Improved maximum likelihood estimation in heteroscedastic nonlinear regression models

F. Giummolè, L. Ventura

2001.15

Dipartimento di Scienze Statistiche
Università degli Studi
Via C. Battisti 241-243
35121 Padova

Ottobre 2001

IMPROVED MAXIMUM LIKELIHOOD ESTIMATION IN HETEROSCEDASTIC NONLINEAR REGRESSION MODELS

FEDERICA GIUMMOLÈ[a] and LAURA VENTURA[b]
[a]*Department of Mathematics, Politecnico of Torino, Italy;*
[b]*Department of Statistics, University of Padova, Italy*

Abstract

Nonlinear heteroscedastic regression models are a widely used class of models in applied statistics, with applications especially in biology, medicine or chemistry. Nonlinearity and variance heterogeneity can make likelihood estimation for a scalar parameter of interest rather inaccurate for small or moderate samples. In this paper, we suggest a new approach to point estimation based on estimating equations obtained from higher-order pivots for the parameter of interest. In particular, we take as an estimating function the modified directed likelihood. This is a higher-order pivotal quantity that can be easily computed in practice for nonlinear heteroscedastic models with normally distributed errors, using a recently developed S-PLUS library (HOA, 2000). The estimators obtained from this procedure are a refinement of the maximum likelihood estimators, improving their small sample properties and keeping equivariance under reparameterisation. Two applications to real data sets are discussed.

Keywords: Likelihood asymptotics; Modified directed likelihood; Nonlinear heteroscedastic models; Parameterisation equivariance

1 INTRODUCTION

Nonlinear regression models are widely used in applied statistics. Here we consider the *normal heteroscedastic nonlinear model* of the form

$$y_{ij} = \mu(x_i;\beta) + \epsilon_{ij}\ , \quad i = 1,\ldots,q\ , \quad j = 1,\ldots,m_i\ , \tag{1}$$

where q is the number of design points, m_i the number of replicates at the design point x_i, y_{ij} represents the response of the jth experimental unit in the ith group, and the errors ϵ_{ij} are independent $N(0,\sigma_i^2)$ random variables. The mean response of the model is given by the nonlinear function $\mu(x;\beta)$, called the *mean function*, which depends on an unknown regression coefficient β. The definition of the model is completed by assuming $\sigma_i^2 = \sigma^2 V^2(x_i;\beta,\rho)$, where $V^2(\cdot)$, called the *variance function*, represents the variance of the observations at different levels of covariates and σ^2 and ρ are variance parameters. If $V^2(\cdot)$ is constant, we obtain as a special case the homoscedastic nonlinear regression model. However, in many applications, the variability of the

error is not constant; this is often the case in immunological and biological studies where the correct specification of the variance function becomes crucial for calibration purposes and for determination of limits of detection.

Several methods have been proposed to fit model (1); they are summarised in Seber and Wild (1989, Chap. 2). The most frequently used estimating procedures are generalised nonlinear least squares (Carroll and Ruppert, 1988, Chap. 2) and maximum likelihood estimation. However, as shown in Bellio and Brazzale (1999), nonlinearity of the mean and variance heterogeneity can lead to substantial inaccuracies of these methods, if the sample size is small or moderate. The problem is mostly present when the parameter of interest is one of the variance parameters.

To overcome this drawback, in this paper we discuss an alternative practical approach to point estimation when inference about an arbitrary one-dimensional parameter of interest is desired in the context of nonlinear heteroscedastic regression models. Point estimators can be defined as zero-level confidence intervals obtained from a class of two-sided equi-tailed rejection regions. A natural implementation of this idea is to consider estimating functions based on higher-order pivotal quantities for the parameter of interest (see Pace and Salvan, 1999, and Giummolè and Ventura, 2001). In particular, we consider the modified directed likelihood r^* which is a higher-order adjustment of the signed square root r of the usual generalised loglikelihood ratio statistic. Since the estimating equations based on r give the maximum likelihood estimator (MLE), the proposed solution is essentially a correction to the MLE that improves its small sample properties, respecting the requirement of parameterisation equivariance.

The paper is organised as follows. Section 2 gives a short review of point estimation based on the modified directed likelihood in the context of nonlinear heteroscedastic models. In Section 3 the proposed method is illustrated on two application examples, showing that the resulting estimators are accurate and improve on the MLE.

It is important noticing that the statistic r^* and consequently the proposed estimators can be easily computed in practice for nonlinear heteroscedastic regression models with normal errors. In fact, a recently developed library for higher-order asymptotics (`HOA`, 2000) implements likelihood-based solutions, including r^*, for a wide class of models of practical interest in many areas of statistics.

2 BACKGROUND RESULTS

Let us assume that the d-dimensional parameter $\omega = (\beta, \sigma^2, \rho)$ in model (1) is of the form $\omega = (\psi, \lambda)$, where ψ is a scalar parameter of interest and λ a $(d-1)$-dimensional nuisance parameter. In many applications, ψ is a unidimensional component of the variance parameters or a scalar combination of the parameters of the model, like the value of the mean function or its inverse (in range where μ is monotone) at a fixed point x_0 (see Schwenke and Milliken, 1991, for an example of such kind of problems). Sometimes, as in Examples 1 and 2 of Section 3, one is interested in particular scalar functions such as the proportion of maximum (see (6)) or the function ED_t (see Bellio et al. (2000), Sec. 4).

Let $\hat{\omega} = (\hat{\psi}, \hat{\lambda})$ be the MLE of ω and $\tilde{\omega} = (\psi, \hat{\lambda}_\psi)$, with $\hat{\lambda}_\psi$ the MLE of λ for a given value of ψ. The profile loglikelihood function for ψ is $\ell_P(\psi) = \ell_P(\tilde{\omega}) = \ell(\psi, \hat{\lambda}_\psi)$, where $\ell(\omega)$ is the loglikelihood function for ω. As in Barndorff-Nielsen and Cox (1994), a tilde over a likelihood quantity is used when the quantity is evaluated at $(\psi, \hat{\lambda}_\psi)$, while a hat denotes evaluation at $(\hat{\psi}, \hat{\lambda})$. Usually,

inference about the scalar parameter of interest ψ is based on the directed likelihood

$$r(\psi) = \mathrm{sgn}(\hat{\psi} - \psi) \left[2\{\ell_P(\hat{\psi}) - \ell_P(\psi)\} \right]^{1/2} ,$$

which null distribution is standard normal with first order accuracy. A higher-order modification of the directed likelihood has been proposed by Barndorff-Nielsen (1991) and in useful equivalent versions by Skovgaard (1996) and Severini (1998, 1999). The modified directed likelihood can be written as

$$r^*(\psi) = r(\psi) - \frac{\log(\gamma(\psi))}{r(\psi)} , \tag{2}$$

where $\gamma(\psi)$ is a suitable higher-order correction term (see e.g. Skovgaard, 1996). The normal approximation for the null distribution of r^* is exact up to third order and, moreover, its expression is parameterisation invariant (see e.g. Barndorff-Nielsen and Cox, 1994, § 6.6).

As explained in Giummolè and Ventura (2001), the modified directed likelihood gives rise to a simple estimating equation of the form

$$r^*(\psi) = 0 . \tag{3}$$

A numerical procedure is usually required in order to solve (3). The existence and uniqueness of the solution, denoted by $\hat{\psi}^*$, is asymptotically guaranteed, at least in a neighbourhood of $\hat{\psi}$. The estimator $\hat{\psi}^*$ is a refinement of $\hat{\psi}$, with the estimating equation (3) giving implicitly a higher-order correction to the MLE. An explicit close form for $\hat{\psi}^*$ is given in Giummolè and Ventura (2001, Sec. 3). Nevertheless, a recent library for higher-order asymptotics (`HOA`, 2000) gives directly the statistic $r^*(\psi)$ for model (1), making it easier to numerically solve equation (3) than to implement the closed form expression of $\hat{\psi}^*$.

In view of the properties of r^*, the estimating equation (3) is mean unbiased as well as median unbiased at the third-order of accuracy. The median unbiasedness property also holds for the corresponding estimator $\hat{\psi}^*$, under the condition that the estimating equation is a monotone function of the parameter of interest. For nonlinear heteroscedastic regression models, monotonicity of r^* is verified in many practical situations, independently of the sample size (Bellio, 2000). Moreover, since the modified directed likelihood r^* is invariant under interest respecting reparameterisations, $\hat{\psi}^*$ is an equivariant estimator of ψ. Numerical investigations (see Giummolè and Ventura, 2001) show that the estimators based on r^* improve on the MLE.

In the context of nonlinear heteroscedastic regression models with normal errors, the modified directed likelihood (2) is rather simple to calculate. In fact, it is possible to rewrite these models as curved exponential families, for which the expression for the $\gamma(\psi)$ term in (2) remarkably simplifies. The loglikelihood function for the parameter $\omega = (\beta, \sigma^2, \rho)$ corresponding to model (1) is

$$\ell(\beta, \sigma^2, \rho) = -\sum_{i=1}^{q} \frac{m_i}{2} \log \sigma^2 - \sum_{i=1}^{q} \frac{m_i}{2} \log V^2(x_i; \beta, \rho) - \sum_{i=1}^{q} \sum_{j=1}^{m_i} \frac{(y_{ij} - \mu(x_i; \beta))^2}{2\sigma^2 V^2(x_i; \beta, \rho)} , \tag{4}$$

and can be written as

$$\ell(\beta, \sigma, \rho) = \theta(\omega) \cdot t(y) - K\left(\theta(\omega)\right) ,$$

with

$$\begin{aligned}
\theta(\omega) &= (\theta_{11}(\omega), \theta_{12}(\omega), \ldots, \theta_{q1}(\omega), \theta_{q2}(\omega)) \\
&= \left(\frac{\mu(x_1;\beta)}{\sigma^2 V^2(x_1;\beta,\rho)}, -\frac{1}{2\sigma^2 V^2(x_1;\beta,\rho)}, \ldots, \frac{\mu(x_q;\beta)}{\sigma^2 V^2(x_q;\beta,\rho)}, -\frac{1}{2\sigma^2 V^2(x_q;\beta,\rho)} \right) , \\
t(y) &= (t_{11}(y), t_{12}(y), \ldots, t_{q1}(y), t_{q2}(y)) \\
&= \left(\sum_{j=1}^{m_1} y_{1j}, \sum_{j=1}^{m_1} y_{1j}^2, \ldots, \sum_{j=1}^{m_q} y_{qj}, \sum_{j=1}^{m_q} y_{qj}^2 \right) , \\
K(\theta(\omega)) &= -\sum_{i=1}^{q} \frac{m_i}{2} \log(-2\theta_{i2}) - \sum_{i=1}^{q} \frac{m_i}{4} \frac{\theta_{i1}^2}{\theta_{i2}} \\
&= \sum_{i=1}^{q} \frac{m_i}{2} \log(\sigma^2) + \sum_{i=1}^{q} \frac{m_i}{2} \log(V^2(x_i;\beta,\rho)) + \sum_{i=1}^{q} \frac{m_i}{2} \frac{\mu^2(x_i;\beta)}{\sigma^2 V^2(x_i;\beta,\rho)} .
\end{aligned}$$

Following Skovgaard (1996), the correction term $\gamma(\psi)$ in (2) can be written as

$$\gamma(\psi) = \frac{|\tilde{j}_{\lambda\lambda}|^{1/2} \, |\hat{i}| \, r}{|\hat{S}_{\lambda\lambda}| |\hat{j}|^{1/2} \left(\hat{q}_\psi - \hat{S}_{\psi\lambda} \hat{S}_{\lambda\lambda}^{-1} \hat{q}_\lambda \right)} . \tag{5}$$

In (5), i and j denote respectively the expected and observed information matrices,

$$\begin{aligned}
\hat{S} &= \left(\frac{\partial\theta(\hat{\omega})}{\partial\hat{\theta}} \right)^T \Sigma(\hat{\omega}) \left(\frac{\partial\theta(\tilde{\omega})}{\partial\tilde{\theta}} \right) , \\
\hat{q} &= \left(\frac{\partial\theta(\hat{\omega})}{\partial\hat{\theta}} \right)^T \Sigma(\hat{\omega}) \left(\theta(\hat{\omega}) - \theta(\tilde{\omega}) \right) ,
\end{aligned}$$

and the notation $j_{\lambda\lambda}$, e.g., identifies the observed information matrix once the row and columns corresponding to ψ are deleted. Σ is the variance-covariance matrix of the sufficient statistic $t(Y)$ and can be easily computed by calculating the second order derivatives of the function $K(\theta)$ with respect to the components of θ. Σ is block-diagonal and each block Σ^i, $i = 1, \ldots q$, is a 2×2 matrix with

$$\begin{aligned}
\Sigma_{11}^i(\omega) &= \mathrm{var}_\omega(t_{i1}(Y)) = m_i \sigma^2 V^2(x_i;\beta,\rho) , \\
\Sigma_{12}^i(\omega) &= \Sigma_{21}^i(\omega) = \mathrm{cov}_\omega(t_{i1}(Y), t_{i2}(Y)) = 2m_i \mu(x_i;\beta) \sigma^2 V^2(x_i;\beta,\rho) , \\
\Sigma_{22}^i(\omega) &= \mathrm{var}_\omega(t_{i2}(Y)) = 2m_i \sigma^4 V^4(x_i;\beta,\rho) + 4m_i \mu^2(x_i;\beta) \sigma^2 V^2(x_i;\beta,\rho) .
\end{aligned}$$

Using the preceding derivation, one can easily make his own program for calculating r^*. However, higher-order asymptotics for the class of models (1), can be directly computed using the `nlreg` section of the S-PLUS library `HOA` (2000) (see Example 1 in Section 3, and Brazzale, 2000, Chap. 5, for examples and applications). In the next section we discuss two practical examples of the use of this library for point estimation purposes.

3 APPLICATION EXAMPLES

3.1 Example 1: Weed Data

Nonlinear models are widely used in herbicide bioassays. In a typical experiment a dose-response relation is established by measuring the growth of a plant corresponding to different levels of doses of herbicide. Figure 1 shows the logarithm of the callus area of a tissue culture of *Brassica napus* corresponding to different doses of a sulfonylurea herbicide, metsulfuron methyl. In the following, we will refer to this data set as to the `weed` data. This data set is presented in Seiden et al. (1998, data set M_2) and is also discussed in Bellio et al. (2000). The experiment consists of $q = 8$ doses and 5 replications at each level. The mean response of the model is given by the four parameter log-logistic function

$$\mu(x, \beta) = \log\left[\beta_1 + \frac{\beta_2 - \beta_1}{\{1 + (x/\beta_4)^{\beta_3}\}}\right] .$$

Moreover, the model assumes that the error term ϵ in (1) follows a centered normal distribution with variance function

$$\sigma^2 V^2(x; \beta, \rho) = \sigma^2 (1 + x^{\rho})^2 ,$$

with σ^2, $\rho > 0$. This function is usually called the power of the x (POX) variance function. Other variance functions could also be used (see Carroll and Ruppert, 1988).

(FIGURE 1 here)

The S-PLUS code for analysing this data set through the statistics r and r^* is given in the library `HOA` (2000). In particular, the `nlreg` fitting routine of the `nlreg` section allows to easily fit model (1) by maximum likelihood. The S-PLUS commands to fit the model are:

```
weed.nl <- nlreg( formula = log(y) ~ log(b1+(b2-b1)/(1+(x/b4)^b3)),
          weights=~(1+x^exp(g))^2, hoa=T, data=m2.data,
          start=c(b1=130,b2=2474,b3=1.7,b4=0.07,g=log(0.2)))
```

where, for computational reasons, we have used $\rho = \exp(g)$. The starting values given in `start` have been fixed following the analysis of Bellio et al. (2000).

The `profile` method of the `nlreg` section gives the directed likelihood r and the modified directed likelihood (2). Denoting by ψ the scalar parameter of interest, we can obtain r and r^* as:

```
weed.profs <- profile(weed.nl,offset=psi)
```

Figure 2 gives the plots of the directed likelihood statistic and of the higher-order solution r^* (on the logarithmic scale for the variance parameters), for the six parameters of the nonlinear heteroscedastic model for the `weed` data. Every single plot can be easily obtained as:

```
plot(weed.profs$r,type="l")
lines(weed.profs$rs,lty=6)
abline(h=0,lty=2)
```

parameter	r^* estimate	MLE
β_1	139.148	138.92
β_2	2482.106	2474.00
β_3	1.697	1.707
β_4	0.076	0.077
σ^2	0.026	0.022
ρ	0.288	0.285

Table 1: Point estimates based on r^* and MLE for the parameters of the nonlinear regression model for the `weed` data.

(FIGURE 2 here)

Figure 2 is straightforward to interpret. If a parameter is estimated with good accuracy by maximum likelihood, the corresponding r and r^* profiles are almost linear and very close to each other. It is noticeable from Figure 2 how first-order methods are satisfactory for those parameters that enter the regression function in a linear fashion, like β_1 and β_2. On the other hand, a big difference between r and r^* profiles indicates poor behaviour of the first-order methods. For the `weed` data, this gap is more significant for the variance parameter σ^2. The profile of r^* is shifted rightwards from that of r. This means that the MLE for σ^2 is biased and that the modified directed likelihood corrects for most of this bias.

Point estimates $\hat{\psi}$ and $\hat{\psi}^*$ for the parameter of interest ψ can be easily read off from the plots in Figure 2 as the values of the x-coordinate in the intersection point of the r and r^* functions with the horizontal line $y = 0$. To find the numerical solution to $r^*(\psi) = 0$, i.e. the point estimate for the parameter of interest ψ, we suggest the following S-PLUS procedure:

1. Compute $r^*(\psi)$ for a range of values around the MLE.
2. Interpolate the points by a smoothing method.
3. Invert the interpolating function and find the corresponding value in 0.

Step 1 is directly solved for nonlinear regression models by the `profile` command. Step 2 and 3 can be implemented by means of the following S-PLUS commands:

```
smoother <- smooth.spline(rstar,rstar.range)
estim.star <- predict.smooth.spline(smoother,0)$y
```

where `rstar` is the vector of values of r^* calculated in an appropriate range `rstar.range`. For the `weed` data the S-PLUS commands are:

```
rstarfun <- weed.profs$rs
smoother <- smooth.spline(rstarfun$y, rstarfun$x)
estim.star <- predict.smooth.spline(smoother, x=0)$y
```

Table 1 gives the values of the MLE and of the point estimates based on r^* for the six parameters of the model for the `weed` data.

$\hat{\sigma}^2$			$\hat{\sigma}^{2*}$		
BI	MSE	PU	BI	MSE	PU
-0.0715	0.0407	0.6457	-0.0113	0.0360	0.5029

Table 2: Comparison for the nonlinear regression model for the `weed` data. The parameter of interest is the scale parameter σ^2.

To assess the finite-sample properties of the proposed estimators in comparison with the MLE, a Monte Carlo experiment (based on 10,000 trials) has been performed, being σ^2 the parameter of interest. The estimators are compared in terms of the probability of underestimation (PU), that gives median bias. Moreover, the usual centering and dispersion measures, i.e. bias (BI) and mean square error (MSE), are considered even if they depend on the parameterisation of the model. The results are presented in Table 2, showing the improvement of $\hat{\sigma}^{2*}$ on the MLE.

3.2 Example 2: Calcium Data

A simple example of nonlinear regression model is discussed in Davison and Hinkley (1997, Example 7.7). The data concern the calcium uptake of cells y as a function of time x, after being suspended in a solution of radioactive calcium (Rawlings, 1988, p. 403). For these data, called `calcium` data, a suitable mean function is

$$\mu(x;\beta) = \beta_0\{1 - \exp(-\beta_1 x)\} \ ,$$

where β_0 and β_1 are unknown regression coefficients. Moreover, we have $q = 27$ and $m = 1$, and the error term ϵ is assumed to follow a centered normal distribution with unknown variance σ^2. The data are available in the `calcium` data frame of the S-PLUS library `boot` (Davison and Hinkley, 1997, Chap. 11).

(FIGURE 3 here)

The S-PLUS code for analysing this data set through the statistics r and r^* is given in the library `HOA` (2000). Again, the `nlreg` fitting routine of the `nlreg` section allows to easily fit model by maximum likelihood. Moreover, the `profile` method gives the directed likelihood r and the modified directed likelihood r^*. Figure 4 gives the plots of the directed likelihood statistic and of the higher-order solution r^* (on the logarithmic scale for the variance parameter), for the three parameters of the considered model. As in the previous example, for the variance parameter the profile of r^* is shifted rightwards from that of r. This means that the MLE for σ^2 is biased and that the estimator based on the modified directed likelihood corrects for most of this bias.

(FIGURE 4 here)

Table 3 gives the values of the MLE and of the point estimates based on r^* for the three parameters of the model for the `calcium` data.

Davison and Hinkley (1997, p. 356) consider not only inference on the mean function $\mu(x;\beta)$, but also on other aspects of the model such as the *proportion of maximum*,

$$\pi = 1 - \exp(-\beta_1 x) \ . \tag{6}$$

parameter	r^* estimate	MLE
β_0	4.29	4.31
β_1	0.21	0.20
σ^2	0.307	0.275

Table 3: Point estimates based on r^* and MLE for the parameters of the nonlinear regression model for the `calcium` data.

	$\hat{\psi}$			$\hat{\psi}^*$		
	BI	MSE	PU	BI	MSE	PU
π	0.401	1.94	0.40	0.292	1.67	0.497
σ^2	-0.12	0.10	0.647	-0.01	0.03	0.502

Table 4: Comparison for the nonlinear regression model for the `calcium` data. The parameters of interest are the proportion of maximum π and the variance parameter σ^2.

For the time $x = 15$ they give the estimate $\hat{\pi} = 0.95$. We may obtain the higher-order analogue both by using the equivariance property of the point estimate based on r^* and by reparameterising the mean curve into (β_0, π) and re-computing the point estimate for π. From the estimating equation $r^*(\psi) = 0$, we obtain $\hat{\pi}^* = 0.96$.

To assess the finite-sample properties of the MLE and the estimates based on r^* a Monte Carlo study has been performed, considering π or σ^2 as the parameter of interest. The results are presented in Table 4, showing the improvement of $\hat{\pi}^*$ and $\hat{\sigma}^{2*}$ on the corresponding MLE.

4 FINAL REMARKS

This paper presents a new approach to point estimation when inference about an arbitrary one-dimensional parameter of interest is desired in the context of nonlinear heteroscedastic regression models. This method is generally applicable to nonlinear regression models with normal errors, using the recently developed S-PLUS library `HOA` (2000). The application examples considered show that the proposed estimators improve on the MLE.

For cases when the dimension of the parameter of interest is greater than one, Skovgaard (2001) proposed a generalisation of r^* useful for inference on multidimensional parameters. This is a modification of the usual loglikelihood ratio test improving its small sample properties. However, the use of this new statistic for point estimation purposes requires further investigation. The resulting estimators could be computationally very intensive, since the evaluation of the statistic on a large grid of values is required.

Acknowledgements

We wish to thank A.R. Brazzale and R. Bellio for their developing the S-PLUS library `HOA` with which the data were analysed. This work has been partially supported by grants from Ministero dell'Università e della Ricerca Scientifica e Tecnologica, Italy.

References

Barndorff-Nielsen, O.E. (1991) Modified signed log likelihood ratio. *Biometrika*, **78**, 557–564.

Barndorff-Nielsen, O.E., Cox, D.R. (1994) *Inference and Asymptotics.* Chapman & Hall, London.

Bellio, R. (2000) *Likelihood Asymptotics: Applications in Biostatistics.* Ph.D. Thesis, Department of Statistics, University of Padova.

Bellio, R., Brazzale, A.R. (1999) Higher-order asymptotics in nonlinear regression. In *Proceedings of the 14th International Workshop on Statistical Modelling*, Graz, July 19–23, 1999, 440–443.

Bellio, R., Jensen, J.E., Seiden, P. (2000) Applications of likelihood asymptotics for nonlinear regression in herbicide bioassays, *Biometrics*, **56**, 1204–1212.

Brazzale, A.R. (2000) *Practical Small-Sample Parametric Inference.* Ph.D Thesis, N. 2230, Department of Mathematics, Swiss Federal Institute of Technology Lausanne.

Carroll, R.J., Ruppert, D. (1988) *Transformation and Weighting in Regression.* Chapman & Hall, London.

Davison, A.C., Hinkley, D.V. (1997) *Bootstrap Methods and Their Application.* Cambridge University Press, Cambridge.

Giummolè, F., Ventura, L. (2001) Practical point estimation from higher-order pivots, *J. Statist. Comput. Simul.*, to appear.

HOA (2000) *Higher-Order Asymptotics.* A library of S-PLUS functions (and data) for higher-order asymptotic inference written by A.R. Brazzale and R. Bellio, `http://statwww.epfl.ch/people/brazzale/lib.shtml`.

Pace, L., Salvan, A. (1999) Point estimation based on confidence intervals: exponential families. *J. Statist. Comput. Simul.*, **64**, 1–21.

Rawlings, J.O. (1988) *Applied Regression Analysis: A Research Tool.* Pacific Grove, Wadsworth Brooks/Cole, California.

Schwenke, J.R., Milliken, G. (1991) On the calibration problem extended to nonlinear models. *Biometrics*, **47**, 563–574.

Seber, G.A.F., Wild, C.J. (1989) *Nonlinear Regression.* Wiley, New York.

Seiden, P., Kappel, D., Streibig, J.C. (1998) Response of *Brassica napus* L. tissue culture to metsulfuron methyl and chlorsulfuron. *Weed Research*, **38**, 221–228

Severini, T.A. (1998) An approximation to the modified profile likelihood function. *Biometrika*, **85**, 403–411.

Severini, T.A. (1999) An empirical adjustment to the likelihood ratio statistic. *Biometrika*, **86**, 235–247.

Skovgaard, I.M. (1996) A general large deviation approximation to one-parameter tests. *Bernoulli*, **2**, 145–165.

Skovgaard, I.M. (2001) Likelihood asymptotics. *Scand. J. Statist.*, **28**, 3–2.

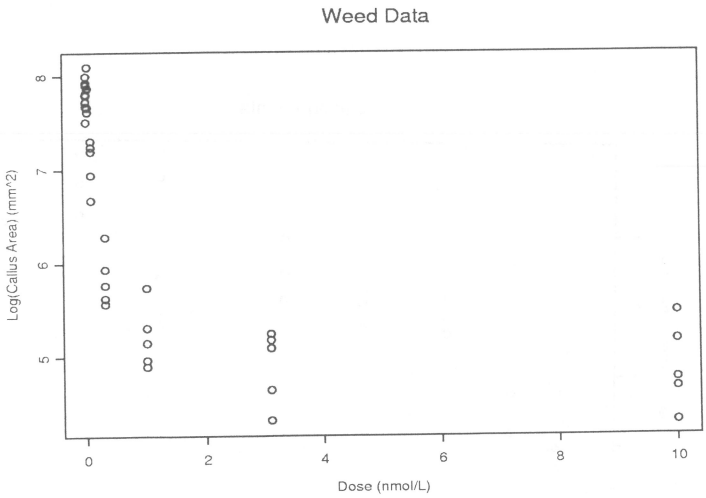

Figure 1: Plot of the weed data.

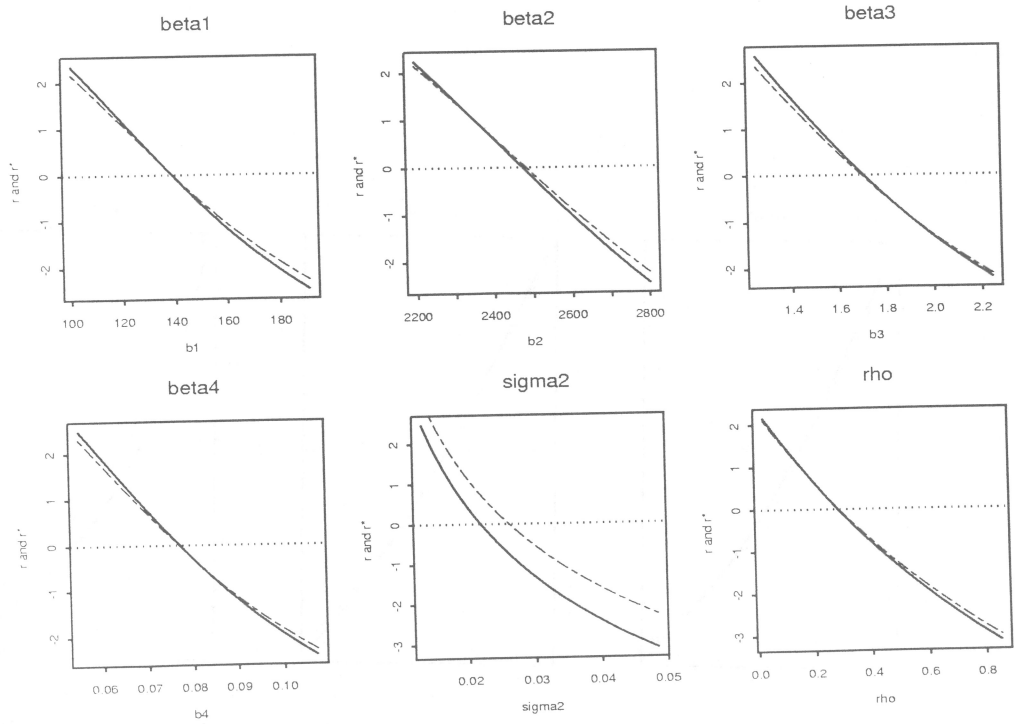

Figure 2: Plot of the directed likelihood (solid line) and of the modified directed likelihood (dashed line) for the weed data.

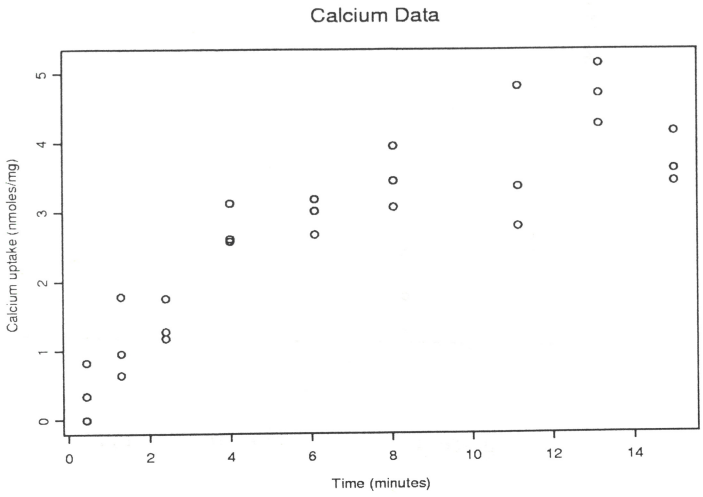

Figure 3: Plot of the `calcium` data.

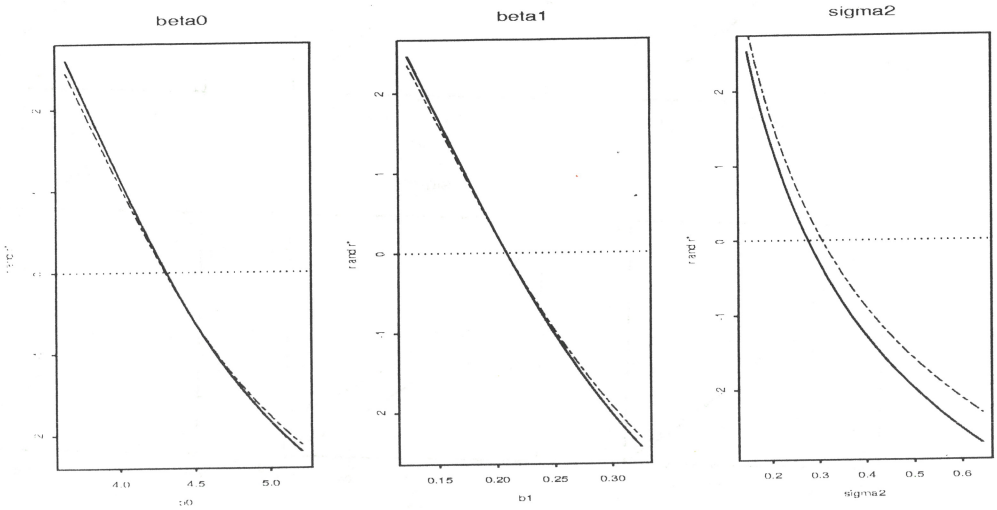

Figure 4: Plot of the directed likelihood (solid line) and of the modified directed likelihood (dashed line) for the `calcium` data.