

Research Article

A Variance Shift Model for Detection of Outliers in the Linear Measurement Error Model

Babak Babadi,¹ Abdolrahman Rasekh,¹ Ali Akbar Rasekhi,² Karim Zare,³ and Mohammad Reza Zadkarami¹

¹ Department of Statistics, Shahid Chamran University, Ahvaz, Iran

² Department of Biostatistics, Tarbiat Modares University, Tehran, Iran

³ Department of Statistics, Islamic Azad University, Science and Research Branch, Fars, Iran

Correspondence should be addressed to Babak Babadi; b.babadi@scu.ac.ir

Received 27 May 2014; Accepted 5 August 2014; Published 14 September 2014

Academic Editor: Allan Peterson

Copyright © 2014 Babak Babadi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We present a variance shift model for a linear measurement error model using the corrected likelihood of Nakamura (1990). This model assumes that a single outlier arises from an observation with inflated variance. The corrected likelihood ratio and the score test statistics are proposed to determine whether the i th observation has an inflated variance. A parametric bootstrap procedure is used to obtain empirical distributions of the test statistics and a simulation study has been used to show the performance of proposed tests. Finally, a real data example is given for illustration.

1. Introduction

Outliers are observations that appear inconsistent with the rest of the data set and can have a profound destructive influence on the statistical analysis. To detect these kinds of observations in the linear models, different approaches have been suggested, among those one can refer to the case-deletion and variance shift models. The first approach is based on the assumption that outliers result from a shift in the mean of contaminated observations (see Barnett and Lewis [1] or Weisberg [2]) and the second procedure takes into account the assumption that an outlier arises from an error term with an increased variance (Cook and Weisberg [3]). Cook et al. [4] indicated that the maximum likelihood estimates for the position of the outlier under two methods could be different, unless the largest absolute studentized residual corresponds to the largest absolute residual. Using the residual maximum likelihood (REML), Thompson [5] showed that the residual variance and outlier position are the same under both methods.

In the linear mixed models, case deletion method, variance shift outlier model, and related diagnostics are studied widely by different authors. Christensen et al. [6] presented case deletion diagnostics for both fixed effects and variance

components models. Banerjee and Frees [7] introduced case deletion diagnostics for both fixed effects and random subject effect in linear longitudinal models. Xuping and Bocheng [8] presented a unified diagnostic method for linear mixed models based upon the joint likelihood given by Robinson [9]. They showed that the estimates of parameters in case deletion method are equivalent to those in mean shift outlier model. Haslett and Dillane [10] proved a “delete = replace” identity in linear models and applied it to deletion diagnostics for estimators of variance components. Zewotir and Galpin [11] provided routine diagnostic tools for fixed effects, random effects, and variance components, which are computationally inexpensive. Li et al. [12] considered subset deletion diagnostics for fixed effects, random effects and one variance component in varying coefficient mixed models. Gumede et al. [13] extended the variance shift outlier model (VSOM) to the linear mixed model.

In linear regression models, independent variables are often susceptible to nonnegligible errors, and then it will be more appropriate to consider the measurement error models (see Fuller [14] and Stefanski [15]). However, in measurement error models ordinary maximum likelihood (ML) estimates lose the consistency. In order to correct

the bias on parameter estimation, a method in which the score function itself is corrected for measurement errors, is available for the estimation of parameters. This method is based on the corrected log-likelihood of Nakamura [16] (see also Giménez and Bolfarine [17] for more discussion).

On diagnostic methods for measurement error models, some previous works are due to Kelly [18] and Wellman and Gunst [19]. Zhong et al. [20] obtained case deletion and mean shift outlier model for linear measurement error models based upon the corrected likelihood of Nakamura [16]. Rasekh [21] studied multiple outlier detection in multivariate functional measurement error models based on the suitable definition of standardized residuals. Giménez and Galea [22] studied influence measures on corrected score estimators in functional heteroscedastic measurement error models and a local influence study on functional comparative calibration models with replicated data is developed by Giménez and Patat [23].

In this paper, we concentrate on the variance shift model of Cook et al. [4], for the linear measurement error model, using the corrected likelihood [16]. In Section 2, we present the basis of the corrected score method and obtain the estimates of parameters of the model. In Section 3, a variance shift model for the linear measurement error model is derived and the joint corrected maximum likelihood estimates are characterized. In Section 4, we develop the likelihood ratio and the score test statistics. Furthermore, a parametric bootstrap procedure is used to generate the empirical distribution of these statistics. In Section 5, to verify the performance of the proposed test statistics, a simulation study is reported and finally, an illustrative example is given in Section 6.

2. Corrected Log-Likelihood of Measurement Error Models

Consider the linear measurement error models:

$$\begin{aligned} y &= Z\beta + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I_n), \\ X &= Z + U, \quad U \sim N(0, I_n \otimes \Lambda), \end{aligned} \quad (1)$$

where $y = (y_1, y_2, \dots, y_n)'$, Z is an $n \times p$ matrix of unobservable regressors, β is a $p \times 1$ vector of unobservable parameters, and σ^2 is the unknown common variance. The matrix X is the observed value of Z with the measurement error U . Furthermore, ε and U are independent, Λ is a $p \times p$ matrix of known values with nonnegative diagonal elements (Fuller [14]) and I_n is an $n \times n$ identity matrix. The model (1) is known as functional linear measurement error model.

The log-likelihood of y is given by

$$\begin{aligned} l(\beta, \sigma^2, Z, y) &= -\frac{n}{2} \log(2\pi\sigma^2) \\ &\quad - \frac{1}{2\sigma^2} (y - Z\beta)' (y - Z\beta). \end{aligned} \quad (2)$$

If we replace Z by X without considering the measurement errors, then the ML estimates are not consistent in general. To correct for the effects of measurement errors on parameter

estimation, we use the corrected score method proposed by Nakamura [16]. This method proposes a corrected log-likelihood $l^*(\beta, \sigma^2, X, y)$, which satisfies

$$\begin{aligned} E^* \left[\frac{\partial}{\partial \beta} l^*(\beta, \sigma^2, X, y) \right] &= \frac{\partial}{\partial \beta} l(\beta, \sigma^2, Z, y), \\ E^* \left[\frac{\partial}{\partial \sigma^2} l^*(\beta, \sigma^2, X, y) \right] &= \frac{\partial}{\partial \sigma^2} l(\beta, \sigma^2, Z, y), \end{aligned} \quad (3)$$

where E^* denotes the conditional mean with respect to X given y . For model (1), the appropriate corrected log-likelihood is suggested by Nakamura [16] as

$$\begin{aligned} l^*(\beta, \sigma^2, X, y) &= -\frac{n}{2} \log(2\pi\sigma^2) \\ &\quad - \frac{1}{2\sigma^2} \left[(y - X\beta)' (y - X\beta) - n\beta' \Lambda \beta \right]. \end{aligned} \quad (4)$$

By solving the equations $(\partial/\partial\beta)l^*(\beta, \sigma^2, X, y) = 0$ and $(\partial/\partial\sigma^2)l^*(\beta, \sigma^2, X, y) = 0$, the corrected score estimates of β and σ^2 , respectively, are given by (see [16])

$$\begin{aligned} \hat{\beta} &= (X'X - n\Lambda)^{-1} X'y, \\ \hat{\sigma}^2 &= \frac{1}{n} \left[(y - X\hat{\beta})' (y - X\hat{\beta}) - n\hat{\beta}' \Lambda \hat{\beta} \right]. \end{aligned} \quad (5)$$

3. A Variance Shift Model in the Linear Measurement Error

Suppose the i th observation is considered with inflated error variance. A variance shift model for this observation takes the form

$$y = Z\beta + b_i d_i + \varepsilon, \quad (6)$$

$$X = Z + U,$$

where d_i is an $n \times 1$ vector with value 1 in the i th element and zero elsewhere and b_i is an unknown random coefficient of the form $N(0, \alpha_i)$. Model (6) can be considered as a linear mixed measurement error model in which b_i is a random effect with the variance α_i and the covariance matrix of the data are as follows:

$$\begin{aligned} \text{Var}(y) &= \alpha_i d_i d_i' + \sigma^2 I_n = \sigma^2 \left[(w_i - 1) d_i d_i' + I_n \right] \\ &= \sigma^2 H_i, \end{aligned} \quad (7)$$

where $w_i = \alpha_i/\sigma^2 + 1$ and $\text{Var}(y_i) = w_i \sigma^2$, for $w_i \geq 1$.

The log-likelihood and the corrected log-likelihood for this model are given by

$$\begin{aligned} l_i(\beta, \sigma^2, w_i, Z, y) &= -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2} \log |H_i| \\ &\quad - \frac{1}{2\sigma^2} \left[(y - Z\beta)' H_i^{-1} (y - Z\beta) \right], \end{aligned}$$

$$\begin{aligned}
 l_i^* (\beta, \sigma^2, w_i, X, y) &= -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2} \log |H_i| \\
 &\quad - \frac{1}{2\sigma^2} [(y - X\beta)' H_i^{-1} (y - X\beta) - \text{tr}(H_i^{-1}) \beta' \Lambda \beta],
 \end{aligned} \tag{8}$$

respectively, which have the following properties:

$$\begin{aligned}
 E \left[\frac{\partial}{\partial \beta} l_i^* (\beta, \sigma^2, w_i, X, y) \right] &= \frac{\partial}{\partial \beta} l_i (\beta, \sigma^2, w_i, Z, y), \\
 E \left[\frac{\partial}{\partial \sigma^2} l_i^* (\beta, \sigma^2, w_i, X, y) \right] &= \frac{\partial}{\partial \sigma^2} l_i (\beta, \sigma^2, w_i, Z, y), \tag{9} \\
 E \left[\frac{\partial}{\partial w_i} l_i^* (\beta, \sigma^2, w_i, X, y) \right] &= \frac{\partial}{\partial w_i} l_i (\beta, \sigma^2, w_i, Z, y).
 \end{aligned}$$

Now, for w_i fixed, the corrected score estimates of β and σ^2 will be obtained with differentiating from the corrected log-likelihood of the variance shift model given in (8) with respect to β and σ^2 [24]. Then we have

$$\begin{aligned}
 \widehat{\beta}_i(w_i) &= (X' H_i^{-1} X - \text{tr}(H_i^{-1}) \Lambda)^{-1} X' H_i^{-1} y,
 \end{aligned} \tag{10}$$

$$\begin{aligned}
 \widehat{\sigma}_i^2(w_i) &= \frac{1}{n} [(y - X\widehat{\beta}_i)' H_i^{-1} (y - X\widehat{\beta}_i) - \text{tr}(H_i^{-1}) \widehat{\beta}_i' \Lambda \widehat{\beta}_i],
 \end{aligned} \tag{11}$$

respectively. In the following theorem we derive the asymptotic expressions of the $\widehat{\beta}_i(w_i)$ and $\widehat{\sigma}_i^2(w_i)$ as functions of $\widehat{\beta}$ and $\widehat{\sigma}^2$ given no outliers.

Theorem 1. For model (6), we have

$$\begin{aligned}
 \widehat{\beta}_i(w_i) &= \widehat{\beta} - \frac{w_i - 1}{1 + r_{ii}(w_i - 1)} (X'X - n\Lambda)^{-1} x_i \widehat{v}_i \\
 &\quad + O_p(n^{-1}), \\
 \widehat{\sigma}_i^2(w_i) &= \widehat{\sigma}_v^2 \left(1 - \frac{r_{ii}(w_i - 1)}{1 + r_{ii}(w_i - 1)} \frac{t_i^2}{n} \right) - \widehat{\beta}' \Lambda \widehat{\beta} + O_p(n^{-1}),
 \end{aligned} \tag{12}$$

where r_{ii} is the i th diagonal element of $R = I - X(X'X - n\Lambda)^{-1} X'$, $\widehat{v}_i = y_i - x_i' \widehat{\beta}$ is the i th residual and $t_i = \widehat{v}_i / \widehat{\sigma}_v \sqrt{r_{ii}}$ is the i th studentized residual of the model, in which $\widehat{\sigma}_v^2 = \widehat{\sigma}^2 + \widehat{\beta}' \Lambda \widehat{\beta}$ [25].

Proof. From the estimate of β given in (10), we have

$$\begin{aligned}
 &(X' H_i^{-1} X - \text{tr}(H_i^{-1}) \Lambda)^{-1} \\
 &= \left[X'X - \text{tr}(H_i^{-1}) \Lambda - \frac{w_i - 1}{w_i} X' d_i d_i' X \right]^{-1}
 \end{aligned}$$

$$\begin{aligned}
 &= (X'X - \text{tr}(H_i^{-1}) \Lambda)^{-1} \\
 &\quad + (X'X - \text{tr}(H_i^{-1}) \Lambda)^{-1} X' d_i \\
 &\quad \cdot \left[\left(\frac{w_i - 1}{w_i} \right)^{-1} - d_i' X (X'X - \text{tr}(H_i^{-1}) \Lambda)^{-1} X' d_i \right]^{-1} \\
 &\quad \cdot d_i' X (X'X - \text{tr}(H_i^{-1}) \Lambda)^{-1}.
 \end{aligned} \tag{13}$$

On the other hand, we know that, $X'X = O_p(n)$, $X'y = O_p(n)$, $d_i' X X' d_i = O_p(1)$, $X' d_i d_i' X = O_p(1)$, $(X'X - n\Lambda)^{-1} = O_p(n^{-1})$, and $[I + O_p(n^{-1})]^{-1} = I + O_p(n^{-1})$ (see the appendix for more details). Therefore, the first term in the right hand side of (13) will be

$$\begin{aligned}
 &(X'X - \text{tr}(H_i^{-1}) \Lambda)^{-1} \\
 &= \left(X'X - n\Lambda + \frac{w_i - 1}{w_i} \Lambda \right)^{-1} \\
 &= (X'X - n\Lambda)^{-1} - \frac{w_i - 1}{w_i} (X'X - n\Lambda)^{-1} \\
 &\quad \cdot \left[I + \frac{w_i - 1}{w_i} \Lambda (X'X - n\Lambda)^{-1} \right]^{-1} \Lambda (X'X - n\Lambda)^{-1} \\
 &= (X'X - n\Lambda)^{-1} + O_p(n^{-1}) [I + O_p(n^{-1})]^{-1} O_p(n^{-1}) \\
 &= (X'X - n\Lambda)^{-1} + O_p(n^{-2}),
 \end{aligned} \tag{14}$$

and the second term is

$$\begin{aligned}
 &(X'X - \text{tr}(H_i^{-1}) \Lambda)^{-1} X' d_i \\
 &\quad \cdot \left[\left(\frac{w_i - 1}{w_i} \right)^{-1} - d_i' X (X'X - \text{tr}(H_i^{-1}) \Lambda)^{-1} X' d_i \right]^{-1} \\
 &\quad \cdot d_i' X (X'X - \text{tr}(H_i^{-1}) \Lambda)^{-1} \\
 &= [(X'X - n\Lambda)^{-1} + O_p(n^{-2})] X' d_i \\
 &\quad \cdot \left[\left(\frac{w_i - 1}{w_i} \right)^{-1} \right. \\
 &\quad \left. - d_i' X [(X'X - n\Lambda)^{-1} + O_p(n^{-2})] X' d_i \right]^{-1} \\
 &\quad \cdot d_i' X [(X'X - n\Lambda)^{-1} + O_p(n^{-2})] \\
 &= \frac{w_i - 1}{1 + r_{ii}(w_i - 1)} (X'X - n\Lambda)^{-1} X' d_i d_i' X (X'X - n\Lambda)^{-1} \\
 &\quad + O_p(n^{-3}).
 \end{aligned} \tag{15}$$

Combining both terms of (13), we have

$$\begin{aligned} & (X'H_i^{-1}X - \text{tr}(H_i^{-1})\Lambda)^{-1} \\ &= (X'X - n\Lambda)^{-1} \\ &+ \frac{w_i - 1}{1 + r_{ii}(w_i - 1)}(X'X - n\Lambda)^{-1}X'd_i d_i'X(X'X - n\Lambda)^{-1} \\ &+ O_p(n^{-2}). \end{aligned} \tag{16}$$

Consequently, we have

$$\begin{aligned} & \hat{\beta}_i(w_i) \\ &= (X'H_i^{-1}X - \text{tr}(H_i^{-1})\Lambda)^{-1}X'H_i^{-1}y \\ &= \left[(X'X - n\Lambda)^{-1} + \frac{w_i - 1}{1 + r_{ii}(w_i - 1)} \right. \\ &\quad \cdot (X'X - n\Lambda)^{-1}X'd_i d_i'X(X'X - n\Lambda)^{-1} + O_p(n^{-2}) \left. \right] \\ &\quad \cdot X' \left(I - \frac{w_i - 1}{w_i}d_i d_i' \right) y. \end{aligned} \tag{17}$$

Multiplying the above matrix expression out and simplifying them, $\hat{\beta}_i(w_i)$ will be derived. Next, substituting $\hat{\beta}_i(w_i)$ in the estimate of σ^2 given in (11), we can write

$$\begin{aligned} & n\hat{\sigma}_i^2(w_i) \\ &= (y - X\hat{\beta}_i)'H_i^{-1}(y - X\hat{\beta}_i) - \text{tr}(H_i^{-1})\hat{\beta}_i'\Lambda\hat{\beta}_i \\ &= y'H_i^{-1}y - \hat{\beta}_i'X'H_i^{-1}y \\ &= y'y - \frac{w_i - 1}{w_i}y'd_i d_i'y \\ &\quad - \left[\hat{\beta}_i - \frac{w_i - 1}{1 + r_{ii}(w_i - 1)}(X'X - n\Lambda)^{-1}x_i\hat{v}_i + O_p(n^{-1}) \right]' \\ &\quad \cdot \left(X'y - \frac{w_i - 1}{w_i}X'd_i d_i'y \right) \\ &= y'y - \frac{w_i - 1}{w_i}y'd_i d_i'y - \hat{\beta}'X'y \\ &\quad + \frac{w_i - 1}{w_i}\hat{\beta}'X'd_i d_i'y \\ &\quad + \frac{w_i - 1}{1 + r_{ii}(w_i - 1)}d_i'X(X'X - n\Lambda)^{-1}X'y\hat{v}_i \\ &\quad - \frac{(w_i - 1)^2}{w_i[1 + r_{ii}(w_i - 1)]}d_i'X(X'X - n\Lambda)^{-1}X'd_i d_i'y\hat{v}_i \\ &+ O_p(1) \end{aligned}$$

$$\begin{aligned} &= n\hat{\sigma}^2 - \frac{w_i - 1}{1 + r_{ii}(w_i - 1)}\hat{v}_i^2 + O_p(1) \\ &= n\hat{\sigma}_v^2 \left(1 - \frac{r_{ii}(w_i - 1)}{1 + r_{ii}(w_i - 1)} \frac{t_i^2}{n} \right) - n\hat{\beta}'\Lambda\hat{\beta} + O_p(1), \end{aligned} \tag{18}$$

and hence, $\hat{\sigma}_i^2(w_i)$ will be obtained. \square

In the rest of paper, we define $\tilde{\beta}_i(w_i) = \hat{\beta} - ((w_i - 1)/(1 + r_{ii}(w_i - 1)))(X'X - n\Lambda)^{-1}x_i\hat{v}_i$ and $\tilde{\sigma}_i^2(w_i) = n\hat{\sigma}_v^2(1 - (r_{ii}(w_i - 1)/(1 + r_{ii}(w_i - 1)))(t_i^2/n)) - n\hat{\beta}'\Lambda\hat{\beta}$. It is obvious that for $w_i = 1$ or $r_{ii} = 0$, $\tilde{\beta}_i(w_i) = \hat{\beta}$, $\tilde{\sigma}_i^2(w_i) = \hat{\sigma}^2$ and for $r_{ii} > 0$ we have

$$\begin{aligned} & \lim_{w_i \rightarrow +\infty} \tilde{\beta}_i(w_i) = \hat{\beta} - r_{ii}^{-1}(X'X - n\Lambda)^{-1}x_i\hat{v}_i = \hat{\beta}_{mi}, \\ & \lim_{w_i \rightarrow +\infty} \tilde{\sigma}_i^2(w_i) = \hat{\sigma}_v^2 \left(1 - \frac{t_i^2}{n} \right) - \hat{\beta}'\Lambda\hat{\beta} = \hat{\sigma}_{mi}^2. \end{aligned} \tag{19}$$

Remark 2. $\hat{\beta}_{mi}$ and $\hat{\sigma}_{mi}^2$ are the corrected score estimates of β and σ^2 , respectively, in a mean shift outlier model for i th observation, given by [20, 25].

Now $l_i^*(\beta, \sigma^2, w_i, X, y)$ evaluated at $(\beta, \sigma^2) = (\hat{\beta}_i(w_i), \hat{\sigma}_i^2(w_i))$ is, except for an additive constant, proportional to

$$h(w_i) = -n \log [\hat{\sigma}_i^2(w_i)] - \log(w_i). \tag{20}$$

Using Taylor series expansion, an approximate to the $h(w_i)$ analogues to one given by Cook et al. [4] will be as

$$\tilde{h}(w_i) = -n \log \left[\hat{\sigma}^2 - \frac{w_i - 1}{1 + r_{ii}(w_i - 1)} \frac{\hat{v}_i^2}{n} \right] - \log(w_i). \tag{21}$$

The existence of the value of w_i , say \tilde{w}_i , over the range of $[1, +\infty)$ that maximizes $\tilde{h}(w_i)$, was proved by Cook et al. [4].

4. Analogue of Likelihood Ratio Test and Score Test Statistics

In this section we derive analogues of likelihood ratio test and score test statistics for testing the null hypothesis that the observation is not unusual ($w_i = 1$) against the alternative that it has an inflated variance, $w_i > 1$.

4.1. Corrected Likelihood Ratio Test. Let $l_i^*(\hat{\beta}_i(\tilde{w}_i), \hat{\sigma}_i^2(\tilde{w}_i), \tilde{w}_i, X, y)$ and $l^*(\hat{\beta}, \hat{\sigma}^2, X, y)$ be the corrected log-likelihood evaluated at $(\hat{\beta}_i(\tilde{w}_i), \hat{\sigma}_i^2(\tilde{w}_i), \tilde{w}_i)$ and under the null hypothesis, respectively, for testing. Consider

$$H_0 : w_i = 1 \text{ versus } H_1 : w_i > 1. \tag{22}$$

The corrected likelihood ratio test is defined as

$$\begin{aligned} & \text{CLRT}_i \\ &= -2 [l^*(\hat{\beta}, \hat{\sigma}^2, X, y) - l_i^*(\hat{\beta}_i(\tilde{w}_i), \hat{\sigma}_i^2(\tilde{w}_i), \tilde{w}_i, X, y)]. \end{aligned} \tag{23}$$

4.2. *Score Test Statistic.* In order to derive the score test statistic, we only require the estimate of parameters under the null hypothesis. In the following theorem based on the observed information matrix we obtain this test statistic under the hypothesis of $w_i = 1$.

Theorem 3. *The score test statistic for the i th observation (SC_i), based on the observed information matrix for testing $H_0 : w_i = 1$, is given by*

$$SC_i(w_i = 1) = \frac{nA^2(r_{ii}t_i^2 - 1)^2}{2\{(n-1)[2A(r_{ii}t_i^2 - 1) + 1] - A^2(r_{ii}t_i^2 - 1)^2\}} \quad (24)$$

if $r_{ii}t_i^2 > 1$,

where $A = \hat{\sigma}_v^2/\hat{\sigma}^2$.

Proof. Let the corrected observed information matrix of y for σ^2 and w_i be $J(\sigma^2, w_i)$, then the score test statistic (see [26]) for testing $H_0 : w_i = 1$ against $H_1 : w_i > 1$ is

$$SC_i = \left[\frac{\partial}{\partial w_i} l_i^*(\beta, \sigma^2, w_i, X, y) \right]^2 J^{ww}, \quad (25)$$

where J^{ww} is the lower right corner of $J^{-1}(\sigma^2, w_i)$. Substituting $w_i = 1$, $\beta = \hat{\beta}$, and $\sigma^2 = \hat{\sigma}^2$ into the elements of (25), we have

$$\begin{aligned} & \frac{\partial}{\partial w_i} l_i^*(\beta, \sigma^2, w_i, X, y) \\ &= -\frac{1}{2} + \frac{1}{2\hat{\sigma}^2} (\hat{v}_i^2 - \hat{\beta}'\Lambda\hat{\beta}) = \frac{1}{2}A(r_{ii}t_i^2 - 1), \\ & J(\sigma^2, w_i) \\ &= \begin{bmatrix} \frac{n}{2\hat{\sigma}^4} & \frac{1}{2\hat{\sigma}^4}(\hat{v}_i^2 - \hat{\beta}'\Lambda\hat{\beta}) \\ \frac{1}{2\hat{\sigma}^4}(\hat{v}_i^2 - \hat{\beta}'\Lambda\hat{\beta}) & -\frac{1}{2} + \frac{1}{\hat{\sigma}^2}(\hat{v}_i^2 - \hat{\beta}'\Lambda\hat{\beta}) \end{bmatrix} \\ &= \begin{bmatrix} \frac{n}{2\hat{\sigma}^4} & \frac{1}{2\hat{\sigma}^2}[A(r_{ii}t_i^2 - 1) + 1] \\ \frac{1}{2\hat{\sigma}^2}[A(r_{ii}t_i^2 - 1) + 1] & A(r_{ii}t_i^2 - 1) + \frac{1}{2} \end{bmatrix}, \end{aligned} \quad (26)$$

where $J^{ww} = 2n/\{(n-1)[2A(r_{ii}t_i^2 - 1) + 1] - A^2(r_{ii}t_i^2 - 1)^2\}$, and then substituting in (25) the result is achieved. \square

Because the null hypothesis is on the boundary of the parameter space, the standard asymptotic theory does not apply in this case [27]. Therefore, a parametric bootstrap procedure can be used to approximate the distributions of the likelihood ratio and the score test statistic (see Section 4.3 and [13]).

4.3. *Empirical Distribution.* Based on the Gumedze et al. [13] the following parametric bootstrap procedures, for test

statistics $CLRT_i$ and SC_i , can be used to derive the empirical distributions of these statistics under the hypothesis of no outliers exist in the observations:

Step 1. Fit model (1) to the data and calculate estimates \widehat{Z} (see Zare and Rasekh [28]), $\widehat{\beta}$ and $\widehat{\sigma}^2$, where

$$\widehat{Z}' = X' + \widehat{\sigma}_v^{-2}\Lambda\widehat{\beta}\widehat{v}'. \quad (27)$$

Step 2a. Generate a new data vector from

$$\begin{aligned} y^* &= \widehat{Z}\widehat{\beta} + \varepsilon^*, \\ X^* &= \widehat{Z} + U, \end{aligned} \quad (28)$$

where ε^* is randomly generated as $N(0, \widehat{\sigma}^2 I_n)$ and U is randomly generated as $N(0, I_n \otimes \Lambda)$. Fit model (1) to y^* .

Step 2b. Compute the test statistic ($CLRT_i$ or SC_i) for $i = 1, 2, 3, \dots, n$, by fitting a variance shift model to simulated data y^* for each observation in turn and save the order statistics of the set $\{CLRT_i$ or $SC_i : i = 1, 2, 3, \dots, n\}$.

Step 3. Repeat Steps 2a and 2b R times, for R acceptably large, for example, $R = 10000$. Therefore, an empirical distribution of size R is generated for each order statistic.

Step 4. Calculate the $100(1 - \alpha)$ percentile for each order statistic for the level of size α .

The percentile of the k th order statistic can be considered as a threshold for the k th largest value of the test statistic from the original data and if the k largest values of the test statistic from the original data all exceed their respective thresholds, then it is concluded that these are all outliers.

5. Simulation Study

A parametric bootstrap simulation study is carried out to demonstrate the empirical performance of the various proposed test statistics in terms of the probability of a type I error and power on a single unit.

The response variable y_j is generated from the model $y_j = Z\beta + \varepsilon_j$; $j = 1, 2, \dots, 1100$, where $y_j = (y_{1j}, y_{2j}, \dots, y_{nj})'$, $Z = (\mathbf{1}_n, z)$, $\mathbf{1}_n$ is a $n \times 1$ vector all of whose elements are 1's, $z' = (z_1, z_2, \dots, z_n)$ and ε_j is rewritten in accordance with y_j . We consider the following combinations for simulation: $n = 40$ or 100 where $\beta = (\beta_0, \beta_1) = (10, 0.3)$ or $(\beta_0, \beta_1) = (20, 1)$, $z_i \sim U(100, 900)$, $\varepsilon_{ij} \sim N(0, \sigma^2)$, $i = 1, 2, \dots, n$, $\sigma^2 = 2$ or 8 and $\Lambda = \text{diag}(\lambda_1, \lambda_2) = \text{diag}(0, 0.25)$ or $\Lambda = \text{diag}(0, 1)$. The simulation study was conducted using the R software and the codes are available from the second author upon request.

For each simulated data set, the CLRT and the score test statistics were calculated for the first observation. The choice of the first observation was arbitrary. To generate an empirical distribution of the test statistic under the null hypothesis, data sets for $k = 1, 2, \dots, 2500$ were simulated as $y_{jk}^* = \widehat{Z}_j\widehat{\beta}_j + \varepsilon_{jk}^*$, where $\varepsilon_{jk}^* \sim N(0, \widehat{\sigma}_j^2 I_n)$, $\widehat{\beta}_j$, \widehat{Z}_j , and $\widehat{\sigma}_j^2$ are the corrected estimate of β , Z and σ^2 from y_j . The probability

TABLE 1: The probability of a type I error ($\alpha = 0.05$) and power of CLRT and score test statistics for a variance shift model in a linear measurement error with combination different of parameters β , σ^2 , Λ , and $b = 1$.

n	β	σ^2	(λ_1, λ_2)	CLRT		Score test	
				Sig level	Power	Sig level	Power
40	(20, 1)	2	(0, 1)	0.0457	0.0620	0.0459	0.0621
			(0, 0.25)	0.0478	0.0870	0.0477	0.0873
		8	(0, 1)	0.0478	0.0511	0.0477	0.0513
	(0, 0.25)		0.0411	0.0600	0.0414	0.0603	
	(10, 0.3)	2	(0, 1)	0.0443	0.0909	0.0441	0.0909
			(0, 0.25)	0.0388	0.0944	0.0387	0.0945
8		(0, 1)	0.0388	0.0600	0.0387	0.0603	
	(0, 0.25)	0.0354	0.0619	0.0360	0.0621		
100	(20, 1)	2	(0, 1)	0.0450	0.0801	0.0450	0.0801
			(0, 0.25)	0.0520	0.0950	0.0522	0.0954
		8	(0, 1)	0.0520	0.0649	0.0522	0.0648
	(0, 0.25)		0.0529	0.0668	0.0531	0.0666	
	(10, 0.3)	2	(0, 1)	0.0520	0.0950	0.0522	0.0954
			(0, 0.25)	0.0520	0.0989	0.0522	0.0990
8		(0, 1)	0.0529	0.0631	0.0531	0.0630	
	(0, 0.25)	0.0520	0.0636	0.0522	0.0639		

TABLE 2: The probability of a type I error ($\alpha = 0.05$) and power of CLRT and score test statistics for a variance shift model in a linear measurement error with combination different of parameters β , σ^2 , Λ , and $b = 3$.

n	β	σ^2	(λ_1, λ_2)	CLRT		Score test	
				Sig level	Power	Sig level	Power
40	(20, 1)	2	(0, 1)	0.0455	0.2840	0.0453	0.2844
			(0, 0.25)	0.0476	0.4311	0.0478	0.4311
		8	(0, 1)	0.0478	0.1385	0.0475	0.1386
	(0, 0.25)		0.0419	0.1558	0.0419	0.1557	
	(10, 0.3)	2	(0, 1)	0.0413	0.4768	0.0417	0.4770
			(0, 0.25)	0.0364	0.4832	0.0364	0.4833
8		(0, 1)	0.0367	0.1592	0.0363	0.1593	
	(0, 0.25)	0.0368	0.1608	0.0366	0.1611		
100	(20, 1)	2	(0, 1)	0.0450	0.3660	0.0453	0.3663
			(0, 0.25)	0.0525	0.4823	0.0529	0.4824
		8	(0, 1)	0.0520	0.1584	0.0521	0.1584
	(0, 0.25)		0.0537	0.1629	0.0534	0.1629	
	(10, 0.3)	2	(0, 1)	0.0524	0.5239	0.0522	0.5239
			(0, 0.25)	0.0531	0.5394	0.0528	0.5391
8		(0, 1)	0.0529	0.1653	0.0534	0.1656	
	(0, 0.25)	0.0525	0.1672	0.0525	0.1674		

of a type I error estimate for a given test statistic and $\alpha = 0.05$ was calculated as the number of data sets for which the test statistic exceeded the 95th percentile of the empirical distribution, divided by 1100 [13].

The CLRT and the score test statistics were performed for a variance shift model for the first observation of each simulated data and 95th percentiles from the empirical distribution of each test statistics were used as threshold values for the test statistics observed on the original data set y_j . The empirical probability of type I errors for thresholds derived from the empirical distribution under the null hypothesis are

calculated for the corrected likelihood ratio and score test statistics for $\alpha = 0.05$ (Tables 1, 2, and 3). A glance at the results of these tables indicate that in general the probability of a type I errors of both CLRT and score test statistics are close to the nominal value of 0.05.

In order to access the relative sensitivity of the CLRT and score test statistics, we introduce the shift values 1, 3, and 5 for the first observation and again for each combination of parameters, 1100 data sets are generated from the following model:

$$y_j = Z\beta + bd_1 + \varepsilon_j, \quad (29)$$

TABLE 3: The probability of a type I error ($\alpha = 0.05$) and power of CLRT and score test statistics for a variance shift model in a linear measurement error with combination different of parameters β, σ^2, Λ , and $b = 5$.

n	β	σ^2	(λ_1, λ_2)	CLRT		Score test	
				Sig level	Power	Sig level	Power
40	(20, 1)	2	(0, 1)	0.0457	0.7172	0.0459	0.7174
			(0, 0.25)	0.0478	0.8755	0.0475	0.8758
		8	(0, 1)	0.0476	0.3169	0.0474	0.3168
	(10, 0.3)	2	(0, 0.25)	0.0411	0.3536	0.0414	0.3537
			(0, 1)	0.0449	0.9002	0.0446	0.9001
		8	(0, 0.25)	0.0389	0.9082	0.0387	0.9082
100	(20, 1)	2	(0, 1)	0.0388	0.3600	0.0387	0.3600
			(0, 0.25)	0.0366	0.3650	0.0364	0.3654
		8	(0, 1)	0.0457	0.7893	0.0453	0.7894
	(10, 0.3)	2	(0, 0.25)	0.0523	0.9035	0.0521	0.9037
			(0, 1)	0.0525	0.3547	0.0528	0.3546
		8	(0, 0.25)	0.0535	0.3897	0.0532	0.3897
(10, 0.3)	2	(0, 1)	0.0525	0.9229	0.0522	0.9226	
		(0, 0.25)	0.0521	0.9272	0.0521	0.9271	
	8	(0, 1)	0.0536	0.4024	0.0535	0.4023	
			(0, 0.25)	0.0520	0.4059	0.0522	0.4059

for $b = 1, 3$, or 5 , where d_1 is an $n \times 1$ vector with value 1 in the first element and zero elsewhere. The CLRT, the score test statistic, and their empirical distribution were calculated as for the probability of a type I error given above, consequently the power of the test statistics are also derived. Results of Tables 1–3 show that with increase of the displacement, b , the power of the CLRT and score test statistic, increases in general. Moreover, we can see that power of the test statistics also increase as sample size increases. These tables also show the result of the CLRT and score test statistics are nearly identical in the empirical probability of a type I errors and power.

6. Example: Concrete Compressive Strength Data

These data were given by Wellman and Gunst [19] and contain comprehensive strength measurements of 41 sample of concrete. It was desired to use a linear regression model to predict comprehensive strength of concrete 28 days after pouring from the strength measurements taken two days after pouring. Zhong et al. [20] analyzed this data set using the linear measurement error model with $\Lambda = \text{diag}(0, 1)$. The zero diagonal element in Λ corresponds to the constant term, a predictor variable measured without error. They indicated that the sample 21 exhibits a strong influence on the fitted model. Here we consider a variance shift measurement error model for this data set.

Figures 1, 2, and 3 show plots of the square of studentized residual t_i^2 of the data under model (1), estimates of the variance shift parameter \tilde{w}_i , and estimated variance $\hat{\sigma}_i^2$ under model (6), versus case numbers, respectively. From these figures it is obvious that case 21 stands out as a possible outlier

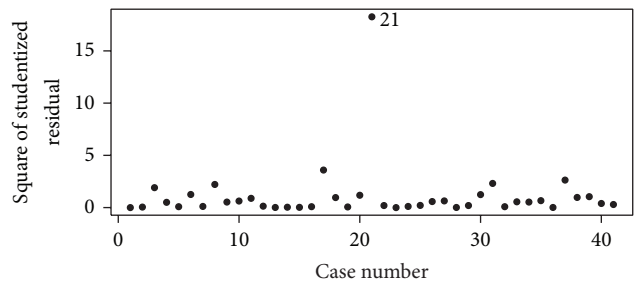


FIGURE 1: Plot of square of the studentized residual (t_i^2) versus case number.

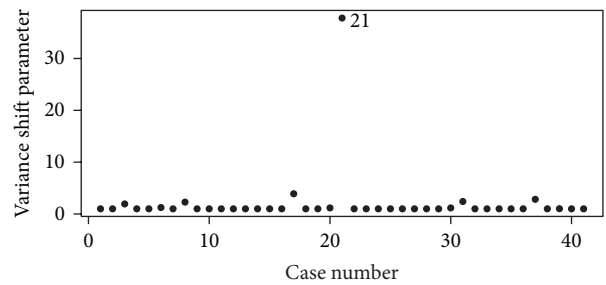


FIGURE 2: Plot of variance shift estimates (\tilde{w}_i) versus case number.

with relatively large values of t_i^2 and \tilde{w}_i and a small estimated $\hat{\sigma}_i^2$.

Next, the corrected likelihood ratio and score test statistics were calculated for each observation under model (6), and then 10000 simulated data sets were generated from the fitted model under the null hypothesis ($H_0 : w_i = 1$). In each simulation, a variance shift model was fitted for each observation and the test statistics were sorted and used

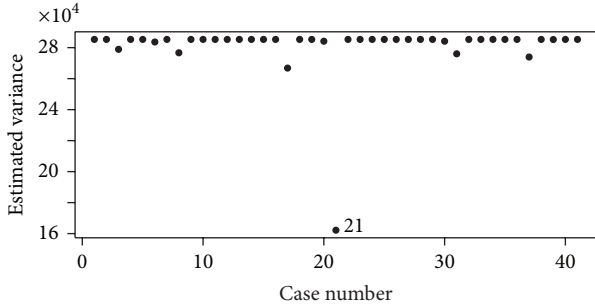


FIGURE 3: Plot of error variance estimates ($\hat{\sigma}_i^2$) versus case number.

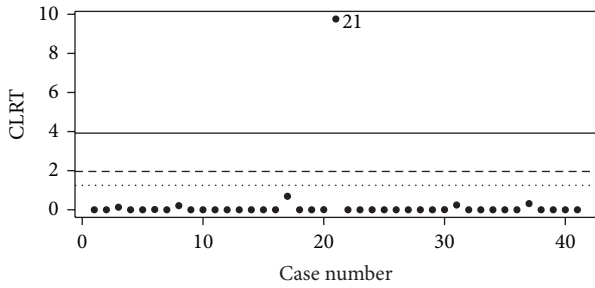


FIGURE 4: Corrected likelihood ratio test statistic for each case, with 95th percentile of the empirical distribution under H_0 shown for the first k order statistics for corrected likelihood ratio test: $k = 1$ (solid line), $k = 2$ (dashed line), and $k = 3$ (dotted line).

to generate the empirical distribution of the order statistics for each test [13]. Figures 4 and 5 give plots of the test statistics from the real data and 95th percentile from the empirical distribution of the first, second and third largest values for each test statistic. These figures show that the statistics for observation 21 is larger than the 95th percentile of the distribution of the corresponding order statistics.

Finally, the fitted regression line measurement error model and the variance shift measurement error model for case 21 are shown in Figure 6.

7. Conclusions

We extended the variance shift model to the linear measurement error models based on the corrected likelihood of Nakamura [16]. We derived the approximate estimate of parameters of the proposed model under the variance shift and indicated that if the variance shift parameter tends to infinity, these estimates will be the same as those obtained from a mean shift outlier model. Also, we proposed a corrected likelihood ratio test and derived the score test statistic for testing that an observation stands out as possible outlier and it is shown that the score test statistic is function of studentized residuals of model. The performance of both the corrected likelihood ratio and the score test statistics is studied using a parametric bootstrap simulation, and it was found out that with the increase of b or sample size, the power of both test statistics increases.

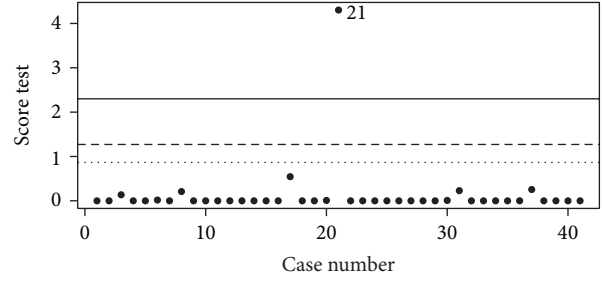


FIGURE 5: Score test statistic for each case, with 95th percentile of the empirical distribution under H_0 shown for the first k order statistics for score test: $k = 1$ (solid line), $k = 2$ (dashed line), and $k = 3$ (dotted line).

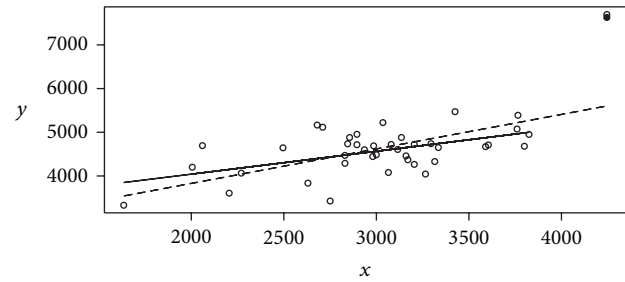


FIGURE 6: Concrete compressive strengths, in pounds per square inch, at 2 and 28 days, measurement error model (dashed line); a variance shift model (solid line).

Appendix

We assume that as n tends to infinity, the limit of $n^{-1}Z'Z$ exists ($n^{-1}Z'Z = O(1)$). The existence of this limit is assured in Lee and Nelder [29]. Since $E(X'X) = Z'Z + n\Lambda$ and $E(X'y) = Z'Z\beta$, by the law of large numbers, it is easy to get $n^{-1}X'X = n^{-1}Z'Z + \Lambda + O_p(n^{-1/2})$ and $n^{-1}X'y = n^{-1}Z'Z\beta + O_p(n^{-1/2})$. Consequently, we have

$$Z'Z = O(n),$$

$$X'X = Z'Z + n\Lambda + O_p(n^{1/2}) = O_p(n),$$

$$(X'X - n\Lambda)^{-1} = O_p(n^{-1}), \quad (\text{A.1})$$

$$X'y = Z'Z\beta + O_p(n^{1/2}) = O_p(n),$$

Moreover, $E(d'_i X) = d'_i Z$, and then

$$d'_i X = d'_i Z + O_p(1),$$

$$X' d_i d'_i X = Z' d_i d'_i Z + O_p(1), \quad (\text{A.2})$$

$$d'_i X X' d_i = d'_i Z Z' d_i + O_p(1),$$

also, $[I + O_p(n^{-1})]^{-1} = I + O_p(n^{-1})$ is obtained from Taylor series expansion.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

- [1] V. Barnett and T. Lewis, *Outliers in Statistical Data*, John Wiley & Sons, New York, NY, USA, 1978.
- [2] S. Weisberg, *Applied Linear Regression*, John Wiley & Sons, New York, NY, USA, 1980.
- [3] R. D. Cook and S. Weisberg, *Residuals and Influence in Regression*, Chapman & Hall, London, UK, 1982.
- [4] R. D. Cook, N. Holschuh, and S. Weisberg, "A note on an alternative outlier model," *Journal of the Royal Statistical Society. Series B*, vol. 44, no. 3, pp. 370–376, 1982.
- [5] R. Thompson, "A note on restricted maximum likelihood estimation with an alternative outlier model," *Journal of the Royal Statistical Society B*, vol. 47, no. 1, pp. 53–55, 1985.
- [6] R. Christensen, L. M. Pearson, and W. Johnson, "Case-deletion diagnostics for mixed models," *Technometrics*, vol. 34, no. 1, pp. 38–45, 1992.
- [7] M. Banerjee and E. W. Frees, "Influence diagnostics for linear longitudinal models," *Journal of the American Statistical Association*, vol. 92, no. 439, pp. 999–1005, 1997.
- [8] Z. Xuping and W. Bocheng, "Influence analysis on linear models with random effects," *Applied Mathematics*, vol. 14, no. 2, pp. 169–176, 1999.
- [9] G. K. Robinson, "That BLUP is a good thing: the estimation of random effects," *Statistical Science*, vol. 6, no. 1, pp. 15–51, 1991.
- [10] J. Haslett and D. Dillane, "Application of "delete=replace" to deletion diagnostics for variance component estimation in the linear mixed model," *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, vol. 66, no. 1, pp. 131–143, 2004.
- [11] T. Zewotir and J. S. Galpin, "Influence diagnostics for linear mixed models," *Journal of Data Science*, vol. 3, pp. 153–177, 2005.
- [12] Z. Li, W. Xu, and L. Zhu, "Influence diagnostics and outlier tests for varying coefficient mixed models," *Journal of Multivariate Analysis*, vol. 100, no. 9, pp. 2002–2017, 2009.
- [13] N. F. Gumede, S. J. Welham, B. J. Gogel, and R. Thompson, "A variance shift model for detection of outliers in the linear mixed model," *Computational Statistics & Data Analysis*, vol. 54, no. 9, pp. 2128–2144, 2010.
- [14] W. A. Fuller, *Measurement Error Models*, John Wiley & Sons, New York, NY, USA, 1987.
- [15] L. A. Stefanski, "Unbiased estimation of a nonlinear function of a normal mean with application to measurement error models," *Communications in Statistics: Theory and Methods*, vol. 18, no. 12, pp. 4335–4358, 1989.
- [16] T. Nakamura, "Corrected score function for errors-in-variables models: methodology and application to generalized linear models," *Biometrika*, vol. 77, no. 1, pp. 127–137, 1990.
- [17] P. Giménez and H. Bolfarine, "Corrected score functions in classical error-in-variables and incidental parameter models," *Australian Journal of Statistics*, vol. 39, no. 3, pp. 325–344, 1997.
- [18] G. E. Kelly, "The influence function in the errors in variables problem," *The Annals of Statistics*, vol. 12, no. 1, pp. 87–100, 1984.
- [19] J. M. Wellman and R. F. Gunst, "Influence diagnostics for linear measurement error models," *Biometrika*, vol. 78, no. 2, pp. 373–380, 1991.
- [20] X.-P. Zhong, B.-C. Wei, and W.-K. Fung, "Influence analysis for linear measurement error models," *Annals of the Institute of Statistical Mathematics*, vol. 52, no. 2, pp. 367–379, 2000.
- [21] A. R. Rasekh, "Local influence in measurement error models with ridge estimate," *Computational Statistics & Data Analysis*, vol. 50, no. 10, pp. 2822–2834, 2006.
- [22] P. Giménez and M. Galea, "Influence measures on corrected score estimators in functional heteroscedastic measurement error models," *Journal of Multivariate Analysis*, vol. 114, pp. 1–15, 2013.
- [23] P. Giménez and M. L. Patat, "Local influence for functional comparative calibration models with replicated data," *Statistical Papers*, vol. 55, pp. 431–454, 2014.
- [24] K. Zare, A. Rasekh, and A. A. Rasekhi, "Estimation of variance components in linear mixed measurement error models," *Statistical Papers*, vol. 53, no. 4, pp. 849–863, 2012.
- [25] K. Zare and A. Rasekh, "Diagnostic measures for linear mixed measurement error models," *SORT*, vol. 35, no. 2, pp. 125–144, 2011.
- [26] D. R. Cox and D. V. Hinkley, *Theoretical Statistics*, Chapman & Hall, London, UK, 1974.
- [27] S. G. Self and K.-Y. Liang, "Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions," *Journal of the American Statistical Association*, vol. 82, no. 398, pp. 605–610, 1987.
- [28] K. Zare and A. Rasekh, "Residuals and leverages in the linear mixed measurement error models," *Journal of Statistical Computation and Simulation*, vol. 84, no. 7, pp. 1427–1443, 2014.
- [29] Y. Lee and N. A. Nelder, "Hierarchical generalized linear models," *Journal of the Royal Statistical Society B*, vol. 58, no. 4, pp. 619–678, 1996.