Metrika manuscript No.
(will be inserted by the editor)

A new bounded log-linear regression model

HaiYing Wang · Nancy Flournoy · Eloi Kpamegan

Received: date / Accepted: date

Abstract In this paper we introduce a new regression model in which the response variable is bounded by two unknown parameters. A special case is a bounded alternative to the four parameter logistic model. The four parameter model which has unbounded responses is widely used, for instance, in bioassays, nutrition, genetics, calibration and agriculture. In reality, the responses are often bounded although the bounds may be unknown, and in that situation, our model reflects the data-generating mechanism better.

Complications arise for the new model, however, because the likelihood function is unbounded, and the global maximizers are not consistent estimators of unknown parameters. Although the two sample extremes, the smallest and the largest observations, are consistent estimators for the two unknown boundaries, they have slow convergence rate and are asymptotically biased. Improved estimators are developed by correcting for the asymptotic biases of the two sample extremes in the one sample case; but even these consistent estimators do not obtain the optimal convergence rate. To obtain efficient estimation, we suggest using the local maximizers of the likelihood function, i.e., the solution to the likelihood equations. We prove that, with probability approaching one as the sample size goes to infinity, there exists a solution to the likelihood equation that is consistent at the rate of the square root of the sample size and it is asymptotically normally distributed.

Keywords Asymptotics; Consistency; Linear model; Logistic model; Maximum likelihood estimation; Parameter dependent support.

HaiYing Wang
Department of Statistics, University of Missouri, 146 Middlebush Hall, Columbia, MO 65203 USA, E-mail: hwzq7@mail.missouri.edu

Nancy Flournoy
Department of Statistics, University of Missouri, 146 Middlebush Hall, Columbia, MO 65203 USA, E-mail: flournoyn@missouri.edu

Eloi Kpamegan
Clinical & Nonclinical Biostatistics, Novavax, 9920 Belward Campus Dr, Rockville, MD 20850 USA, E-mail: ekpamegan@Novavax.com

1 Introduction

Consider the following regression model

$$\log\left(\frac{B-Y}{Y-A}\right)=\mathbf{x}^{\mathrm{T}}\boldsymbol{\beta}+\varepsilon, \tag{1}$$

where Y is the response variable and A and B are the two unknown boundaries of the responses; $\mathbf{x}$ denotes a p dimensional covariate vector; $\boldsymbol{\beta}$ is an unknown p dimensional regression coefficient vector; ε is the error term having a normal distribution with mean 0 and variance σ^2. This model does not belong to the class of generalized linear models because the transformation of Y contains unknown parameters. It also does not belong to the class of non-regular regression model studied by Smith (1994) who focused on linear regression models with error terms having parameter dependent support. We call this model the bounded log-linear regression model.

Let $\boldsymbol{\theta}=(\boldsymbol{\beta}^{\mathrm{T}},\sigma,A,B)^{\mathrm{T}}$ and $\boldsymbol{\theta}_0$ be the true value of $\boldsymbol{\theta}$. The likelihood function of $\boldsymbol{\theta}$ based on an independent random sample $\{(\mathbf{x}_i,Y_i),i=1,...,n\}$ from model (1) is

$$\begin{aligned}L_n(\boldsymbol{\theta})=&\frac{(B-A)^nI(A<Y_{(1)}<Y_{(n)}<B)}{(2\pi)^{\frac{n}{2}}\sigma^n\prod_{i=1}^n(B-Y_i)(Y_i-A)}\times\\&\exp\left[-\frac{\sum_{i=1}^n\left\{\log\left(\frac{B-Y_i}{Y_i-A}\right)-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right\}^2}{2\sigma^2}\right],\end{aligned} \tag{2}$$

where $I(\cdot)$ is the indicator function. This likelihood function is unbounded and may become infinite along some paths; for example, let $\sigma^2=\sum_{i=1}^n\left\{\log\left(\frac{B-Y_i}{Y_i-A}\right)\right\}^2$ and $\boldsymbol{\beta}=\mathbf{0}$, a p-dimensional vector of zeros; then $\sigma^n\prod_{i=1}^n(B-Y_i)(Y_i-A)$ goes to 0 if A approaches $Y_{(1)}$ from the left or B approaches $Y_{(n)}$ from the right. So the likelihood function in (2) goes to infinity as $\boldsymbol{\theta}$ goes to $\left(\mathbf{0}^{\mathrm{T}},+\infty,Y_{(1)},Y_{(n)}\right)^{\mathrm{T}}$ along some paths. Thus the global maximizer of the likelihood function is not a consistent estimator. If $\boldsymbol{\beta}$ and σ are known, the likelihood function is bounded because it is continuous and goes to 0 as A approaches $Y_{(1)}$ or $-\infty$, or as B approaches $Y_{(n)}$ or ∞.

We are motivated by a special case of model (1) in which $\boldsymbol{\beta}=(a,b)^{\mathrm{T}}$:

$$\log\left(\frac{B-Y}{Y-A}\right)=a+bx+\varepsilon. \tag{3}$$

Model (3) can be presented as

$$Y=B-\frac{B-A}{1+e^{-(a+bx+\varepsilon)}},$$

by which one can see its connection to the four parameter logistic model,

$$Y=B-\frac{B-A}{1+e^{-(a+bx)}}+\varepsilon. \tag{4}$$

The four parameter logistic model for continuous responses is also called the $E_{\max}$ model, and it is widely used for curve-fitting, for instance, in bioassays, nutrition, genetics, calibration and agriculture. See, for example, DeLean *et al.* (1978), V∅lund (1978), Holford and Sheiner (1981), Ratkowsky and Reedy (1986), Finke *et al.* (1987), Finke *et al.* (1989), Ernst *et al.* (1997), Triantafilis *et al.* (2000), Nix and Wild (2001), Menon and Bhandarkar (2004), MacDougall (2006), Dragalin *et al.* (2007), Vedenov and Pesti (2008), Sebaugh (2011), Feng *et al.* (2011) and the references therein. The $E(Y|x)$ of model (4) is often used in phase I clinical trials to model the mean response for Bernoulli random variables. The applications considered here, and in the aforementioned references, focus on continuous random variables. We take ε to have a normal distribution $N(0, \sigma^2)$ for models (3) and (4).

A drawback of the four parameter logistic model is that parameters A and B are often interpreted as the minimum and maximum of possible responses, although model (4) allows the responses to be unbounded. Another inadequacy of model (4) is that the responses Y have the same variance for all possible values of the covariate x, which is often violated in practice. Leonov and Miller (2009) tackled this problem by letting the variance of the model error depend on the covariate, but the range of possible responses remained to be unbounded. Our model (3) has bounded responses and the distribution of the response for a given dose is skewed analogous to a beta distribution. Figure 1 displays simulated data from models (3) and (4). Note that observations from model (4) may fall far outside the two hypothetical bounds, while data from model (3) always stays between the two boundaries. Additionally, data from model (4) still has large variation in the two tails, whereas variation in the tails is skewed and very small for model (3), and this scenario is observed frequently in real data.

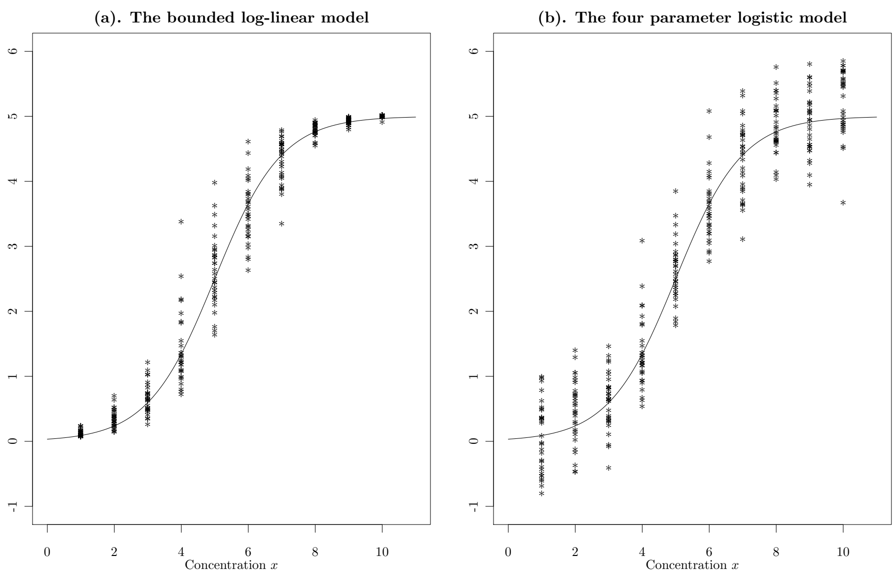

Fig. 1 Simulated data from models (3) and (4) with $a = 5$, $b = -1$, $\sigma = 0.5$, $A = 0$ and $B = 5$. The sigmoid curve is the median response for model (3) while it is the mean response for model (4).

A special case of model (3) is, by setting $b = 0$ and replacing a by μ,

$$\log\left(\frac{B-Y}{Y-A}\right) = Z, \tag{5}$$

where $Z \sim N(\mu, \sigma^2)$, and μ, σ, A and B are unknown parameters. Model (5) has some similarities to the three parameter log-normal distribution, in which

$$\log(Y - A) = Z \sim N(\mu, \sigma^2),$$

and which also has an unbounded likelihood function (see Hill, 1963). Although the three parameter log-normal distribution has been studied by many, including Cohen (1951), Hill (1963), Harter and Moore (1966), Giesbrecht and Kempthorne (1976) and Cohen and Whitten (1980), the theoretical properties of the proposed methods were not addressed rigorously in these papers. Cheng and Amin (1983) proposed an estimation method called maximum product spacings and proved the asymptotic normality of the proposed estimator and the local maximum likelihood estimator for the log-normal distribution. However, a rigorous proof for the consistency of the local maximum likelihood estimator was not provided.

There is a large body of literature rigorously developing methods of statistical inference for models with parameter-dependent support, including Woodroofe (1972), Weiss and Wolfowitz (1973), Woodroofe (1974), Hall (1982), Smith (1985), Cheng and Iles (1987), Smith (1994) and Hall and Wang (2005). More references on non-regular models and estimation approaches for them can be found in Cheng and Traylor (1995) and the references and discussions therein.

This paper is closely related to the work of Smith (1985), in which instead of using the global maximizer of the likelihood function, the solution to the likelihood equation is used to estimate unknown parameters. This idea was originally proposed by Harter and Moore (1966). The theory of local maximum likelihood estimation was established for a broad class of non-regular models without covariates in Smith (1985) by an elegant mathematical derivation. A key requirement in their proof is that the difference between the sample minimum and the lower bound of the support of a distribution has a non-degenerate distribution asymptotically. However, as is shown in Section 2, this quantity for a sample from model (5) always converges to a constant. Smith (1994) extended the results in Smith (1985) to a class of non-regular regression models, but model (3) does not meet the assumptions required in their analysis. In this paper, we provide another technique for proving the existence of a consistent maximum likelihood estimator. With small modifications, this technique applies to the consistency of local maximum likelihood estimator for the well known three parameter log-normal distribution, for which a rigorous proof has been missing for a long time. Uniqueness of the local maximizer of the likelihood function is also investigated and a theorem similar to Theorem 2 in Smith (1985) is formulated for a general regression model.

The rest of this paper is organized as follows. In Section 2, the one sample case is addressed and properties of estimators based on the sample extremes are derived. In Section 3, we present the results for the local maximum likelihood estimator for the regression problem. Results of simulation experiments that are designed to investigate the finite sample properties are contained in Section 4. Technical details are given in the appendix.

2 Estimation based on the extreme order statistics for the one sample case

2.1 Naive estimators

Suppose an independent sample $\{Y_1, ..., Y_n\}$ is taken from model (5). If parameters A and B are estimated in advance, μ and σ can be estimated simply by ordinary least-squares. An naive approach is then to use the two sample extremes, $Y_{(1)}$ and $Y_{(n)}$, to estimate A and B, respectively, and then remove them from the sample and use the rest of the sample to estimate μ and σ. We call such estimators naive. Why do the two sample extremes not perform well? It is not difficult to show that the two sample extremes are consistent, but their convergence rate is very slow. The proposition below gives asymptotic properties for these two statistics.

Proposition 1 *Let*

$$r_n = \{2\log n\}^{1/2} - \frac{\log\log n + \log(4\pi)}{\{8\log n\}^{1/2}} \qquad s_n = \frac{1}{\{2\log n\}^{1/2}}.$$

The following convergence results hold in distribution as $n \to \infty$:

$$\begin{aligned} \frac{e^{\mu_0+\sigma_0 r_n}}{\sigma_0 s_n}\left(\frac{1}{e^{\mu_0+\sigma_0 r_n}} - \frac{Y_{(1)} - A_0}{B_0 - A_0}\right) &\to G_1, \\ \frac{e^{-\mu_0+\sigma_0 r_n}}{\sigma_0 s_n}\left(\frac{1}{e^{-\mu_0+\sigma_0 r_n}} - \frac{B_0 - Y_{(n)}}{B_0 - A_0}\right) &\to G_2, \end{aligned} \tag{6}$$

where μ_0, σ_0, A_0 and B_0 are the true values of the parameters, and G_1 and G_2 are two independent random variables having the same distribution function $F(t) = e^{-e^{-t}}$.

Proof In Appendix A.1.

From this convergence result, it follows that, as $n \to \infty$,

$$\begin{aligned} e^{\mu_0+\sigma_0 r_n}\left(Y_{(1)} - A_0\right) &\to B_0 - A_0, \\ e^{-\mu_0+\sigma_0 r_n}\left(B_0 - Y_{(n)}\right) &\to B_0 - A_0, \end{aligned} \tag{7}$$

in distribution, which gives the rate of convergence as $e^{-\sigma_0 r_n}$. Since, for any $\alpha > 0$, $e^{-\sigma_0 r_n} n^{\alpha} \to \infty$, the rate of convergence is slower than $n^{-\alpha}$ for any $\alpha > 0$. But it is still faster than $1/\log n$ because $e^{-\sigma_0 r_n}\log n \to 0$. This proposition also tells us that there does not exist a constant sequence $r_n^* \to \infty$ such that $r_n^*(Y_{(1)} - A_0)$ or $r_n^*(B_0 - Y_{(n)})$ converges to a non-degenerate distribution.

2.2 Bias adjusted estimators

Estimation based on the two sample extreme values can be improved by adjusting for their asymptotic biases. From (6) and (7), better estimators of A and B are obtained:

$$\begin{aligned} \widehat{A}_{adj} &= Y_{(1)} - \frac{(1 - \gamma\widehat{\sigma}^* s_n)(Y_{(n)} - Y_{(1)})}{\exp(\widehat{\mu}^* + \widehat{\sigma}^* r_n)}, \\ \widehat{B}_{adj} &= Y_{(n)} + \frac{(1 - \gamma\widehat{\sigma}^* s_n)(Y_{(n)} - Y_{(1)})}{\exp(-\widehat{\mu}^* + \widehat{\sigma}^* r_n)}, \end{aligned} \tag{8}$$

where $\widehat{\mu}^*$ and $\widehat{\sigma}^*$ are two consistent estimates of μ and σ, respectively, and $\gamma \approx 0.577$ is the Euler-Mascheroni constant. Their asymptotic sampling properties are given by the following convergence results:

$$\frac{e^{\mu_0+\sigma_0 r_n}}{\sigma_0(B_0 - A_0)s_n}(\widehat{A}_{adj} - A_0) \to \gamma - G_1,$$
$$\frac{e^{\mu_0+\sigma_0 r_n}}{\sigma_0(B_0 - A_0)s_n}(\widehat{B}_{adj} - B_0) \to \gamma - G_2,$$

in distribution. By adjusting for the asymptotic biases of the two sample extremes, the estimators in (8) improve the rate of convergence from $e^{-\sigma_0 r_n}$ to $s_n e^{-\sigma_0 r_n}$. Although this rate is also between $1/\log n$ and $n^{-\alpha}$ for any $\alpha > 0$, simulation results show that these estimators are much more efficient than the two sample extremes.

3 Maximum likelihood estimators

The estimators given in Section 2 do not posses the optimal convergence rate and their properties are difficult to derive when the model involves covariates. Thus we evaluate the method of maximum likelihood estimation in this section focusing on model (1).

Denote the log-likelihood function by $\ell_n(\boldsymbol{\theta})$. The likelihood equations are

$$\begin{aligned}
\frac{\partial \ell_n(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} &= \sum_{i=1}^{n} \frac{\left\{\log\left(\frac{B-Y_i}{Y_i-A}\right) - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right\}}{\sigma^2}\mathbf{x}_i = 0,\\
\frac{\partial \ell_n(\boldsymbol{\theta})}{\partial \sigma} &= -\frac{n}{\sigma} + \sum_{i=1}^{n} \frac{\left\{\log\left(\frac{B-Y_i}{Y_i-A}\right) - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right\}^2}{\sigma^3} = 0,\\
\frac{\partial \ell_n(\boldsymbol{\theta})}{\partial A} &= \frac{1}{B-A}\sum_{i=1}^{n}\frac{B-Y_i}{Y_i-A} - \sum_{i=1}^{n}\frac{\log\left(\frac{B-Y_i}{Y_i-A}\right) - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}}{\sigma^2(Y_i-A)} = 0,\\
\frac{\partial \ell_n(\boldsymbol{\theta})}{\partial B} &= -\frac{1}{B-A}\sum_{i=1}^{n}\frac{Y_i-A}{B-Y_i} - \sum_{i=1}^{n}\frac{\log\left(\frac{B-Y_i}{Y_i-A}\right) - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}}{\sigma^2(B-Y_i)} = 0.
\end{aligned} \tag{9}$$

In this section, following the idea of Smith (1985), we study the properties of local maximizer of the likelihood function, i.e., the solution to the likelihood equations. We prove the existence and consistency of the resultant estimator.

From calculations in Appendix A.2, the Fisher information matrix based on the sample is

$$\mathcal{I}_n(\boldsymbol{\theta}) = \sum_{i=1}^{n} \begin{pmatrix} \frac{\mathbf{x}_i\mathbf{x}_i^{\mathrm{T}}}{\sigma^2} & 0 & \frac{-1-c_i d}{\sigma^2(B-A)}\mathbf{x}_i^{\mathrm{T}} & \frac{-1-\frac{d}{c_i}}{\sigma^2(B-A)}\mathbf{x}_i^{\mathrm{T}} \\ 0 & \frac{2}{\sigma^2} & \frac{-2c_i d}{\sigma(B-A)} & \frac{2\frac{d}{c_i}}{\sigma(B-A)} \\ \frac{-1-c_i d}{\sigma^2(B-A)}\mathbf{x}_i & \frac{-2c_i d}{\sigma(B-A)} & \frac{c_i^2 d^4}{(B-A)^2} + \frac{1+2c_i d+c_i^2 d^4}{\sigma^2(B-A)^2} & \frac{-1}{(B-A)^2} + \frac{2+c_i d+\frac{d}{c_i}}{\sigma^2(B-A)^2} \\ \frac{-1-\frac{d}{c_i}}{\sigma^2(B-A)}\mathbf{x}_i & \frac{2\frac{d}{c_i}}{\sigma(B-A)} & \frac{-1}{(B-A)^2} + \frac{2+c_i d+\frac{d}{c_i}}{\sigma^2(B-A)^2} & \frac{\frac{d^4}{c_i^2}}{(B-A)^2} + \frac{1+2\frac{d}{c_i}+\frac{d^4}{c_i^2}}{\sigma^2(B-A)^2} \end{pmatrix},$$

where $c_i = e^{\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}}$ and $d = e^{\sigma^2/2}$.

The following assumptions are required for the asymptotic results in this section.

Assumption 1 $\sup_i \|\mathbf{x}_i\| < \infty$, *where* $\|\cdot\|$ *denotes the Euclidean norm.*

Assumption 2 *The following quantities converge as* $n \to \infty$*:*
$n^{-1}\sum_{i=1}^{n}\mathbf{x}_i$, $n^{-1}\sum_{i=1}^{n}\mathbf{x}_i\mathbf{x}_i^{\mathrm{T}}$, $n^{-1}\sum_{i=1}^{n}e^{\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}}$, $n^{-1}\sum_{i=1}^{n}e^{-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}}$, $n^{-1}\sum_{i=1}^{n}\mathbf{x}_i e^{\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}}$, $n^{-1}\sum_{i=1}^{n}\mathbf{x}_i e^{-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}}$, $n^{-1}\sum_{i=1}^{n}e^{2\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}}$ *and* $n^{-1}\sum_{i=1}^{n}e^{-2\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}}$.

Assumption 3 $(\mathbf{x}_1^{\mathrm{T}}, ..., \mathbf{x}_n^{\mathrm{T}})^{\mathrm{T}}$ *is full rank.*

If the Assumptions 1-3 below hold, then $\mathcal{I}_n(\boldsymbol{\theta})/n$ converges to a positive-definite matrix, say $\mathcal{I}(\boldsymbol{\theta})$.

The following theorems present the properties of the local maximum likelihood estimator, the solution to (9). Proofs of these theorems are given in Appendix A.3.

Theorem 1 (Existence) *If assumptions 1-3 hold, then with probability approaching 1, there exists a sequence of solutions* $\widehat{\boldsymbol{\theta}}_n$ *to the likelihood equations in* (9) *that is a local maximizer of the likelihood function and is* $n^{1/2}$*-consistent for* $\boldsymbol{\theta}$.

Theorem 2 (Uniqueness) *Assume assumptions 1-3 hold. Let* δ *be some fixed value and* $\delta_n = n^{-\alpha}$ *for some* $\alpha > 0$. *Denote by* $S_\delta = \{\boldsymbol{\theta} :\ A \le A_0 - \delta \text{ and } B \ge B_0 + \delta\}$ *and* $T_{\delta,n} = \{\boldsymbol{\theta} :\ A_0 - \delta \le A \le A_0 + \delta_n, B_0 - \delta_n \le B \le B_0 + \delta \text{ and } \|\boldsymbol{\beta} - \boldsymbol{\beta}_0\| + |\sigma - \sigma_0| > \delta\}$. *Then, for any compact set* $K \subset \mathbb{R}^{p+3}$,

$$\lim_{n\to\infty} Pr\left\{\sup_{S_\delta \cap K} \ell_n(\boldsymbol{\theta}) < \ell_n(\boldsymbol{\theta}_0)\right\} = 1, \quad \lim_{n\to\infty} Pr\left\{\sup_{T_{\delta,n} \cap K} \ell_n(\boldsymbol{\theta}) < \ell_n(\boldsymbol{\theta}_0)\right\} = 1.$$

Theorem 3 (Asymptotic normality) *If assumptions 1-3 hold, the* $n^{1/2}$*-consistent estimator* $\widehat{\boldsymbol{\theta}}_n$ *in Theorem 1 satisfies*

$$n^{1/2}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \to N\left\{0, \mathcal{I}^{-1}(\boldsymbol{\theta}_0)\right\}$$

in distribution.

4 Numerical examples

4.1 Simulation

In this subsection, simulation results are reported that examine the finite sample performance of the biased adjusted and local maximum likelihood estimators given in Sections 2 and 3. The computation was carried out using R and a package called BB was used to find the solutions to the likelihood equations. Although multiple solutions to the likelihood equations may exist, according to our Theorem 2, the solution that yields the largest value of the likelihood function should be chosen as the estimate. From equation (9), all other parameters can be written as functions of the two boundary parameters A and B explicitly. So there are only two non-linear equations to solve to obtain estimates of A and B, and these estimates can be inserted into the first two equations in (9) to solve for $\boldsymbol{\beta}$ (or μ for the one sample model) and σ. All the simulations results are based on 1000 iterations.

Tables 1 and 2 give the relative mean square errors for model (5),

$$\log\left(\frac{B-Y}{Y-A}\right) \sim N(\mu, \sigma^2),$$

the one sample case without covariates. The relative mean square errors are the ratios of the mean square errors of a given estimator calculated from simulated sample to that of the local maximum likelihood estimator defined in Section 3. So a value of relative mean square error greater than unity indicates that the given estimator is less efficient than the local maximum likelihood estimator, and vise-versa. Table 1 reports results when μ and σ are assumed to be known while they are unknown in Table 2. A consistent solution to the likelihood equations always exists in our simulation studies if μ and σ are known. When μ and σ are unknown, a consistent solution to the likelihood equations occasionally did not exist for small sample sizes. From Table 2, one can see that this occurs rarely; in the worst case, consistent solutions were not found in 6 iterations out of 1000. When a solution to the likelihood equations was not found, the bias adjusted estimator was used instead.

It is seen in Tables 1 and 2 that all the relative mean square errors are greater than unity, which means both the naive and the bias adjusted estimators are dominated by the local maximum likelihood estimator. Furthermore, the relative performance of these two estimators deteriorates as the sample size grows. The bias adjusted estimator outperforms the naive estimator uniformly, and its performance relative to that of the naive estimator improves as the sample size increases. It is also observed that improvement of the local maximum likelihood estimator compared to the other estimators in Table 1 with known μ and σ is more significant than in Table 2, in which μ and σ are unknown.

For the regression model (3),

$$\log\left(\frac{B-Y}{Y-A}\right) = a + bx + \varepsilon.$$

the covariate values, x, were generated from a discrete uniform distribution on $(1, 2, ..., 10)$. There was only one case in 1000 iterations where the consistent solution was not found when $n = 50$. When n is larger than 50, consistent solutions are always found in our studies. Table 3 gives the biases, standard errors, estimates of standard errors and the coverage probabilities of confidence intervals with a nominal level of 95%. The biases and the standard errors are calculated from the estimates based on the 1000 simulated samples, while the estimates of standard errors, $\widehat{SE}$, are calculated from the Hessian matrix of the likelihood function. The confidence intervals are constructed by $\widehat{\boldsymbol{\theta}} \pm \mathcal{Z}_{0.975}\widehat{SE}$, where $\widehat{\boldsymbol{\theta}}$ is the local maximum likelihood estimator and $\mathcal{Z}_{0.975}$ is the 97.5% normal quantile. Figure 2 presents the box plots of the estimates obtained from the 1000 simulated samples.

It is seen that both the biases and standard errors are small, and they decrease as the sample size increases, reflecting the consistency of the local likelihood estimator. Although it is evident that the standard errors are underestimated for small sample sizes and the coverage probabilities are lower than the nominal level, this situation ameliorates as the sample size increases.

4.2 An application

The data presented in Table 4, which was generated by Kpamegan and Jani (2013) during the qualification of an Anti-F IgG ELISA based assay (FDA, 2010). The study was about the F protein nanoparticle vaccine. A total of 736 absorbances measured at OD450-630 nanometers (nm) were taken at 8 different concentrations in ELISA units (EU) from 46 plates. There were two replicates at each concentration for each plate, so we used the average as the responses in our analysis. There are three plates for which data were recorded incorrectly, so we removed observations from these plates and only used the remaining 688 observations. In the qualification study, the four parameter logistic model (4) was used to fit data from each plate to determine the reference standard. Data from plates that produced estimates of B that were more than 2 standard deviations less than the average of all its estimates and data from plates that gave estimates of A that were more than 2 standard deviations greater than the average of all its estimates were excluded from the reference data set. Here we use all the data from that study except the data recorded incorrectly.

Figure 3 presents the data as a plot of the response against log concentration. The data points marked with a "$-$" sign are observations that were removed from the reference set. Clearly the variances of the responses at middle concentrations are larger than that of the responses at either low or high concentrations, and the distribution of absorbance readings for a given concentration skew away from the boundaries as the responses asymptote. The pattern observed in Figure 3 is very similar to the pattern for data generated from our model (3) as depicted in (a) of Figure 1. The parameter estimates for Model (3) based on this data are $\widehat{A} = 0.050$, $\widehat{B} = 3.963$, $\widehat{a} = 0.126$, $\widehat{b} = -1.320$ and $\widehat{\sigma} = 0.659$ with standard errors $\widehat{SE}_A = 0.022$, $\widehat{SE}_B = 0.044$, $\widehat{SE}_a = 0.808$, $\widehat{SE}_b = 0.424$ and $\widehat{SE}_\sigma = 0.647$, respectively.

Figure 4 shows the linearized data using the parameter estimates in Model (3): $\log\{(\widehat{B} - Y_i)/(Y_i - \widehat{A})\}$, $i = 1, 2, \ldots, n$, where $\widehat{A}$ and $\widehat{B}$ are maximum likelihood estimates of A and B, respectively, as proposed in Section 3. The dotted lines are obtained by replacing ε with $\pm\mathcal{Z}_{0.975}\widehat{\sigma}$ in the MLE of equation (3).

5 Summary

In this paper, a new model was proposed which is an alternative to the commonly used four parameter logistic model. Although this is a non-regular model, we have proved that a local maximizer of the likelihood function is consistent, asymptotically normal and asymptotically efficient. Additionally, our result about the uniqueness of the MLE can help to choose the local maximizer in practice.

When the responses of a model have unknown boundaries, one may intuitively want to use the smallest and the largest observation to estimate them. Both our theoretical and numerical results showed that these two statistics were not efficient for the model proposed. Actually, since the extreme order statistics are always biased in general, one should always correct the bias to gain efficiency. This is also true for more general models.

Acknowledgement

We thank Dr. Valerii Fedorov for suggesting to us the importance of model (3).

Appendix A. Technical details

Appendix A.1. Proof of Proposition 1

Proof First, following the idea in Section 2.3 of Galambos (1978), for any t,

$$\lim_{n\to\infty} \Pr\left[\left\{\frac{\log(B_0 - Y) - \log(Y - A_0) - \mu_0}{\sigma_0}\right\}_{(n)} < r_n + s_n t\right] = e^{-e^{-t}}$$
$$= \lim_{n\to\infty} \Pr\left\{\left(\frac{B_0 - Y}{Y - A_0}\right)_{(n)} < e^{\mu_0 + \sigma_0 r_n + \sigma_0 s_n t}\right\}$$
$$= \lim_{n\to\infty} \Pr\left\{\frac{B_0 - Y_{(1)}}{Y_{(1)} - A_0} < e^{\mu_0 + \sigma_0 r_n} + \sigma_0 s_n e^{\mu_0 + \sigma_0 r_n}(1 + v\sigma_0 s_n t)t\right\},$$

where $|v| \le 1$. Since $v\sigma_0 s_n t \to 0$ as $n \to \infty$, it follows that

$$\lim_{n\to\infty} \Pr\left(\frac{B_0 - Y_{(1)}}{Y_{(1)} - A_0} < c_n + d_n t\right) = e^{-e^{-t}},$$

where $c_n = e^{\mu_0 + \sigma_0 r_n}$ and $d_n = \sigma_0 s_n e^{\mu_0 + \sigma_0 r_n}$.

Then, for any $t \neq 0$,

$$\Pr\left(\frac{B_0 - Y_{(1)}}{Y_{(1)} - A_0} < c_n + d_n t\right) = \Pr\left(\frac{Y_{(1)} - A_0}{B_0 - A_0} > \frac{1}{1 + c_n + d_n t}\right)$$
$$= \Pr\left[\frac{Y_{(1)} - A_0}{B_0 - A_0} > \frac{c_n - 1}{c_n^2} - \left\{\frac{d_n}{c_n^2} - \frac{(1 + d_n t)}{c_n^2(1 + c_n + d_n t)t}\right\} t\right].$$

It can be shown that $[(1 + d_n t)/\{c_n^2(1 + c_n + d_n t)t\}]/(d_n/c_n^2) \to 0$ and $(1/c_n^2)/(d_n/c_n^2) \to 0$ as $n \to \infty$. So from Lemma 2.2.2 in Galambos (1978),

$$\lim_{n\to\infty} \Pr\left\{\frac{B_0 - Y_{(1)}}{Y_{(1)} - A_0} < c_n + d_n t\right\} = \lim_{n\to\infty} \Pr\left\{\frac{Y_{(1)} - A_0}{B_0 - A_0} > \frac{1}{c_n} - \frac{d_n}{c_n^2} t\right\}.$$

When $t = 0$, the result can be verified by using the properties of the extreme order statistics of normal distribution directly. The second equation can be proved similarly.

Appendix A.2. Derivation of the Fisher information

The lemma below is useful in deriving the Fisher information.

Lemma 1 *From Lemma 2 of Stein (1981), we obtain that, if $E|h'(Z)| < \infty$ for a normal random variable $Z \sim N(\mu, \sigma^2)$ and some differentiable function h. Then*

$$E\{(Z-\mu)h(Z)\} = \sigma^2 E\{h'(Z)\}.$$

The log-likelihood function of model (1) based on one observation $(\mathbf{x}, Y)$ is

$$\begin{aligned}\ell(\boldsymbol{\theta}, \mathbf{x}, Y) = &-\frac{\log(2\pi)}{2} - \log\sigma + \log(B-A) - \log(Y-A)\\ &- \log(B-Y) - \frac{\{\log(B-Y) - \log(Y-A) - \mathbf{x}^{\mathrm{T}}\boldsymbol{\beta}\}^2}{2\sigma^2}\end{aligned}$$

for $Y \in (A, B)$ and 0 otherwise. Let $Z = \log(B-Y) - \log(Y-A)$. By direct calculation,

$$\begin{aligned}\frac{\partial \ell(\boldsymbol{\theta}, \mathbf{x}, Y)}{\partial A} &= \frac{1}{Y - A + \frac{(Y-A)^2}{B-Y}} - \frac{\{\log(B-Y) - \log(Y-A) - \mathbf{x}^{\mathrm{T}}\boldsymbol{\beta}\}}{\sigma^2(Y-A)};\\ \frac{\partial^2 \ell(\boldsymbol{\theta}, \mathbf{x}, Y)}{\partial A^2} &= \frac{1 + 2\frac{Y-A}{B-Y}}{\left\{Y - A + \frac{(Y-A)^2}{B-Y}\right\}^2} - \frac{\{\log(B-Y) - \log(Y-A) - \mathbf{x}^{\mathrm{T}}\boldsymbol{\beta}\}}{\sigma^2(Y-A)^2} - \frac{1}{\sigma^2(Y-A)^2}\\ &= \frac{e^{2Z} + 2e^Z}{(B-A)^2} - \frac{(Z - \mathbf{x}^{\mathrm{T}}\boldsymbol{\beta})(1+e^Z)^2}{\sigma^2(B-A)^2} - \frac{(1+e^Z)^2}{\sigma^2(B-A)^2}.\end{aligned}$$

Then from Lemma 1,

$$\begin{aligned}E\left\{\frac{\partial^2 \ell(\boldsymbol{\theta}, \mathbf{x}, Y)}{\partial A^2}\right\} &= -E\left\{\frac{e^{2Z}}{(B-A)^2} + \frac{1 + 2e^Z + e^{2Z}}{\sigma^2(B-A)^2}\right\}\\ &= -\frac{e^{2\mathbf{x}^{\mathrm{T}}\boldsymbol{\beta} + 2\sigma^2}}{(B-A)^2} - \frac{1 + 2e^{\mathbf{x}^{\mathrm{T}}\boldsymbol{\beta} + \frac{\sigma^2}{2}} + e^{2\mathbf{x}^{\mathrm{T}}\boldsymbol{\beta} + 2\sigma^2}}{\sigma^2(B-A)^2}.\end{aligned}$$

Other elements of the Fisher information can be derived similarly.

Appendix A.3. Proof of Theorems for the Regression Model

The proof of Theorem 1 begins with some lemmas.

Lemma 2 *For constant sequences $v_n \downarrow v$ and $w_n \uparrow w$ as $n \to \infty$, let $\xi_{v_n} \in (v_{n+1}, v_n)$ and $\xi_{w_n} \in (w_n, w_{n+1})$. If a continuous function sequence $f_n(\cdot) > 0$, which is decreasing in n, satisfies $n^{1+\alpha} f_n(\xi_{v_n}) \to 0$ and $n^{1+\alpha} f_n(\xi_{w_n}) \to 0$ for $\alpha > 0$ as $n \to \infty$, then*

$$\limsup_n \int_{v_n}^{w_n} f_n(x)\mathrm{d}x < \infty.$$

Proof Let $S_n = \int_{v_n}^{w_n} f_n(x)\mathrm{d}x$. Then

$$\begin{aligned}S_n - S_{n-1} = &\int_{v_n}^{w_n} f_n(x)\mathrm{d}x - \int_{v_{n-1}}^{w_{n-1}} f_{n-1}(x)\mathrm{d}x\\ \leq &(v_{n-1} - v_n) f_{n-1}(\xi_{v_{n-1}}) + (w_n - w_{n-1}) f_{n-1}(\xi_{w_{n-1}})\\ = &\frac{(v_{n-1} - v_n) n^{1+\alpha} f_{n-1}(\xi_{v_{n-1}}) + (w_n - w_{n-1}) n^{1+\alpha} f_{n-1}(\xi_{w_{n-1}})}{n^{1+\alpha}} = o\left(\frac{1}{n^{1+\alpha}}\right).\end{aligned}$$

So $\limsup_n S_n = \limsup_n \sum_{i=1}^n (S_n - S_{n-1})$ is finite.

Lemma 3 *For any $\alpha > 0$, let $\delta_n = n^{-\alpha}$. Then for any $k_1 \geq 0$ and $k_2 \geq 0$, there exists a constant M such that*

$$\lim_{n\to\infty} Pr\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{|\log(B-Y_i)|^{k_1}}{(Y_i-A)^{k_2}} < M\right\} = 1, \quad \lim_{n\to\infty} Pr\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{|\log(Y_i-A)|^{k_1}}{(B-Y_i)^{k_2}} < M\right\} = 1 \tag{10}$$

uniformly in A and B so long as $|A - A_0| < \delta_n$ and $|B - B_0| < \delta_n$.

Proof We gives the details of proof for the first quantity in (10). The proof for the other one is similar.

$$\begin{aligned}
\frac{|\log(B-Y_i)|^{k_1}}{(Y_i-A)^{k_2}} =& \frac{|\log(B-B_0+B_0-Y_i)|^{k_1}}{(Y_i-A_0+A_0-A)^{k_2}} I(Y_i-A_0>2\delta_n, B_0-Y_i>2\delta_n)+o_p(1)\\
=& \frac{|\log(B-B_0+B_0-Y_i)|^{k_1}}{(Y_i-A_0+A_0-A)^{k_2}} I(Y_i-A_0>2\delta_n, 1-2\delta_n>B_0-Y_i>2\delta_n)\\
&+\frac{|\log(B-B_0+B_0-Y_i)|^{k_1}}{(Y_i-A_0+A_0-A)^{k_2}} I(Y_i-A_0>2\delta_n, B_0-Y_i>1)+o_p(1)\\
<& \frac{I(Y_i-A_0>2\delta_n, 1-2\delta_n>B_0-Y_i>2\delta_n)}{(B_0-Y_i-\delta_n)^{k_1}(Y_i-A_0-\delta_n)^{k_2}}\\
&+\frac{(B_0-Y_i+\delta_n)^{k_1}}{(Y_i-A_0-\delta_n)^{k_2}} I(Y_i-A_0>2\delta_n, B_0-Y_i>1)+o_p(1)\\
<& \frac{I(B_0-2\delta_n>Y_i>A_0+2\delta_n)}{(B_0-Y_i-\delta_n)^{k_1}(Y_i-A_0-\delta_n)^{k_2}}\\
&+\frac{(B_0-A_0+1)^{k_1}}{(Y_i-A_0-\delta_n)^{k_2}} I(B_0-1>Y_i>A_0+2\delta_n)+o_p(1)\\
=& C_{in1}+C_{in2}+o_p(1).
\end{aligned}$$

Note that

$$\begin{aligned}
&(2\pi)^{\frac{1}{2}}E(C_{in1})\\
&=\frac{1}{\sigma_0}\int_{A_0+2\delta_n}^{B_0-2\delta_n}\frac{1}{(B_0-y-\delta_n)^{k_1}(y-A_0-\delta_n)^{k_2}}\times\\
&\qquad\frac{B_0-A_0}{(B_0-y)(y-A_0)}\exp\left[-\frac{\left\{\log\left(\frac{B_0-y}{y-A_0}\right)-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}_0\right\}^2}{2\sigma_0^2}\right]\mathrm{d}y\\
&\leq\frac{1}{\sigma_0}\int_{A_0+2\delta_n}^{B_0-2\delta_n}\frac{1}{(B_0-y-\delta_n)^{k_1+1}(y-A_0-\delta_n)^{k_2+1}}\times\\
&\qquad\exp\left[-\frac{\frac{1}{4}\left\{\log\left(\frac{B_0-y}{y-A_0}\right)\right\}^2-(\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}_0)^2}{2\sigma_0^2}\right]\mathrm{d}y\\
&=\frac{\exp\left\{\frac{(\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}_0)^2}{2\sigma_0^2}\right\}}{\sigma_0}\int_{A_0+2\delta_n}^{B_0-2\delta_n}\frac{1}{(B_0-y-\delta_n)^{k_1+1}(y-A_0-\delta_n)^{k_2+1}}\times\\
&\qquad\exp\left[-\frac{\left\{\log\left(\frac{B_0-y}{y-A_0}\right)\right\}^2}{8\sigma_0^2}\right]\mathrm{d}y.
\end{aligned}$$

From Lemma 2, $\limsup_n E(C_{in1})$ is bounded by a finite constant, say C_1. Similarly, it can also be shown that $\limsup_n E(C_{in2})$ is bounded by a finite constant, say C_2. So using the formula

$X_n = E(X_n) + O_P\{\text{var}(X_n)^{1/2}\}$, we have

$$\frac{1}{n}\sum_{i=1}^{n}\frac{|\log(B-Y_i)|^{k_1}}{(Y_i-A)^{k_2}} = \frac{1}{n}\sum_{i=1}^{n}E(C_{in1}) + \frac{1}{n}\sum_{i=1}^{n}E(C_{in2}) + O_P\left(n^{-\frac{1}{2}}\right) + o_P(1).$$

Thus any M that is greater than $C_1 + C_2$ satisfys the requirement.

Lemma 4 *If assumptions 1- 3 hold, then* $-n^{-1}\partial^2\ell_n(\boldsymbol{\theta})/(\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^{\mathrm{T}}) \to \mathcal{I}(\boldsymbol{\theta}_0)$ *in probability uniformly over* $\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\| < \delta_n$.

Proof The first element of $\partial^2\ell_n(\boldsymbol{\theta})/(\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^{\mathrm{T}})$ is

$$\frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial A^2} = \sum_{i=1}^{n}\left[\frac{1-\frac{1}{\sigma^2}}{(Y_i-A)^2} - \frac{1}{(B-A)^2} - \frac{\left\{\log\left(\frac{B-Y_i}{Y_i-A}\right) - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right\}}{\sigma^2(Y_i-A)^2}\right].$$

So it is straightforward to get

$$\begin{aligned}
&\frac{1}{n}\left|\frac{\partial^2\ell_n(\boldsymbol{\theta})}{\partial A^2} - \frac{\partial^2\ell_n(\boldsymbol{\theta}_0)}{\partial A^2}\right| \\
\leq&\frac{1}{n}\sum_{i=1}^{n}\left|\frac{1}{(B-A)^2} - \frac{1}{(B_0-A_0)^2}\right| + \frac{1}{n}\sum_{i=1}^{n}\left|\frac{1+\frac{1}{\sigma^2}}{(Y_i-A)^2} - \frac{1+\frac{1}{\sigma_0^2}}{(Y_i-A_0)^2}\right| \\
&+\frac{1}{n}\sum_{i=1}^{n}\left|\frac{\left\{\log\left(\frac{B-Y_i}{Y_i-A}\right) - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right\}}{\sigma^2(Y_i-A)^2} - \frac{\left\{\log\left(\frac{B_0-Y_i}{Y_i-A_0}\right) - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}_0\right\}}{\sigma_0^2(Y_i-A_0)^2}\right| \\
=&\Delta_1 + \Delta_2 + \Delta_3.
\end{aligned}$$

Δ_1 goes to 0 as δ_n goes to 0. By straightforward but tedious calculation, we obtain

$$\begin{aligned}
\Delta_3 \leq&\frac{1}{n}\sum_{i=1}^{n}\left|\log\left(\frac{B-Y_i}{Y_i-A}\right) - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right|\left|\frac{1}{\sigma^2(Y_i-A)^2} - \frac{1}{\sigma_0^2(Y_i-A_0)^2}\right| \\
&+\frac{1}{n}\sum_{i=1}^{n}\left|\frac{\log\left(\frac{B-Y_i}{Y_i-A}\right) - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}}{\sigma_0^2(Y_i-A_0)^2} - \frac{\log\left(\frac{B_0-Y_i}{Y_i-A_0}\right) - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}_0}{\sigma_0^2(Y_i-A_0)^2}\right| \\
\leq&\left|\frac{1}{\sigma^2} - \frac{1}{\sigma_0^2}\right| \times \frac{1}{n}\sum_{i=1}^{n}\frac{\left|\log\left(\frac{B-Y_i}{Y_i-A}\right) - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right|}{(Y_i-A)^2} \\
&+\frac{1}{n}\sum_{i=1}^{n}\frac{\left|\log\left(\frac{B-Y_i}{Y_i-A}\right) - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right|}{\sigma_0^2} \times \left|\frac{1}{(Y_i-A)^2} - \frac{1}{(Y_i-A_0)^2}\right| \\
&+\frac{1}{n}\sum_{i=1}^{n}\frac{\left|\log\left(\frac{B-Y_i}{B_0-Y_i}\right)\right|}{\sigma_0^2(Y_i-A_0)^2} + \frac{1}{n}\sum_{i=1}^{n}\frac{\left|\log\left(\frac{Y_i-A}{Y_i-A_0}\right)\right|}{\sigma_0^2(Y_i-A_0)^2} + \frac{1}{n}\sum_{i=1}^{n}\frac{\mathbf{x}_i^{\mathrm{T}}(\boldsymbol{\beta}-\boldsymbol{\beta}_0)}{\sigma_0^2(Y_i-A_0)^2} \\
\leq&\frac{1}{n}\sum_{i=1}^{n}\frac{\left|\log\left(\frac{B-Y_i}{Y_i-A}\right) - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right|}{(Y_i-A)^2} \times \left|\frac{1}{\sigma^2} - \frac{1}{\sigma_0^2}\right| + \frac{4B|A-A_0|}{\sigma_0^2} \times \frac{1}{n}\sum_{i=1}^{n}\frac{\left|\log\left(\frac{B-Y_i}{Y_i-A}\right) - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right|}{(Y_i-A)^2(Y_i-A_0)^2} \\
&+\frac{1}{n}\sum_{i=1}^{n}\frac{|B-B_0|}{\sigma_0^2(B^*-Y_i)(Y_i-A_0)^2} + \frac{1}{n}\sum_{i=1}^{n}\frac{|A-A_0|}{\sigma_0^2(Y_i-A^*)(Y_i-A_0)^2} + \frac{1}{n}\sum_{i=1}^{n}\frac{\mathbf{x}_i^{\mathrm{T}}(\boldsymbol{\beta}-\boldsymbol{\beta}_0)}{\sigma_0^2(Y_i-A_0)^2} \\
=&\Delta_{3.1} + \Delta_{3.2} + \Delta_{3.3} + \Delta_{3.4} + \Delta_{3.5},
\end{aligned}$$

where A^* is between A and A_0 and B^* is between B and B_0. Now we look into each term in the last equation above.

$$\Delta_{3.2} \leq \frac{2B|A-A_0|}{\sigma_0^2} \times \frac{1}{n}\sum_{i=1}^{n}\left|\log\left(\frac{B-Y_i}{Y_i-A}\right) - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right|\left\{\frac{1}{(Y_i-A)^4} + \frac{1}{(Y_i-A_0)^2}\right\}. \quad (11)$$

The right hand side term in (11) goes to 0 in probability uniformly since the second factor is bound with probability tending to 1 by Lemma 3 and the boundedness of x_i. Similarly, $\Delta_{3.1}$, $\Delta_{3.3}$, $\Delta_{3.4}$ and $\Delta_{3.5}$ can be shown to go to 0 in probability uniformly which implies Δ_3 goes to 0 in probability uniformly. Similarly but more easily, Δ_1 and Δ_2 can be showen to converge to 0 in probability uniformly, which implies $n^{-1}\left|\partial^2\ell_n(\boldsymbol{\theta})/\partial A^2 - \partial^2\ell_n(\boldsymbol{\theta}_0)/\partial A^2\right| \to 0$ in probability uniformly. By similiar arguments, other components of $\partial^2\ell_n(\boldsymbol{\theta})/(\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^{\mathrm{T}})$ can be shown to have the same property. This implys that $-n^{-1}\partial^2\ell_n(\boldsymbol{\theta})/(\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^{\mathrm{T}}) \to \mathcal{I}(\boldsymbol{\theta}_0)$ in probability uniformly over $\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\| < \delta_n$.

The following lemma is the Lemma 5 of Smith (1985). We state it for integrity and skip the proof.

Lemma 5 *Let h be a continuously differentiable real-valued function of $p+1$ real variables and let H denote the gradient vector of h. Suppose that the scalar product of $\mathbf{u}$ and $H(\mathbf{u})$ is negative whenever $\|\mathbf{u}\| = 1$. Then h has a local maximum, at which $H = 0$, for some $\mathbf{u}$ with $\|\mathbf{u}\| < 1$.*

Proof (of Theorem 1) It suffices to show for any ϵ, there exists a constant c such that

$$\Pr\left\{\mathbf{u}^{\mathrm{T}}\frac{\partial\ell_n\left(\boldsymbol{\theta}_0+n^{-1/2}c\mathbf{u}\right)}{\partial\boldsymbol{\theta}} < 0\right\} > 1-\epsilon \tag{12}$$

for any vector $\mathbf{u}$ such that $\|\mathbf{u}\| = 1$. Using Taylor's expansion,

$$\begin{aligned}\frac{\partial\ell_n\left(\boldsymbol{\theta}_0+n^{-1/2}c\mathbf{u}\right)}{\partial\boldsymbol{\theta}} &= \frac{\partial\ell_n\left(\boldsymbol{\theta}_0\right)}{\partial\boldsymbol{\theta}} + cn^{-1/2}\frac{\partial^2\ell_n\left(\boldsymbol{\theta}_0+n^{-1/2}c\mathbf{u}^*\right)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^{\mathrm{T}}}\mathbf{u}\\ &= \frac{\partial\ell_n\left(\boldsymbol{\theta}_0\right)}{\partial\boldsymbol{\theta}} - cn^{1/2}\mathcal{I}(\boldsymbol{\theta}_0)\mathbf{u} + n^{1/2}\epsilon_{n,\mathbf{u}},\end{aligned}$$

where $\mathbf{u}^*$ is a vector satisfying $\|\mathbf{u}^*\| \le 1$ and, by Lemma 3, $\epsilon_{n,\mathbf{u}} \to 0$ in probability uniformly over $\|\mathbf{u}\| \le 1$ as $n \to \infty$. It follows that

$$n^{-1/2}\mathbf{u}^{\mathrm{T}}\frac{\partial\ell_n\left(\boldsymbol{\theta}_0+n^{-1/2}\mathbf{u}\right)}{\partial\boldsymbol{\theta}} = n^{-1/2}\mathbf{u}^{\mathrm{T}}\frac{\partial\ell_n\left(\boldsymbol{\theta}_0\right)}{\partial\boldsymbol{\theta}} - c\mathbf{u}^{\mathrm{T}}\mathcal{I}(\boldsymbol{\theta}_0)\mathbf{u} + \mathbf{u}^{\mathrm{T}}\epsilon_{n,\mathbf{u}}. \tag{13}$$

Note that $n^{-1/2}\mathbf{u}^{\mathrm{T}}\partial\ell_n\left(\boldsymbol{\theta}_0\right)/\partial\boldsymbol{\theta}$ is $O_P(1)$. So the second term dominates the first term in (13) for large enough c. This proves equation (12) and the result follows from Lemma 5.

Proof (of Theorem 2, part 1) For any $\boldsymbol{\theta}_1 \in S$, $E\ell_n(\boldsymbol{\theta}_1) < \infty$, so $E[\ell_n(\boldsymbol{\theta}_1) - \ell_n(\boldsymbol{\theta}_0)] < 0$ by Jensen's inequality. This implies the existence of $\xi_{\boldsymbol{\theta}_1}$ such that

$$\lim_{n\to\infty}\Pr\left\{\ell_n(\boldsymbol{\theta}_1) - \ell_n(\boldsymbol{\theta}_0) < -\xi_{\boldsymbol{\theta}_1}\right\} = 1.$$

For $\boldsymbol{\theta}$ and η such that $|\boldsymbol{\theta}-\boldsymbol{\theta}_1| < \eta < |\boldsymbol{\theta}_1-\boldsymbol{\theta}_0| - \delta$,

$$\begin{aligned}&|\ell_n(\boldsymbol{\theta}) - \ell_n(\boldsymbol{\theta}_1)|\\ \le&|\log\sigma - \log\sigma_1| + \frac{1}{n}\sum_{i=1}^{n}\left|\log\left(\frac{1}{B-Y_i}+\frac{1}{Y_i-A}\right) - \log\left(\frac{1}{B_1-Y_i}+\frac{1}{Y_i-A_1}\right)\right|\\ &+\frac{1}{n}\sum_{i=1}^{n}\left|\frac{\left\{\log\left(\frac{B-Y_i}{Y_i-A}\right)-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right\}^2}{\sigma^2} - \frac{\left\{\log\left(\frac{B_1-Y_i}{Y_i-A_1}\right)-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}_1\right\}^2}{\sigma_1^2}\right|\\ =&\Delta_4+\Delta_5+\Delta_6.\end{aligned}$$

Δ_4 can be made smaller than $\xi_{\boldsymbol{\theta}_1}/4$ by choosing η small enough. By the mean value theorem,

$$\begin{aligned}\Delta_5 =&\frac{1}{n}\sum_{i=1}^{n}\left|\frac{1}{B^*-Y_i}\frac{Y_i-A^*}{B^*-A^*}(B-B_1) + \frac{1}{Y_i-A^*}\frac{B^*-Y_i}{B^*-A^*}(A-A_1)\right|\\ \le&\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{B_0-A_1+\eta}{B_0-A_0}\frac{|B-B_1|}{B_0-Y_i} + \frac{B_1-A_0+\eta}{B_0-A_0}\frac{|A-A_1|}{Y_i-A_0}\right\},\end{aligned}$$

for some A^* between A_0 and A_1 and B^* between B_0 and B_1. So $E(\Delta_5)$ can be make arbitrarily small by choosing small enough η, which implies

$$\lim_{n\to\infty} \Pr\left(\Delta_5 < \frac{\xi_{\boldsymbol{\theta}_1}}{4}\right) = 1$$

for small enough η.

$$\begin{aligned}
\Delta_6 \leq & \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\left\{\log\left(\frac{B-Y_i}{Y_i-A}\right)-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right\}^2}{\sigma^2}-\frac{\left\{\log\left(\frac{B_1-Y_i}{Y_i-A_1}\right)-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}_1\right\}^2}{\sigma^2}\right| \\
&+\frac{1}{n}\sum_{i=1}^{n}\left\{\log\left(\frac{B_1-Y_i}{Y_i-A_1}\right)-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}_1\right\}^2\left|\frac{1}{\sigma^2}-\frac{1}{\sigma_1^2}\right| \\
\leq & \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\log\left(\frac{B-Y_i}{Y_i-A}\right)-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}+\log(\frac{B_1-Y_i}{Y_i-A_1})-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}_1}{\sigma^2}\right| \\
&\times\left\{\frac{|A-A_0|}{A_0-Y_i}+\frac{|B-B_0|}{B_0-Y_i}+|\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}_1|\right\} \\
&+\frac{1}{n}\sum_{i=1}^{n}\left\{\log\left(\frac{B_1-Y_i}{Y_i-A_1}\right)-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}_1\right\}^2\left|\frac{1}{\sigma^2}-\frac{1}{\sigma_1^2}\right|.
\end{aligned}$$

So, for small enough η, we obtain

$$\lim_{n\to\infty} \Pr\left(\Delta_6 < \frac{\xi_{\boldsymbol{\theta}_1}}{4}\right) = 1.$$

Combining results for Δ_4, Δ_5 and Δ_6,

$$\lim_{n\to\infty} \Pr\left\{\sup_{|\boldsymbol{\theta}-\boldsymbol{\theta}_1|<\eta} \ell_n(\boldsymbol{\theta})-\ell_n(\boldsymbol{\theta}_0) < -\frac{\xi_{\boldsymbol{\theta}_1}}{4}\right\} = 1.$$

For any compact set K, $S_\delta \cap K$ can be covered by a finite number of neighborhoods of points in S_δ, so it follows that

$$\lim_{n\to\infty} \Pr\left\{\sup_{S_\delta\cap K} \ell_n(\boldsymbol{\theta})-\ell_n(\boldsymbol{\theta}_0) < -\xi_m\right\} = 1.$$

Proof (of Theorem 2, part 2) First, if A_0 and B_0 are known, model (1) can be transformed to a linear model with normal random error with unknown mean and variance. It follows that

$$\lim_{n\to\infty} \Pr\left\{\sup_{\|\boldsymbol{\beta}-\boldsymbol{\beta}_0\|>\delta\ |\sigma-\sigma_0|>\delta} \ell_n(\boldsymbol{\beta},\sigma,A_0,B_0)-\ell_n(\boldsymbol{\theta}_0) < -\xi\right\} = 1. \quad (14)$$

For $\boldsymbol{\beta}_1$, σ_1, η and $(\boldsymbol{\beta},\sigma,A,B)\in T$ such that $(\boldsymbol{\beta}_1,\sigma_1,A,B)\in T$, $\|\boldsymbol{\beta}-\boldsymbol{\beta}_1\|<\eta$, $|\sigma-\sigma_1|<\eta$ and $\delta<\eta$,

$$\begin{aligned}
&|\ell_n(\boldsymbol{\beta},\sigma,A,B)-\ell_n(\boldsymbol{\beta}_1,\sigma_1,A_0,B_0)| \\
&\leq|\log\sigma-\log\sigma_1|+|\log(B-A)-\log(B_0-A_0)| \\
&\quad+\frac{1}{n}\sum_{i=1}^{n}|\log(B-Y_i)-\log(B_0-Y_i)|+\frac{1}{n}\sum_{i=1}^{n}|\log(Y_i-A)-\log(Y_i-A_0)| \\
&\quad+\frac{1}{2n}\sum_{i=1}^{n}\left|\frac{\left\{\log\left(\frac{B-Y_i}{Y_i-A}\right)-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right\}^2}{\sigma^2}-\frac{\left\{\log\left(\frac{B_0-Y_i}{Y_i-A_0}\right)-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}_1\right\}^2}{\sigma_1^2}\right| \\
&=\Delta_7+\Delta_8+\Delta_9+\Delta_{10}+\Delta_{11}.
\end{aligned} \quad (15)$$

The terms Δ_7 and Δ_8 can be made smaller than $\xi/8$ by choosing η small enough. By the mean value theorem,

$$\Delta_9 = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{B-B_0}{B^*-Y_i}\right| \leq \frac{|B-B_0|}{n}\sum_{i=1}^{n}\frac{1}{\min(B,B_0)-Y_i}$$

with probability tending to 1. If $B \geq B_0$,

$$n^{-1}\sum_{i=1}^{n}\frac{1}{|\min(B,B_0)-Y_i|} \leq n^{-1}\sum_{i=1}^{n}\frac{1}{(B_0-Y_i)},$$

and the right hand side of the upper inequality goes to the limit of

$$\frac{1+n^{-1}\sum_{i=1}^{n}e^{-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}+\sigma^2/2}}{(B_0-A_0)}$$

in probability. If $B_0-\delta_n < B < B_0$, Lemma 3 provides that there exists some constant M^* such that

$$\lim_{n\to\infty}\Pr\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{|B-Y_i|} < M^*\right) = 1$$

for small enough η. This implies that for small enough η,

$$\lim_{n\to\infty}\Pr\left(\Delta_9 < \frac{\xi}{8}\right) = 1. \tag{16}$$

The same result can be found for Δ_{10} using similar arguments.

$$\begin{aligned}\Delta_{11} \leq & \frac{1}{2n}\sum_{i=1}^{n}\left|\frac{\left\{\log\left(\frac{B-Y_i}{Y_i-A}\right)-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}\right\}^2}{\sigma^2} - \frac{\left\{\log\left(\frac{B_0-Y_i}{Y_i-A_0}\right)-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}_1\right\}^2}{\sigma^2}\right| \\ & + \frac{1}{2n}\sum_{i=1}^{n}\left\{\log\left(\frac{B_0-Y_i}{Y_i-A_0}\right)-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}_1\right\}^2\left|\frac{1}{\sigma^2}-\frac{1}{\sigma_1^2}\right| \\ \leq & \frac{1}{2n}\sum_{i=1}^{n}\left|\frac{\log\left(\frac{B-Y_i}{Y_i-A}\right)-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}+\log\left(\frac{B_0-Y_i}{Y_i-A_0}\right)-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}_1}{\sigma^2}\right| \times |\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}_1| \\ & + \frac{1}{2n}\sum_{i=1}^{n}\left\{\log\left(\frac{B_0-Y_i}{Y_i-A_0}\right)-\mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}_1\right\}^2\left|\frac{1}{\sigma^2}-\frac{1}{\sigma_1^2}\right|.\end{aligned}$$

So we obtain, for small enough η,

$$\lim_{n\to\infty}\Pr\left(\Delta_{11} < \frac{\xi}{8}\right) = 1. \tag{17}$$

Combining (14), (15), (16) and (17), we have

$$\lim_{n\to\infty}\Pr\left\{\sup \ell_n(a,b,\sigma,A,B) - \ell_n(\boldsymbol{\theta}_0) < -\frac{3\xi}{8}\right\} = 1,$$

where the supermum is taken over all $\boldsymbol{\theta}$ satisfying $(\boldsymbol{\beta}_1,\sigma_1,A,B) \in T$, $\|\boldsymbol{\beta}-\boldsymbol{\beta}_1\| < \eta$ and $|\sigma-\sigma_1| < \eta$ for fixed $\boldsymbol{\beta}_1$ and σ_1. This result can be extended directly to any finite set of values of $\boldsymbol{\beta}_1$ and σ_1, and then to any compact sets of values of $\boldsymbol{\beta}_1$ and σ_1.

Proof (of Theorem 3) By Taylor expansion,

$$0 = \frac{\partial\ell_n(\widehat{\boldsymbol{\theta}}_n)}{\partial\boldsymbol{\theta}} = \frac{\partial\ell_n(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\theta}} + \frac{\partial^2\ell_n(\widehat{\boldsymbol{\theta}}^*)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^{\mathrm{T}}}(\widehat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}_0),$$

where $\widehat{\boldsymbol{\theta}}^*$ is between $\boldsymbol{\theta}_0$ and $\widehat{\boldsymbol{\theta}}_n$. From Lemma 4, $n^{-1}\partial^2\ell_n(\widehat{\boldsymbol{\theta}}^*)/(\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^{\mathrm{T}}) \to -\mathcal{I}(\boldsymbol{\theta}_0)$ in probability. So

$$n^{1/2}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) = \{\mathcal{I}(\boldsymbol{\theta}_0)\}^{-1}\, n^{-1/2}\frac{\partial\ell_n\left(\boldsymbol{\theta}_0\right)}{\partial\boldsymbol{\theta}} + o_P(1). \tag{18}$$

Note $n^{-1/2}\partial\ell_n\left(\boldsymbol{\theta}_0\right)/\partial\boldsymbol{\theta} = n^{-1/2}\sum_{i=1}^n \partial\ell(\boldsymbol{\theta}_0, \mathbf{x}_i, Y_i)/\partial\boldsymbol{\theta}$ is summation of independent random vectors and its variance converges to $\mathcal{I}(\boldsymbol{\theta}_0)$. Also we have for $t > 0$,

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n E\left[\left\|\frac{\partial\ell(\boldsymbol{\theta}_0, \mathbf{x}_i, Y_i)}{\partial\boldsymbol{\theta}}\right\|^2 I\left\{\left\|\frac{\partial\ell(\boldsymbol{\theta}_0, \mathbf{x}_i, Y_i)}{\partial\boldsymbol{\theta}}\right\| > n^{1/2}\epsilon\right\}\right] \\
&\leq \frac{1}{n}\frac{1}{(n^{1/2}\epsilon)^t}\sum_{i=1}^n E\left[\left\|\frac{\partial\ell(\boldsymbol{\theta}_0, \mathbf{x}_i, Y_i)}{\partial\boldsymbol{\theta}}\right\|^{2+t} I\left\{\left\|\frac{\partial\ell(\boldsymbol{\theta}_0, \mathbf{x}_i, Y_i)}{\partial\boldsymbol{\theta}}\right\| > n^{1/2}\epsilon\right\}\right] \\
&\leq \frac{1}{n}\frac{1}{(n^{1/2}\epsilon)^t}\sum_{i=1}^n E\left[\left\|\frac{\partial\ell(\boldsymbol{\theta}_0, \mathbf{x}_i, Y_i)}{\partial\boldsymbol{\theta}}\right\|^{2+t}\right] \to 0 \text{ as } n \to \infty.
\end{aligned}$$

By the multivariate central limit theorem (cf. Rao, 1973; Serfling, 1980),

$$n^{-1/2}\frac{\partial\ell_n\left(\boldsymbol{\theta}_0\right)}{\partial\boldsymbol{\theta}} \to N\{0, \mathcal{I}(\boldsymbol{\theta}_0)\} \tag{19}$$

in distribution. Combining (18), (19) and applying Slutsky's theorem, the result follows.

References

Cheng, R. C. H. and Amin, N. A. K. (1983). Estimating parameters in continuous univariate distributions with a shifted origin. *Journal of the Royal Statistical Society. Series B (Methodological)* **45**, 394–403.

Cheng, R. C. H. and Iles, T. C. (1987). Corrected maximum likelihood in non-regular problems. *Journal of the Royal Statistical Society. Series B (Methodological)* **49**, 95–101.

Cheng, R. C. H. and Traylor, L. (1995). Non-regular maximum likelihood problems. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**, 3–44.

Cohen, A. C. (1951). Estimating parameters of logarithmic-normal distributions by maximum likelihood. *Journal of the American Statistical Association* **46**, 206–212.

Cohen, A. C. and Whitten, B. J. (1980). Estimation in the three-parameter lognormal distribution. *Journal of the American Statistical Association* **75**, 399–404.

DeLean, A., Munson, P. J., and Rodbard, D. (1978). Simultaneous analysis of families of sigmoidal curves: application to bioassay, radioligand assay, and physiological dose-response curves. *American Journal of Physiology: Endocrinology and Metabolism* **235**, E97–E102.

Dragalin, V., Hsuan, F., and Padmanabhan, S. K. (2007). Adaptive designs for dose-finding studies based on sigmoid $e_{\max}$ model. *Journal of Biopharmaceutical Statistics* **17**, 1051–1070.

Ernst, A. A., Nick, T. G., Weiss, S. J., Houry, D., and Mills, T. (1997). Domestic violence in an inner-city ed. *Annals of Emergency Medicine* **30**, 190–197.

FDA (2010). Characterization and qualification of cell substrates and other biological materials used in the production of viral vaccines for infectious disease indications. *Guidance for Industry* U.S. Department of Health and Human Services, Food and Drug Administration, Center for Biologics Evaluation and Research.

Feng, F., Sales, A. P., and Kepler, T. B. (2011). A Bayesian approach for estimating calibration curves and unknown concentrations in immunoassays. *Bioinformatics* **27**, 5, 707–712.

Finke, M. D., DeFoliart, G. R., and Benevenga, N. J. (1987). Use of a four-parameter logistic model to evaluate the protein quality of mixtures of mormon cricket meal and corn gluten meal in rats. *Journal of Nutrition* **117**, 1740–1750.

Finke, M. D., DeFoliart, G. R., and Benevenga, N. J. (1989). Use of a four-parameter logistic model to evaluate the quality of the protein from three insect species when fed to rats. *Journal of Nutrition* **119**, 864–871.

Galambos, J. (1978). *The asymptotic theory of extreme order statistics*. John Wiley & Sons, New York.

Giesbrecht, F. and Kempthorne, O. (1976). Maximum likelihood estimation in the three-parameter lognormal distribution. *Journal of the Royal Statistical Society. Series B (Methodological)* **38**, 257–264.

Hall, P. (1982). On estimating the endpoint of a distribution. *The Annals of Statistics* **10**, 556–568.

Hall, P. and Wang, J. Z. (2005). Bayesian likelihood methods for estimating the end point of a distribution. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* **67**, 717–729.

Harter, H. L. and Moore, A. H. (1966). Local-maximum-likelihood estimation of the parameters of three-parameter lognormal populations from complete and censored samples. *Journal of the American Statistical Association* **61**, 842–851.

Hill, B. M. (1963). The three-parameter lognormal distribution and Bayesian analysis of a point-source epidemic. *Journal of the American Statistical Association* **58**, 72–84.

Holford, N. H. and Sheiner, L. B. (1981). Understanding the dose-effect relationship: Clinical application of pharmacokinetic-pharmacodynamic models. *Clinical Pharmacokinetics* **6**, 429–453.

Kpamegan, E. and Jani, D. (2013). Anti-F IgG ELISA based assay data from a qualification study. *personal communication,* Novavax Inc.

Leonov, S. and Miller, S. (2009). An adaptive optimal design for the $E_{\max}$ model and its application in clinical trials. *Journal of Biopharmaceutical Statistics* **19**, 360–385.

MacDougall, J. (2006). Analysis of dose-response studies $E_{\max}$. In N. Ting, ed., *Dose Finding in Drug Development*, 127–145. Springer, New York.

Menon, A. and Bhandarkar, S. (2004). Predicting polymorphic transformation curves using a logistic equation. *International Journal of Pharmaceutics* **286**, 125–129.

Nix, B. and Wild, D. (2001). Calibration curve-fitting. In D. Wild, ed., *The Immunoassay Handbook*, 198–210. Nature Publishing Group, New York.

Rao, C. R. (1973). *Linear Statistical Inference and its Applications*. John Wiley & Sons, 2nd edn.

Ratkowsky, D. A. and Reedy, T. J. (1986). Choosing near-linear parameters in the four-parameter logistic model for radioligand and related assays. *Biometrics* **42**, 575–582.

Sebaugh, J. L. (2011). Guidelines for accurate EC50/IC50 estimation. *Pharmaceutical Statistics* **10**, 128–134.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons, New York.

Smith, R. L. (1985). Maximum likelihood estimation in a class of nonregular cases. *Biometrika* **72**, 67–90.

Smith, R. L. (1994). Nonregular regression. *Biometrika* **81**, 173–183.

Stein, C. M. (1981). Estimation of the mean of a multivariate normal distribution. *The Annals of Statistics* **9**, 1135–1151.

Triantafilis, J., Laslett, G., and McBratney, A. B. (2000). Calibrating an electromagnetic induction instrument to measure salinity in soil under irrigated cotton. *Soil Science Society of America Journal* **64**, 1009–1017.

Vedenov, D. and Pesti, G. M. (2008). A comparison of methods of fitting several models to nutritional response data. *Journal of animal science* **86**, 500–507.

V∅lund, A. (1978). Application of the four-parameter logistic model to bioassay: Comparison with slope ratio and parallel line models. *Biometrics* **34**, 357–365.

Weiss, L. and Wolfowitz, J. (1973). Maximum likelihood estimation of a translation parameter of a truncated distribution. *The Annals of Statistics* **1**, 944–947.

Woodroofe, M. (1972). Maximum likelihood estimation of a translation parameter of a truncated distribution. *The Annals of Mathematical Statistics* **43**, 113–122.

Woodroofe, M. (1974). Maximum likelihood estimation of translation parameter of truncated distribution ii. *The Annals of Statistics* **2**, 474–488.

Table 1 Relative mean square errors for the one sample case when μ and σ are known.

Parameter		$\boldsymbol{\theta}_0 = (1, 0.5, 0, 10)$		$\boldsymbol{\theta}_0 = (-1, 0.5, 0, 5)$	
		Adjusted	Naive	Adjusted	Naive
$n = 50$	A	7.12	20.65	3.82	36.41
	B	3.62	35.03	6.64	19.29
$n = 100$	A	9.85	33.20	5.43	64.44
	B	5.13	62.55	9.15	31.13
$n = 400$	A	19.44	88.55	11.06	199.12
	B	10.69	191.09	17.94	81.85

In this table, Adjusted is the bias adjusted estimator relative to the local maximum likelihood estimator; Naive is the naive estimator relative to the local maximum likelihood estimator.

Table 2 Relative mean square errors for the one sample case when μ and σ are unknown.

Parameter		$\boldsymbol{\theta}_0 = (1, 0.5, 0, 10)$		$\boldsymbol{\theta}_0 = (-1, 0.5, 0, 5)$	
		Adjusted	Naive	Adjusted	Naive
$n = 50$		NS= 6		NS= 3	
	μ	1.22	1.38	1.30	1.45
	σ	1.63	2.21	1.67	2.26
	A	1.37	1.93	1.47	1.84
	B	1.38	1.75	1.27	1.81
$n = 100$		NS= 2		NS= 0	
	μ	1.47	1.73	1.66	1.92
	σ	2.68	4.30	2.91	4.67
	A	1.81	2.63	1.90	2.43
	B	1.66	2.15	1.77	2.60
$n = 400$		NS= 0		NS= 0	
	μ	2.01	2.62	2.66	3.46
	σ	4.08	8.53	6.01	12.57
	A	2.72	4.53	3.10	4.38
	B	2.14	3.05	3.42	5.73

In this table, Adjusted is the bias adjusted estimator relative to the local maximum likelihood estimator; Naive is the naive estimator relative to the local maximum likelihood estimator; NS is the number of cases out of 1000 iterations that consistent solutions to the likelihood equation are not found.

Table 3 Biases ($\times 10^3$), standard errors ($\times 10^3$), estimates of standard errors ($\times 10^3$) and coverage probabilities ($\times 10^2$) for the regression model.

		$\boldsymbol{\theta}_0 = (5, -1, 0.5, 0, 10)$				
		a	b	σ	A	B
$n = 50$	Bias	234	40	53	42	15
	SE	270	46	65	53	19
	SEE	248	42	61	42	16
	CP	90.3	90.2	92.5	84.2	84.3
$n = 100$	Bias	142	24	35	25	9
	SE	170	29	44	31	11
	SEE	170	29	43	29	11
	CP	93.8	93.6	93.3	89.2	89.6
$n = 400$	Bias	67	11	17	12	4
	SE	83	14	22	14	5
	SEE	82	14	21	14	5
	CP	94.5	95	94.5	93.3	94.4

In this table, SE is standard errors; $\widehat{SE}$ is estimates of standard errors; CP is coverage probabilities.

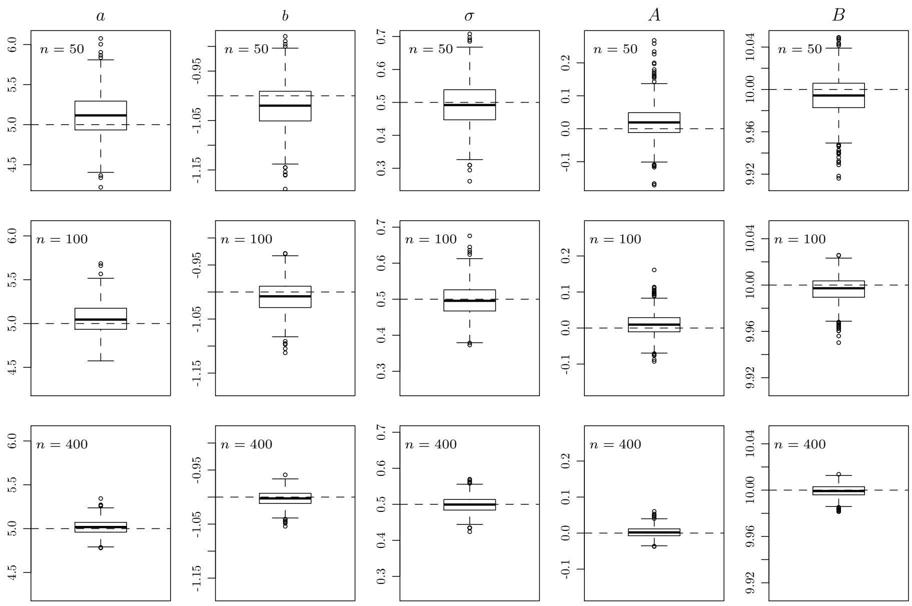

Fig. 2 Box plots of maximum likelihood estimates for model (3); the dashed reference line is the true of the parameter.

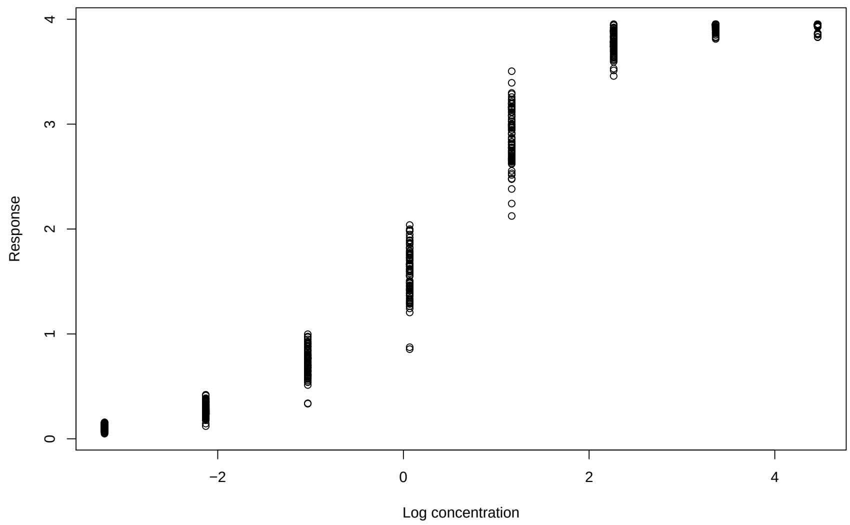

Fig. 3 Anti-F IgG ELISA based assay data

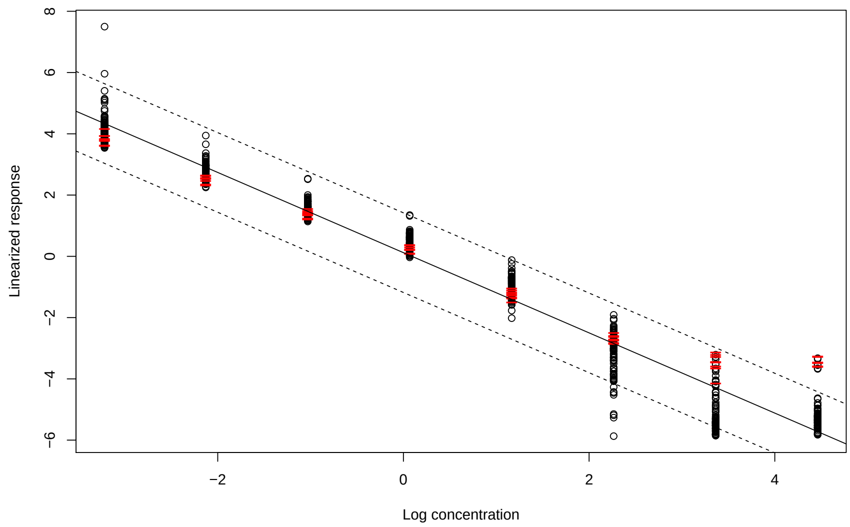

Fig. 4 Linearized dose-response data from a bioassay qualification study (The data points marked with the "–" sign are observations that were removed from the reference set)

Table 4 Anti-F IgG ELISA based assay data: Absorbance readings at OD450-630 nm at different concentrations.

Plate #	Concentrations (EU)							
	86.6600	28.8870	9.6290	3.2100	1.0700	0.3570	0.1190	0.0400
1	3.9498	3.9178	3.7808	2.6208	1.4098	0.6134	0.2355	0.0868
	3.9432	3.9089	3.7777	2.6401	1.2864	0.6428	0.2290	0.0890
2	3.9434	3.9446	3.7718	2.8291	1.5930	0.6122	0.2167	0.0693
	3.9388	3.9449	3.6962	2.7780	1.5519	0.5964	0.2152	0.0642
3	3.9445	3.9479	3.8967	3.1427	1.6654	0.7894	0.2966	0.1152
	3.9396	3.9506	3.8935	2.9896	1.7527	0.7792	0.2700	0.1120
4	3.9434	3.9354	3.8076	2.8573	1.6144	0.7380	0.2949	0.1116
	3.9423	3.9472	3.7921	2.7797	1.3637	0.6991	0.2768	0.1113
5	3.9445	3.9533	3.8347	2.9934	1.2432	0.5697	0.2075	0.0928
	3.9499	3.9533	3.7819	2.5160	1.3078	0.6097	0.2565	0.1192
6	3.9398	3.9383	3.9105	3.2344	1.7645	0.7713	0.2896	0.1185
	3.9373	3.9452	3.8991	3.1122	1.7533	0.7623	0.2992	0.1277
7	3.9455	3.9470	3.9224	3.1622	1.7718	0.7677	0.2715	0.1159
	3.9513	3.9510	3.8956	3.2977	1.7734	0.7479	0.2895	0.1048
8	3.9434	3.9521	3.9419	3.2581	2.0384	0.9671	0.3695	0.1374
	3.9489	3.9514	3.9428	3.2305	1.7904	0.9055	0.3459	0.1449
9	3.9504	3.9230	3.7490	2.7035	1.6418	0.6990	0.2486	0.1033
	3.9499	3.9191	3.6420	2.7721	1.4487	0.7338	0.2720	0.1125
10	3.9445	3.9467	3.6925	2.8302	1.4905	0.6952	0.2373	0.0993
	3.9433	3.9262	3.7413	2.6530	1.2919	0.5754	0.2308	0.0860
11	3.9505	3.9485	3.8821	3.0268	1.6438	0.7672	0.3102	0.1187
	3.9461	3.9470	3.8860	2.8729	1.7461	0.7641	0.3297	0.1287
12	3.9428	3.9479	3.7934	2.7472	1.5931	0.6437	0.2392	0.0973
	3.9412	3.9478	3.7516	2.6484	1.4227	0.6250	0.2392	0.1011
13	3.9507	3.9529	3.8987	3.1587	1.6701	0.8375	0.3091	0.1331
	3.9482	3.9529	3.8195	3.0025	1.5046	0.7113	0.3197	0.1163
14	3.9381	3.9474	3.8828	3.1992	1.9188	0.9271	0.3795	0.1473
	3.9378	3.9453	3.8621	3.0693	1.8931	0.9122	0.3862	0.1533
15	3.9487	3.9475	3.9197	3.5049	1.9841	0.8774	0.3870	0.1480
	3.9500	3.9504	3.9444	3.3941	1.8866	0.9199	0.3846	0.1348
16	3.9438	3.9458	3.8895	3.1650	1.8222	0.9471	0.3637	0.1473
	3.9390	3.9461	3.9182	3.1691	1.8237	0.8997	0.3816	0.1498
17	3.9492	3.9002	3.7236	2.6768	1.5957	0.7657	0.2908	0.1062
	3.9476	3.8967	3.7771	2.4807	1.3907	0.6842	0.2472	0.1073
18	3.9465	3.9477	3.7658	2.7472	1.6613	0.6813	0.2672	0.0947
	3.9474	3.9486	3.7936	2.8124	1.4808	0.7995	0.2952	0.1107
19	3.9276	3.9192	3.5302	2.3820	1.2046	0.5128	0.1899	0.0783
	3.9458	3.9272	3.6707	2.5309	1.3012	0.5897	0.2613	0.0995
20	3.9520	3.9518	3.8276	2.9789	1.7208	0.8234	0.3346	0.1293
	3.9457	3.9500	3.8370	2.9050	1.7264	0.7922	0.3064	0.1152
21	3.9477	3.9437	3.7822	2.7311	1.4848	0.7137	0.2549	0.1098
	3.9467	3.9440	3.7843	2.8297	1.4640	0.6983	0.2570	0.1090
22	3.9529	3.9528	3.7861	2.7965	1.5530	0.6820	0.2556	0.0954
	3.9528	3.9526	3.8368	2.6643	1.3649	0.6840	0.2783	0.0990

Table 4 (continued) Anti-F IgG ELISA based assay data: absorbances measured at OD450-630 nm at different concentrations.

Plate #	Concentrations (EU)							
	86.6600	28.8870	9.6290	3.2100	1.0700	0.3570	0.1190	0.0400
23	3.9461	3.9482	3.8405	3.1912	1.9987	0.8885	0.3562	0.1398
	3.9464	3.9469	3.8826	3.1616	1.8199	0.8455	0.3449	0.1378
24	3.9475	3.9505	3.9535	3.1945	1.9474	0.8780	0.3552	0.1292
	3.9491	3.9468	3.8679	3.1380	1.8158	0.8374	0.3624	0.1421
25	3.9413	3.9488	3.8746	3.0827	1.6729	0.7299	0.3223	0.1329
	3.9447	3.9500	3.8498	2.9775	1.6568	0.7677	0.3476	0.1423
26	3.9345	3.9057	3.6507	2.6681	1.4546	0.6705	0.2498	0.0921
	3.9475	3.8984	3.6092	2.6873	1.3826	0.6762	0.2395	0.1047
27	3.9427	3.9367	3.7872	2.6949	1.4249	0.5884	0.2237	0.0803
	3.9459	3.9513	3.7349	2.7381	1.4122	0.7695	0.2756	0.1153
28	3.9433	3.9182	3.6337	2.4758	1.3178	0.6008	0.2368	0.1067
	3.9455	3.9260	3.6322	2.4784	1.3250	0.6101	0.2560	0.1023
29	3.9499	3.9513	3.8665	2.9811	1.6988	0.7713	0.3286	0.1274
	3.9490	3.9503	3.8928	2.9663	1.6165	0.8068	0.3339	0.1286
30	3.9443	3.9403	3.6609	2.9135	1.4633	0.6113	0.2615	0.1068
	3.9472	3.9451	3.5953	2.6724	1.3873	0.5557	0.2433	0.1056
31	3.9530	3.9517	3.7453	2.9452	1.4906	0.7973	0.1998	0.1139
	3.9527	3.9527	3.7931	3.0859	1.5577	0.6035	0.2407	0.0926
32	3.9482	3.9503	3.8863	3.1558	1.6623	0.7236	0.2731	0.0968
	3.9489	3.9493	3.8743	3.0259	1.5583	0.7047	0.2817	0.0946
33	3.9437	3.9469	3.8255	2.9104	1.5772	0.6540	0.2618	0.1030
	3.9420	3.9397	3.7621	2.8048	1.4306	0.6434	0.2569	0.1000
34	3.9267	3.9339	3.7424	2.7722	1.6535	0.6739	0.2664	0.1019
	3.9321	3.9316	3.7097	2.8640	1.5051	0.7018	0.2844	0.1100
35	3.9469	3.9487	3.7412	2.6572	1.3382	0.5416	0.1890	0.0699
	3.9451	3.9496	3.6200	2.5546	1.2652	0.5788	0.1951	0.0709
36	3.9458	3.8754	3.4593	2.2431	0.8734	0.3389	0.1211	0.0488
	3.9490	3.8867	3.5123	2.1245	0.8547	0.3345	0.1451	0.0567
37	3.9501	3.9488	3.8125	2.9623	1.7317	0.7581	0.3166	0.1177
	3.9471	3.9527	3.7610	2.9454	1.6028	0.7928	0.3081	0.1290
38	3.9346	3.9425	3.7578	2.7169	1.6213	0.6545	0.2392	0.0958
	3.9499	3.9477	3.7383	2.6758	1.4361	0.6393	0.2544	0.0906
39	3.9525	3.9349	3.6950	2.7245	1.2856	0.5398	0.1769	0.0724
	3.9524	3.9530	3.8259	2.6267	1.4229	0.6532	0.2311	0.0877
Observations below were removed from the reference set								
40	3.8676	3.8508	3.7417	2.9959	1.7066	0.8438	0.3455	0.1348
	3.8661	3.8228	3.7411	3.0831	1.7275	0.8090	0.3260	0.1276
41	3.8659	3.8531	3.7660	3.1472	1.9807	0.9972	0.4113	0.1572
	3.8552	3.8715	3.7583	3.2118	1.9869	0.9752	0.4200	0.1553
42	3.8284	3.8661	3.7081	3.1984	1.8703	0.8813	0.3482	0.1376
	3.8324	3.9077	3.7342	3.1170	1.8644	0.9029	0.3536	0.1412
43	3.8553	3.8124	3.7165	3.2861	1.8536	0.8579	0.3730	0.1271
	3.8532	3.8323	3.6846	3.0458	1.7864	0.7754	0.3221	0.1105