Statistica Sinica 4(1994), 187-198

RESAMPLING-BASED ESTIMATOR IN NONLINEAR REGRESSION

Jie Mong and Xueren Wang

Yunnan University

Abstract: In this paper we suggest a resampling-based estimator (RSE) of nonlinear regression by using Wu's (1986) resampling idea. The RSE is bias-reducing without increasing the variance. Some examples are given using the data of Ratkowsky (1983), and several simulations are also presented.

Key words and phrases: Absolute relative bias, bias reduction, Jackknife-based estimator, nonlinear regression, resampling-based estimator.

1. Introduction

Nonlinear regression models have wide application in a variety of contexts, and much of the theory is established. The usual model has the form

$$y_i = f(x_i, \theta) + \epsilon_i \quad (i = 1, \ldots, n), \tag{1}$$

where f is a nonlinear function of the p-dimensional parameter θ and assumed twice continuously differentiable in θ, and the errors ϵ_i are assumed to be independent and identically distributed random variables with mean 0 and unknown variance σ^2. For this model, the least square estimator (LSE) $\hat{\theta}$ of θ can be obtained by minimizing the objective function

$$J(\theta) = \sum_{i=1}^{n} \{y_i - f(x_i, \theta)\}^2 \tag{2}$$

or

$$\sum_{i=1}^{n} \frac{\partial f_i}{\partial \theta}(y_i - f(x_i, \theta))\|_{\theta=\hat{\theta}} = 0. \tag{3}$$

The estimator of σ^2 is $\hat{\sigma}^2 = J(\hat{\theta})/(n-p)$. Let $f_i = f(x_i, \theta)$ $(i = 1, \ldots, n)$, V denote the $n \times p$ matrix with elements

$$f_i^r = \frac{\partial f_i}{\partial \theta_r} \quad (i = 1, \ldots, n, \quad r = 1, \ldots, p).$$

W denote the $n \times p \times p$ array with elements

$$f_i^{rs} = \frac{\partial^2 f_i}{\partial\theta_r \partial\theta_s} \quad (i = 1, \ldots, n, \quad r, s = 1, \ldots, p)$$

and e denote the residual vector with elements

$$e_i = y_i - f(x_i, \hat{\theta}).$$

Then, Equation (3) can be written as

$$V^T(\hat{\theta})e = 0. \tag{4}$$

2. Bias of $\hat{\theta}$ and Jackknife-Based Estimators

The LSE of the parameters in the model (1) are generally biased estimators of the true parameter values. The bias $\text{Bias}(\hat{\theta}) = E(\hat{\theta} - \theta)$ can be expressed as

$$\text{Bias}(\hat{\theta}) \approx -\frac{\sigma^2}{2}(V^T V)^{-1} V^T \text{tr}\{(V^T V)^{-1} W\} \tag{5}$$

where $\text{tr}\{(V^TV)^{-1}W\}$ is an $n \times 1$ vector with elements $\text{tr}\{(V^TV)^{-1}W_i\}(i = 1, \ldots, n)$ and W_i, is the ith face of W, a $p \times p$ matrix with elements

$$f_i^{rs} = \frac{\partial^2 f_i}{\partial\theta_r \partial\theta_s} \quad (r, s = 1, \ldots, p).$$

Bates and Watts (1980) pointed out that the bias of the LSE (5) is only a function of the parameter-effects curvature array and thus could be reduced or eliminated by reparameterization. For the single-parameter case, Hougaard (1982) showed that there is a transformation which eliminates bias, but the appropriate transformation does not exist in general for the multiparameter case (Hougaard (1984)).

Many Jackknife-based estimators are supposed to construct confidence regions in nonlinear regression, but, these estimators are not valid in bias-reduction as shown in Simonoff and Tsai (1986).

The linear Jackknife was developed by Fox et al. (1980). With pseudo-values defined as

$$LP_i = \hat{\theta} + (n(\hat{V}^T\hat{V})^{-1}\hat{V}_i e_i)/(1 - \hat{h}_{ii}) \tag{6}$$

where $\hat{V}$ is the values of V evaluated at $\hat{\theta}, \hat{V}_i$ is the ith row of $\hat{V}$, and $\hat{h}_{ii}$ is the ith diagonal element of the matrix $\hat{V}(\hat{V}^T\hat{V})^{-1}\hat{V}^T$, the linear Jackknife estimator $\hat{\theta}_{LP}$ is defined as the average of the pseudo-values

$$\hat{\theta}_{LP} = \frac{1}{n}\sum_{i=1}^{n} LP_i. \tag{7}$$

By extending the work of Hinkley (1977) to nonlinear regression, Fox et al. (1980) proposed a weighted linear Jackknife estimator to improve the bias reduction. Weighted pseudo-values are given by

$$LQ_i = \hat{\theta} + n(\hat{V}^T\hat{V})^{-1}\hat{V}_i e_i \tag{8}$$

and replacing LP_i by LQ_i in (7) yields the weighted linear Jackknife estimator. The linear Jackknife estimator and the weighted linear Jackknife estimator do not have the first-order bias term eliminated; this is due to using a linear approximation and to lack of design balance.

Simonoff and Tsai (1986) proposed three new Jackknife-based estimators. The pseudo-values are given by

$$\begin{aligned} MLP_i &= \hat{\theta} + nT_i^{-1}\hat{V}_i e_i/(1 - h_{ii}^*) && (9) \\ MLQ_i &= \hat{\theta} + nT_i^{-1}\hat{V}_i e_i && (10) \\ RLQ_i &= \hat{\theta} + n(\hat{V}^T\hat{V})^{-1}\hat{V}_i e_i(1 - \hat{h}_{ii}) && (11) \end{aligned}$$

where $T_i = \{\hat{V}^T\hat{V} - [e_{(i)}][\hat{W}_{(i)}]\}$, $h_{ii}^* = \hat{V}_i^T T_i \hat{V}_i, \hat{V}$, $\hat{W}$ are all evaluated at $\hat{\theta}, e_{(i)}$ and $W_{(i)}$ are respectively the $(n-1)\times 1$ vector and $(n-1)\times P\times P$ array with the ith component removed from the $n\times 1$ vector e and the $n\times p\times p$ array $\hat{W}$. If the term $[e_{(i)}][W_{(i)}]$ is ignored, then (9) and (10) are identical to (6) and (8), respectively. These Jackknife-based estimators do not reduce the bias of the LSE because of the linear approximation in producing the pseudo-values.

In this article we suggest an alternative weighted scheme and quadratic approximation whereby the bias reduction properties hold, and the first-order bias term of the LSE is eliminated. Some examples are presented to illustrate this property.

3. Resampling-Based Estimator (RSE)

Let $s = \{i_1, \ldots, i_r\}$ be a subset of $\{1, \ldots, n\}$, and $\hat{\theta}_s$ be the LSE of Model (1) based on the (x_i, y_i) in s, i.e. $\hat{\theta}_s$ can be obtained by minimizing the objective function,

$$J_s(\theta) = \sum_{i\in s}\{y_i - f(x_i, \theta)\}^2.$$

For a linear model, $\hat{\theta}_s$ can be expressed as

$$\hat{\theta}_s = (V_s^T V_s)^{-1} V_s^T y_s \tag{12}$$

where $y_s = (y_{i_1}, \ldots, y_{i_r})^T$ and $V_s = (V_{i_1}, \ldots, V_{i_r})$, V_{i_j} is the i_jth row of V. The following results are based on Wu (1986).

Result. For any $r \geq p$, we have

$$\binom{n-p}{r-p}|V^TV| = \sum_r |V_s^TV_s| \tag{13}$$

where $\sum_r$ denotes the summation over all the subsets of size r, and when the model (1) is linear, the LSE can be expressed as

$$\hat{\theta} = (V^TV)^{-1}V^Ty = \sum_r \omega_s\hat{\theta}_s, \tag{14}$$

where ω_s is a weight defined as

$$\omega_s \propto |V_s^TV_s| \qquad \sum_r \omega_s = 1.$$

For the nonlinear model (1), we suggest a resampling-based estimator (RSE). The RSE is given by

$$\hat{\theta}_{rs} = \hat{\theta} + \frac{r-p+1}{n-r}\sum_r \hat{\omega}_s(\hat{\theta} - \hat{\theta}_s), \tag{15}$$

where $\hat{\omega}_s$ are evaluated at $\hat{\theta}$ to reduce the computation of RSE. We construct a Taylor series approximation to $\hat{\theta}_s$, yielding

$$\hat{\theta}_s = \hat{\theta} + (\hat{V}_s^T\hat{V}_s)^{-1}\hat{V}_s^Te_s - \frac{1}{2}(\hat{V}_s^T\hat{V}_s)^{-1}\hat{V}_s^T[\Delta\theta_{1s}^T\hat{W}_s\Delta\theta_{1s}] \tag{16}$$

where $\hat{W}_s = (\hat{W}_{i_1}, \ldots, \hat{W}_{i_r})^T$ and $\hat{V}_s$ are evaluated at $\hat{\theta}$, and $\Delta\theta_{1s} = (\hat{V}_s^T\hat{V}_s)^{-1}\hat{V}_s^Te_s$. Substituting $\hat{\theta}_s$ into (15), and utilizing (12) and (14), yields

$$\begin{aligned}\hat{\theta}_{rs} &= \hat{\theta} - \frac{r-p+1}{n-r}(\hat{V}^T\hat{V})^{-1}\hat{V}^Te \\ &+ \frac{r-p+1}{2(n-r)}\sum_r \hat{\omega}_s(\hat{V}_s^TV_s)^{-1}\hat{V}_s^T[\Delta\theta_{1s}^TW_s\Delta\theta_{1s}];\end{aligned} \tag{17}$$

and from (4), we have

$$\hat{\theta}_{rs} = \hat{\theta} + \frac{r-p+1}{2(n-r)}\sum_r \hat{\omega}_s(\hat{V}_s^T\hat{V}_s)^{-1}\hat{V}_s^T[\Delta\theta_{1s}^T\hat{W}_s\Delta\theta_{1s}]. \tag{18}$$

In the remainder of this paper, the RSE is defined as (18).

Theorem. *For $r = n-1$, and under some appropriate regularity conditions (Mong (1988))*

$$\begin{aligned}\text{(i) Bias}(\hat\theta_{rs}) &= (n-p+1)\text{Bias}(\hat\theta) \\ &+ \frac{\sigma^2}{2}\sum_r \frac{|V_s^T V_s|(V_s^T V_s)^{-1}V_s^T \text{tr}[(V_s^T V_s)^{-1}W_s]}{|V^T V|} + O(n^{-3/2}) \qquad (19)\\ &= O(n^{-3/2}) \qquad (20)\end{aligned}$$

$$\text{(ii) Var}(\hat\theta_{rs}) = \text{Var}(\hat\theta) + O(n^{-3/2}) \qquad (21)$$

i.e. *the RSE eliminates the first-order bias term of the LSE, and the variance of the RSE and LSE have the same first-order term.*

Proof.

(i)

$$\text{Bias}(\hat\theta_{rs}) = \text{Bias}(\hat\theta) + E\frac{n-p}{2}\sum_r \hat\omega_s(\hat V_s^T \hat V_s)^{-1}\hat V_s^T[\Delta\theta_{1s}^T \hat W_s \Delta\theta_{1s}]. \qquad (22)$$

Let

$$\Delta = E\frac{n-p}{2}\sum_r \hat\omega_s(\hat V_s^T \hat V_s)^{-1}V_s^T[\Delta\theta_{1s}^T \hat W_s \Delta\theta_{1s}]. \qquad (23)$$

Since $\hat\theta = \theta + O_p(n^{-1/2})$, substituting into (23) yields

$$\Delta = \frac{n-p}{2}\sum_r \omega_s(V_s^T V_s)^{-1}V_s^T \text{tr}\Big[W_s(V_s^T V_s)^{-1}V_s^T E e_s e_s^T V_s(V_s^T V_s)^{-1}\Big] + O(n^{-3/2}). \qquad (24)$$

Note that

$$\begin{aligned} e_s &= y_s - f(x_s,\theta) \\ &= \epsilon_s - V_s^T(V^T V)^{-1}V^T\epsilon + O_p(n^{-1}). \end{aligned}$$

We have

$$\Delta = \frac{\sigma^2}{2}(n-p)\sum_r \omega_s(V_s^T V_s)^{-1}V_s^T \text{tr}\Big[W_s((V_s^T V_s)^{-1} - (V^T V)^{-1})\Big] + O(n^{-3/2}). \qquad (25)$$

From (14), Δ can be described as

$$\Delta = \frac{\sigma^2(n-p)}{2}\sum_r \omega_s(V_s^T V_s)^{-1}V_s^T \text{tr}[W_s(V_s^T V_s)^{-1}] + (n-p)\text{Bias}(\hat\theta) + O(n^{-3/2}). \qquad (26)$$

Now, Equation (19) follows from (26) and (22). Using (25), another expression can be obtained:

$$\Delta = \frac{\sigma^2}{2}\sum_r (V_s^T V_s)^{-1} V_s^T \mathrm{tr}\Big[W_s (V^T V)^{-1} V_i V_i^T (V^T V)^{-1}\Big] + O(n^{-3/2}) \tag{27}$$

since $|V_s^T V_s| = |V^T V|(1-h_{ii})$ and $(V_s^T V_s)^{-1} - (V^T V)^{-1} = \frac{(V^T V)^{-1} V_i V_i^T (V^T V)^{-1}}{1-h_{ii}}$. Replacing $(V_s^T V_s)^{-1}$ by $(V^T V)^{-1}$ in (27), the difference can be ignored. Thus,

$$\begin{aligned}\Delta &= \frac{\sigma^2}{2}\sum_r (V^T V)^{-1} V_s^T \mathrm{tr}\Big[W_s (V^T V)^{-1} V_i V_i^T (V^T V)^{-1}\Big] + O(n^{-3/2}) \\ &= \frac{\sigma^2}{2}(V^T V)^{-1} V^T \mathrm{tr}[W (V^T V)^{-1}] \\ &- \frac{\sigma^2}{2}\sum_r (V^T V)^{-1} V_i \mathrm{tr}\Big[W_i (V^T V)^{-1} V_i V_i^T (V^T V)^{-1}\Big] + O(n^{-3/2}) \\ &= -\mathrm{Bias}(\hat\theta) + O(n^{-3/2}). \end{aligned} \tag{28}$$

From (28), (22) and (2) Equation (20) is obtained immediately.

(ii) Under some regularity conditions, the RSE can be written as

$$\hat\theta_{rs} = \hat\theta + O_p(n^{-1})$$

and it can be proved that

$$\mathrm{Var}(\hat\theta_{rs}) = \mathrm{Var}(\hat\theta) + O(n^{-3/2}).$$

The details can be found in Mong (1988).

The theorem can be established when $n-r$ is bounded. More information is available in Mong (1988).

4. Examples

In this section several examples are presented to illustrate the result of Section 3. The biases, given in following Tables are defined as

$$\%\mathrm{Bias}(\hat\theta) = \frac{\mathrm{Bias}(\hat\theta)}{\hat\theta} \times 100\% \tag{29}$$

$$\%\mathrm{Bias}(\hat\theta_{rs}) = \frac{\mathrm{Bias}(\hat\theta_{rs})}{\hat\theta_{rs}} \times 100\%. \tag{30}$$

These examples were all analyzed by Ratkowsky (1983). The first example is the Gompertz model, with the form

$$y = \alpha \exp\{-\exp(\beta - rX)\}. \tag{31}$$

The example involves two sets of data, based on data sets 2 and 4 in Table 4 of Ratkowsky (1983). The parameter estimators, with associated bias and parameter-effects curvature, appear in Table 1. For data set 1, the LSE's biases of the two parameters β and γ are serious, but for RSE the biases of parameters β and γ are reduced. Data set 2 also shows that the bias reduction property holds, and in the two data sets, the parameter-effects curvature for LSE and RSE are quite close.

Table 1. LSE and RSE (Model (31)), with associated estimated biases and parameter-effects curvature (PE)

	Data 1				Data 2			
	α	β	γ	PE	α	β	γ	PE
LSE	723.1	2.5	0.45	0.70	22.51	2.106	0.388	0.880
%Bias	0.185	1.088	1.007		0.271	1.066	0.976	
RSE	721.539	2.4881	0.446	0.71	22.441	2.086	0.384	0.896
%Bias	−0.077	0.027	0.0116		−0.112	0.066	0.0177	

The second example is based on the model

$$y = \theta_1 + \theta_2(x - \theta_4) + \theta_3\left\{(x - \theta_4)^2 + \theta_5\right\}^{1/2} \tag{32}$$

and data set 3 in Table 6.18 of Ratkowsky (1983). Ratkowsky (1983) showed that the most effective method to reduce the bias of the LSE is reparameterization. However, it is difficult to find an appropriate transformation. We propose a transformation for Model (32). Let $\theta_5' = \theta_5^{1/2}$,

$$y = \theta_1 + \theta_2(x - \theta_4) + \theta_3\left\{(x - \theta_4)^2 + {\theta_5'}^2\right\}^{1/2}. \tag{33}$$

The results for models (32)and (33) are presented in Table 2. The calculations indicate that the transformation has reduced the bias of LSE, but it is still serious. The bias of RSE is the only one that can be ignored.

Table 2. LSE and RSE (Models (32) and (33)), with associated estimated biases and parameter-effects curvature (PE)

	LSE	%Bias(θ)	RSE	%Bias(θ_{rs})	LSE*	%Bias(θ)
θ_1	136.822	0.124	136.501	0.060	136.822	0.124
θ_2	0.696	−0.996	0.713	.016	0.696	−0.996
θ_3	−0.587	1.761	−0.566	.007	−0.587	1.761
θ_4	18.752	0.330	18.609	.004	18.752	0.330
θ_5	6.727	20.521	4.627	.750	2.594	1.609
PE	1.812		1.893		4.497	

*Note: The LSE and biases of parameters in Model (30)

For the third example, we consider the Morgan-Mercer-Flodin model (Ratkowsky (1983)),

$$y = \frac{\beta\gamma + \alpha x^\delta}{\gamma + x^\delta}. \tag{34}$$

Two data sets are used to fit this model; the results are shown in Table 3. This example also shows that RSE is an effective bias-reducing estimator.

Table 3. LSE and RSE (Model (34)), with associated estimated biases

	Data 1				Data 2			
	α	β	γ	δ	α	β	γ	δ
LSE	723.9	33.35	6266	4.641	22.08	1.653	5586	4.56
%Bias	0.313	–2.468	86.600	1.424	0.237	–0.794	55.327	0.942
RSE	722.218	33.546	3714.59	4.592	22.006	1.663	2238.86	4.510
%Bias	–0.02	0.836	0.432	0.053	–0.034	0.771	–0.078	0.005

5. Simulation

In this section, several simulations are provided to examine the properties of RSE, LSE and other Jackknife-based estimators.

Suppose that $\hat{\theta}$ and $\hat{\sigma}^2$ are treated as the true values of θ, σ^2 and M sets of size-n samples are simulated from

$$Y_i = f(x_i, \hat{\theta}) + \epsilon_i \quad (i = 1, \ldots, n)$$

where ϵ_i are i.i.d. $N(0, \hat{\sigma}^2)$. For each set we obtain an estimator $\tilde{\theta}^{(m)}$ $(m = 1, \ldots, M)$. We consider the estimation characteristics in regard to bias, variance, skewness and kurtosis. They are defined as follows:

$$\begin{aligned}
\tilde{\theta}_i^{(\cdot)} &= \sum_{m=1}^{M} \tilde{\theta}_i^{(m)}/M, \\
\text{Bias}_i &= \tilde{\theta}_i^{(\cdot)} - \hat{\theta}_i, \\
\text{Var}_i &= \frac{1}{M}\sum_{m=1}^{M}(\tilde{\theta}_i^{(m)} - \tilde{\theta}_i^{(\cdot)})^2, \\
g_{1i} &= \frac{M^{1/2}\sum_{m=1}^{M}\left(\tilde{\theta}_i^{(m)} - \tilde{\theta}_i^{(\cdot)}\right)^3}{\left(\sum_{m=1}^{M}\left(\tilde{\theta}_i^{(m)} - \tilde{\theta}_i^{(\cdot)}\right)^2\right)^{3/2}}, \\
g_{2i} &= \frac{M\sum_{m=1}^{M}\left(\tilde{\theta}_i^{(m)} - \tilde{\theta}_i^{(\cdot)}\right)^4}{\left(\sum_{m=1}^{M}\left(\tilde{\theta}_i^{(m)} - \tilde{\theta}_i^{(\cdot)}\right)^2\right)^2} - 3.
\end{aligned}$$

The absolute relative biases (ARB) of LSE and RSE are also presented, they are computed as:

$$\text{ARB}_i = \|\text{Bias}_i\|/S.E._{\cdot i},$$

where $S.E._{\cdot i} = \sqrt{v_i/M}$, v_i is a linear approximation variance of LSE and RSE. Let $C = \hat{\sigma}^2(V^TV)^{-1}$; then $v_i = C_{ii}$.

The first simulation set is based on the Gompertz model (Section 4, Model (31)). The simulation results (Table 4) show that simulation biases and variance of LSE are very close to the theory values, and the results also quantify the bias-reduction property of RSE. The simulation variances of RSE are close to the variances of LSE. This supports the theory of theorem part (ii); RSE is less biased without increasing variance and has lower ARB.

The second simulation is based on Duncan's (1978) example. The model is

$$Y = \frac{\theta_1}{\theta_1 - \theta_2}(e^{-\theta_2 x} - e^{-\theta_1 x}). \tag{35}$$

All of the estimators are quite close as regards bias, variance, ARB, skewness and kurtosis, and RSE is still a good estimator, with lower bias, variance and ARB. (Table 5)

The final simulation is based on a chemical-reaction rate model from Meyer and Roth (1972, Example 1) previously analyzed by Rotkowsky (1983, Sec. 6.3) and Simonoff and Tsai (1986). The model is:

$$Y = \frac{\theta_1\theta_3 x_1}{1 + \theta_1 x_1 + \theta_2 x_2}. \tag{36}$$

The results of simulation are presented in Table 6. There are some differences in this example; Only $\hat{\theta}, \hat{\theta}_{RLQ}$ and $\hat{\theta}_{rs}$ perform well. Although the biases of $\hat{\theta}_{LP}$ are close to those of LSE, the variances of $\hat{\theta}_{LP}$ are much bigger than the variances of LSE. Both the biases and variances of $\hat{\theta}_{MLP}$ and $\hat{\theta}_{MLQ}$ are much larger than those of LSE. The RSE is still good at bias-reducing and yielding lower variance.

In conclusion, RSE is effective as a bias-reducing estimator without increasing the variance. In addition, RSE could be applied to construct confidence regions.

Table 4. Simulation results for the Gompertz Model

		LSE	RSE	LP	RLQ	MLQ	MLP
Data set 1.							
Bias	θ_1	2.3211	0.9883	2.4441	2.2495	0.1231	-0.6779
(%Bias)		(0.3210)	(0.1367)	(0.3380)	(0.3111)	(0.0170)	(-0.0937)
	θ_2	0.0275	0.0010	0.0236	0.0294	0.0603	0.0885
		(1.0983)	(0.0396)	(0.9435)	(1.1855)	(2.4126)	(3.5390)
	θ_3	0.0041	-0.0003	0.0034	0.0044	0.0102	0.0150
		(0.8998)	(-0.0733)	(0.7598)	(0.9798)	(2.2662)	(3.3418)
ARB	θ_1	3.3271	1.4167				
	θ_2	3.0502	0.1100				
	θ_3	2.4739	0.2016				
Var	θ_1	500.8908	493.9358	505.5539	501.9933	507.8207	548.6551
	θ_2	0.0866	0.0839	0.0884	0.0873	0.0926	0.1128
	θ_3	0.0029	0.0028	0.0029	0.0029	0.0031	0.0037
Ske.	θ_1	0.5193	0.5266	0.5313	0.5130	0.5225	0.4847
	θ_2	0.4669	0.4660	0.5843	0.4277	0.5033	0.8462
	θ_3	0.3583	0.3550	0.4370	0.3376	0.3786	0.6463
Kur.	θ_1	0.6589	0.6841	0.6802	0.6419	0.6649	0.7686
	θ_2	0.7583	0.7715	1.1560	2.5605	2.5887	2.5012
	θ_3	0.5082	0.5153	0.7781	0.4094	0.5308	1.3310
Data set 2.							
Bias	θ_1	0.0934	0.0313	0.0979	0.0905	0.0111	0.0162
(%Bias)		(0.4150)	(0.1393)	(0.4349)	(0.4023)	(0.0491)	(0.0719)
	θ_2	0.0247	0.0030	0.0221	0.0262	0.0498	0.0643
		(1.1703)	(0.1405)	(1.0502)	(1.2415)	(2.3620)	(3.0528)
	θ_3	0.0037	0.0001	0.0033	0.0040	0.0088	0.0111
		(0.9533)	(0.0180)	(0.8399)	(1.0203)	(2.2545)	(2.8548)
ARB	θ_1	3.5281	1.1837				
	θ_2	3.3154	0.3981				
	θ_3	2.5472	0.0482				
Var	θ_1	0.7520	0.7333	0.7610	0.7558	0.7707	0.9054
	θ_2	0.0590	0.0574	0.0598	0.0593	0.0635	0.0762
	θ_3	0.0023	0.0022	0.0023	0.0023	0.0024	0.0029
Ske.	θ_1	0.5934	0.5998	0.6058	0.5907	0.6171	0.7189
	θ_2	0.4696	0.4711	0.5690	0.4299	0.5054	0.7560
	θ_3	0.3530	0.3525	0.4150	0.3296	0.3606	0.4797
Kur.	θ_1	0.6613	0.6940	0.6709	0.6571	0.7305	1.1844
	θ_2	0.6832	0.6956	1.0212	0.5500	2.1706	2.1092
	θ_3	0.5166	0.5282	0.7601	0.4159	0.5385	1.1274

Table 5. Simulation results for Duncan's Model

		LSE	RSE	LP	RLQ	MLQ	MLP
Bias	θ_1	0.0017	−0.0001	0.0016	0.0017	0.0051	0.0051
(%Bias)		(0.7798)	(−0.0567)	(0.7656)	(0.7940)	(2.4103)	(2.4103)
	θ_2	0.0024	−0.0002	0.0023	0.0025	0.0141	0.0167
		(0.5335)	(−0.0448)	(0.5089)	(0.5515)	(3.1585)	(3.7436)
ARB	θ_1	2.0871	0.1518				
	θ_2	1.4598	0.1227				
Var	θ_1	6.2285	6.1741	6.2660	6.2158	6.3399	6.4061
$\times 10^4$	θ_2	29.7200	29.2228	29.7699	29.7606	32.7239	33.5404
Ske.	θ_1	0.2478	0.2502	0.2457	0.2497	0.2452	0.2437
	θ_2	0.4529	0.4566	0.4546	0.4501	0.4605	0.4613
Kur.	θ_1	0.0551	0.0615	0.0284	0.0749	0.0386	0.0033
	θ_2	0.1951	0.2089	0.1939	0.1936	0.1925	0.1955

Table 6. Simulation results for Model (36)

		LSE	RSE	LP	RLQ	MLQ	MLP
Bias	θ_1	0.0171	−0.0167	0.1565	0.0186	−2.8677	−0.2644
(%Bias)		(0.5451)	(−0.5333)	(4.9960)	(0.5927)	(−91.5756)	(−8.4426)
	θ_2	0.0373	0.0102	0.8352	0.0379	−0.5020	−10.6552
		(0.2461)	(0.0670)	(5.5097)	(0.2501)	(−3.3114)	(−70.2881)
	θ_3	0.0401	0.0009	0.0489	0.0397	0.6719	−0.6768
		(5.1342)	(0.1141)	(6.2700)	(5.0932)	(86.1280)	(−86.7651)
ARB	θ_1	0.6678	0.6533				
	θ_2	1.8673	0.5086				
	θ_3	8.3329	0.1852				
Var	θ_1	0.6344	0.6204	60.7894	0.6425	4710.9600	7916.0800
	θ_2	0.3630	0.3601	958.1200	0.3660	163.9922	4884.8700
	θ_3	0.0336	0.0274	0.6649	0.0338	204.4682	843.0529
Ske.	θ_1	0.2984	0.2994	0.0684	0.3197	−30.2997	29.0962
	θ_2	0.1650	0.1668	−0.0857	0.1839	−29.8790	−22.4977
	θ_3	1.5195	1.3416	1.5660	1.5162	28.6021	−31.1751
Kur.	θ_1	0.1830	0.1992	0.8167	0.2638	938.9550	899.5855
	θ_2	−0.0073	−0.0348	−0.0857	0.0556	920.8717	593.3932
	θ_3	3.7300	3.2149	8.3050	3.7577	852.1255	978.6693

References

Bates, D. M. and Watts, D. G. (1980). Relative curvature measures of nonlinearity. *J. Roy. Statist. Soc. Ser.B* **42**, 1–25.

Duncan, G. T. (1978). An empirical study of Jackknife-constructed confidence regions in nonlinear regression. *Technometrics* **20**, 123–129.

Fox, T., Hinkley, D. and Larntz, K. (1980). Jackknifing in nonlinear regression. *Technometrics* **22**, 29–33.

Hinkley, D. V. (1977). Jackknifing in unbalanced situations. *Technometrics* **19**, 285–292.

Hougaard, P. (1982). Parametrizations of non-linear models. *J. Roy. Statist. Soc. Ser.B* **44**, 244–252.

Hougaard, P. (1984). Parameter transformation in multiparameter nonlinear regression models. Technical Report 2, University of Copenhagen, Institute of Mathematical Statistics.

Meyer, R. R. and Roth, P. M. (1972). Modified damped least squares. *J. Inst. Math. Appl.* **9**, 218.

Miller, R. G. (1974). An unbalanced Jackknife. *Ann. Statist.* **2**, 880–891.

Mong, J. (1988). RSE in nonlinear regression. Technical Report, Yunnan Univ. (In Chinese).

Simonoff, J. S. and Tsai, C. L. (1986). Jackknife-based estimators and confidence regions in nonlinear regression. *Technometrics* **28**, 103–112.

Ratkowsky, D. A. (1983). *Nonlinear Regression Modeling.* Marcel Dekker, New York.

Wu, C. F. J. (1986). Jackknife, Bootstrap and other resampling methods in regression analysis. *Ann. Statist.* **14**, 1261–1295.

Institute of Applied Statistics, Yunnan University, Kunming 650091.

(Received March 1992; accepted August 1993)