized linear models to the theophylline data from bronchial asthma study. They were originally analyzed in Upton et al. [22] and are available from the Resource Facility for Population Kinetics. Initially this data set was used for Bayesian analysis of linear and nonlinear population models by using the Gibbs sampler. Details about the related design, methods, and medical implications of the asthma study have been described by Upton et al. [22]. Although all the individuals were scheduled to have clinical visits followed the time schedule, due to the various reasons, some individuals missed scheduled visits, which resulted in unequal number of measurements and different measurement times across individuals. It has been also used in Bayesian analysis of linear and nonlinear population models employing Gibbs sampler with normal errors. Wakefield et al. [8] suggested the generalized linear mixed model with the gamma error utilize for it. Based on this model, he had obtained a Bayesian result better than the compartment model result [8]. A Bayesian analysis of the dataset was completed by Salway and Wakefied [3] and Yan et al. [23].

Table 2: Comparing parameter estimates for theophylline data in terms of accuracy and efficiency. EL: empirical likelihood; GEE: general estimation equation; CP: compartment model.

	EL	95% CI	GEE	95% CI	CP	95% CI
t_{max}	2.1676	(2.03, 2.30)	2.1676	(2.02, 2.38)	2.26	(1.73, 2.96)
c_{max}	8.3182	(7.53, 9.24)	8.3182	(7.52, 9.28)	9.06	(8.15, 10.1)
$t_{1/2}$	7.3862	(7.29, 7.97)	7.3862	(7.00, 7.96)	7.75	(7.06, 8.51)
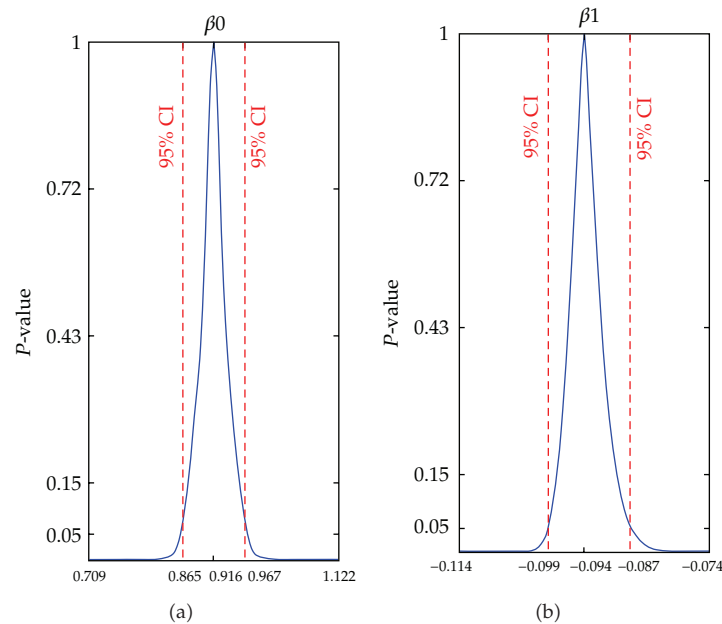
Figure 1: The estimated value of β and its confidence region.

Here we applied our method to the data of 12 subjects who were administered an oral dose of the antiasthmatic agent theophylline, with 10 concentration measurements obtained from each individual over the subsequent 25 hours. Confidence regions for each parameters are shown in Figure 1, respectively. The ranges of the confidence regions are small, which suggests a good fit for the parameters.

A comparable representation with the GEE methods completed for the derived parameters can be seen in Table 2. We compared the proposed empirical likelihood method with GEE method and the typically compartment model results. It concludes that these methods are comparable in terms of accuracy and efficiency. Table 2 indicates that EL-based confidence region is smaller than the GEE-based region and compartment results. Moreover, these findings are similar to the ones of Salway and Wakefield [3] and Yan et al. [23].

5. Conclusions

This paper shows that empirical likelihood is an effective method for analyzing the PPK model. The empirical log-likelihood ratios are proven to be asymptotically chi-squared, and then the corresponding confidence regions for the parameters of interest are constructed. A simulation study is used to investigate the performance of the empirical likelihood method. The results show that the empirical likelihood could obtain accurate estimator and construct

the effective confidence regions. We also analyze the theophylline data from the Upton et al. [22], and the results are similar and comparable to the ones of Wakefield et al. [8] and Yan et al. [23].

At last, we have to mention the two disadvantages of this method. For one, the likelihood function in the empirical likelihood method is completely determined by the data, while avoiding the estimation of the variance, so the estimation of individual effect may not be obtained. But actually, it is an important parameter in PPK study and directly impacts doctor's recommendation of individual administration. The following two steps are advised to solve this problem. First, estimate the parameters in the model through empirical likelihood function and then calculate the random effect with the estimators.

For another, in this paper, we did not consider the correlation structure of within-subject observation. However, in the current literature on empirical likelihood methods, the potential correlation structure of within-subject observation was first taken into consideration by Wang et al. [15]. They also demonstrated that it was easy to incorporate the correlation structure. It offered a promising direction to study this problem. Furthermore, we have only accomplished the complete data set at present and incomplete data will be further studied in future. From the application point of view, missing data and censoring data are worth being studied.

Appendix

Lemma A.1. *Assume that $E[(Z_i(\beta)Z_i^T(\beta))]$ is positive definite, $(\partial Z_i(\beta))/\partial\beta$ is continuous in a neighborhood if the true value β_0, $\|\partial Z_i(\beta)/\partial\beta\|$和$\|Z_i(\beta)\|^3$ are bounded by some integrable function $Z(x)$ in this neighborhood, and the rank of $E[\partial Z(\beta)/\partial\beta]$ is p. Then, as $n \to \infty$, with probability 1 $l(\beta)$ attains its minimum value at some point $\widetilde{\beta}$ in the interior of the ball $\|\beta - \beta_0\| \leq n^{-1/3}$, and $\widetilde{\theta}$ and $\widetilde{t}(\widetilde{\beta})$ satisfy*

$$Q_{1n}\left(\widetilde{\beta},\widetilde{t}\right) = 0, \qquad Q_{2n}\left(\widetilde{\beta},\widetilde{t}\right) = 0, \tag{A.1}$$

where

$$Q_{1n}(\beta,t) = \frac{1}{n}\sum_{i=1}^{n}\frac{Z_i(\beta)}{1+t(\beta)^T Z_i(\beta)}, \tag{A.2}$$

$$Q_{2n}(\beta,t) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{1+t(\beta)^T Z_i(\beta)}\left(\frac{\partial Z_i(\beta)}{\partial\beta}\right)^T t. \tag{A.3}$$

Proof of Lemma A.1. Denote $\beta = \beta_0 + un^{-1/3}$, for $\beta \in \{\|\beta - \beta_0\|\} = n^{-1/3}$, where $\|u\| = 1$. First, we give a lower bound for $l(\beta)$ on the surface of the ball. Similar to the proof of Owen (1990) [24], when $E\|Z(\beta)\|^3 < \infty$ and $\|\beta - \beta_0\| < n^{-1/3}$, we have

$$\begin{aligned} t(\beta) &= \left[\frac{1}{n}\sum_{i=1}^{n}Z_i(\beta)Z_i^T(\beta)\right]^{-1}\left[\frac{1}{n}\sum_{i=1}^{n}Z_i(\beta)\right] + o\left(n^{-1/3}\right)(\text{a.s.}) \\ &= O\left(n^{-1/3}\right)(\text{a.s.}) \end{aligned} \tag{A.4}$$

uniformly about $\beta \in \{\beta \mid \|\beta - \beta_0\| \leq n^{-1/3}\}$.

By this and Taylor expansion, we have (uniformly for u)

$$
\begin{aligned}
l(\beta) &= \sum_{i=1}^{n} t^{T}(\beta) Z_{i}(\beta) - \frac{1}{2} \sum_{i=1}^{n} \left[t^{T}(\beta) Z_{i}(\beta) \right]^{2} + o\left(n^{-1/3}\right) \text{(a.s.)} \\
&= \frac{n}{2} \left[\frac{1}{n} \sum_{i=1}^{n} Z_{i}(\beta) \right]^{T} \left[\frac{1}{n} \sum_{i=1}^{n} Z_{i}(\beta) Z_{i}^{T}(\beta) \right]^{-1} \left[\frac{1}{n} \sum_{i=1}^{n} Z_{i}(\beta) \right] + o\left(n^{-1/3}\right) \text{(a.s.)} \\
&= \frac{n}{2} \left[\frac{1}{n} \sum_{i=1}^{n} Z_{i}(\beta) + \frac{1}{n} \sum_{i=1}^{n} \frac{\partial Z_{i}(\beta_{0})}{\partial} \beta u n^{-1/3} \right]^{T} \left[\frac{1}{n} \sum_{i=1}^{n} Z_{i}(\beta) Z_{i}^{T}(\beta) \right]^{-1} \\
&\quad \times \frac{n}{2} \left[\frac{1}{n} \sum_{i=1}^{n} Z_{i}(\beta) + \frac{1}{n} \sum_{i=1}^{n} \frac{\partial Z_{i}(\beta_{0})}{\partial \beta} u n^{-1/3} \right] + o\left(n^{-1/3}\right) \text{(a.s.)} \\
&= \frac{n}{2} \left[O\left(n^{-1/2} (\log \log n)^{1/2}\right) + E\left(\frac{\partial Z_{i}(\beta_{0})}{\partial \beta} \right) u n^{-1/3} \right]^{T} \left[E\left(Z_{i}(\beta_{0}) Z_{i}^{T}(\beta_{0}) \right) \right]^{-1} \\
&\quad \times \left[O\left(n^{-1/2} (\log \log n)^{1/2}\right) + E\left(\frac{\partial Z_{i}(\beta_{0})}{\partial \beta} u n^{-1/3} \right) \right] + o\left(n^{-1/3}\right) \text{(a.s.)} \\
&\geq (c - \varepsilon) n^{1/3}, \text{(a.s.)},
\end{aligned}
\tag{A.5}
$$

where, $c - \varepsilon > 0$, and c is the smallest eigenvalue of

$$
E\left(\frac{\partial Z_{i}(\beta_{0})}{\partial \beta} \right)^{T} \left[E\left(Z_{i}(\beta_{0}) Z_{i}^{T}(\beta_{0}) \right) \right]^{-1} E\left(\frac{\partial Z_{i}(\beta_{0})}{\partial \beta} \right).
\tag{A.6}
$$

Similarly,

$$
\begin{aligned}
l(\beta_{0}) &= \frac{n}{2} \left[\frac{1}{n} \sum_{i=1}^{n} Z_{i}(\beta_{0}) \right]^{T} \left[\frac{1}{n} \sum_{i=1}^{n} Z_{i}(\beta_{0}) Z_{i}^{T}(\beta_{0}) \right]^{-1} \\
&\quad \times \left[\frac{1}{n} \sum_{i=1}^{n} Z_{i}(\beta_{0}) \right] + o(1) \text{(a.s.)} \\
&= O(\log \log n), \text{(a.s.)}.
\end{aligned}
\tag{A.7}
$$

Since $l(\beta)$ is a continuous function about β as β belongs to the ball $\|\beta - \beta_{0}\| < n^{-1/3}$, $l(\beta)$ has minimum value in the interior if this ball, and $\tilde{\beta}$ satisfies

$$
\begin{aligned}
\left. \frac{\partial l(\beta)}{\partial(\beta)} \right|_{\theta=\tilde{\theta}} &= \sum_{i=1}^{n} \left. \frac{\left(\partial t^{T}(\beta)/\partial \beta \right) Z_{i}(\beta) + \left(\partial Z_{i}(\beta)/\partial \beta \right)^{T} t(\beta)}{1 + t^{T}(\beta) Z_{i}(\beta)} \right|_{\beta=\tilde{\beta}} \\
&= \sum_{i=1}^{n} \left. \frac{1}{1 + t^{T}(\beta) Z_{i}(\beta)} \left(\frac{\partial Z_{i}(\beta)}{\partial \beta} \right)^{T} t(\theta) \right|_{\beta=\tilde{\beta}} \\
&= 0.
\end{aligned}
\tag{A.8}
$$

$\square$

Theorem A.2. *In addition to Lemma A.1. one assumes that $\partial^2 Z_i(\beta)/\partial\beta\partial\beta^T$ is continuous in β in a neighborhood of the true value β_0. Then if $\|\partial^2 Z_i(\beta)/\partial\beta\partial\beta^T\|$ can bounded by some integrable function $Z(x)$ in the neighborhood, then*

$$\sqrt{n}\left(\widetilde{\beta} - \beta_0\right) \longrightarrow N(0, V), \tag{A.9}$$

where $V = [E(\partial Z/\partial\beta)^T(EZZ^T)^{-1}E(\partial Z/\partial\beta)]^{-1}$.

Proof of Theorem A.2. Taking derivatives about β, t^T, we have

$$\frac{\partial Q_{1n}(\beta,0)}{\partial\beta} = \frac{\partial Z_i(\beta)}{\partial\beta}, \qquad \frac{\partial Q_{1n}(\beta,0)}{\partial t^T} = -\frac{1}{n}\sum_{i=1}^{n} Z_i(\beta)Z_i^T(\beta),$$

$$\frac{\partial Q_{2n}(\beta,0)}{\partial\beta} = 0, \qquad \frac{\partial Q_{2n}(\beta,0)}{\partial t^T} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\partial Z_i(\beta)}{\partial\beta}\right)^T. \tag{A.10}$$

Expanding $Q_{1n}(\widetilde{\beta},\widetilde{t}), Q_{2n}(\widetilde{\beta},\widetilde{t})$ at $(\beta_0, 0)$, by the condition of the Theorem 2.1 and Lemma A.1, we have

$$0 = Q_{1n}\left(\widetilde{\beta},\widetilde{t}\right)$$

$$= Q_{1n}(\beta_0, 0) + \frac{\partial Q_{1n}(\beta_0, 0)}{\partial\beta}\left(\widetilde{\beta} - \beta_0\right) + \frac{\partial Q_{1n}(\beta_0, 0)}{\partial t^T}\left(\widetilde{t} - 0\right) + o_p(\delta_n),$$

$$0 = Q_{2n}\left(\widetilde{\beta},\widetilde{t}\right)$$

$$= Q_{2n}(\beta_0, 0) + \frac{\partial Q_{2n}(\beta_0, 0)}{\partial\beta}\left(\widetilde{\beta} - \beta_0\right) + \frac{\partial Q_{2n}(\beta_0, 0)}{\partial t^T}\left(\widetilde{t} - 0\right) + o_p(\delta_n), \tag{A.11}$$

where $\delta_n = \|\widetilde{\beta} - \beta_0\| + \|\widetilde{t}\|$.

We have

$$\begin{pmatrix} \widetilde{t} \\ \widetilde{\beta} - \beta_0 \end{pmatrix} = S_n^{-1}\begin{pmatrix} -Q_{1n}(\beta_0, 0) + o_p(\delta_n) \\ o_p(\delta_n) \end{pmatrix}, \tag{A.12}$$

where

$$S_n = \begin{pmatrix} \dfrac{\partial Q_{1n}}{\partial t^T} & \dfrac{\partial Q_{1n}}{\partial\beta} \\ \dfrac{\partial Q_{2n}}{\partial t^T} & 0 \end{pmatrix}_{(\beta_0, 0)} \longrightarrow \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & 0 \end{pmatrix} = \begin{pmatrix} -E(ZZ^T) & E\left(\dfrac{\partial Z}{\partial\beta}\right) \\ E\left(\dfrac{\partial Z}{\partial\beta}\right)^T & 0 \end{pmatrix}. \tag{A.13}$$

From this and $Q_{1n}(\beta_0, 0) = \dfrac{1}{n} \sum_{n=1}^{n} Z_i(\beta) = O_p(n^{-1/2})$, we know that $\delta_n = O_p(n^{-1/2})$. Easily we have

$$\sqrt{n}\left(\widetilde{\beta} - \beta_0\right) = S_{22.1}^{-1} S_{21} S_{11}^{-1} \sqrt{n} Q_{1n}(\beta_0, 0) + o_p(1) \longrightarrow N(0, V), \tag{A.14}$$

where, $V = S_{22.1}^{-1} = \{E(\partial Z/\partial \beta)^T E(ZZ^T)^{-1} E(\partial Z/\partial \beta)\}^{-1}$. $\qquad \square$

Theorem A.3. *If β is the true value of the parameter and the last component of β is a positive number, similar to Qin and Lawless [11], then one has*

$$l_n(\beta) \longrightarrow \chi_p^2, \tag{A.15}$$

where $\longrightarrow$ stands for convergence in distribution, and χ_p^2 is a chi-squared distribution with p degrees of freedom.

Proof of Theorem A.3. The log-empirical likelihood ratio test statistics is

$$W(\beta_0) = 2\left\{ \sum_{i=1}^{n} \log\left[1 + t^T Z_i(\beta_0)\right] - \sum_{i=1}^{n} \log\left[1 + \widetilde{t}^T Z_i\left(\widetilde{\beta}\right)\right] \right\}. \tag{A.16}$$

Note that

$$l\left(\widetilde{\beta}, \widetilde{t}\right) = \sum_{i=1}^{n} \log\left[1 + \widetilde{t}^T Z_i\left(\widetilde{\beta}\right)\right] = -\frac{n}{2} Q_{1n}^T(\beta_0, 0) A Q_{1n}(\beta_0, 0) + o_p(1), \tag{A.17}$$

where $A = S_{11}^{-1}\{I + S_{12} S_{22.1}^{-1} S_{21} S_{11}^{-1}\}$.

Also under H_0:

$$\frac{1}{n} \sum_{i=1}^{n} \frac{1}{1 + t_0^T Z_i(\beta_0)} Z_i(\beta_0) = 0 \longrightarrow t_0 = -S_{11}^{-1} Q_{1n}(\beta_0, 0) + o_p(1), \tag{A.18}$$

and

$$\frac{1}{n} \sum_{i=1}^{n} \log\left[1 + t_0^T Z_i(\beta_0)\right] = -\frac{n}{2} Q_{1n}^T(\beta_0, 0) S_{11}^{-1} Q_{1n}(\beta_0, 0) + o_p(1). \tag{A.19}$$

Thus

$$
\begin{aligned}
W(\beta_0) &= nQ_{1n}^T(\beta_0, 0)\left(A - S_{11}^{-1}\right)Q_{1n}(\beta_0, 0) + o_p(1) \\
&= nQ_{1n}^T(\beta_0, 0)S_{11}^{-1}S_{12}S_{22.1}^{-1}S_{21}S_{11}^{-1}Q_{1n}(\beta_0, 0) + o_p(1) \\
&= \left[\left(-S_{11}^{-1}\right)^{-1/2}\sqrt{nQ_{1n}}(\beta_0, 0)\right]^T\left[\left(-S_{11}^{-1}\right)^{-1/2}S_{12}S_{22.1}^{-1}S_{21}\left(-S_{11}^{-1}\right)^{-1/2}\right] \\
&\quad \times \left[\left(-S_{11}^{-1}\right)^{-1/2}\sqrt{n}Q_{1n}(\beta_0, 0)\right] + o_p(1).
\end{aligned}
\tag{A.20}
$$

Note that $\left(-S_{11}^{-1}\right)^{-1/2}\sqrt{nQ_{1n}}(\beta_0, 0)$ converges to a standard multivariate normal distribution and that $\left(-S_{11}^{-1}\right)^{-1/2}S_{12}S_{22.1}^{-1}S_{21}(-S_{11}^{-1})^{-1/2}$ is symmetric and idempotent, with trace equal to p. Hence the empirical likelihood ratio statistic $W(\beta_0)$ converges to $\chi^2_{(p)}$. □

Acknowledgments

References

[1] N. M. Laird and J. H. Ware, "Random-effects models for longitudinal data," *Biometrics*, vol. 38, no. 4, pp. 963–974, 1982.

[2] L. B. Sheiner, B. Rosenberg, and V. V. Marathe, "Estimation of population characteristics of pharmacokinetic parameters from routine clinical data," *Journal of Pharmacokinetics and Biopharmaceutics*, vol. 5, no. 5, pp. 445–479, 1977.

[3] R. Salway and J. Wakefield, "Gamma generalized linear models for pharmacokinetic data," *Biometrics*, vol. 64, no. 2, pp. 620–626, 2008.

[4] J. Wakefield, "Non-linear regression modeling," in *Methods and Models in Statistics*, pp. 119–153, Imperial College Press, 2004.

[5] E. F. Vonesh, "Non-linear models for the analysis of longitudinal data," *Statistics in Medicine*, vol. 11, no. 14-15, pp. 1929–1954, 1992.

[6] M. Davidian and D. Giltinan, "Nonlinear models for repeated measurement data," *Journal of the American Statistical Association*, vol. 5, no. 13, pp. 1462–1463, 1995.

[7] C. E. McCulloch, "Maximum likelihood algorithms for generalized linear mixed models," *Journal of the American Statistical Association*, vol. 92, no. 437, pp. 162–170, 1997.

[8] J. Wakefield, A. F. M. Smith, A. R. Poon, and A. E. Gelfand, "Bayesian analysis of linear and non-linear population models by using the gibbs sampler," *Journal of the Royal Statistical Society*, vol. 43, no. 1, pp. 201–221, 1994.

[9] A. B. Owen, "Empirical likelihood ratio confidence intervals for a single functional," *Biometrika*, vol. 75, no. 2, pp. 237–249, 1988.

[10] A. B. Owen, *Empirical Likelihood*, Chapman & Hall, New York, NY, USA, 2001.

[11] J. Qin and J. Lawless, "Empirical likelihood and general estimating equations," *The Annals of Statistics*, vol. 22, no. 1, pp. 300–325, 1994.

[12] J. Qin and J. Lawless, "Estimating equations, empirical likelihood and constraints on parameters," *The Canadian Journal of Statistics*, vol. 23, no. 2, pp. 145–159, 1995.

[13] G. Li, "Nonparametric likelihood ratio estimation of probabilities for truncated data," *Journal of the American Statistical Association*, vol. 90, no. 431, pp. 997–1003, 1995.

[14] B. Y. Jing, "Two-sample empirical likelihood method," *Statistics & Probability Letters*, vol. 24, no. 4, pp. 315–319, 1995.

[15] S. J. Wang, L. F. Qian, and R. J. Carroll, "Generalized empirical likelihood methods for analyzing longitudinal data," *Biometrika*, vol. 97, no. 1, pp. 79–93, 2010.

[16] K. Y. Liang and S. L. Zeger, "Longitudinal data analysis using generalized linear models," *Biometrika*, vol. 73, no. 1, pp. 13–22, 1986.

[17] C. N. Morris, "Natural exponential families with quadratic variance functions," *The Annals of Statistics*, vol. 10, no. 1, pp. 65–80, 1982.

[18] J. A. Nelder and D. Pregibon, "An extended quasilikelihood function," *Biometrika*, vol. 74, no. 2, pp. 221–232, 1987.

[19] E. D. Kolaczyk, "Empirical likelihood for generalized linear models," *Statistica Sinica*, vol. 4, no. 1, pp. 199–218, 1994.

[20] L. G. Xue and L. X. Zhu, "Empirical likelihood semiparametric regression analysis for longitudinal data," *Biometrika*, vol. 94, no. 4, pp. 921–937, 2007.

[21] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, Cambridge University Press, New York, NY, USA, 2nd edition, 1992.

[22] R. A. Upton, J. F. Thiercelin, and T. W. Guentert, "Intraindividual variability in theophylline pharmacokinetics: statistical verification in 39 of 60 healthy young adults," *Journal of Pharmacokinetics and Biopharmaceutics*, vol. 10, no. 2, pp. 123–134, 1982.

[23] F. R. Yan, Y. Huang, and T. Lu, "Bayesian inferencefor generalized linear mixed modeling based on the multivariate t distribution in population pharmacokinetic study," *Journal of Pharmacokinetics and Pharmacodynamics*. In press.

[24] A. B. Owen, "Empirical likelihood confidence regions," *Annals of Statistics*, vol. 19, no. 1, pp. 90–120, 1990.