

Empirical process based on the recursive residuals in functional measurement error models

A. R. RASEKH*

Statistical Research Center and Department of Statistics, Ahwaz University, Ahwaz, Iran

(Received 13 October 2003; in final form 4 January 2006)

Recursive residuals are a linear transformation of ordinary residuals. They have frequently been suggested for testing model fit and model assumptions in linear regression. In this paper, we generalize the theory of the empirical process based on the residuals in the measurement error models to the recursive residuals in the measurement error models. We prove that recursive residuals in these models are asymptotically independent and identically distributed and show that, in this case, the weak convergence properties of the standardized residuals hold for the studentized recursive residuals. Furthermore, we look at some tests for goodness-of-fit based on the weak convergence of the empirical distribution of the recursive residuals. We justify the use of goodness-of-fit tests based on recursive residuals by carrying out a parametric bootstrap simulation study.

Keywords: Errors in variables; Kolmogorov–Smirnov test; Normality test; Weak convergence; Goodness-of-fit

1. Introduction

There has been considerable interest in recent years in the behaviour of empirical distribution functions based on regression residuals. Much of this interest stems from the desire to use residuals in place of unobservable experimental errors in goodness-of-fit tests, often with a specific interest in tests of normality.

The early work of Durbin [1], Rao and Sethurannan [2] and Neuhaus [3] set the stage by examining empirical processes when parameters of the underlying distribution function were estimated. Mugantseua [4] and Pierce and Kopecky [5] were the first to derive the limiting distribution of the sample empirical process based on least squares residuals. Shorack [6] extended these results to the general class of residuals based on $n^{1/2}$ -consistent estimators of the regression coefficients [see also ref. 7, section 4.6]. Wood [8] had earlier examined the special case of ridge regression.

Portnoy [9] had examined the situation in which the number of regression parameters is large and Loynes [10] looked at the weak convergence of processes based on generalized residuals defined by Cox and Snell [11]. Miller [12, 13] was the first to derive the limiting

*Email: rasekh_a@cua.ac.ir

distribution of empirical process based upon residuals from errors in variables models. In all these papers, the authors did not assume normally distributed equation errors.

Recursive residuals are a linear transformation of ordinary residuals. They have frequently been suggested for testing model fit and model assumptions in linear regression [14–18]. Loynes [19] considered recursive residuals in non-linear models and McGilchrist and Matawie [20] introduced recursive residuals in generalized linear models. Magnus and Sinha [21] have a fresh look at recursive residuals and BLUS residuals in regression models.

Unlike ordinary residuals, they are independently and identically distributed and do not have the problem of deficiencies in one part of the data being smeared over all the residuals. Galpin and Hawkins [22] proposed the use of normal plot of the recursive residuals to check the model assumption of normality. In fact, as they are all of the same scale, indications of non-normality from normal probability plots do indicate non-normality but not differences in the influences of the observations. An excellent review of developments and applications of recursive residuals in linear models is given in the work of Kianifard and Swallow [23].

Miller [12, 13] considered the residuals in measurement error models. He studied the theory of the empirical process based on the residuals and derived the weak convergence of the residuals as a basis for goodness-of-fit tests. In this article, we generalize these results to the recursive residuals. In the remaining part of this section, we present the measurement error model. In section 2, we define the recursive residual and show that they are asymptotically independent and identically distributed. Furthermore, in section 3, we generalize the theory of the empirical process to the recursive residuals and show that the main theorem of the Miller [12, 13], for the weak convergence of the standardized residuals, holds for the studentized recursive residuals in this case. In section 4, we look at some goodness-of-fit tests based on weak convergence of the empirical distribution of the recursive residuals. In section 5, we give an example of real data. Finally, in section 6, we perform parametric bootstrap simulation studies to examine the power of the proposed goodness-of-fit tests in detecting normality of the errors for the small and medium sample sizes.

The functional measurement error models will be defined as

$$Y_i = y_i + e_i, \quad X_i = x_i + u_i, \quad y_i = \beta'x_i, \quad i = 1, \dots, n \tag{1}$$

In this model, β is the vector of coefficients and x_i is the vector of unobservable fixed values, both with k dimension. For each unobservable x_i and y_i , we have observable random variables X_i and Y_i , respectively. Let $\varepsilon'_i = (e_i, u_i)'$, $i = 1, \dots, n$, be independent and normally distributed with zero mean and covariance matrix. $\Sigma_{\varepsilon\varepsilon} = \begin{bmatrix} \sigma_{ee} & \Sigma_{eu} \\ \Sigma_{ue} & \Sigma_{uu} \end{bmatrix}$. We allow $\Sigma_{\varepsilon\varepsilon}$ to be a singular matrix. An example involving singular matrix $\Sigma_{\varepsilon\varepsilon}$ is when the model contains an intercept. We can write model (1) as

$$Y_i = \beta'X_i + v_i, \quad v_i = e_i - \beta'u_i, \quad i = 1, \dots, n \tag{2}$$

and hence v_i 's are independent and normally distributed with zero mean and variances, $\sigma_{vv} = \sigma_{ee} - \beta' \Sigma_{eu} - \Sigma_{eu} \beta + \beta' \Sigma_{uu} \beta$. Generally, we would anticipate that replicate observations are available on this model and hence an estimate of the $\Sigma_{\varepsilon\varepsilon}$ exists and is given by $S_{\varepsilon\varepsilon}$. As the replicates do not directly affect the central themes of our work, we have simplified the notation by referring only to single observations. If we define $\eta'_i = (Y_i, X'_i)$, $i = 1, \dots, n$, then an estimate of β will be $\hat{\beta} = (M_{XX} - \tilde{\tau} S_{uu})^{-1} (M_{XY} - \tilde{\tau} S_{ue})$, in which S_{uu} and S_{ue} are estimates of Σ_{uu} and Σ_{ue} , respectively, and $M_{\eta\eta} = n^{-1} \sum_{i=1}^n \eta_i \eta'_i = \begin{bmatrix} M_{YY} & M_{YX} \\ M_{XY} & M_{XX} \end{bmatrix}$. $\tilde{\tau}$ is the smallest

eigenvalue of $M_{\eta\eta}$ in the metric of $S_{\varepsilon\varepsilon}$ [24, 25]. Fuller [24] refers to v_i as an unobservable error of the model and hence

$$\hat{v}_i = Y_i - \hat{\beta}' X_i \quad i = 1, \dots, n \tag{3}$$

as the residuals of the model. These residuals have been suggested by Fuller [24] for model checking, and the limiting behaviour of several diagnostic procedures based on equation (3) have been examined by Miller [12]. Furthermore, Miller [12, 13] examined the limiting behaviour of empirical process based on the residuals of equation (3).

2. Recursive residuals and asymptotic properties

2.1 Definition and assumptions

Suppose that $\hat{\beta}_i$ is an estimator of β using the first i th observations of the model. We define recursive residuals as $s_i = Y_i - \hat{\beta}' X_i, i = k + 1, \dots, n$. It is obvious that if $i = n$ then $\hat{\beta}'_i$ will be equal to $\hat{\beta}'$. Another representation of s_i is

$$s_i = (Y_i - \beta' X_i) - (\hat{\beta}'_i - \beta') X_i = v_i - (\hat{\beta}'_i - \beta) X_i, \quad i = k + 1, \dots, n \tag{4}$$

We introduce the random vector $\ddot{x}_i = X_i - v_i \sigma_{vv}^{-1} \Sigma_{vu}$ in which $\Sigma_{vu} = \Sigma_{eu} - \Sigma_{uu} \beta$. It is constructed by subtracting from X_i the best predictor of u_i given v_i under the normal model [24]. Substituting \ddot{x}_i in equation (4), we have

$$\begin{aligned} s_i &= v_i - (\hat{\beta}'_i - \beta)' (\ddot{x}_i + v_i \sigma_{vv}^{-1} \Sigma_{vu}) \\ &= v_i [1 - (\hat{\beta}'_i - \beta)' \sigma_{vv}^{-1} \Sigma_{vu}] - (\hat{\beta}'_i - \beta)' \ddot{x}_i \quad i = k + 1, \dots, n \end{aligned} \tag{5}$$

Our assumptions on the estimators $\hat{\beta}_i$ and the limiting behaviour of i with respect to n are as follows:

- (a) One complication with the limiting behaviour of the recursive residuals that did not occur in the ordinary residuals is that some attention must be paid to the size of i . We shall assume from now on that i is of the same order as n , or in a limited context, we would assume that i/n is bounded away from zero. The need for this is obvious if we note that $\hat{\beta}_i$ depends on i observations and asymptotic considerations could not apply if i were too small.
- (b) $\hat{\beta}_i, i = k + 1, \dots, n$, are any estimators such that $\hat{\beta}_i - \beta = o_P(n^{(-1/2)})$. Methods for producing $n^{(-1/2)}$ -consistent estimators for $\hat{\beta}$, can be found in Fuller [24], Amemiya and Fuller [26] and Gleaser [27]. Under assumption (a), we can extend these results to $\hat{\beta}_i$.

2.2 Asymptotic properties

In this subsection, we concentrate on the asymptotic properties of the recursive residuals and studentized recursive residuals.

THEOREM 1 *Under the assumption of normality of the errors of model (1) and from assumption (a), recursive residuals, s_i defined by equation (4) are asymptotically independent.*

Proof Suppose that $\hat{\beta}_i$ is a maximum likelihood estimate of β using first i observations. According to Cox and Snell [11], the first-order expansion of the log maximum likelihood

equation $L'(\hat{\beta}_i) = 0$ is $L'(\beta) + (\hat{\beta}_i - \beta)'L''(\beta) = 0$ in which $L''(\beta)$ will be the derivative of $L'(\beta)$ with respect to β . If we write

$$U^{(j)} = \frac{\partial \log p_j(v_j, \beta)}{\partial \beta}, \quad V_{rs}^{(j)} = \frac{\partial^2 \log p_j(v_j, \beta)}{\partial \beta_r \partial \beta_s}$$

in which $p_j(v_j, \beta)$ is the probability distribution function of v_j and replace $-L''(\beta)$ by its expectation, $I_i = \sum_{j=1}^i (-V^{(j)})$, where $V^{(j)} = \{V_{rs}^{(j)}\}$ and I_i is the information using first i observations, then, we have $(\hat{\beta}_i - \beta) = I_i^{-1}U^{*(i)}$, where $U^{*(i)} = \sum_{j=1}^i (U^{(j)})$. On the other hand, we can write the recursive residual as $s_i = v_i - U'^{* (i)} I_i^{-1} X_i + o(n^{-1/2})$. Now suppose that $i < j$, then

$$\begin{aligned} E(s_i s_j) &= E(v_i v_j) - E(v_i U'^{* (j)} I_j^{-1} X_j) - E(v_j U'^{* (i)} I_i^{-1} X_i) \\ &\quad + E(U'^{* (i)} I_i^{-1} X_i U'^{* (j)} I_j^{-1} X_j) + o(n^{-1}) \end{aligned}$$

However, the first and third terms will be equal to zero, and for the second term, we have

$$\begin{aligned} E(v_i U'^{* (j)} I_j^{-1} X_j) &= \sum_{i=1}^j E(v_i U'^{(l)} I_j^{-1} X_j) = E(v_i U'^{(i)} I_j^{-1} X_j) \\ &= -E(v_i U'^{(i)}) I_j^{-1} E(X_j) \end{aligned}$$

For the fourth term, we have

$$\begin{aligned} E[X_i' I_i^{-1} U^{*(i)} U'^{* (j)} I_j^{-1} X_j] &= E \left[X_i' I_i^{-1} U^{*(i)} \left(U^{*(i)} + \sum_{l=i+1}^j U^{(l)} \right)' I_j^{-1} X_j \right] \\ &= E[X_i' I_i^{-1} U^{*(i)} U'^{* (i)} I_j^{-1} X_j] \\ &\quad + E \left[X_i' I_i^{-1} U^{*(i)} \left(\sum_{l=i+1}^j U^{(l)} \right)' I_j^{-1} X_j \right] \\ &= E[X_i' I_i^{-1} I_i I_j^{-1} X_j] \\ &\quad + E[X_i' I_i^{-1} U^{*(i)}] E \left[\left(\sum_{l=i+1}^j U^{(l)} \right)' I_j^{-1} X_j \right] \end{aligned}$$

However, $E[X_i' I_i^{-1} U^{*(i)}]$, and $E[(\sum_{l=i+1}^j U^{(l)})' I_j^{-1} X_j]$ are $o(n^{-1/2})$ and consequently multiplication of these two terms will be $o(n^{-1})$. Therefore, the fourth term will be equal to $E(X_i' I_j^{-1} X_j) + o(n^{-1})$, which in turn implies that

$$E(s_i s_j) = E(X_j) I_j^{-1} E(-v_i U^{(i)} + X_i) + o(n^{-1})$$

However, it can be shown that $E(-v_i U^{(i)} + X_i) = (\partial/\partial \beta) E(v_i)$. Therefore, $E(s_i s_j) = 0$ up to $o(n^{-1})$ and consequently s_i 's are uncorrelated. Under the normality of the errors, they will be asymptotically independent. ■

An estimator of σ_{vv} based on all observations will be

$$\hat{\sigma}_{vv} = \sigma_{ee} - \hat{\beta}' \Sigma_{eu} - \Sigma_{eu} \hat{\beta} + \hat{\beta}' \Sigma_{uu} \hat{\beta}$$

If $\Sigma_{\varepsilon\varepsilon}$ is unknown, then we can estimate it using the replicated observations and substitute the estimators in $\hat{\sigma}_{vv}$. Miller [28] showed that

$$\hat{\sigma}_{vv} = \sigma_{vv} \left[1 - \sigma_{vv}^{-1} (\hat{\beta} - \beta)' \Sigma_{vu} \right] o_p(n^{-1})$$

Another estimator of σ_{vv} using first i th observations will be

$$\hat{\sigma}_{vvi} = \sigma_{ee} - \hat{\beta}'_i \Sigma_{eu} - \Sigma_{eu} \hat{\beta}_i + \hat{\beta}'_i \Sigma_{uu} \hat{\beta}_i$$

Under assumption (b) and by the approach similar to Miller's [28], we find that

$$\hat{\sigma}_{vvi} = \sigma_{vv} \left[1 - \sigma_{vv}^{-1} (\hat{\beta}_i - \beta)' \Sigma_{vu} \right] o_p(n^{-1})$$

Therefore, we define $Z_i = \hat{\sigma}_{vvi}^{-(1/2)} s_i, i = k + 1, \dots, n$, as the studentized recursive residuals of model (1). In the following lemma, we derive their asymptotic properties.

LEMMA 1 For the studentized recursive residuals $Z_i = \hat{\sigma}_{vvi}^{-(1/2)} s_i, i = k + 1, \dots, n$, we have

$$Z_i = \hat{\sigma}_{vvi}^{-(1/2)} v_i + o_p(n^{-(1/2)}).$$

Proof If we substitute s_i from equation (1) into Z_i , then we have

$$\begin{aligned} Z_i &= \hat{\sigma}_{vvi}^{-(1/2)} v_i \left(1 - \sigma_{vv}^{-1} (\hat{\beta}_i - \beta)' \Sigma_{vu} \right) - \hat{\sigma}_{vvi}^{-(1/2)} (\hat{\beta}_i - \beta)' \ddot{x}_i \\ &= \sigma_{vv}^{-(1/2)} \hat{\sigma}_{vvi}^{1/2} \hat{\sigma}_{vvi}^{-(1/2)} v_i \left[1 - \sigma_{vv}^{-1} (\hat{\beta}_i - \beta)' \Sigma_{vu} \right] - \hat{\sigma}_{vvi}^{-(1/2)} (\hat{\beta}_i - \beta)' \ddot{x}_i \quad (6) \\ &= \sigma_{vv}^{-(1/2)} v_i (1 + a_{i,n}) + B_{i,n} \ddot{x}_i \end{aligned}$$

in which $a_{i,n} = \sigma_{vv}^{1/2} \hat{\sigma}_{vvi}^{-(1/2)} v_i [1 - \sigma_{vv}^{-1} (\hat{\beta}_i - \beta)' \Sigma_{vu}] - 1$ and $B_{i,n} = \hat{\sigma}_{vvi}^{-(1/2)} (\hat{\beta}_i - \beta)'$. By assumption (b), we have $B_{i,n} = o_p(n^{-(1/2)})$. However,

$$\begin{aligned} 1 + a_{i,n} &= \left(\sigma_{vv}^{-(1/2)} v_i \left[1 - \sigma_{vv}^{-1} (\hat{\beta}_i - \beta)' \Sigma_{vu} \right]^{-1} \right) \\ &\quad \times \left(\sigma_{vv}^{1/2} v_i \left[1 - \sigma_{vv}^{-1} (\hat{\beta}_i - \beta)' \Sigma_{vu} \right] \right) + o_p(n^{-(1/2)}) \end{aligned}$$

and hence $a_{i,n} = o_p(n^{-(1/2)})$. Therefore, $Z_i = \sigma_{vv}^{-(1/2)} v_i + o_p(n^{-(1/2)})$. ■

3. Limiting behaviour of the empirical process

Before presenting our main theorem, we introduce the Gaussian process briefly. Let R_1, \dots, R_n be independent random variables, which are normally distributed with zero mean and variance σ^2 . We define

$$B_n(\omega) = n^{-(1/2)} \sum_{i=1}^n \{1[\Phi(Z_i) \leq \omega] - \omega\}$$

$$Z_i = s^{-1}(R_i) \quad i = 1, \dots, n$$

for $\omega \in [0, 1]$, where $1(\cdot)$ is the indicator function, Φ is the standard normal distribution function and s^2 is an estimate of σ^2 . As n tends to infinity, B_n will converge in Skorohod space $D[0, 1]$ to W , which is a Gaussian process on $[0, 1]$, with mean $E[W(\omega)] = 0$, for $\omega \in [0, 1]$ and covariance function given by

$$E[W(\omega_1)W(\omega_2)] = \min(\omega_1, \omega_2) - \omega_1\omega_2$$

$$- \left[1 + \frac{1}{2} \Phi^{-1}(\omega_1)\Phi^{-1}(\omega_2) \right] \phi[\Phi^{-1}(\omega_1)]\phi[\Phi^{-1}(\omega_2)]$$

for $\omega_1, \omega_2 \in [0, 1]$ where ϕ is the standard normal density function. This result can be found in Kac *et al.* [29] and it will be useful if the random variables R_1, \dots, R_n are directly observable. In the case of regression residuals, this result does not apply because the v_i 's are not directly observable. Furthermore, the observable residuals are not independent and identically distributed normal random variables.

Miller [12, 13] generalized the previously given result to the class of empirical process based on certain non-independent and non-identical sequences of random variables. The class also contains the standardized residuals presented by Miller [12, 13]. In the following theorem, we show that our definition of the studentized recursive residuals and its limiting behaviour satisfies the conditions of the Miller's generalization.

THEOREM 2 Let v_1, \dots, v_n be independent $N(\mu_v, \sigma_{vv})$ random variables with $\sigma_{vv} > 0$, which are independent of the random vectors $\ddot{x}_1, \dots, \ddot{x}_n$ and $m_{\ddot{x}\ddot{x}} \xrightarrow{a.s.} m_{\ddot{x}\ddot{x}}$, where $m_{\ddot{x}\ddot{x}} = 1/(n-1) \sum_{i=1}^n (\ddot{x}_i - \bar{\ddot{x}})(\ddot{x}_i - \bar{\ddot{x}})'$, and $\bar{m}_{\ddot{x}\ddot{x}}$ is non-negative definite matrix. If we define

$$Z_i = \sigma_{vv}^{-(1/2)}(1 + a_n) + B_n \ddot{x}_i, \quad i = k+1, \dots, n,$$

where $a_n = o_p(n^{-(1/2)})$, $B_n = O_p(n^{-(1/2)})$, and

$$\hat{W}_n(\omega) = (n-k)^{-(1/2)} \sum_{i=k+1}^n [I(\Phi(Z_i) \leq \omega) - \omega]$$

$$\hat{W}_n(\omega) = (n-k)^{-(1/2)} \sum_{i=k+1}^n [I(\Phi(\sigma_{vv}^{-(1/2)} v_i) \leq \omega) - \omega] \quad \omega \in [0, 1]$$

then

$$\sup_{\omega \in [0, 1]} \left| \hat{W}_n(\omega) - W_n(\omega) \right| = o_p(1)$$

Proof First, we note that $\hat{W}_n(\omega)$ can be written as

$$\begin{aligned} \hat{W}_n(\omega) &= (n - k)^{-(1/2)} \sum_{i=k+1}^n [I(Z_i \leq \Phi^{-1}(\omega)) - \omega] \\ &= (n - k)^{-(1/2)} \sum_{i=k+1}^n [I(\sigma_{vv}^{-1/2} v_i \leq (1 + c_n)\Phi^{-1}(\omega) + D_n \ddot{x}_i) - \omega] + o_p(1) \end{aligned}$$

where $c_n = -(1 + a_n)^{-1} a_n$ and $D_n = -(1 + a_n)^{-1} B_n$.

If we define

$$\begin{aligned} W_{1,n}(\omega) &= (n - k)^{-(1/2)} \sum_{i=k+1}^n \{\Phi[R_{in}^*(\omega)] - \Phi[V_n(\omega)]\} \\ W_{2,n}(\omega) &= (n - k)^{-(1/2)} \sum_{i=k+1}^n \{I[U_i \leq \Phi(R_{in}^*(\omega))] - I(U_i \leq \omega) - \Phi(R_{in}^*(\omega)) + \omega\} \\ W_{3,n}(\omega) &= (n - k)^{-(1/2)} \sum_{i=k+1}^n \{I[U_i \leq \Phi(V_n(\omega))] - I(U_i \leq \omega) - \Phi(V_n(\omega)) + \omega\} \end{aligned}$$

in which

$$\begin{aligned} V_n(\omega) &= \sigma_{vv}^{-1/2} \Phi^{-1}(\omega) \\ R_{in}^*(\omega) &= V_n(\omega) + [\Phi^{-1}(\omega)c_n + D_n(\ddot{x}_i - \bar{\ddot{x}})]\sigma_{vv}^{-1/2} \\ U_i &= \Phi[\sigma_{vv}^{-1} v_i] \end{aligned}$$

After some mathematical manipulations, we can write

$$\hat{W}_n(\omega) = W_n(\omega) + W_{1,n}(\omega) + W_{2,n}(\omega) - W_{3,n}(\omega) + o_p(1)$$

Then, following the method used by Miller [28, Theorem A], we have

$$\sup_{\omega \in [0,1]} |W_{2,n}(\omega)| = o_p(1), \quad \sup_{\omega \in [0,1]} |W_{1,n}(\omega)| = o_p(1), \quad \sup_{\omega \in [0,1]} |W_{3,n}(\omega)| = o_p(1)$$

Therefore

$$\sup_{\omega \in [0,1]} |\hat{W}_n(\omega) - W_n(\omega)| = o_p(1) \quad \blacksquare$$

COROLLARY 1 Let $\hat{Z}_i = \Phi(\hat{\sigma}_{vvi}^{-1/2} s_i)$, $i = k + 1, \dots, n$, and $\hat{F}_n(\omega)$, $\hat{W}_n(\omega)$, for $\omega \in [0, 1]$, be the empirical distribution function and empirical process, respectively, based on $\hat{Z}_{k+1}, \dots, \hat{Z}_n$ such that $\hat{W}_n(\omega) = n^{-(1/2)}[\hat{F}_n(\omega) - \omega]$. Also, let $Z_i^* = \Phi(\sigma_{vv}^{-1/2} v_i)$, $i = k + 1, \dots, n$, and $\hat{F}_n^*(\omega)$, $W_n(\omega)$ be the empirical distribution function and empirical process, respectively, based on Z_{k+1}^*, \dots, Z_n^* such that $W_n(\omega) = n^{-(1/2)}[F_n(\omega) - \omega]$. Then, under the assumptions (a) and (b), $\sup_{\omega \in [0,1]} |\hat{W}_n(\omega) - W_n(\omega)| = o_p(1)$

Proof The result follows immediately from Theorem 2. \blacksquare

The consequence of Theorem 2 and Corollary 1 is that, by Kac *et al.* [29] $W_n \xrightarrow{L} W$, where W is the limiting Gaussian process given previously. Hence by Theorem 2, \hat{W}_n has the same limiting distribution as W_n .

The empirical process based on recursive residuals given in Theorem 2 and the one based on standardized residuals given by Miller [12, 13] have the same limiting behaviour. One may question the justification of using recursive residuals instead of standardized residuals in goodness-of-fit tests. The results of simulation study in section 6 will justify the use of recursive residuals in detecting normality of the errors in measurement error models.

4. Goodness-of-fit tests of normality

In the previous section, we found conditions under which \hat{W}_n and W_n converge to the same limiting Gaussian process W . A consequence of this result is that we can construct functions of W_n or \hat{W}_n , which are continuous with respect to the Skorohod topology (except at possibly a set of probability zero with respect to the distribution of the limiting process W). Then, we can use these functions to examine the limiting behaviour of goodness-of-fit statistics of normality. Fortunately, many tests of normality as functions of sample processes are available. The idea was originally proposed by Doob [30], which provides a way to describe the limiting distribution of goodness-of-fit statistics. Miller [12, 13] extended these results to the measurement error models using ordinary residuals. We consider some of the goodness-of-fit tests of normality based on recursive residuals in functional measurement error models.

The one-sided Kolmogrov–Smirnov statistics are defined by $\hat{K}_n^+ = (n - k)^{-(1/2)} \sup_{\omega \in [0, 1]} \hat{W}_n(\omega)$ and $\hat{K}_n^- = (n - k)^{-(1/2)} \inf_{\omega \in [0, 1]} \hat{W}_n(\omega)$. In terms of \hat{Z}_i 's, \hat{K}_n^+ and \hat{K}_n^- become $\hat{K}_n^+ = \max_{k+1 \leq i \leq n} (i / (n - k) - \hat{Z}_{i:n-k})$ and $\hat{K}_n^- = \max_{k+1 \leq i \leq n} (\hat{Z}_{i:n-k} - ((i - 1) / (n - k)))$, where $\hat{Z}_{i:n-k}$ stands for the i th order statistic from a sample of size $n - k$. The one-sided Kolmogrov–Smirnov statistics have not been widely used. Instead, the function $\hat{K}_n = (n - k)^{-(1/2)} \sup_{\omega \in [0, 1]} |\hat{W}_n(\omega)|$ has been used. \hat{K}_n is a continuous function of \hat{W}_n with respect to the Skorohod metric on $D[0, 1]$ (for more details see [7, Theorem 2.3.5]). Therefore, the limit laws of $(n - k)^{1/2} \hat{K}_n$ are given by $K = \sup_{\omega \in [0, 1]} |W_n(\omega)|$, which is also a continuous function of W_n . Stephen [31] has tabulated the distribution of K .

The Anderson–Darling statistic, originally studied by Anderson and Darling [32], is defined by $\hat{A}_{(n-k)}^2 = \int_0^1 [\omega(1 - \omega)]^{-2} \hat{W}_{(n-k)}^2(\omega) d\omega$. Durbin [1] showed that this function is continuous in the Skorohod metric on $D[0, 1]$, and hence under the conditions of Theorem 2, the limiting law of $\hat{A}_{(n-k)}^2$ is given by the law of the random variable $A^2 = \int_0^1 [\omega(1 - \omega)]^{-2} W^2(\omega) d\omega$. The percentage points for the distribution of A^2 have been tabled by Stephen [31].

5. Example

In this section, we consider examples from real data and examine the ability of goodness-of-fit statistics based on standardized residuals and recursive residuals in detecting the normality of the errors in measurement error models.

Briefly, the data set in this example arises from an extensive archaeological survey of pottery production and distribution in the ancient Egyptian city of Al-Amarna. The data consist of measurements of chemical contents (mineral elements) made on many samples of pottery, using two different techniques, neutron activation analysis (NAA) and inductively coupled plasma (ICP) (see [33] for description of techniques). The set of pottery has been collected from different locations around the city, and each pottery has its own

Table 1. Goodness-of-fit tests of normality of Na and Co measured with ICP and NAA.

Test		Co				Na			
		ICP	NAA	Z_i	$\sigma_{vv}^{-(1/2)} \hat{v}_i$	ICP	NAA	Z_i	$\sigma_{vv}^{-(1/2)} \hat{v}_i$
Kolmogorov–Smirnov	Statistic	0.11	0.129	0.09	0.25	0.15	0.20	0.14	0.186
	<i>p</i> -value	>0.15	>0.15	>0.15	<0.01	>0.15	0.10	>0.15	>0.15
Anderson–Darling	Statistic	0.33	0.47	0.24	2.84	0.48	0.56	0.19	0.59
	<i>p</i> -value	0.50	0.23	0.76	0.00	0.20	0.13	0.88	0.11

‘fabric code’. Among different mineral elements, we are interested in the relation between Co and Na measured with ICP against the same elements measured with NAA. Archaeologists believe that observations from pottery with the same fabric code and from the same provenance can essentially be regarded as replicated observations. Our preliminary analysis using the multivariate analysis of variance (not reported here) also confirmed this idea. Consequently, the Co and Na data sets have been divided into 28 and 15 groups, respectively, and in each group, there are replicated observations. Owing to limitation of space, the original data set is not given here, but it is available from the author on request.

We fitted a functional measurement error model to both data sets and measured the standardized residuals and recursive residuals. Then, we used Kolmogorov–Smirnov and Anderson–Darling goodness-of-fit statistics to test the normality of the data and errors of the model. The results of these tests are given in table 1. Results of this table show that although Na and Co measured with ICP and NAA have normal distribution and tests based on recursive residuals do not reject normality of the errors, tests based on standardized residuals will reject the hypothesis of normality of errors in Co data.

Figure 1 shows plot of recursive residuals and standardized residuals of Co data vs. groups number and indicates that standardized residuals are affected by deficiency within data. Removing the unusual data and fitting the model, the normality tests based on standardized residuals will no longer be rejected (*p*-value >0.15).

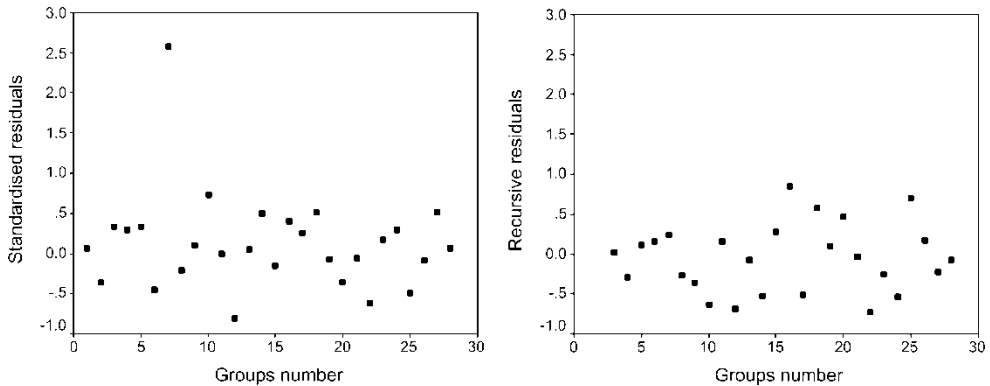


Figure 1. Plot of recursive residuals and standardized residuals of Co data vs. groups number.

6. Parametric bootstrap study

In this section, we perform parametric bootstrap simulation studies to examine the power of the proposed goodness-of-fit tests based on recursive residuals and standardized residuals in detecting normality of the errors for the small and medium sample sizes. The bootstrap studies are based on the Egyptian pottery data, which have been considered in the previous section. We refer to the Co and Na data sets as medium and small sample sizes, respectively. We used the Co data after removing unusual data.

We simulated a total of 1000 data sets for each case of Co and Na data sets in accordance with the model

$$\begin{aligned} Y_{ij} &= \hat{\beta}_0 + \hat{\beta}_1 \hat{x}_i + e_{ij} \quad i = 1, \dots, n \\ X_{ij} &= \hat{x}_i + u_{ij} \quad j = 1, \dots, r_i \end{aligned} \quad (7)$$

in which the groups number are 27 and 15, respectively, and the number of replications in each group (r_i) is the same as those for the original data set. In this model, $\hat{\beta}_0$, $\hat{\beta}_1$ and \hat{x}_i , are estimates of β_0 , β_1 and x_i , respectively, calculated from the original data. In addition, we assumed that $\varepsilon'_{ij} = (e_{ij}, u_{ij})$ has normal distribution with zero mean and covariance matrix $\hat{\Sigma}_{\varepsilon\varepsilon}$ based on the original data. Then, we fitted the functional measurement error model (7) to the each simulated data set and calculated Kolmogorov–Smirnov and Anderson–Darling goodness-of-fit test statistics based on studentized recursive residuals and standardized residuals. We selected different significant levels and compared the test statistics with the corresponding values given by Stephen [31].

Table 2 shows the percentage of times the goodness-of-fit test statistics, based on recursive residuals, for each significance level are less than the corresponding values, tabulated by Stephen [30]. Table 3 shows the same results based on standardized residuals. Results of table 2 show that for the Co, both tests give the type I errors that are near to the significant levels, whereas for the Na, Anderson–Darling gives the type I error larger than the significant levels, which indicates that for the small sample sizes, Anderson–Darling goodness-of-fit test based on recursive residuals in measurement error models is less accurate than Kolmogorov–Smirnov test.

Table 2. Simulation results of the goodness-of-fit test based on recursive residuals, using Co and Na from Egyptian pottery data.

Mineral elements	Test statistic	$p > 0.15$	$p > 0.10$	$p > 0.05$	$p > 0.025$	$p > 0.01$
Co	Kolmogorov–Smirnov	0.865	0.912	0.957	0.978	0.992
	Anderson–Darling	0.842	0.890	0.946	0.976	0.995
Na	Kolmogorov–Smirnov	0.852	0.902	0.947	0.984	0.993
	Anderson–Darling	0.791	0.851	0.916	0.953	0.982

Table 3. Simulation results of the goodness-of-fit test based on standardized residuals, using Co and Na from Egyptian pottery data.

Mineral elements	Test statistic	$p > 0.15$	$p > 0.10$	$p > 0.05$	$p > 0.025$	$p > 0.01$
Co	Kolmogorov–Smirnov	0.825	0.874	0.921	0.942	0.967
	Anderson–Darling	0.811	0.865	0.907	0.936	0.955
Na	Kolmogorov–Smirnov	0.798	0.836	0.883	0.914	0.926
	Anderson–Darling	0.784	0.829	0.874	0.896	0.917

A glance at the results of table 3 reveals that the type I error of tests for both Co and Na models are considerably larger than the significant levels. The overall results from both tables indicate that for small and medium sample size data sets, tests based on recursive residuals are more accurate than those based on standardized residuals.

Acknowledgements

I would like to thank the anonymous referee for his valuable comments and suggestions, which improved the presentation of this article considerably. This research is partially supported by the statistical Research center.

References

- [1] Durbin, J., 1973, Weak convergence of the sample distribution function when parameters are estimated. *Annals of Statistics*, **1**, 279–290.
- [2] Rao, J.S. and Sethurannan, J., 1975, Weak convergence of empirical distribution functions of random variables subject to perturbations and scale factors. *Annals of Statistics*, **3**, 299–313.
- [3] Neuhaus, G., 1976, Weak convergence under contiguous alternatives of the empirical process when parameters are estimated: the D_k approach. *Empirical Distributions and Processes. Lecture Notes in Math.*, Vol. 566 (New York: Springer), pp. 68–82.
- [4] Mugantseva, L.A., 1977, Testing normality in one-dimensional and multi-dimensional linear regression. *Theory of Probability Applications*, **22**, 591–602.
- [5] Pierce, D. and Kopecky, K., 1979, Testing goodness-of-fit for the distribution of errors in regression models. *Biometrika*, **66**, 1–6.
- [6] Shorack, G.R., 1984, Empirical and rank processes of observations and residuals. *Canadian Journal of Statistics*, **12**, 319–332.
- [7] Shorack, G.R. and Wellner, J.A., 1986, *Empirical Processes with Applications to Statistics* (New York: John Wiley).
- [8] Wood, C.L., 1984, On test of normality of experimental error in ridge regression. *Journal of Statistical Planning and Inference*, **9**, 367–374.
- [9] Portnoy, S., 1986, Asymptotic behavior of the empiric distribution of M-estimated residuals from a regression model with many parameters. *Annals of Statistics*, **14**, 1152–1170.
- [10] Loynes, R.M., 1980, The empirical distribution function of residuals from generalized regression. *Annals of Statistics*, **8**, 285–298.
- [11] Cox, D.R. and Snell, E.J., 1968, A general definition of residuals (with discussion). *Journal of Royal Statistical Society, B*, **30**, 248–275.
- [12] Miller, S.M., 1986, The limiting behavior of residuals from measurement error regressions. PhD dissertation, Iowa State University.
- [13] Miller, S.M., 1989, Empirical processes based upon residuals from errors in variables regressions. *Annals of Statistics*, **17**, 282–292.
- [14] Brown, R.L., Durbin, J. and Evans, J.M., 1975, Techniques for testing the constancy of regression relationships over time. *Journal of Royal Statistical Society, B*, **37**, 149–163.
- [15] Phillips, G.D.A. and Harvey, A.C., 1974, A simple test for serial correlation in regression analysis. *Journal of the American Statistical Association*, **69**, 935–939.
- [16] Riddell, W.C., 1980, Estimating switching regressions: a computational note. *Journal of Statistical Computation and Simulation*, **10**, 95–101.
- [17] Hawkins, D.M., 1991, Diagnostics for use with regression recursive residuals. *Technometrics*, **33**, 221–234.
- [18] Tobing, H. and McGilchrist, C.A., 1992, Recursive residuals for multivariate regression models. *Australian Journal of Statistics*, **34**, 217–232.
- [19] Loynes, R.M., 1987, Cross-validatory, recursive and partial residuals for non-linear models. *Communitons in Statistics, Part A—Theory and Methods*, **16**, 997–1011.
- [20] McGilchrist, C.A. and Matawie, K.M., 1998, Recursive residuals in generalised linear models. *Journal of Statistical Planning and Inference*, **70**, 335–344.
- [21] Magnus, J.R. and Sinha, A.K., 2005, On the Theil's errors. *Econometrics Journal*, **8**, 39–54.
- [22] Gaplin, J.S. and Hawkins, D.M., 1984, The use of recursive residuals in checking model fit in linear regression. *The American Statistician*, **38**, 94–105.
- [23] Kianifard, F. and Swallow, W.H., 1996, A review of the development and application of recursive residuals in linear models. *Journal of American Statistical Association*, **433**, 391–400.
- [24] Fuller, W.A., 1987, *Measurement Error Models* (New York: John Wiley).
- [25] Cheng, C.L. and Van Ness, J.W., 1999, *Statistical Regression with Measurement Error* (London: Arnold).
- [26] Amemiya, Y. and Fuller, W.A., 1984, Estimating for the multivariate errors in variables model with estimated errors covariance matrix. *Annals of Statistics*, **12**, 497–509.

- [27] Gleser, L.J., 1981, Estimation in a multivariate errors in variables regression model: large sample results. *Annals of Statistics*, **9**, 24–44.
- [28] Miller, S.M., 1990, Analysis of residuals from measurement error models. In statistical analysis of measurement error models and applications. In: P. Brown and W.A. Fuller (Eds). Proceedings of the AMS–IMS–SIAM Joint Summer Research Conference, June 10–16, Iowa State University, USA, 1989.
- [29] Kac, M., Kiefer, J. and Wolfowitz, J., 1955, On tests of normality and other tests of goodness of fit based on distance methods. *Annals of Mathematical Statistics*, **26**, 189–211.
- [30] Doob, J.L., 1949, Heuristic approach to the sample distribution function when parameters are estimated. *Annals of Mathematical Statistics*, **20**, 393–403.
- [31] Stephen, M.A., 1974, EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*, **69**, 730–737.
- [32] Anderson, T.W. and Darling, D.A., 1952, Asymptotic theory of certain goodness of fit criteria based on stochastic processes. *Annals of Mathematical Statistics*, **23**, 192–212.
- [33] Smith, D.M., Hart, F.A., Symond, R.D. and Walsh, J.N., 1988, Analysis of Roman pottery from Colchester by inductively coupled plasma spectrometry. In: E.A. Slater and J.O. Tate (Eds) *Science and Archaeology Glasgow 1987*, Vol. 196(i) (Oxford: B.A.R.), pp. 41–55.