CHAPTER VI

TEST FOR SWITCHING IN AUTOREGRESSIVE (SAR) AND THRESHOLD IN TIME SERIES MODELS

Increasing attention has been devoted to the problem of parameter variation in regression and time series models. The first of these approaches allow for an infinite number of possible parameter values and for random parameter variations. This refers to the situation of random coefficient regression model or in time series, random coefficient autoregressive model. Alternatively, the number of possible changes may be finite (usually small) where each possible state of the parameter vector may be called a regime. In time series analysis these regimes may be associated with fundamental structural changes. This time series technique may be referred as switching autoregressive (SAR) model.

6.1 GENERAL CONCEPT

The simplest switching autoregressive model may be formulated as follows. Let n observations be available on a time series \( z_t \) (\( t = 1, 2, \ldots, n \)). There may be reason to believe that observation on \( z_t \) were generated by two distinct autoregressive models or regimes, i.e.

\[
\begin{align*}
    z_t &= \mu_1 + \phi_1 z_{t-1} + a_{1t} \quad t \leq t_0 \\
    z_t &= \mu_2 + \phi_2 z_{t-1} + a_{2t} \quad t > t_0
\end{align*}
\]
\( a_{1t} \) and \( a_{2t} \) are error terms assumed to be distributed as \( N(0, \sigma^2_1) \), and \( N(0, \sigma^2_2) \) and \( \sigma^2_1 \) is not equal to \( \sigma^2_2 \) in general.

Several econometric approaches have been introduced to deal with switching regression models under a variety of conditions. The principle difference among condition is whether nature between the two regimes is assumed to be stochastic, i.e. depend on unknown probabilities \( \lambda \) and \( 1-\lambda \) respectively or deterministic in the sense that it depends on the comparison of an observable variable \( z \) with cut off value \( z_o \) which is unknown. A special case of this mechanism is one in which the variable \( z \) is the time index of the observations.

6.2 DETERMINISTIC SWITCHING BASED ON TIME INDEX

Quandt (1960) tackled the problem of the estimation of unknown switching point \( z_o \) for testing the null hypothesis i.e. no switch has occurred against the alternative that one switch took place. Assume that \( n \) observations are available on a dependent variable \( y \) such that

\[
\begin{align*}
y_i &= x_i \beta_{1t} + u_{1t} & i \leq i^* \\
y_i &= x_i \beta_{2t} + u_{2t} & i > i^*
\end{align*}
\]

on a time index. He gave the likelihood ratio

\[
\mu = \frac{\hat{\sigma}^2_i \hat{\sigma}^{(n-i^*)} / \hat{\sigma}^n}{\hat{\sigma}^2_i / \hat{\sigma}^n} \]

(3)
choosing the estimate \( \hat{\beta} \) which maximises likelihood

\[
L(y_i/\hat{\beta}) = \left(2\pi\right)^{-n/2} \sigma^{-n/2} \exp\left(-\frac{1}{2\sigma^2} \sum(y_i - \hat{\beta} \cdot x_i)^2\right) \tag{4}
\]

where \( \hat{\sigma} \) is the estimated standard deviation of the residuals from a single regression over the entire sample. This technique provides a method both for estimation and testing.

The distribution of the relevant likelihood ratio \( \mu \) is such that \(-\log \mu\) has the \( \chi^2 \) distribution with appropriate degrees of freedom is rejected and an empirical table of percentage points is obtained.

6.3 STOCHASTIC CHOICE OF REGIMES

It is assumed that the error terms are normally distributed, the probability density function corresponding to (1) and (2) are

\[
f_1(y_i) = \frac{1}{\sigma_1 \left(2\pi\right)^{1/2}} \exp\left(-\frac{1}{2\sigma_1^2} (y_i - \hat{\beta}_1 \cdot x_i)^2\right) \tag{5}
\]

\[
f_2(y_i) = \frac{1}{\sigma_2 \left(2\pi\right)^{1/2}} \exp\left(-\frac{1}{2\sigma_2^2} (y_i - \hat{\beta}_2 \cdot x_i)^2\right) \tag{6}
\]

It has been suggested by Quandt (1971, 1972) that one may think of
nature choosing regime (1) and (2) with unknown probability \( \lambda \) and \( 1-\lambda \). The regime switching problem then clearly becomes a problem of mixture distribution. The probability density function of \( y \) then becomes

\[
h(y_i) = \lambda f_{1i} + (1-\lambda) f_{2i}.
\]

The likelihood function of which can be written as

\[
L = \prod_{i=1}^{n} \left( \frac{\lambda}{\sigma_1 (2\pi)^{-1/2}} \exp\left( -\frac{1}{2\sigma_1^2} (y_i - x \beta_{1i})^2 \right) + \frac{1-\lambda}{\sigma_2 (2\pi)^{-1/2}} \exp\left( -\frac{1}{2\sigma_2^2} (y_i - x \beta_{2i})^2 \right) \right)
\]

or

\[
\ln(L) = \sum \ln(\lambda f_{1i} + (1-\lambda) f_{2i})
\]

Log(L) is to be maximised with respect to the parameters of (5) and (6) and \( \lambda \). Tests of null hypothesis again may employ the natural likelihood ratio.

6.4 TEST FOR SWITCHING IN AUTOREGRESSIVE (SAR) MODELS

Let,

\[
z_t = \mu_1 + \phi_1 z_{t-1} + a_{1t}, \quad t \leq t_0 \tag{8}
\]

\[
z_t = \mu_2 + \phi_2 z_{t-1} + a_{2t}, \quad t > t_0 \tag{9}
\]
be the two AR models which the $z_t$ follows in the two separate regimes. \( \{a_j\}; j=1,2 \) is the sequence of identically and independently distributed random variables with mean zero and variance $\sigma^2_a$. It may be mentioned here that we have assumed similar noise (the distribution is same) in different models. This assumption is usually true in practice but theoretically there can be different noise sequences in different models.

Our objective is to develop suitable small test for testing the hypothesis that no switch has occurred in the parameters of an autoregressive system against the alternative that one switch took place. In other words, the variance of the autoregressive equations and the parameter values are same in both the regimes. Therefore,

Null hypothesis, $H_0$: no switch occurred

Alternate hypothesis, $H_1$: one switch took place.

6.4.1 TEST-1

Let the cutoff value of series $z_t$, $t_0$ is known which divides the whole series into two regimes. Let $\hat{\phi}_1$ and $\hat{\phi}_2$ are the estimates of $\phi_1$ and $\phi_2$ respectively. The residuals of the two equations can be obtained in the following way

$$ a_{1t} = z_t - \mu_2 - \hat{\phi}_2 z_{t-1}, \quad t \leq t_0 $$

$$ a_{2t} = z_t - \mu_1 - \hat{\phi}_1 z_{t-1}, \quad t > t_0 $$
The procedure consists of obtaining the mean of the residuals, say

\[ \overline{a}_1 = \frac{1}{n_1} \sum_{t=1}^{n_1} (z - \mu_2 - \phi_2 z_{t-1} - \overline{a}_1), \quad t \leq t_0 \]

\[ \overline{a}_2 = \frac{1}{n_2} \sum_{t=1}^{n_2} (z - \mu_2 - \phi_2 z_{t-1} - \overline{a}_2), \quad t > t_0 \]

where \( n_1 \) and \( n_2 \) are the number of observations in the two regimes according to \( t \leq t_0 \) and \( t > t_0 \). Under null hypothesis the terms summed-up above are normally and independently distributed. The variances of the residuals in both equations are defined as

\[ s_1^2 = \frac{n_1}{n_1 - 1} \left( \frac{\sum_{t=1}^{n_1} (z - \mu_2 - \phi_2 z_{t-1} - \overline{a}_1)^2}{n_1} \right) \]

and

\[ s_2^2 = \frac{n_2}{n_2 - 1} \left( \frac{\sum_{t=1}^{n_2} (z - \mu_2 - \phi_2 z_{t-1} - \overline{a}_2)^2}{n_2} \right) \]

The quantities

\[ \frac{(n_1)^{1/2} (\overline{a}_1 - E(\overline{a}_1))}{s_1} \quad \text{and} \quad \frac{(n_2)^{1/2} (\overline{a}_2 - E(\overline{a}_2))}{s_2} \]
would appear to have t-distribution. $H_0$ will be rejected if either of the mean residuals is significantly different from zero.

One major drawback of this test is that distribution of test statistic will be affected by cutting off point, since $n_1$ and $n_2$ depend upon cutoff point, therefore, the use of students t-distribution with $n_1 - 1$ and $n_2 - 1$ degrees of freedom is incorrect.

6.4.2 TEST-2

In order to avoid difficulty mentioned in the TEST-1, the observations here are divided in two groups of $t$ and $n-t$ observations where $t=n/2$ if $n$ is even and $t=(n-1)/2$ or $t=(n+1)/2$ if $n$ is odd. Let the residuals be defined as earlier then the quantities

$$\frac{(n-t-1)^{1/2}}{\hat{\sigma}_1} \quad \text{and} \quad \frac{(t-1)^{1/2}}{\hat{\sigma}_2}$$

have the standard t-distribution with $n-t-1$ and $t-1$ degree of freedom respectively. Where

$$\hat{\sigma}_1^2 = \frac{\sum_{i=1}^{t} (z_i - \hat{\mu}_1 - \hat{\phi}_1 z_i) z_i^2}{t}$$

and

$$\hat{\sigma}_2^2 = \frac{\sum_{i=1}^{n-t} (z_i - \hat{\mu}_2 - \hat{\phi}_2 z_i) z_i^2}{n-t}$$
This test might generally be more powerful than TEST-1 in which the denominator of the test statistic is $s_1$ (or $s_2$) rather than the standard error of the estimate $\hat{\sigma}_1^2$ (or $\hat{\sigma}_2^2$). One would generally expect the standard error of the estimate to be less than the standard deviation of the residuals if the null hypothesis is false. Hence for given values of $\tilde{a}_1$ and $\tilde{a}_2$ the absolute value of the test statistic would tend to larger under TEST-2 than under TEST-1, thus making it easier to reject the null hypothesis when it is false.

6.4.3 TEST-3

Under null hypothesis the quantities

$$F_1 = \frac{\sum_{t=1}^{n_1} (z - \hat{\mu}_2 - \hat{\phi}_2 z_{t-1})^2}{\sigma^2}$$

and

$$F_2 = \frac{\sum_{t=n_1+1}^{n_1+n_2} (z - \hat{\mu}_2 - \hat{\phi}_2 z_{t-1})^2}{\sigma^2}$$

of the two regimes are independently distributed as $\chi^2$ with $n_1-1$ and $n_2-1$ degree of freedom respectively. Their ratios

$$W = \frac{\max(F_1, F_2)}{\min(F_1, F_2)}$$

will follow F-distribution with degree of freedom of $\max(F_1, F_2)$ and degree of freedom of $\min(F_1, F_2)$ respectively. An analogous F-ratio can be evaluated by utilizing the calculated left hand
regression in the following way:

\[ F_1 = \sum_{t=1}^{n_1} (z_t - \hat{\mu}_1 - \phi_1 z_{t-1})^2 / \sigma^2 \]

and

\[ F_2 = \sum_{t=n_1+1}^{n_1+n_2} (z_t - \hat{\mu}_2 - \phi_2 z_{t-1})^2 / \sigma^2 \]

of the two regimes are independently distributed as \( \chi^2 \) with \( n_1 - 1 \) and \( n_2 - 1 \) degree of freedom respectively. Their ratios

\[ W = \frac{\max(F_1, F_2)}{\min(F_1, F_2)} \]

will also follow \( F \)-distribution with degree of freedom of \( \max(F_1, F_2) \) and degree of freedom of \( \min(F_1, F_2) \) respectively. The null hypothesis can be rejected if either of the \( F \)-ratios is larger than the characteristic value of \( F \) at respective degree of freedom.

6.5 TEST FOR THRESHOLD IN TIME SERIES MODELS

6.5.1 Durbin's Method (based on AR representation)

Suppose,

\[ x_t = e_t + \beta e_{t-k}, \quad t=1,2,...,n \quad (7) \]

where \( e_t \) follows \( N(0, \sigma^2) \) and \( |\beta| < 1 \). It follows an infinite autoregressive representation
\[ x + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \ldots = e_t \tag{8} \]

where \( \alpha = (-\beta)^i \). The remainder after \((k+1)\) terms of the series (8) is

\[ (-\beta)^{k+1} (x_{t-k-1} - \beta x_{t-k-2} + \ldots) = -\beta^{k+1} e_{t-k-1} \]

which has variance \( \beta^{2k+2} \sigma^2 \). This tends to zero as \( k \) tends to infinity, since \( |\beta| < 1 \). So the finite representation remains as

\[ x + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \ldots + \alpha_k x_{t-k} = e_t. \tag{9} \]

This can be made as accurate as we please by taking \( k \) sufficiently large.

If \( b \) is the estimate of \( \beta \) then the asymptotic variance matrix of \( b_1, b_2, b_3, \ldots, b_h \), to the first order in \( n \) is \( U/n \), where

\[
U^{-1} = \begin{bmatrix}
\Sigma \alpha_i^2_{i=0} & \Sigma \alpha_i \alpha_{i+1} & \ldots & \sum_{i=0}^{k-1} \alpha_i \alpha_{i+k} \\
\Sigma \alpha_i \alpha_{i+1} & \Sigma \alpha_i \alpha_{i+2} & \ldots & \sum_{i=0}^{k-2} \alpha_i \alpha_{i+k-1} \\
\vdots & \vdots & \ddots & \vdots \\
\Sigma \alpha_i \alpha_{i+k-1} & \sum_{i=0}^{k-3} \alpha_i \alpha_{i+k-2} & \ldots & \Sigma \alpha_i^2_{i=0}
\end{bmatrix}
\tag{10}
\]

The values of \( U \) for \( k=1,2,3 \) are,
h=1 : 
\[
[1 - \beta_2^2]
\]

h=2 :
\[
\begin{bmatrix}
1 - \beta_2^2 & \beta_1^2 - \beta_1 \beta_2 \\
\beta_1^2 - \beta_1 \beta_2 & 1 - \beta_2^2
\end{bmatrix}
\]

h=3 :
\[
\begin{bmatrix}
1 - \beta_3^2 & \beta_1^2 - \beta_1 \beta_3 & \beta_2^2 - \beta_2 \beta_3 \\
\beta_1^2 - \beta_1 \beta_3 & 1 - \beta_3^2 & \beta_2^2 - \beta_2 \beta_3 \\
\beta_2^2 - \beta_2 \beta_3 & \beta_1^2 - \beta_1 \beta_3 & 1 - \beta_3^2
\end{bmatrix}
\]

To test the hypothesis

\[H_0 : \beta = \beta_0\text{ in equation (7)}\]

we have
\[
z = \frac{\sqrt{n} (b - \beta_0)}{(1 - \beta_0^2)^{1/2}} \sim SNV
\]

as standard normal deviate.

Similarly to the hypothesis

\[H_0 : \beta_j = \beta_{0j} \quad j=1,2,\ldots,h\]

where \(h\) denotes the order of moving average model. We test

\[W = (b - \beta)'Uq(b - \beta)\]

which gives
\[
v = n \sum_i \sum_j u^{ij} (b_i - \beta_{a_i}) (b_j - \beta_{a_j})
\]

\[
\chi^2_h \approx \chi^2_h.
\]

In other words, \( w \) follows a \( \chi^2 \) distribution with \( h \) degrees of freedom, where \( u^{ij} = U^{ij} \) and \( U \) is calculated from \( \beta_{a_i}, \beta_{a_2}, \ldots, \beta_{a_h} \).

6.5.2 Proposed Test: The test for Threshold Time Series Models

A test is being proposed here in which a null hypothesis of no threshold in a time series model, say, MA(1) will be tested against the alternative that there is a threshold in the time series model, i.e.,

null hypothesis, \( Z_t = \theta_1 e_{t-1} + e_t \), and

alternate hypothesis, it follows a threshold moving average model

\[
Z_t = \mathcal{O}^{(1)}_1 e_{t-1} + e_t \quad \text{if} \quad Z_{t-1} \leq c
\]

\[
Z_t = \mathcal{O}^{(2)}_2 e_{t-1} + e_t \quad \text{if} \quad Z_{t-1} > c
\]

that is, it follows a threshold moving average model where \( e_t \) follows \( N(0, \sigma_e^2) \) and \( Z_t \)'s are measures from their mean.
Empirical studies suggest that a time series, \( \{Z_t\} \) of length \( n = 1, 2, \ldots, n \) if divided into equal parts around some central value (say, median or at zero), the behaviour of ACF in two parts remains more or less the same even for large series.

If the observations follow an ordinary moving average model, then the estimate of the parameter in two regimes will not be significantly different. While if the observations follow a threshold moving average the estimates \( \theta_1^{(4)} \) and \( \theta_1^{(2)} \) may be different.

Let \( \hat{\theta}_1^{(4)} \) and \( \hat{\theta}_1^{(2)} \) be the estimates of the parameters \( \theta_1^{(4)} \) and \( \theta_1^{(2)} \). For a sufficiently large sample

under \( H_0 : \theta_1^{(4)} = \theta_1^{(2)} = \theta_1 \)

\[
\hat{\theta}_1^{(4)} \sim N \left( \theta_1^{(4)}, \frac{1 - \hat{\theta}_1^{(4)2}}{n/2} \right)
\]

\[
\hat{\theta}_1^{(2)} \sim N \left( \theta_1^{(2)}, \frac{1 - \hat{\theta}_1^{(2)2}}{n/2} \right)
\]  

(15)

under \( H_0 : (\hat{\theta}_1^{(4)} - \hat{\theta}_1^{(2)}) \sim N(0, \frac{2(1-\theta_1^2)}{n} + \frac{2(1-\theta_1^2)}{n}) \)
Therefore, under $H_0$

$$\begin{align*} Z & = \frac{\hat{\theta}^{(1)} - \hat{\theta}^{(2)}}{\sqrt{\frac{4(1-\hat{\theta}^2)}{n}}}, \\ \text{SNV} & \geq 1.96 \end{align*}$$

If the value of $Z$ in equation (17) is more than 1.96 then the null hypothesis is said to be rejected at 5 percent level of significance.

This idea can also be extended for AR(1) models on the similar lines as MA(1). The test statistic will be same as equation (17). The only difference will be to replace the parameters of the MA model by the parameters of the AR model.

In other words, it can be said that the test is applicable for testing threshold in time series models.

6.6 SIMULATION RESULTS

Example 1: A total of 400 observations were generated from AR(1) model

$$Z_t = 0.8Z_{t-1} + a_t \quad t = 1, 2, \ldots, 400$$

Last 100 observations, $t = 301$ to 400 were taken into
consideration. These one hundred observations were divided into two parts according as $Z_{t-4} \leq y(350)$ and $Z_{t-4} > y(350)$ where $y(350)$ is the 350th observation, i.e. median of all the observations. The estimate of parameters $\phi_4$ and $\phi_j^{(j)}, j=1,2$ are as follows:

For full model, $z_t = \phi_4 z_{t-4} + a_t$, $t=301$ to 400

$\hat{\phi}_4 = 0.7509$, $SE(\hat{\phi}_4) = 0.0668$

For

$z_t = \phi_4^{(4)} z_{t-4} + a_t$, if $z_{t-4} \leq y(350)$

$\hat{\phi}_4^{(4)} = 0.7404$, $SE(\hat{\phi}_4^{(4)}) = 0.0911$

and

$z_t = \phi_4^{(2)} z_{t-4} + a_t$, if $z_{t-4} > y(350)$

$\hat{\phi}_4^{(2)} = 0.5827$, $SE(\hat{\phi}_4^{(2)}) = 0.1274$

Testing the difference between the estimator $\hat{\phi}_4^{(4)}$ and $\hat{\phi}_4^{(2)}$. Then

$$Z = \frac{\hat{\phi}_4^{(4)} - \hat{\phi}_4^{(2)}}{(4(1-\phi_4^{(2)})/n)^{1/2}}$$

$$= \frac{0.7404-0.5827}{(4(1-.7509^2)/100)^{1/2}}$$

$$= 1.5/1.32 = 1.13.$$
threshold autoregressive, SETAR(2;1,1) model

\[ Z_t = 0.87 z_{t-4} + a_t \quad \text{if } z_{t-4} \leq 0 \]

\[ Z_t = -0.7 z_{t-4} + a_t \quad \text{if } z_{t-4} > 0 \]

Last 100 observations were taken into account. The 100 observations were again divided into two parts according as \( z_{t-4} \leq 0 \) and \( z_{t-4} > 0 \).

In estimate of parameter \( \phi_4 \) in various cases are as follows:

If we consider all the observations coming from a single model

\[ Z_t = \phi_4 z_{t-4} + a_t \]

then \( \hat{\phi}_4 = 0.6295, \quad \text{SD}(\hat{\phi}_4) = 0.0782 \)

\[ z_t = \phi_4^{(4)} z_{t-4} + a_t \quad \text{if } z_{t-4} \leq 0 \]

\( \hat{\phi}_4^{(4)} = 0.7211, \quad \text{SD}(\hat{\phi}_4^{(4)}) = 0.0775 \)

\[ z_t = \phi_4^{(2)} z_{t-4} + a_t \quad \text{if } z_{t-4} > 0 \]

\( \hat{\phi}_4^{(2)} = -0.2798, \quad \text{SD}(\hat{\phi}_4^{(2)}) = 0.2264 \)

Applying the same test for testing the difference between the estimator \( \hat{\phi}_4^{(4)} \) and \( \hat{\phi}_4^{(2)} \).
\[
Z = \frac{\hat{\phi}_4^{(4)} - \hat{\phi}_4^{(2)}}{(4(1 - \hat{\phi}_4^2)/n)^{1/2}}
\]

\[
= \frac{(0.7211 + 0.2798)}{(4(1 - 0.6295^2)/100)^{1/2}}
\]

\[
= \frac{10.009}{1.50}
\]

\[
= 6.41
\]

Since this value of \( Z > 1.96 \) this shows that the difference between the \( \phi \) values of the two regimes is significant.

6.7 CONCLUSION

Various approaches for testing the hypothesis that no switch occurred in the parameters of a linear regression system against the alternative that one switch took place are presented. Each of the tests depends upon dividing the observations into two groups. A student-t test suffers from the relatively minor disadvantage that the expected value of the test statistic can equal zero under the alternative hypothesis as well, albeit with probability zero. Finally, an F-test is suggested for testing the null hypothesis.

To sum up it can be said that test proposed for testing
a switch in autoregressive model appears to be simple in use and easy to understand.

The attempt of testing a threshold in time series models is made here. The simulation exercise in section 6.6 validates the utility of the proposed test for testing a threshold in the model.
CHAPTER VII

APPROXIMATING REAL DATA SETS

Simulation studies in chapter III, IV and V have shown that the regression errors, which are serially correlated, can be easily modelled by Pade Approximation method. It has also been shown that the estimate of parameters when the error model is duly considered in the regression equation improves i.e. their standard error reduces, as the time series error structure are taken into consideration while estimating a regression model and the parameters become more efficient. In the present chapter, the usefulness of Pade Approximation method will be validated with the help of three real data sets.

The procedure for estimating the parameters, however, will be the same as discussed in chapter IV. In the present chapter it will be shown that how the error structure of the OLS residuals, obtained by the C-table of Pade Approximation method when incorporated in the regression model, helps in deciding adequacy of the model in real data sets. This will be confirmed if the C-table of the residuals of modified regression model shows ARMA(0,0) structure. It is, however, noted that in case, the residuals have nonstationary property (if any) then the differencing may be required before the C-table is produced.

The three examples considered in this chapter are
(i) Daily food consumption of rats, 
(ii) Body weight of monkey and 
(iii) Herbst (1965) residuals.

7.1 DAILY FOOD CONSUMPTION OF RATS

Toxicity studies in the drug development provides new lead in the usefulness of compounds. This is usually done by treating animals with various doses of the compound to see the long period and short period effects of compound at various concentrations. The data in this example is the outcome of one such study of 'Kutkin' conducted in the Toxicology Division of Central Drug Research Institute, Lucknow, India. The full experiment was originally conducted using an untreated control and three dose levels viz; low dose, intermediate dose and high dose. The daily food intake of rats was recorded for 88 consecutive days after which the experiment was terminated. The food consumption of female rats treated with high dose of Kutkin has been considered for the purpose of the present study. The response profile in Fig. 7.1 suggests that a nonlinear regression model should be considered.

The model considered for fitting the food intake \( y_t \) data is

\[
y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + u_t \quad (7.1)
\]
PROFILE OF FOOD INTAKE BY RATS TREATED WITH HIGH DOSE

FOOD INTAKE (GMS.)

TIME (DAYS)

-OBSERVED

(FIG. 7.1)
where \( t \) is the day \((t=1,2,3,\ldots,88)\) and \( u_t \) is the residuals of its OLS estimates. This second degree polynomial is fitted by ordinary least squares and OLS residuals are obtained. These residuals are shown in Fig 7.2. The estimates of the parameters are shown in Table 7.3. The quantity of food consumed by rats after drug treatment varies with the day to day condition of rats. This variation may be due to the quantity of food consumed on the previous days. So there is a need to identify the effect of previous day's food consumption. This will be performed on the OLS residuals of equation 7.1.

The OLS residuals are plotted in Fig 7.2, which shows some systematic patterns. A high order regression model, AR(10) was fitted to these residuals and a C-table has been constructed. In order to check for stationarity the ACF was plotted. Since the ACF of this series appeared to die off after few lags, it is said that the series of residuals is stationary and differencing is not required. The corresponding C-table for the OLS residuals is shown in Table 7.1.

Inspecting this C-table it can be observed that the values corresponding to column \( p=1 \) in Table 7.1 are non-zero, while the other succeeding columns of higher \( p \) level are almost zero. There is also a possibility to consider the non zero cell entries against row, \( q=1 \) in addition to \( p=1 \). This gives rise to two possibilities of error structure AR(1) and ARMA(1,1).

The model 7.1 has now been modified to incorporate the
RESIDUALS (OLS) BEFORE CORRECTION OF THE MODEL

(FIG. 7.2)
Table 7.1: C-table for the OLS residuals of daily food consumption of rats.

<table>
<thead>
<tr>
<th>q</th>
<th>p=0</th>
<th>p=1</th>
<th>p=2</th>
<th>p=3</th>
<th>p=4</th>
<th>p=5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>1</td>
<td>1.000</td>
<td>-0.549</td>
<td>0.138</td>
<td>-0.096</td>
<td>-0.056</td>
<td>-0.176</td>
</tr>
<tr>
<td>2</td>
<td>-1.000</td>
<td>-0.163</td>
<td>0.034</td>
<td>-0.017</td>
<td>0.014</td>
<td>-0.035</td>
</tr>
<tr>
<td>3</td>
<td>-1.000</td>
<td>0.110</td>
<td>0.012</td>
<td>-0.002</td>
<td>-0.007</td>
<td>0.007</td>
</tr>
<tr>
<td>4</td>
<td>1.000</td>
<td>0.147</td>
<td>-0.010</td>
<td>0.005</td>
<td>-0.005</td>
<td>0.002</td>
</tr>
<tr>
<td>5</td>
<td>1.000</td>
<td>-0.288</td>
<td>0.056</td>
<td>-0.011</td>
<td>0.001</td>
<td>0.001</td>
</tr>
</tbody>
</table>

two kind of errors, AR(1) and ARMA(1,1) individually. This new model has been estimated by CLS. The table 7.2 has been prepared to compare the two error models.

Table 7.2: Results of modified model to identify an appropriate error model for the OLS residuals of daily food consumption of rats.

<table>
<thead>
<tr>
<th>Error Model</th>
<th>Variance</th>
<th>AIC</th>
<th>$G^*$</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARMA(1,0)</td>
<td>1.2925</td>
<td>30.58</td>
<td>12.207</td>
<td>0.275</td>
</tr>
<tr>
<td>ARMA(1,1)</td>
<td>1.2985</td>
<td>32.99</td>
<td>14.866</td>
<td>0.144</td>
</tr>
</tbody>
</table>

Table 7.2 suggests that the variance of modified model with ARMA(1,0) or AR(1) model is smaller than the variance of the modified model with ARMA(1,1) model of residuals. The probability of Ljung Box $G^*$ statistic is greater than $\alpha=0.05$ in
Table 7.3: Parameter estimates before and after correction of AR(1) errors in the daily food consumption of rat.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Before Correction</th>
<th>After Correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_0 \pm SE(\beta_0)$</td>
<td>15.686±0.3909</td>
<td>15.670±0.3917</td>
</tr>
<tr>
<td>$\beta_1 \pm SE(\beta_1)$</td>
<td>0.00011±0.0202</td>
<td>0.00047±0.0200</td>
</tr>
<tr>
<td>$\beta_2 \pm SE(\beta_2)$</td>
<td>0.00125±0.0002</td>
<td>0.00125±0.0002</td>
</tr>
<tr>
<td>AIC</td>
<td>37.32</td>
<td>30.58</td>
</tr>
<tr>
<td>$\phi \pm SE(\phi)$</td>
<td>NA</td>
<td>0.3252±0.1035</td>
</tr>
<tr>
<td>$Q^*$</td>
<td>36.996</td>
<td>12.207</td>
</tr>
<tr>
<td>p-value$^\dagger$</td>
<td>&lt;0.001</td>
<td>0.275</td>
</tr>
</tbody>
</table>

$Q^*$: is computed on lag 10.
$^\dagger$: Probability of $Q^*$ statistic.
NA: Not applicable

both the models AR(1) and ARMA(1,1), which means that the final errors of both models are following white noise. But only one has to be selected.

The probability of $Q^*$ statistic of the model with AR(1) errors is much higher from 0.05 than its probability in the model with ARMA(1,1) errors. Also, the AIC value of the former is lower than the latter. Since both the criterions are in favour of AR(1) error structure, hence, it has been selected definitely by Padé Approximation method and include it in the equation to modify. Thus the error structure that has been identified on the OLS residuals of food consumption of rats is AR(1).

The effects of the inclusion of AR(1) model on the
parameters of equation (7.1) will now be studied here. The equation of the corrected model is

\[ y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \phi u_{t-4} + \epsilon_t \]

for \( t=1,2, \ldots, 88 \), conditioned upon \( u_1 = 0 \). The parameters of this modified model have been estimated by CLS. The results of the equation 7.1 and the corrected model have been given in Table 7.3.

The parameter estimates of the model obtained from the two methods are almost similar to each other while the standard error of \( \beta \)'s in the modified model have reduced. Fig 7.3 shows that the residuals of the corrected model are evenly distributed on both sides of zero. This indicates no pattern in the residuals. A sharp reduction in the \( G^* \) statistic of the corrected model to 12.207 make its probability 0.275 which is greater than \( \alpha = 0.05 \). This suggests that the final residuals are random and follows white noise. A smaller AIC value of the corrected model is an indicator of overall improvement in the model form. Therefore, it suggests to select the model (equation 7.1) with the modification that \( u_t = \phi u_{t-4} + \epsilon_t \). Fig 7.4 shows fitted curves of before and after correction of model. The closeness of the corrected model to the observed profile can be revealed easily.

7.2 BODY WEIGHT OF RHESUS MONKEY

In this example the monthly body weight of male rhesus monkey was recorded since birth (i.e. day zero) up to the age of
RESIDUALS (CLS) AFTER CORRECTION OF THE MODEL

(FIG. 7.3)
SHOWING FITTED CURVE BEFORE AND AFTER MODIFICATION OF MODEL

FOOD INTAKE (GMS.)

- OBSERVED
- EXPECTED (OLS)
- EXPECTED (CLS)

TIME (DAYS)

(FIG. 7.4)
seven years. Thus, giving a total of N=85 observations. This data has been given in Bournie (1975). The mean body weight given in grams is the dependent variable ($y_t$). The months in which observations were recorded; i.e. $t=0, 1, 2, ..., 84$ have been taken as X variable. The monthly body weight profile of male rhesus monkey has been shown in Fig 7.5. The body weight was found to be related with months by the following polynomial equation:

$$y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \epsilon_t$$  \hspace{1cm} (7.2)

Equation 7.2 was fitted by OLS procedure and its residuals were recorded. These residuals when plotted (see Fig 7.6) show a very clear systematic pattern initially and an unusually different pattern later on. The C-table for these residuals of body weight data of rhesus monkey has been given in Table 7.4.

**Table 7.4:** C-table for the residuals of body weight data.

<table>
<thead>
<tr>
<th>q</th>
<th>p=0</th>
<th>p=1</th>
<th>p=2</th>
<th>p=3</th>
<th>p=4</th>
<th>p=5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>1</td>
<td>1.000</td>
<td>-0.827</td>
<td>-0.035</td>
<td>-0.106</td>
<td>-0.006</td>
<td>0.032</td>
</tr>
<tr>
<td>2</td>
<td>-1.000</td>
<td>-0.719</td>
<td>0.086</td>
<td>-0.011</td>
<td>-0.003</td>
<td>-0.002</td>
</tr>
<tr>
<td>3</td>
<td>-1.000</td>
<td>0.730</td>
<td>-0.013</td>
<td>0.004</td>
<td>-0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>4</td>
<td>1.000</td>
<td>0.722</td>
<td>0.031</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>5</td>
<td>1.000</td>
<td>-0.672</td>
<td>0.066</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

No clearcut zero entries are there in the Table 7.4 for p=1 and q=1. This gives rise to two possibilities of residuals.
OBSERVED PROFILE OF BODY WEIGHT OF MALE RHESUS MONKEY

WT. IN GMS. (Thousands)

MONTHS

(Fig. 7.5)
RESIDUALS (OLS) BEFORE CORRECTION OF THE MODEL

(FIG. 7.6)
structures, AR(1) or ARMA(1,1). The model (equation 7.2) was modified for both kind of errors and estimated by CLS. A summary of the model with the two kind of error structures is given in table 7.5.

Table 7.5: Results for checking the model of residuals in the monkey body weight data.

<table>
<thead>
<tr>
<th>Error Model</th>
<th>Variance</th>
<th>AIC</th>
<th>$Q^*$</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARMA(1,0)</td>
<td>0.5315</td>
<td>-416.19</td>
<td>15.362</td>
<td>0.125</td>
</tr>
<tr>
<td>ARMA(1,1)</td>
<td>0.4763</td>
<td>-422.45</td>
<td>13.836</td>
<td>0.185</td>
</tr>
</tbody>
</table>

The variance of the model, AIC and the $Q^*$ statistic all are lower for ARMA(1,1) error structure. The probability of the $Q^*$ statistic for ARMA(1,1) model is more farther from $\alpha=0.05$ than the AR(1). There is a mention in Seber and Wild (1989) and Jugde et.al. (1985) that the growth models follow AR(1) errors. But in the present example the situation is more in favour of ARMA(1,1) error structure. Theoretically speaking, the ARMA(1,1) model is better than the AR(1) model, however, the values of the variance, AIC etc. of AR(1) structure are close to ARMA(1,1). The error model, ARMA(1,1) has been considered here.

The estimates of the parameters of the model 7.2 when modified for the inclusion of ARMA(1,1) errors and re-estimated by CLS. are presented in Table 7.6. The modified equation now reduces
\[ y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \phi y_{t-1} + \theta y_{t-1} + \nu_t \]

was fitted with conditioned upon \( u = 0 \) and \( \nu = 0 \).

Table 7.6: Parameter estimates before and after correction of ARMA(1,1) errors in monkey body weight data.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Before Correction</th>
<th>After Correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_0 \pm \text{SE}(\beta_0) )</td>
<td>0.71978±0.0805</td>
<td>0.73834±0.0390</td>
</tr>
<tr>
<td>( \beta_1 \pm \text{SE}(\beta_1) )</td>
<td>0.07508±0.0081</td>
<td>0.07395±0.0038</td>
</tr>
<tr>
<td>( \beta_2 \pm \text{SE}(\beta_2) )</td>
<td>0.00206±0.0002</td>
<td>0.00208±0.0001</td>
</tr>
<tr>
<td>( \beta_3 \pm \text{SE}(\beta_3) )</td>
<td>-0.00002±0.000002</td>
<td>-0.00002±0.000001</td>
</tr>
<tr>
<td>AIC</td>
<td>-285.92</td>
<td>-422.45</td>
</tr>
<tr>
<td>( \phi \pm \text{SE}(\phi) )</td>
<td>NA</td>
<td>0.9590±0.0523</td>
</tr>
<tr>
<td>( \theta \pm \text{SE}(\theta) )</td>
<td>NA</td>
<td>-0.0310±0.0096</td>
</tr>
<tr>
<td>( G^* )</td>
<td>102.924</td>
<td>13.836</td>
</tr>
<tr>
<td>p-value</td>
<td>&lt;0.0001</td>
<td>0.185</td>
</tr>
</tbody>
</table>

\( G^* \): is computed on lag 10.

\( \epsilon \): Probability of \( G^* \) statistic.

NA: Not applicable

Change in the estimates of the parameters before and after the correction of model can be seen. Their standard error has reduced sharply after correction of model. A drastic reduction in the AIC value and \( G^* \) statistic, to the extent that they became not significant at 10 degree of freedom, shows that the resultant errors are IID. These residuals are plotted in Fig 7.7. is in support of satisfactory correction of the model (Eq. 7.2).
RESIDUALS (CLS) AFTER CORRECTION
OF THE MODEL

(FIG. 7.7)
SHOWING FITTED CURVE BEFORE AND AFTER CORRECTION OF THE MODEL

WT. IN GMS. (Thousands)

- OBSERVED
- EXPECTED (OLS)
- EXPECTED (CLS)

MONTHS

(FIG. 7.8)
All the parameters of the model including the error parameters are contributing significantly to explain the correction of errors pattern in the model. Although the contribution of MA coefficient is low, in comparison to the AR coefficient, yet it has been included because it is contributing significantly to explain the model. The closeness of the fitted model (see Fig 7.8) to the observed profile is more with the corrected model than the uncorrected model.

Finally we conclude that the error structure identified by the Pade Approximation Method, is satisfactory and acceptable too since the correction of errors makes the model adequate.

7.3 Herbst's Residuals

The residuals studied in this section are extracted from Herbst's (1965) article. The model that he had fitted had the following set-up: for every 15 minutes interval, observation \( Y_t \) were made on the amount of carbon-di-oxide fixed in the study of photosynthesis in corn. Light intensity \( X_t \) is assumed to be related to \( Y_t \) and the following model was fitted using OLS based on 42 observations

\[
Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_t^2 + u_t
\]

Since the data about \( Y_t \) and \( X_t \) of these residuals are not available, so, the conditional least square estimates of the
corrected model could not be obtained. Hence, the improvement in the model estimation could not be determined. The best that could be done here is to identify the serial dependency of residuals using C-table which was not considered by Herbst's originally.

Forty two observations of carbon-di-oxide fixed in the study of photosynthesis in corn are taken at a regular interval of 15 minutes hence it is time index and the time series theory is applicable on its residuals. The OLS residuals of the above model are given in Herbst's article. It will be used directly to compute the C-table of the residuals.

The residuals appear to be stationary from the time plot. Its ACF dies out after few lags indicating it to be stationary. An AR of of order 16 was fit to prepare C-table (Table 7.7).

Table 7.7: C-table for the Herbst's residuals.

<table>
<thead>
<tr>
<th>q</th>
<th>p=0</th>
<th>p=1</th>
<th>p=2</th>
<th>p=3</th>
<th>p=4</th>
<th>p=5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>1</td>
<td>1.000</td>
<td>-0.375</td>
<td>0.747</td>
<td>-0.418</td>
<td>0.595</td>
<td>-0.314</td>
</tr>
<tr>
<td>2</td>
<td>-1.000</td>
<td>0.607</td>
<td>-0.402</td>
<td>0.270</td>
<td>-0.223</td>
<td>-0.030</td>
</tr>
<tr>
<td>3</td>
<td>-1.000</td>
<td>-0.089</td>
<td>0.003</td>
<td>-0.040</td>
<td>-0.097</td>
<td>-0.007</td>
</tr>
<tr>
<td>4</td>
<td>1.000</td>
<td>-0.018</td>
<td>-0.009</td>
<td>0.007</td>
<td>0.041</td>
<td>-0.007</td>
</tr>
<tr>
<td>5</td>
<td>1.000</td>
<td>0.103</td>
<td>0.016</td>
<td>0.010</td>
<td>0.017</td>
<td>0.006</td>
</tr>
</tbody>
</table>

In this C-table the zero entries are not very clear to indicate single model. The possible error models here could be ARMA(1,2) or ARMA(2,2). A better error model can be determined by AIC and $Q^*$ statistic. The results for this are given in Table 7.8.
Table 7.8: Results for checking a better error model for Herbst's residuals.

<table>
<thead>
<tr>
<th>Error Model</th>
<th>S.E.</th>
<th>AIC</th>
<th>$Q^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARMA(1,2)</td>
<td>0.8942</td>
<td>113.594</td>
<td>0.646</td>
</tr>
<tr>
<td>ARMA(2,2)</td>
<td>0.8724</td>
<td>112.402</td>
<td>0.596</td>
</tr>
</tbody>
</table>

The standard error of ARMA(2,2) is smaller than the standard error of ARMA(1,2). The AIC value is also lesser for ARMA(2,2). The $Q^*$ statistic for this is slightly higher for ARMA(1,2). The $Q^*$ statistic measures only the randomness of error and both have probability greater than $\alpha=0.05$, therefore, both measures the independent errors. Under this condition it may be sufficient to accept ARMA(2,2) as the final error structure.

Having specified the ARMA(2,2) errors the estimates of the parameters were analysed and are shown in Table 7.9.

Table 7.9: Parameter estimates of ARMA(2,2) with respect to Herbst's residuals.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\phi_1$</td>
<td>0.1660</td>
<td>0.2567</td>
<td>0.65</td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>0.5039</td>
<td>0.2583</td>
<td>1.95</td>
</tr>
<tr>
<td>MA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>0.1219</td>
<td>0.1952</td>
<td>0.62</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>0.8781</td>
<td>0.1999</td>
<td>4.39</td>
</tr>
</tbody>
</table>
The ARMA error in this case will be \( u_t = \phi_2 u_{t-2} - \theta_2 e_{t-2} + e_t \).

The findings above are indicative of incorrect model specification by Herbst. The results of Table 7.8 and 7.9 are sufficient to say that the Pade Approximation method can be used in the identification of ARMA errors.

7.4 CONCLUSION

It can be concluded from the performance of the above three examples that Pade Approximation method can be used as a successful tool in the identification of ARMA errors. The results obtained from applying this method to real data sets substantiate our claim of using this method.