TESTING FOR TWO-PHASE MULTIPLE REGRESSIONS

Max Moldovan

Queensland University of Technology, Brisbane, Australia
E-mail: m.moldovan@student.qut.edu.au

Nicholas A. Nechval, Edgars K. Yasermanis, Uldis Rozevskis & Kristine Rozite

Mathematical Statistics Department, University of Latvia
Raina Blvd 19, LV-1050 Riga, Latvia
Ph: (+371)-7034701, Fax: (+371)-7034702, E-mail: nechval@junik.lv

Abstract

This paper describes a technique for detecting departures from constancy of regression relationships over time when regression analysis is applied to time-series data. The problem we consider is a special case of a general class of problems concerned with the detection of changes of model structure over time. The technique proposed here for solving the above problem is based on transforming a set of random variables into a smaller set of random variables that are independently distributed with parameter-free distributions. It allows one to eliminate unknown parameters from the problem and is especially efficient when we deal with small samples of the data. A numerical example is given.

Key words: Switching regression, Segmented regression, Change-points, Detection

1. Introduction

In many areas of scientific endeavor, scientists are often faced with the problem of inference for the point in a sequence of time-series data at which regression relationships change. This problem is a special case of a general class of problems concerned with the detection of changes of model structure over time.

In the present paper we consider two-phase or switching multiple regression models which have many applications. For example, in econometrics a variable such as gross domestic product is assumed to have a normal distribution whose mean is a linear function of several explanatory variables, such as labor input and capital input, and it is suspected that the model parameters have changed after an unknown time. Brown, Durbin, and Evans [1] give three examples involving the growth in the number of local telephone calls, the demand for money, and staff requirements of an organization. In economics the problem is met with frequently. A regression model for a market may show a structural shift, for example, when the demand exceeds supply. This point may, of course, have economic significance. Teräsvirta [2] studied the price elasticities in the demand of post and telephone services in Finland from this point of view. His main aim was to detect over what period the regression structure describing the market had been stable, to get as firm a basis for predictions as possible. This is not an unusual problem in economic control and prediction situations. A common feature of the situations we have in mind is that there exists a natural ordering of the data points, either by time, size, age, or some other variable.

The problem was first studied by Quandt [3], who proposed a maximum likelihood method for estimating the parameters in two separate regression lines that switch at an unknown point τ. The first τ observations in a sample size T follow one regression model, and the last T-τ observations follow another. A likelihood ratio test was proposed by Quandt [4] to test for two separate regression lines as opposed to the null hypothesis that the data follow only one. Except in some special cases, the exact null distribution of the likelihood ratio statistic remains intractable. It depends on the values of the independent variables, and, as Hawkins [5] points out, no limiting distribution is known. Beckman and Cook [6] give some Monte Carlo results for a single independent variable.

The purpose of this paper is to suggest a new technique for solving the above problem, which is especially efficient when we deal with small samples of the data. It is shown that certain conditional distributions, obtained by conditioning on a sufficient statistic, can be used to transform a set of random variables into a smaller set of random variables that, under the null hypothesis (no change), are uniformly and independently distributed on the interval from zero to one. This result is then used to construct an exact distribution-free test for detection of changes in regression relationships. To illustrate the proposed technique, a numerical example is given.

2. Problem Statement

The basic regression model we consider is

$$y_t = \mathbf{x}_t'\mathbf{a}_t + w_t, \quad t = 1(1)T, \tag{1}$$

where at time t, y_t is the observation on the dependent variable and $\mathbf{x}_t$ is the column vector of observations on p regressors. The first regressor, x_{1t}, will be taken to equal unity for all values of t if the model contains a constant. The other regressors are assumed to be non-stochastic so auto-regressive models are excluded from consideration. The column vector of parameters, $\mathbf{a}_t = (a_{1t}, \ldots, a_{pt})'$, is written with the subscript t to indicate that it may vary with time. We assume that the error terms, w_t, are independent and normally distributed with means zero and variances σ_t^2, t = 1(1)T (the number of observations). The problem is to construct a test for constancy of regression relationships over time, which consists of testing the null hypothesis

$$H_0 : y_t \sim N(\mathbf{x}_t'\mathbf{a}_t, \sigma_t^2), \mathbf{a}_t = \mathbf{a}(0), \sigma_t^2 = \sigma^2(0), \quad t = 1(1)T, \tag{2}$$

against the alternative

$$H_1 : y_t \sim N(\mathbf{x}_t'\mathbf{a}_t, \sigma_t^2), \mathbf{a}_t = \mathbf{a}(0), \sigma_t^2 = \sigma^2(0), \quad t = 1(1)\tau;$$

$$\mathbf{a}_t = \mathbf{a}(1),\ \sigma_t^2 = \sigma^2(1), \quad t = \tau + 1(1)T, \tag{3}$$

where the parameters $\mathbf{a}(0) = (a_1(0), \ldots, a_p(0))'$, $\mathbf{a}(1) = (a_1(1), \ldots, a_p(1))'$, $\sigma^2(0)$, $\sigma^2(1)$, and τ, the point after which the change occurs, are unknown. The problem as stated is meaningful only if $1 \le \tau \le n-1$.

2.1. Preliminaries

The log-likelihood ratio technique, described in two papers by Quandt [3-4], is appropriate for solving the above problem when it is believed that the regression relationship may have changed abruptly at an unknown time point t=τ from one constant relationship specified by $\mathbf{a}(0)$, $\sigma^2(0)$, to another constant relationship specified by $\mathbf{a}(1)$, $\sigma^2(1)$. For each τ from τ=p+1 to τ = T-p-1 the program computes and plots

$$LR_\tau = \ln\left(\frac{max\ likelihood\ of\ the\ observations\ given\ H_0}{max\ likelihood\ of\ the\ observations\ given\ H_1}\right), \qquad (4)$$

where H_1 is the hypothesis that the observations in the time segments (1, … ,τ) and (τ+1, … ,T) come from two different regressions. This is the standard likelihood ratio statistic for deciding between the two hypotheses H_0 and H_1, and it is easy to show that

$$LR_\tau = (1/2)[\tau \ln \hat{\sigma}^2(0) + (T-\tau)\ln \hat{\sigma}^2(1) - T\ln \hat{\sigma}^2], \qquad (5)$$

where $\hat{\sigma}^2(0), \hat{\sigma}^2(1)$ and $\hat{\sigma}^2$ are the ratios of the residual sums of squares to number of observations when the regression is fitted to the first τ observations, the remaining T-τ observations and the whole set of T observations, respectively. The estimate of the point at which the switch from one relationship to another has occurred is then the value of τ at which LR_τ attains its minimum. Unfortunately, no test has yet been devised for min LR_τ since its distribution on H_0 is unknown. However, the behaviour of the graph of LR_τ against τ sheds light on the stability of the regression and in particular indicates whether changes have occurred as an abrupt transition or gradually. In this paper, we propose the technique, which has been found to be more effective for deciding between the two hypotheses H_0 and H_1 than the techniques that have been recommended in the literature. This technique has made it possible to construct an exact test for detecting abrupt or gradual changes in regression relationships from the data of small samples. It is based on the method, which leads to the problem of replacing composite hypothesis by equivalent simple one, best described by Prohorov [7], as follows: let ($\mathcal{Y}$,$\mathcal{A}$) be a measurable space and let $\mathcal{P}$ be a set of probability distributions defined on $\mathcal{A}$. Let ($\mathcal{Z}$,$\mathcal{B}$) be another measurable space and let $\mathbf{z}_m = f(\mathbf{y}_n)$, $\mathbf{y}_n \in \mathcal{Y}$, be a measurable mapping of ($\mathcal{Y}$,$\mathcal{A}$) into ($\mathcal{Z}$,$\mathcal{B}$). With this mapping every distribution P induces on $\mathcal{B}$ a corresponding distribution which we shall denote by $G(\mathbf{z}_m;P)$. We will be interested in the mappings (statistics) $\mathbf{z}_m$, $\mathbf{z}_m \in \mathcal{Z}$, which have the following two properties:

(i) $G(\mathbf{z}_m;P)$ is the same for all $P \in \mathcal{P}$; in this case we will simply write $G(\mathbf{z}_m;\mathcal{P})$; (6)

(ii) If for some P1 on A one has G(zm;P1) = G(zm;P), then P1$\in$P. (7)

If $\mathbf{z}_m$ is a statistic satisfying (6) and (7), then it is clear that the hypothesis that the distribution of $\mathbf{y}_n$ belongs to the class $\mathcal{P}$ is equivalent to the hypothesis that the distribution of $\mathbf{z}_m$ is equal to $G(\mathbf{z}_m;\mathcal{P})$. In this paper, we give a certain transformation, which satisfies the properties (6) and (7) and is sufficient to enable an exact test for deciding between the two hypotheses H_0 and H_1 to be constructed. The price paid for the elimination of unknown parameters is the fact that the number of the original variables has been decreased by the number of eliminated unknown parameters through these transformations.

3. Main Lemmas

Assuming H_0 to be true, let $\hat{\mathbf{a}}_t$ be the least-squares and maximum likelihood estimator of $\mathbf{a}_t$ for the regression model (1), if we set $\mathbf{a}_t = \mathbf{a}(0)$, $\sigma_t^2 = \sigma^2(0)$, $\forall t = 1(1)T$, based on the first t observations, i.e.,

$$\hat{\mathbf{a}}_t = (\mathbf{X}_t'\mathbf{X}_t)^{-1}\mathbf{X}_t'\mathbf{y}_t, \quad t = p+1, \ldots, T, \qquad (8)$$

where

$$\mathbf{X}_t' = (\mathbf{x}_1, \ldots, \mathbf{x}_t), \quad \mathbf{y}_t = (y_1, \ldots, y_t)', \qquad (9)$$

and the matrix $\mathbf{X}_t'\mathbf{X}_t$ is assumed to be non-singular. The unbiased estimator of σ_t^2 is given by

$$\hat{\sigma}_t^2 = s_t^2/(t-p), \tag{10}$$

where s_t^2 is the residual sum of squares after fitting the model to the first t observations, i.e.,

$$s_t^2 = (\mathbf{y}_t - \mathbf{X}_t\hat{\mathbf{a}}_t)'(\mathbf{y}_t - \mathbf{X}_t\hat{\mathbf{a}}_t). \tag{11}$$

The estimators $\hat{\mathbf{a}}_t$ and $\hat{\sigma}_t^2$ are independently distributed as follows:

$$\hat{\mathbf{a}}_t \sim N_p(\mathbf{a}_t, \sigma_t^2(\mathbf{X}_t'\mathbf{X}_t)^{-1}), \tag{12}$$

$$\frac{(t-p)\hat{\sigma}_t^2}{\sigma^2} \sim \chi_{t-p}^2 \tag{13}$$

with t-p degrees of freedom.

Lemma 1. (*Recurrence Relations*).

$$(\mathbf{X}_t'\mathbf{X}_t)^{-1} = (\mathbf{X}_{t-1}'\mathbf{X}_{t-1})^{-1} - \frac{(\mathbf{X}_{t-1}'\mathbf{X}_{t-1})^{-1}\mathbf{x}_t\mathbf{x}_t'(\mathbf{X}_{t-1}'\mathbf{X}_{t-1})^{-1}}{1+\mathbf{x}_t'(\mathbf{X}_{t-1}'\mathbf{X}_{t-1})^{-1}\mathbf{x}_t}, \tag{14}$$

$$\mathbf{a}_t' = \mathbf{a}_{t-1}' + (\mathbf{X}_t'\mathbf{X}_t)^{-1}\mathbf{x}_t(y_t - \mathbf{x}_t'\hat{\mathbf{a}}_{t-1}), \tag{15}$$

$$s_t^2 = s_{t-1}^2 + v_t^2, \quad t = p+1, \dots, T, \tag{16}$$

where

$$v_t = (y_t - \mathbf{x}_t'\hat{\mathbf{a}}_{t-1}) \big/ [1+\mathbf{x}_t'(\mathbf{X}_{t-1}'\mathbf{X}_{t-1})^{-1}\mathbf{x}_t]^{1/2}. \tag{17}$$

Proof. The relation (14) was given by Plackett [8] and Bartlett [9]. It is used in the program to avoid having to invert the matrix ($\mathbf{X}'_t\mathbf{X}_t$) directly at each stage of the calculations. It is proved by multiplying the left-hand side by $\mathbf{X}'_t\mathbf{X}_t$ and the right-hand side by $\mathbf{X}'_t\mathbf{X}_t = \mathbf{X}'_{t-1}\mathbf{X}_{t-1}+\mathbf{x}'_t\mathbf{x}_t$. Since $\hat{\mathbf{a}}_t$ is the least-squares estimate it satisfies

$$\mathbf{X}_t'\mathbf{X}_t\hat{\mathbf{a}}_t = \mathbf{X}_t'\mathbf{y}_t = \mathbf{X}_{t-1}'\mathbf{y}_{t-1} + \mathbf{x}_t y_t = \mathbf{X}_{t-1}'\mathbf{X}_{t-1}\hat{\mathbf{a}}_{t-1} + \mathbf{x}_t y_t = \mathbf{X}_t'\mathbf{X}_t\hat{\mathbf{a}}_{t-1} + \mathbf{x}_t(y_t - \mathbf{x}_t'\hat{\mathbf{a}}_{t-1}). \tag{18}$$

This implies (15).

$$s_t^2 = (\mathbf{y}_t - \mathbf{X}_t\hat{\mathbf{a}}_t)'(\mathbf{y}_t - \mathbf{X}_t\hat{\mathbf{a}}_t) = (\mathbf{y}_t - \mathbf{X}_t\hat{\mathbf{a}}_{t-1})'(\mathbf{y}_t - \mathbf{X}_t\hat{\mathbf{a}}_{t-1}) - (\hat{\mathbf{a}}_t - \hat{\mathbf{a}}_{t-1})'\mathbf{X}_t'\mathbf{X}_t(\hat{\mathbf{a}}_t - \hat{\mathbf{a}}_{t-1})$$

$$= s_{t-1}^2 + (y_t - \mathbf{x}_t'\hat{\mathbf{a}}_{t-1})^2 - \mathbf{x}_t'(\mathbf{X}_t'\mathbf{X}_t)^{-1}\mathbf{x}_t(y_t - \mathbf{x}_t'\hat{\mathbf{a}}_{t-1}) \tag{19}$$

which gives (16) on substituting for $(\mathbf{X}'_t\mathbf{X}_t)^{-1}$ from (14). This ends the proof.

Lemma 2 (*Normality and Independence of Prediction Errors*). Under H_0 (assuming that $\mathbf{a}_t = \mathbf{a}(0)$, $\sigma_t^2 = \sigma^2(0)$, $\forall t = 1(1)T$, for the regression model (1)), $v_{p+1}, \dots, v_T$ are independent, $N(0, \sigma^2(0))$, where v_t ($t = p+1, \dots, T$, see (17)) is the standardized prediction error of y_t when predicted from $y_1, \dots, y_{t-1}$.

Proof. The unbiasedness of v_t is obvious and the assertion $\mathrm{Var}\{v_t\} = \sigma_t^2$ follows immediately from the independence of y_t and $\hat{a}_{t-1}$. Also,

$$v_t = \left(w_t - \mathbf{x}_t'(\mathbf{X}_{t-1}'\mathbf{X}_{t-1})^{-1}\sum_{i=1}^{t-1}\mathbf{x}_i w_i \right)[1+\mathbf{x}_t'(\mathbf{X}_{t-1}'\mathbf{X}_{t-1})^{-1}\mathbf{x}_t]^{-1/2}. \tag{20}$$

Since each v_t is a linear combination of the normal variates w_j, the v_t's are jointly normally distributed. Now

$$E\left\{\left(w_t - \mathbf{x}_t'(\mathbf{X}_{t-1}'\mathbf{X}_{t-1})^{-1}\sum_{i=1}^{t-1}\mathbf{x}_i w_i \right)\left(w_j - \mathbf{x}_j'(\mathbf{X}_{j-1}'\mathbf{X}_{j-1})^{-1}\sum_{i=1}^{j-1}\mathbf{x}_i w_i \right)\right\}$$

$$= \sigma^2(0)[0-0-\mathbf{x}_j'(\mathbf{X}_{j-1}'\mathbf{X}_{j-1})\mathbf{x}_t + \mathbf{x}_t'(\mathbf{X}_{t-1}'\mathbf{X}_{t-1})^{-1}(\mathbf{X}_{t-1}'\mathbf{X}_{t-1})(\mathbf{X}_{j-1}'\mathbf{X}_{j-1})^{-1}\mathbf{x}_j'] = 0\ (t<j). \tag{21}$$

It follows that $v_{p+1}, \ldots, v_T$ are uncorrelated and therefore independent in view of their joint normality.

The transformation from the w_t's to the v_t's is a generalized form of the Helmert transformation (Kendall and Stuart [10]).

Lemma 3 ***(First Characterization of Prediction Errors).*** Let yt, t = 1(1)T, be T real independent random variables, with means x′ta(0), t=1(1)T, respectively, and common variance σ2(0) (σ(0)>0). Then vt, t = p+1(1)T, given by (17) are independently and identically distributed normal with mean zero and variance σ2(0) if and only if the yt (t = 1(1)T) are normal with means x′ta(0) and variance σ2(0).

Proof. This rests on Lemma 2 and a result of Cramer (Lukacs [11]).

Lemma 4 (*Student's Random Variables*). Suppose that $v_{p+1}, \ldots, v_T$ are T-p real independent random variables, T-p>1. If v_t ($t = p+1, \ldots, T$) are independently and identically distributed normal variates with zero mean and common standard deviation $\sigma(0)$ ($\sigma(0) > 0$) then $z_{p+2}, \ldots, z_T$ given by

$$z_t = [t-(p+1)]^{1/2} v_t / s_{t-1}, \quad t = p+2, \ldots, T, \tag{22}$$

are random variables independently distributed according to Student's law with 1,2, …, T-(p+1) degrees of freedom respectively.

Proof. This follows by using the fact that $\hat{\mathbf{a}}_t$ and s_t^2, $\forall t = p+1(1)T$, are independently distributed and the result of Basu's lemma [12].

Lemma 5 (*Second Characterization of Prediction Errors*). The necessary and sufficient condition for v_t, $t = p+1, \ldots, T$, to be independently and identically distributed normal with zero mean and common standard deviation $\sigma(0)$ is that $z_{p+2}, \ldots, z_T$ given by (22) are independently distributed according to Student's law with 1, 2, …, T-(p+1) degrees of freedom respectively, and $T \geq p+3$.

Proof. It can be shown, after some algebra (see Lemma 1), that

$$z_t = [t-(p+1)]^{1/2} v_t \left(\sum_{j=p+1}^{t-1} v_j^2 \right)^{-1/2}, \quad t = p+2, \ldots, T. \tag{23}$$

We can therefore apply Kotlarski's result [13] to the random variables $v_{p+1}, \ldots, v_T$ and obtain $z_{p+2}, \ldots, z_T$. Then the proof follows by using this result and Lemma 4.

4. Characterization Theorems

Theorem 1 (*Characterization of a Normal Regression via the Student Distribution*). Let y_t, $t = 1(1)T$, be T real independent random variables ($T \geq p+3$) with means $\mathbf{x}'_t\mathbf{a}(0)$, $t = 1(1)T$, respectively, and common variance $\sigma^2(0)$ ($\sigma(0)>0$). Let v_t, $t = p+1, \ldots, T$, be defined by (17) and let z_t, $t = p+2, \ldots, T$, be defined by (22), then the y_t ($t = 1, \ldots, T$) are $N(\mathbf{x}'_t\mathbf{a}(0), \sigma^2(0))$ if and only if $z_{p+2}, \ldots, z_T$ are independently distributed according to Student's law with 1, 2, … ,T-(p+1) degrees of freedom respectively.

Proof. This follows immediately by applying Lemma 3 and Lemma 5.

Thus we can replace the composite null hypothesis

$$H_0 : y_t \sim N(\mathbf{x}'_t\mathbf{a}(0), \sigma^2(0)), \quad t = 1(1)T \quad (T \geq p+3), \tag{24}$$

by the simple equivalent null hypothesis

$$H_0^{\bullet} : z_{p+2}, \ldots, z_T \tag{25}$$

are independently distributed according to Student's law with 1, 2, … ,T-(p+1) degrees of freedom respectively. It should be also appreciated here that the number of the original variables has only been decreased by the number of eliminated unknown parameters through these transformations.

Theorem 2 (*Characterization of a Normal Regression via the F Distribution*). Let y_t, $t = 1(1)T$, be T real independent random variables ($T \geq p+3$) with means $\mathbf{x}'_t\mathbf{a}(0)$, $t = 1(1)T$, respectively, and common variance $\sigma^2(0)$ ($\sigma(0)>0$).

Let z_t^2, $t = p+2, \ldots, T$, be defined by

$$z_t^2 = [t-(p+1)]\left(\frac{s_t^2}{s_{t-1}^2} - 1\right), \quad t = p+2, \ldots, T, \tag{26}$$

then the y_t ($t = 1, \ldots, T$) are $N(\mathbf{x}'_t\mathbf{a}(0), \sigma^2(0))$ if and only if $z_{p+2}^2, \ldots, z_T^2$ (where $z_t^2 \sim F_{1,t-(p+1)}$) are independently distributed according to the central F distribution with 1 and 1, 2, … ,T-(p+1) degrees of freedom, respectively.

Proof. By using the results of Lemma 4 and Theorem 1, the proof being straightforward is omitted.

Here the following theorem clearly holds.

Theorem 3 (*Characterization of a Normal Regression via the Uniform Distribution*). The y_t ($t = 1, \ldots, T$) are $N(\mathbf{x}'_t\mathbf{a}(0), \sigma^2(0))$ if and only if the random variables $u_{p+2}, \ldots, u_T$ are independently and uniformly distributed on the interval from zero to one (i.i.d. U(0,1)), where

$$u_t = 1 - F_{1,t-(p+1)}(z_t^2), \quad t = p+2(1)T, \tag{27}$$

$F_{1,t-(p+1)}(.)$ is the cumulative distribution function of the central F distribution with 1 and T-(p+1) degrees of freedom, respectively.

Proof. This follows immediately by applying Theorem 2 and the probability integral transformation theorem.

5. Estimating the Point of a Probable Change

An estimation of the point of a probable change in a sequence of observations is based on the so-called C_p criterion [14]. The theoretical derivation of this quantity is beyond the scope of this paper. However, C_p can be expressed as follows:

$$C_p = (N - p - 1)\left(\frac{RMS_p}{RMS} - 1\right) + (p + 1), \qquad (28)$$

where RMS_p is the residual mean square based on p selected parameters, and RMS is the residual mean square derived from the total set of independent parameters. The quantity C_p is the sum of two components, p+1 and the remainder of the expression. While p+1 increases as we choose more independent parameters, the other part of C_p will tend to decrease. When all parameters are chosen, p+1 is at its maximum but the other part of C_p is zero since RMS_p=RMS. Many investigators recommend selecting those independent parameters that minimize the value of C_p. Another recommended criterion for parameter selection is Akaike's information criterion (AIC). This criterion has been shown to be equivalent to C_p. The similarity to C_p can be seen by using a formula for estimating AIC given by Atkinson [15],

$$AIC = \frac{RSS_p}{RMS} + 2(p + 1), \qquad (29)$$

where RSSp is the residual sum of squares based on the p selected parameters. As more parameters are added the fraction part of the formula gets smaller but p+1 gets larger. A minimum value of AIC is desired. The estimate $\hat{\tau}$ of the point τ at which the switch from one regression relationship to another has occurred is given by (28) or (29).

6. Recognition of a Change

Unfortunately, a probability distribution of (27) or (28) is unknown. Therefore, for testing the null hypothesis H_0 against the alternative H_1 we use the results of Theorem 2 and Theorem 3. Now the question is how are we going to test H_0 of (2) against H_1 of (3). It should be noted that the results of Theorem 2 and Theorem 3 can be used in performing an exact test for recognition of the change point in regression relationships over time. The hypothesis H_0 is rejected if there is z_t^2, $t \in \{p+2, \ldots, T\}$, such that $z_t^2 \geq F_{1,t(p+1);\alpha}$ (a critical value for $F_{1,t(p+1)}$ distributed random variable at the significance level α), where the result of Theorem 2 is used.

The result of Theorem 3 can be used to obtain test for the hypothesis of the form H_0: y_t follows $N(\mathbf{x}_t'\mathbf{a}(0), \sigma^2(0))$ versus H_a: y_t does not follow $N(\mathbf{x}_t'\mathbf{a}(0), \sigma^2(0))$, $\forall t = 1(1)T$. The general strategy is to apply Theorem 3 to obtain a set of i.i.d. U(0,1) random variables under H_0. Under H_a this set of random variables will, in general, not be i.i.d. U(0,1). Any statistic, which measures a distance from uniformity in the transformed sample (say, a Kolmogorov-Smirnov statistic), can be used as a test statistic.

7. Numerical Example

To illustrate the suggested technique, we take, as an example, data presented by Beckman and Cook [6] : (x_1 = 1.00, y_1 = 6.4), (x_2 = 1.50, y_2 = 7.0), (x_3 = 1.75, y_3 = 7.4), (x_4 = 2.10, y_4 = 8.8), (x_5 = 2.35, y_5 = 9.0), (x_6 = 2.65, y_6 = 6.4), (x_7 = 3.00, y_7 = 6.6). It is necessary to test the null hypothesis

$$H_0: \; y_t = a_1(0) + a_2(0)x_t + w_t, \;\; t = 1(1)7, \tag{30}$$

against the alternative

$$H_1: \; y_t = a_1(0) + a_2(0)x_t + w_t, \;\; t = 1(1)\tau;$$

$$y_t = a_1(1) + a_2(1)x_t + w_t, \;\; t = \tau + 1(1)7, \tag{31}$$

where w_t, $t = 1(1)7$, is a sequence of independent random variables each is assumed to be $N(0,\sigma^2)$. The parameters $a_1(0)$, $a_2(0)$, $a_1(1)$, $a_2(1)$, σ^2 and k, the point after which the change occurs, are unknown.

In order to detect a possible change in the regression coefficients, we use the technique proposed above. This technique gives

$$\breve{\tau} = 5 \text{ and } z_6^2 = 49.61. \tag{32}$$

Since

$$z_6^2 > F_{1,3;\alpha=0.01} = 34.1, \tag{33}$$

the above null hypothesis is rejected using type I error $\alpha = 0.01$. This result is the same as in Beckman and Cook [6], but no simulation method for determining the critical values for testing for a two-phase regression is used.

8. Conclusions

It has been shown that the technique proposed in this paper provides a sufficient way to make informal references about the change-point of the linear model.

Many other problems can be studied. For example, one could consider the problem of more than one change, or the detection of outliers, or prediction of future observations of a changing linear model.

Acknowledgments

This research was supported in part by Grant No.02.0918 and Grant No.01.0031 from the Latvian Council of Science and the National Institute of Mathematics and Informatics of Latvia.

References

1. R.L. Brown, J. Durbin and J.M. Evans, Techniques for testing the constancy of regression relationships over time, Journal of the Royal Statistical Society, 37, 1975, 149-192 (with discussion).
2. T. Teräsvirta, On Stepwise Regression and Economic Forecasting (Economic Studies XXXI, Helsinki: The Finnish Economic Association, 1970).
3. R.E. Quandt, The estimation of the parameters of a linear regression system obeying two separate regimes, Journal of the American Statistical Association, 53, 1958, 873-880.
4. R.E. Quandt, Test of the hypothesis that a linear regression obeys two separate regimes, Journal of the American Statistical Association, 55, 1960, 324-330.

5. D.M. Hawkins, A note on continuous and discontinuous segmented regressions, Technometrics, 22, 1980, 443-444.

6. R.J. Beckman and R.D. Cook, Testing for two-phase regressions, Technometrics, 21, 1979, 65-69.

7. Yu.V. Prohorov, Some characterization problems in statistics, Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1 (Berkeley: University of California Press, 1966, 341-349).

8. R.L.Plackett, Some theorems in least squares, Biometrika, 1950, 37, 1950, 149-157.

9. M.S. Bartlett, An inverse matrix adjustment arising in discriminant analysis, Ann. Math. Statist., 22, 1951, 107-111.

10. M.G. Kendall and A. Stuart, The Advanced Theory of Statistics (London: Griffin, 1969).

11. E. Lukacs, Characteristic Functions (London: Griffin, 1960).

12. D. Basu, On statistics independent of a complete sufficient statistics, Sankhyā, 15, 1955, 377-380.

13. I. Kotlarski, On characterizing the normal distribution by Student's law, Biometrika, 53, 1966, 603-606.

14. C.L. Mallows, Some comments on Cp, Technometrics, 15, 1973, 661-676.

15. A.C. Atkinson, A note on the generalized information criterion for a choice of a model, Biometrika, 67, 1980, 413-418.