MODEL FITTING FOR YEARLY FERTILITY DATA IN URBAN MYANMAR

KYAING KYAING THET

JULY,1997

MODEL FITTING
FOR YEARLY FERTILITY DATA
IN URBAN MYANMAR

MODEL FITTING FOR YEARLY FERTILITY DATA IN URBAN MYANMAR

by

KYAING KYAING THET

A thesis submitted as a partial fulfilment towards
the Degree of Master of Economics

in the

Institute of Economics
Yangon
Myanmar

July, 1997

INSTITUTE OF ECONOMICS
YANGON
DEPARTMENT OF STATISTICS

MODEL FITTING FOR YEARLY FERTILITY DATA IN URBAN MYANMAR

by

KYAING KYAING THET

This thesis is submitted to the Board of Examiners in Statistics in partial fulfilment of the requirement for the degree of Master of Economics.

Examiner

External Examiner

ဒုတိယညွှန်ကြားရေးမှူးချုပ်
ဗဟိုစာရင်းအင်းအဖွဲ့

Examiner

Professor Maw Than
Rector
INSTITUTE OF ECONOMICS
Yangon.

Chairman
Board of Examiners

1997 July

(KYI WIN)
REGISTRAR (I)
INSTITUTE OF ECONOMICS

ACKNOWLEDGEMENT

I am very grateful to the Professor Daw Khin San Myint, Head of the Department of Statistics, for her kindness to give me permission to continue my thesis and helped me in many ways in preparing this thesis.

I also thank to former Professor Dr. Thet Lwin, Department of Statistics, who give me permission to write this thesis. Again, I feel grately thankful to Associate Professor Dr. Daw Than Toe, from whom I received a lot of guidance.

I convey my heartfelt thanks to my supervisor, Daw Cho Cho Wai, Assistant Lecturer, Department of Statistics, from whom I have received a lot of guidance and wholehearted aid.

I would like to express my thanks to Daw Maw Maw Oo, Head of Department of English, Institute of Economics, for editing the manuscript.

Finally, I would like to express my thanks to my parents who offer me financial and all other necessary support.

CONTENTS

Chapter		Title	Page No

LIST OF TABLES

Table No	Title	Page No

LIST OF FIGURES

ABSTRACT

The main case of error in population projections during the last decades has been due to lack of success in correctly future fertility level. So, it needs to fit a suitable model to the existing fertility data for projecting the fertility level.

In this thesis, a time series model is fitted by the yearly fertility data of the selected towns in Urban Myanmar. The yearly live births data in Myanmar are obtained from the Vital Statistics Reports during the years 1968 to 1988 which were jointly published by the Central Statistical Organization and Department of Health. So, in order to make a reasonable analysis, only 125 common towns are analysed in this thesis.

In fitting a suitable model, first, the type of the model is selected by playing a major role of autocorrelation and partial autocorrelation functions. Second, the model parameters were estimated by using the moment method. Third, the order of fitted model is decided by using the MC Clave Criterion. And lastly, the future trend of fertility level is presented by using the fitted model.

CHAPTER I

INTRODUCTION

1.1 Introduction

Fertility is an essential factor of human reproduction. Fertility constitutes a positive element of fundamental population growth together with another element but negative one, mortality.

While the mortality trend has been obviously on a decrease in many developing countries since the end of world war II, their level of fertility has remained almost unchanged and kept their height almost as high as before. It has been well known that the difference between fertility and mortality trends has brought about a rapid population growth after remarked as a population explosion. The study of fertility has become a great concern of many people in various fields, all among demographers and statisticians of many countries. It has become widely known that the trend of population growth of modern times depends much on the trend of fertility.

Fertility is the measure of the reproductive performance of women as obtained from the statistics of the number of live births. The number of births occurring in any year in a population is determined partly by demographic factors such as the age and sex distribution, the number of married couples and their distribution by age, duration of marriage, and number of children already born. The number is also partly determined by many other factors related to the social and economic, environment of that particular time, such as housing conditions, education, income, religion and current attitudes towards family size.

When studying fertility trends in a given country, or the differences in fertility between different countries, the demographer aims to determine the extent to which differences in the number of births have been caused by difference in these demographic factors and hence to deduce what differences remain to be explained by social and economic conditions. It is very difficult to isolate and measure the effect of each of these factors as they are so closely interrelated. Consequently the study of fertility involves the use of a number of methods of fertility measurement, each with advantages and disadvantages, suitable under certain circumstances and unsuitable under others.

Many demographers have studied the level of fertility from the demographic point of view. But, in this thesis, the fertility data are studied by time series model. First, the model will be fitted by using fertility data from the Vital Statistics Reports published jointly by the Central Statistical Organization (CSO) and Department of Health. The live-births data of the towns with complete data will be used in fitting the model. Using this model, live-births of urban Myanmar will be projected and analysed by suitable time series methods. Demographic techniques will be used to find out the fertility rates. The fertility situations of future urban Myanmar will than be assessed.

1.2 Sourcė of Fertility Data in Myanmar

(The main sources of demographic data in Myanmar are the population censuses and the vital registration and statistics system. But in Myanmar, as in many developing countries, information on the censuses and vital registration and statistics system have defectives for demographic research.)

The first post-war modern nation-wide census was conducted in 1973. One purpose of the census was to obtain reliable population data, but the primary purpose was to prepare the Electoral Rolls for the Constitutional Referendum in 1973 and the Constituent Assembly in 1974 (Nyunt, 1978: 13). According to the 1973 census, it covered about 85.1 per cent of total population (IMD, 1976: 1-3). It was the first census in Myanmar which gave population by single year of age. But, it has no information about the number of births and deaths for the whole country.

The latest census was conducted in 1983. The coverage of this census, the questionnaires were used in two types, a short form and a long form. The short form was given 80 per cent of total population. The long form was addressed the remaining 20 per cent which are selected randomly. The short form includes only seven basic questions: name, relationship to the head of household, sex, age, marital status, race and religion. The long form questionnaire included an additional eleven questions on socio-economic and fertility characteristics such as school attendance, occupation, highest standard passed, literacy, industry, employment status, reason for not working, whether working during last twelve months, children ever born alive, children still living and the date of birth of the last child. Although the 1983 census had included the fertility data, it covered only 20 per cent of the total population. Moreover, the subsequent census was not conducted in 1993 and it is difficult to study population growth and changes in population composition.

Another source of demographic data in Myanmar is the Vital Statistics Reports. Nowadays, the vital statistics are collected under the authority of the Department of Health, and compiled and published jointly by the CSO. In 1968, 152 towns have rendered both births and deaths returns which is about 81.8 per cent of the urban total population. It was gradually extended to the other towns by the end of 1988. According to the 1988 vital statistics report, the vital registration system covered 97.4 per cent of the total urban population but covered only 63 per cent of total rural population.

Therefore, the yearly fertility data for the urban area can be obtained from the Vital Statistics Reports. In this thesis, the live-births data of the towns with most complete data are used from the Vital Statistics Reports in fitting the time series model.

1.3 Requirement of Fitting a Suitable Model for the Fertility Data

The main cause of errors in population projections during the last decades has been due to demographers' lack of success in correctly predicting future births. (Joop, JASA, 1985) Demographers often appear simply to avoid the problem by looking at period fertility rates. Probably the most frequently used to measure the level of fertility is the total fertility rate (TFR), which is equal to the sum of the age-specific fertility rates observed in the same calendar year. Generally, the value of the TFR is interpreted as the average number of children born to the women in the reproductive ages (i.e. from 15 to 49 years of age). A well-known though still often ignored-problem is that the movement of the TFR is affected by changes in the age pattern of the fertility. As a consequence, solely looking at recent changes of the TFR may lead to an incorrect interpretation of recent changes in fertility behaviour.

Obviously, it is necessary to make use of estimates or projections in order to assess recent changes in fertility behaviours. In this thesis, a time series model is presented to project the fertility level of recent situation on the basis of observations already available in the Vital Statistics Reports from 1968 to 1988 for urban areas.

1.4 Compilation of Yearly Fertility Data from Vital Statistics Reports

The main sources of yearly fertility data in Myanmar is the vital statistics reports. In Myanmar, vital statistics were collected by Municipal Health Offices in urban areas and by the village headmen in the rural area during the colonial period. Births, Deaths and Marriages Act was first enacted in 1886 and the registration of vital events (births and deaths) was introduced in Yangon and some parts of Lower Myanmar. Then, it was extended to the towns of Upper Myanmar in 1906 and the villages of Upper Myanmar in 1907. The system covered nearly 80 per cent of the total population. In 1931, about 82.5 per cent of the population was covered by the Vital Registration System. (R. M. Sundrum, 1957)

In the post war period, a new vital registration system starting with Yangon City and 15 other towns was introduced in February, 1962. 1968, during the year, 152 towns have rendered both birth and death returns which is about 81.8 per cent of the urban population. The number of reporting towns are mostly different in each year. They gradually increase in later years. During the years 1982 to 1988 the vital registration system covered 245 towns. (CSO, 1968) Among them, some of the towns rendered births only.

When the new system was initially implemented it had been planned to expand stage by stage starting from the big municipal towns to towns and gradually to reach the rural areas and cover the whole country in 10 year time. The political changes and the limited resources had with held the progress especially between 1973 and 1979. The system now covers 20 per cent of the total population. In fact, this coverage is 90 per cent of the urban population. There are some test area of the rural area and it is hard to determine the exact coverage of the rural area.

According to the 1962 to 1988 Vital Statistics Reports, the coverage of total urban population and the number of towns were varied from year to year. The coverage area for urban area is very low per cent in 1962 to 1967. However, in 1968 to 1988, the coverage of the urban population is more than 80 per cent to 90 per cent. Thus, in this thesis, the fertility data for urban area are analysed from 1968 to 1988.

In the Vital Statistics Reports from 1968 to 1988, the number of reported towns are varied from year to year. The reported towns included in the successive reports are not the same. So in order to make a reasonable analysis, only 125 common towns are analysed in this thesis. The collected towns for 1966 to 1988 by states and divisions are shown in Appendix Table(1).

Although 8 towns were studied in Kachin State between 1968 to 1988, the number of reported towns are varied from year to year. Therefore, only 4 common towns for Kachin State are used in this thesis.

In Kayah State, Demawsoe, Parusoe and Phasaung were collected only 3 to 5 years. But, for Loikaw live births data collected from 1968 to 1988 so it can be used in this thesis.

Similarly, in Karen State, Kawkareik, Kya-in-seikkyi, Pa-an and Thandaung live births data are used in this thesis.

In Chin State , there has 9 towns live-births data were collected between 1968 to 1988. But, among these towns, Kanpetlat, Tunzan and Htantalang could not be collected in the successive years. Therefore, the six towns are used in this thesis.

Although the 13 towns were collected between 1968 to 1988 in Mon State, 9 towns could not be collected in the successive year. So, the remaining 4 towns are compiled in this thesis.

In Rakhine State, kyaukpyu, Sittwe, Myohaung, Kyauktaw, Maungdaw and Thandwe were collected since 1968.. It is gradually extended to the other towns and by the end of 1988, 20 towns can be collected. But, only above six towns can be compiled during the successive years and they are used in this thesis.

Data of 24 towns can be collected in the Shan State. But, only data of 8 towns are used in this thesis because these can be collected in the successive 20 years.

There are 39 townships and 13 towns in Yangon Division. The Yangon City which compiles 27 townships is taken as one town only. Htauk Kyant is a town in the Mingaladon township and is accordingly counted in the number of towns. However, it is merged into Yangon City. Although the number of reporting towns for births and deaths are 11 towns, Yangon

City and the other 8 towns can be compiled in this thesis. Htantabin and Hmawbi can not be collected in the successive year.

In Mandalay Division, 22 towns are collected for births and deaths data. But, only 14 towns are compiled in this thesis because the other towns could not be collected in the successive years.

The number of reported towns for births and deaths data is 26 towns in Magway Division. But the only 11 towns can be collected for this period. And Sagaing Division, it can be compiled 11 towns for 20 successive years. In Bago Division, although the number of reporting towns for births and deaths data was 24 towns, between 1968 and 1988, it can be compiled 21 towns for successive 20 years.

In Ayeyarwaddy Division, the number of reported towns for births and deaths data is 25 towns. But, the only 22 towns can be compiled for successive 20 years.

Similarly, although the number of reported towns for births and deaths data is 6 towns for Tanintharyi Division, 5 towns can be used for successive 20 years.

Therefore, the total number of towns for urban area is 125 towns were used in this thesis.

CHAPTER II

INTER - DEPENDENCE STRUCTURE OF FERTILITY DATA

2.1 Introduction

A time series is a set of observations generated sequentially in time. If the set is continuous, the time series is said to be continuous. If the set is discrete, the time series is said to be discrete. In this thesis it has considered only discrete time series where observations are made at fixed intervals.

A statistical phenomenon that involves in time according to probabilistic law is called a stochastic process.

From a theoretical point of view, an important step in the analysis of the time series data is to fit suitable models for the underlying stochastic process. Therefore, live-births and population data need to fit a suitable model for the underlying stochastic process. In such situations the analysis of the autocovariance function or autocorrelation function and the power spectrum may mainly provide to the choice of a suitable model.

In fitting a suitable model, it is needed to present the following four stages. The first stage is "*selection of the type of model*" and the second is "*identification of the model parameters*". The third stage is "*estimation of the model parameters*" and the last is "*diagnostic checking of* *the model*". In each of these stages the autocovariance function (acvf) and autocorrelation function (acf) play as a major role.

Most of the demographic time series have been found to be represented by a low order autoregressive process or a mixed autoregressive moving average process. In determination of the type of the model, the order of such model and in estimation and diagnostic checking of a chosen model, the autocorrelation function and the partial autocorrelation function are needed.

But, in practice, only a finite number of observation is available and the theoretical autocorrelation function and the partial autocorrelation function have to be estimated from the observations. There are a number of estimations for autocorrelation function and the computational method are discussed in this chapter. Some sampling properties of the sample autocorrelation function and the partial autocorrelation function are also discussed. The sample autocorrelation function and the partial autocorrelation functions are computed for the observed series, standardized series, first differencing series and log-transformed series.

2.2 Sample Autocorrelation and Their Properties

An appropriate model can be defined by examining the sample autocorrelation coefficients. X_t is the observation into a stationary series. Stationarity implies that the autocovariance of observations with fixed intervals are constant through time.

The autocovariance coefficient γ_k, at lag k, measures the covariance between two values X_t and X_{t+k}, a distance k apart. The plot of γ_k versus the lag k is called the autocovariance function $\{\gamma_k\}$ of the process.

$$\gamma_k = Cov[X_t, X_{t+k}] = E|[X_t - E(X_t)][X_{t+k} - E(X_{t+k})]|.$$

Similarly, the plot of the autocorrelation coefficient ρ_k as a function of lag k, is called the autocorrelation function $\{\rho_k\}$ of the process. Since $\rho_k = \rho_{-k}$ the autocorrelation is necessarily symmetric about zero. In the past, the autocorrelation function has been called the correlogram.

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad , \; k = 1, 2, \ldots$$

where

$$\gamma_o = V[X_t] = E[X_t - E(X_t)]^2.$$

The autocorrelation is the dimensionless measurement and it lies between -1 and +1. It is used to measure the degree of linear dependence among of the value of a time series that are of a certain time lag apart.

In practice, It has finite number series $X_1, X_2, X_3, \ldots, X_T$ of T observations, from which it can only obtain estimates of the autocorrelations.

The estimated autocorrelation coefficients are equal to

$$\gamma_k = \frac{\sum_{t=1}^{T} (X_t - \overline{X})(X_{t+k} - \overline{X})}{\sum_{t=1}^{T} (X_t - \overline{X})^2}$$

where $\overline{X} = \frac{1}{T} \sum_{t=1}^{T} X_t$

An indication of the number of AR and MA parameters can be obtained from pattern of the autocorrelation coefficients. MA process characterized by the fact that the autocorrelation functions cut off after the lag that corresponds to the order of the process(hence $\gamma_k=0$ for k>q), whereas the autocorrelation function of AR process decays exponentially. The choice of the model cannot always be determined from the pattern of the sample autocorrelations. Therefore, after the coefficients of the chosen model are estimated, it is important to check the adequacy of the model by examining the autocorrelation pattern of the residuals as well as by fitting more elaborate model.

The obtaining of sample estimates of the autocorrelation function is non-structural approaches, analogous to the representation of an empirical distribution function by a histogram.

Parametric time series models are not necessarily associated with a sample autocorrelation function. Working with either of these non-structural methods, the estimation of many lag-corrrelation is involved. But a parametric model containing only one or two parameters could represent the data. Each correlation is a parameter to be estimated. So that these non-structural approaches might be very prodigal with parameters, when the model could be parsimonious. Initially, we do not know what type of model may be appropriate. These non-structural approaches is necessary to identify the type of the model. The choice of the autocorrelation function depends upon the nature of the models. [Box-Jenkin, 1976]

2.3 Partial Autocorrelation Function

The partial autocorrelation function or partial correlogram is another way of representing the dependence structure of a series or of a given model. It is useful for identification of the type and order of the model when investigating a given sample time series. In characteristic behaviour of autocorrelations, partial autocorrelations for the three classes of processes, autoregressive process, moving average process and the mixed autoregressive moving average process is as shown in the following table.

The autocorrelation ρ_j measures the correlation of terms of the series separated by j terms or j lags apart. The partial autocorrelation ϕ_{kk} measures the linear dependence between ρ_j and ρ_{j-k} for $j \leq k$ or the ϕ_{kk} measures the correlation of the terms of the series k lag apart, irrespective of the other terms of the series. The plot of ϕ_{kk} against the lag value k, k=1,2... is called the 'partial correlogram' and a set of partial autocorrelations ϕ_{kk}, k=1,2,... is known as partial autocorrelation function.

Table (2.1)

CHARACTERISTIC BEHAVIOUR OF ACF, PACF

OF AR, MA AND ARMA PROCESSES

Class of process	Autocorrelations	Partial Autocorrelations
AR(p)	Infinite(damped exponentials and/or damped sine waves.) $e_j = \phi_1 e_{j-1} + \ldots + \phi_p e_{j-p}$	Finite. Spikes at lag 1 through p, then cut off.
MA(q)	Finite. Spikes at lag 1 through q, then cut off.	Infinite(dominated by damped exponentials and/or damped sine waves.) Tail off.
ARMA(p,q)	Infinite(damped exponentials and/or damped sine waves after first q-p lags.) Irregular pattern at lag 1 through q, than tail off according to $\rho_j = \phi_1 \rho_{j-1} + \ldots + \phi_p \rho_{j-p}$	Infinite(dominated by damped exponentials and/or damped sine waves after first q-p lags.) Tail off.

Remarks: ϕ_j is the j^{th} autoregressive parameter and

ρ_j is the j^{th} autocorrelation coefficient.

Since the partial autocorrelations are cut off after the first p values in AR(p) process, ϕ_{kk} can be used to determine the order of AR process.

To compute the partial autocorrelation estimates, let γ_j be the j^{th} sample autocorrelation and $\phi_{i,j}$ be the j^{th} coefficient (parameter) in the i^{th} order AR process. Then, by solving Yule-Walker equations (for k^{th} order AR process, k=1,2,...),

$$\begin{aligned}
\rho_1 &= \phi_{1k} + \phi_{2k}\,\rho_1 + \ldots + \phi_{kk}\,\rho_{k-1} \\
\rho_2 &= \phi_{1k}\,\rho_1 + \phi_{2k} + \ldots + \phi_{kk}\,\rho_{k-2} \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
\rho_k &= \phi_{1k}\,\rho_{k-1} + \phi_{2k}\,\rho_{k-2} + \ldots + \phi_{kk}
\end{aligned}$$

for k=1,2,3,..., using γ_k in place of ρ_k, $\{\hat{\phi}_{kk}\}$, k=1, 2,... can be obtained. [Box-Jenkin, 1976]

On the other hand, the recursive formulae, due to Durbin [Box-Jenkin, 1976], are

$$\hat{\phi}_{k+1,j} = \hat{\phi}_{kj} - \hat{\phi}_{k+1,k+1}\,\hat{\phi}_{k,k-j+1} \quad ; \; j=1,2,\ldots,k$$

and

$$\hat{\phi}_{k+1,k+1} = \frac{\gamma_{k+1} - \sum_{j=1}^{k} \hat{\phi}_{kj}\gamma_{k+1-j}}{1 - \sum_{j=1}^{k} \hat{\phi}_{kj}\gamma_j} \; ; k = 1,2,\ldots$$

By solving these equations the set of partial autocorrelation estimates $\{\hat{\phi}_{kk}\}$, k=1,2,..can be obtained. In these equations ϕ_{11} is set to be r_1.

On the assumption that the process is autoregressive of order p, the estimated partial autocorrelations of order p+1 and higher are not significantly different from zero and they are approximately independently distributed.

Also, if n is the number of observations used in fitting the autoregressive process of order p, then

$$V(\hat{\phi}_{kk}) \approx \frac{1}{n} \quad ; \; k = p+1, \; p+2, \; \ldots$$

Thus, the standard error of the partial autocorrelation estimate $\hat{\phi}_{kk}$ is

$$SE(\hat{\phi}_{kk}) \approx \frac{1}{\sqrt{n}} \quad ; \; k = p+1, \; p+2, \; \ldots$$

Thus, if $\hat{\phi}_{kk}$ lies between the interval $\left(-1.96/\sqrt{n}, +1.96/\sqrt{n}\right)$ for $k \geq p+1$, it can be assumed that the given series obey an autoregressive scheme with order p.

2.4 Inter- Dependence Structure of Live-Births

The live-births records are available for 21 years (1968 to 1988) for Myanmar. Hence, there are 21 yearly records and the sample autocorrelations up to lag 20 are computed by using formula

$$\gamma_k = \frac{\sum_{t}^{T} (X_t - \overline{X})(X_{t+k} - \overline{X})}{\sum_{t}^{T} (X_t - \overline{X})^2} \quad , k = 1, 2, \ldots, 20$$

These sample autocorrelations γ_k and the partial autocorrelation $\hat{\phi}_{kk}$ are presented in Table (2.1a) and Table (2.1b). The partial autocorrelations are computed by the recursive method. The correlogram, the figures of γ_k and $\hat{\phi}_{kk}$ against the lag value k for k=1,2,...,20 are shown in figure (2.1a) and (2.1b).

From table (2.1a) and the figure (2.1a), it can be seen that the lag one (k=1) correlation is highest and the sample autocorrelations decline up tc lag (10).The confidence interval is $(-1.96/\sqrt{21} = -0.4277, +1.96/\sqrt{21} = +0.4277)$ and the sample autocorrelations of lag (1) and lag(2) are out of the confidence interval. Since the sample autocorrelations function is assumed to be tail off, the live-births series obey autocorrelations scheme.

When the partial autocorrelations are examined, it can be seen that the partial correlations are cut off after lag 1. Hence, the AR (1) process can be used to represent the given series appropriately. The complement of coefficient of determination $(1-R^2_k)$, k=1,2,...,20 show the same fact that the series obey AR (1)scheme series it does not change significantly from lag(1) to lag(2). The values $(1-R^2_k)$, k=1,2,...,20 are computed by using the formula,

$$(1-R^2_k) = \prod_{i=1}^{k} (1 - \hat{\phi}^2_{kk}) \quad k = 1, 2, ..., 20$$ and it is

the complement of the coefficient of determination when the underlying process is AR(k).

Table(2.1a)
Sample Autocorrelations for Live-Births of Urban Myanmar

Lag	Corr.	Err.	-1 -.75 -.5 -.25 0 .25 .5 .75 1
1	.797	.218	********.*******
2	.611	.329	***********.
3	.393	.379	********
4	.189	.398	****
5	.019	.402	*
6	-.125	.402	***
7	-.223	.404	****
8	-.319	.410	******
9	-.347	.422	*******
10	-.375	.435	*******
11	-.347	.450	*******
12	-.241	.463	*****
13	-.142	.469	***
14	-.064	.471	*
15	-.036	.471	*
16	-.018	.471	*
17	-.051	.471	*
18	-.106	.472	**
19	-.063	.473	*

Plot Symbols: Autocorrelations * Two Standard Error Limits .
Total cases: 21 Computable first lags: 20

Table(2.1b)
Sample Partial Autocorrelations for Live-Births of Urban Myanmar

Lag	Pr-Aut- Corr.	Stand. Err.	-1 -.75 -.5 -.25 0 .25 .5 .75 1	$(1-R_k^2)$
1	.797	.218	********.*******	0.364
2	-.069	.218	*	0.358
3	-.199	.218	****	0.343
4	-.125	.218	**	0.337
5	-.064	.218	*	0.336
6	-.092	.218	**	0.333
7	-.043	.218	*	0.332
8	-.149	.218	***	0.325
9	.011	.218	*	0.325
10	-.108	.218	**	0.321
11	.012	.218	*	0.321
12	.158	.218	***	0.313
13	-.019	.218	*	0.313
14	-.094	.218	**	0.310
15	-.129	.218	***	0.305
16	-.044	.218	*	0.304
17	-.128	.218	***	0.299
18	-.147	.218	***	0.293
19	.229	.218	*****	0.278

Plot Symbols: Autocorrelations * Two Standard Error Limits .
Total cases: 21 Computable first lags: 20

Figure (2.1a)

Sample Autocorrelations for live-Births of Urban Myanmar

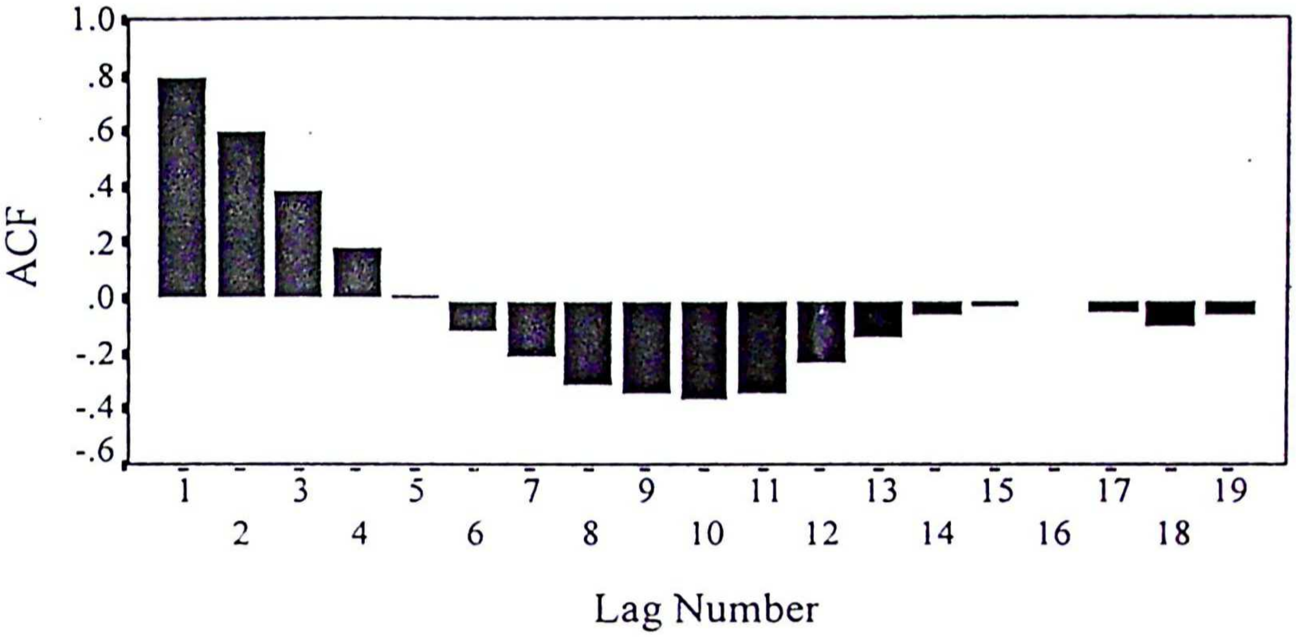
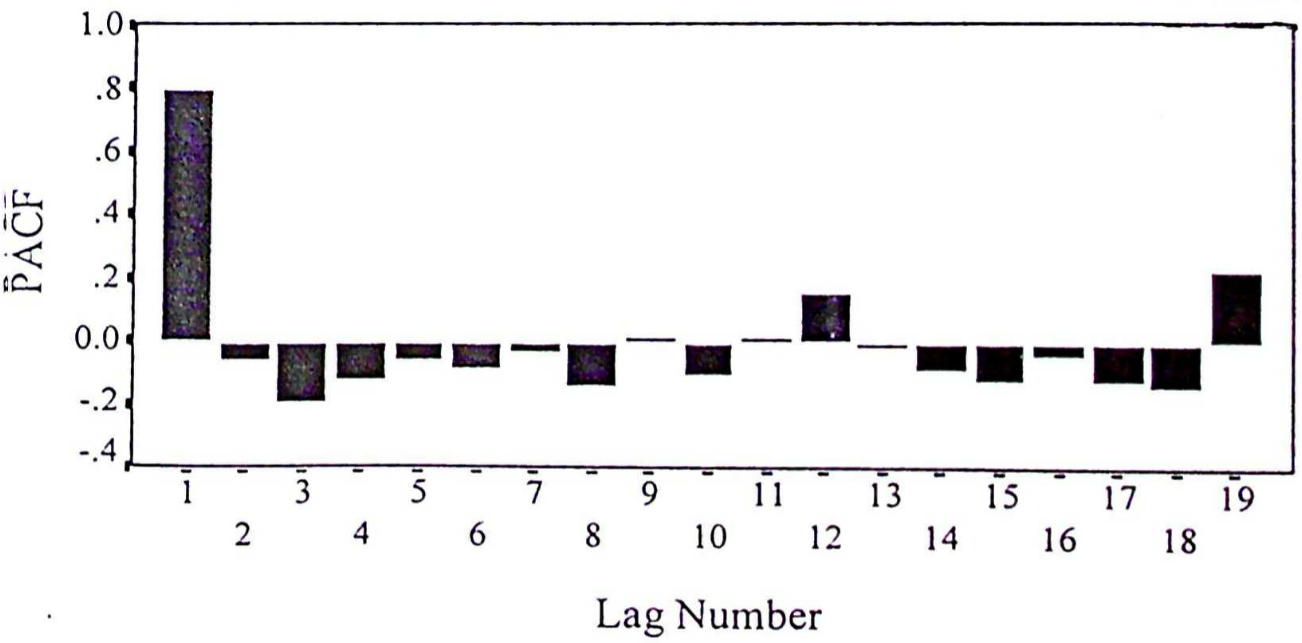


Figure (2.1b)

Sample Partial Autocorrelations for Live-Births of Urban Myanmar



2.5 Inter. Dependence Structure of Mid-year Estimated Population

The sample autocorrelations of lag(1) to (19) are presented in the Table (2.2a). It can be seen that the lag one(k = 1) correlation is the highest and the sample autocorrelations are gradually decline up to fourteen(14). The sample autocorrelations of lag(1) to (4) are out of the confidence intervals(-0.4277,+0.4277)sothat the sample autocorrelations function can be assumed to be tail off.

The sample and partial correlograms are presented in figure (2.2a) and (2.2b). When the partial correlations are examined, the lag(1) is out of the confidence intervals and the partial autocorrelations function cut off after lag(1). The complement of the coefficient of determination $(1-R^2_k)$, k=1,2,...,20 are not significantly different from each other. So it obey AR (1) scheme.

Table(2.2a)
Sample Autocorrelations for Mid-Year Estimated Population

```
Lag  Corr.   Err. -1   -.75   -.5 -.25    0    .25   .5    .75    1
                   +----+----+----+----+----+----+----+----+
 1   .857   .218           .          |********.********
 2   .722   .343        .             |*************
 3   .579   .409       .              |***********
 4   .431   .446     .                |*********        .
 5   .296   .466    .                 |******             .
 6   .195   .475    .                 |****                 .
 7   .056   .478    .                 |*                    .
 8  -.044   .479    .                *|                     .
 9  -.136   .479    .              ***|                     .
10  -.222   .481    .             ****|                     .
11  -.290   .486    .           ******|                     .
12  -.343   .494    .          *******|                     .
13  -.383   .505   .          ********|                      .
14  -.413   .519   .          ********|                      .
15  -.394   .534              ********|
16  -.405   .548              ********|
17  -.366   .562               *******|
18  -.302   .573                ******|
19  -.219   .581                  ****|

Plot Symbols:      Autocorrelations *     Two Standard Error Limits .
Total cases:  21     Computable first lags:  20
```

Table(2.2b)
Sample Partial Autocorrelations for Mid-Year Estimated Population

```
Lag  Corr.   Err. -1   -.75   -.5 -.25    0    .25   .5    .75    1  (1-R_k^2)
                   +----+----+----+----+----+----+----+----+
 1   .857   .218           .          |********.********        0.266
 2  -.046   .218           .         *|        .                0.265
 3  -.111   .218           .        **|        .                0.262
 4  -.108   .218           .        **|        .                0.258
 5  -.054   .218           .         *|        .                0.257
 6   .032   .218           .          |*       .                0.257
 7  -.235   .218           .     *****|        .                0.241
 8   .013   .218           .          *        .                0.241
 9  -.068   .218           .         *|        .                0.240
10  -.082   .218           .        **|        .                0.238
11  -.058   .218           .         *|        .                0.237
12  -.081   .218           .        **|        .                0.235
13  -.013   .218           .          *        .                0.235
14  -.109   .218           .        **|        .                0.232
15   .102   .218           .          |**      .                0.230
16  -.189   .218           .      ****|        .                0.222
17   .102   .218           .          |**      .                0.220
18   .056   .218           .          |*       .                0.219
19   .050   .218           .          |*       .                0.218

Plot Symbols:      Autocorrelations *     Two Standard Error Limits .
Total cases:  21     Computable first lags:  20
```

Figure (2.2a)

Sample Autocorrelations for Mid-Year Estimated Population

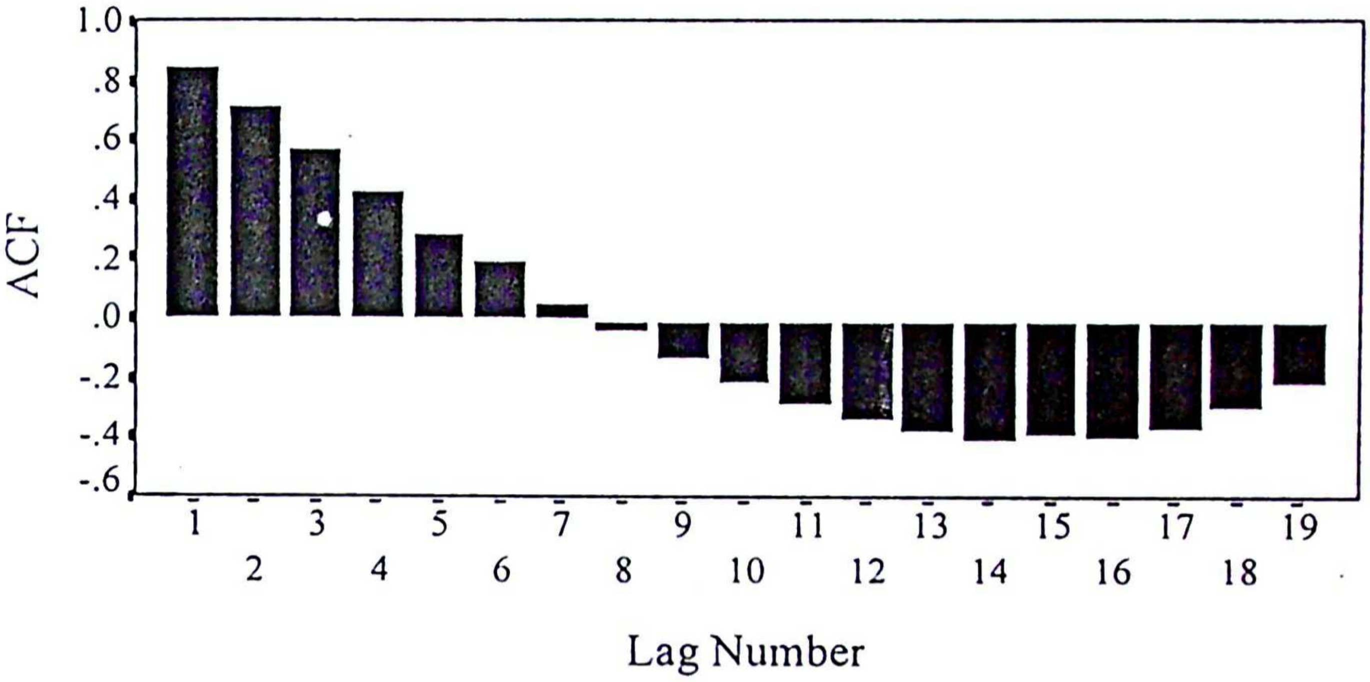
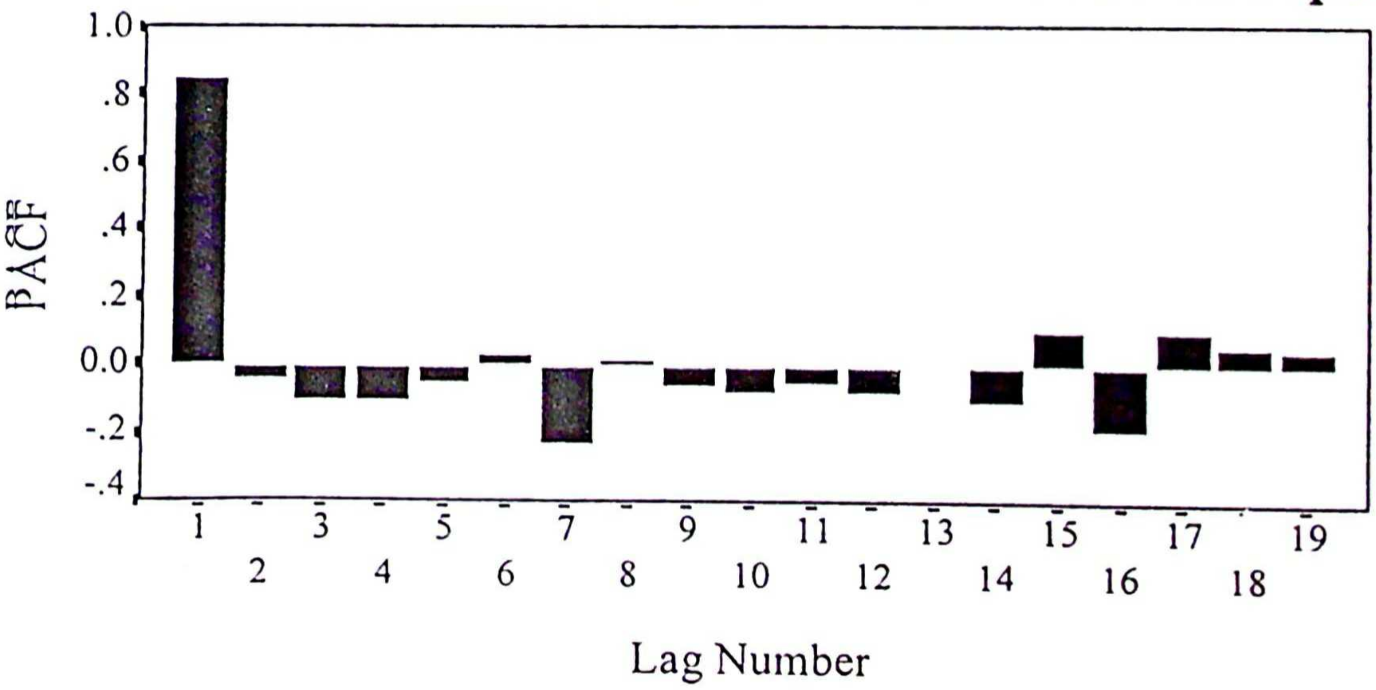


Figure (2.2b)

Sample Partial Autocorrelations for Mid-Year Estimated Population



2.6 Inter - Dependence Structure of Various Transformed Series

In the analysis of time series data, analysis will be carried out on the original series if it satisfies the assumptions needed to use a certain method of analysis or on the transformed series, the transformation being carried out so that the transformed series will satisfy those assumptions.

There are various ways of removing the seasonal component and some other systematic components, such as trend, oscillatory component, etc as well as the high fluctuations in the series.

The following transformations are carried out so that the transformed series no longer have the seasonal component and/or the high fluctuations.

(a) Differencing

When the given series has a trend component and/or high linear dependence between the successive values, the differencing is attempted to eliminate these effects. The differencing can be applied more than once if it is necessary and the seasonal diffencing can be applied to the seasonal time series so that the seasonal variation will be eliminated. Let Y_i be the differenced observation during i^{th} year and also denote Δ_s^r as the operator of γ^{th} difference which is applied to the s^{th} period. Then the successive difference (Δ_1^1)is

$$Y_i = \Delta_1^1 = X_{i+1} - X_i \; ; \; i = 1, 2, \ldots, 21$$

The successive differrencing can also be applied together and more than once.

(b) **Logarithmic Transformation**

Logarithmic are used to stablize the variance if the standard deviation in the original scale varies directly as the mean or if the coefficient of variation is constant [Sendecor-Cohran, 1980]. Similarly by using logarithmic transformation, high fluctuations are reduced and some skewed distributions become symmetric. If the observed series seem to be log-normally distributed, the natural logarithm of the observe series is normally distributed.

Let Y_i be the log-transformed data during the i^{th} year, then

$$Y_i = \ln X_i = \log_e X_i \; ; \; i = 1, 2, \ldots, 21$$

The logarithmic transformations are widely used in the analysis of population data.

(c) **Standardization**

The standardization is attempted to get zero mean and unit standard deviation. The standardization procedure can remove the seasonal variation to a certain extent, but the linear dependence between the successive values are still high.

Let X_t and Y_t be the observed and the standardized data for t^{th} year. Then

$$Y_t = \frac{X_t - \overline{X}}{S_t}; \qquad t = 1, 2, \ldots, 21$$

where $\overline{X} = \frac{1}{n}\sum_{t=1}^{n} X_t$, the mean of live-births or population.

and $S_t = \sqrt{\frac{1}{n-1}\sum_{t=1}^{n}(X_t - \overline{X})^2}$

= the standard deviation of live-births or population series.

In the yearly fertility and population data series of the urban area in the selected towns, there is a high correlation between the successive values. Thus, the following transformations are attempted to eliminate these effects.

(1) First difference

(2) Logarithmic transformation

(3) Standardization

For each of the transformed series, 20 sample autocorrelations γ_k, partial autocorrelations $\hat{\phi}_{kk}$ and the values of $(1-R^2_k)$, $k=1,2,\ldots,20$ are computed to study the internal dependence structure of these series.

The sample auto and partial correlations of lag k can describe the dependence structure of the values which are k lags apart. The values $(1-R^2_k)$, $k=1,2,...,20$ are computed by using the formula.

$$(1-R_k^2)=\prod_{i=1}^{k}(1-\hat{\phi}_{kk}^2);k=1,2,..,20$$ and it is the complement of the coefficient of determination when the underlying process is AR(k).

2.6.1 Inter- Dependence Structure of First Difference Series for Live-Births of Urban Myanmar

The linear dependence between the successive values for the live-births data are high. Thus the first differencing is applied. The sample and partial autocorrelations are presented in Table (2.3a) and (2.3b).

The sample autocorrelations of lag (1) to lag(3) are high and the confidence interval is (-0.4383,+0.4383). So all of them lie between the confidence intervals. By seeing the figure (2.3a), the sample correlation function is assumed to be tail off, since the sample autocorrelations are gradually discreased except for some lag values. The first difference series of live-births can be represented by an autoregressive process.

The correlograms of $\hat{\gamma}_k$ and $\hat{\phi}_{kk}$ are described in the Table (2.3a) and (2.3b). The partial autocorrelations of lag (1) and lag (2) are high and the $(1-R^2_k)$, $k=1,2,...,19$ are not significantly different from lag(1) and lag(2). Therefore the series obeys AR(1) scheme.

Table(2.3a)
Sample Autocorrelations for First Difference Series of Live-Births

```
Lag   Corr.   Err.  -1  -.75 -.5 -.25   0   .25  .5   .75   1
                    +----+----+----+----+----+----+----+----+
  1    .179   .224             .        |****    .
  2    .271   .231             .        |*****   .
  3    .196   .246            .         |****     .
  4   -.025   .254            .        *|         .
  5   -.132   .254            .      ***|         .
  6   -.167   .257            .      ***|         .
  7    .022   .263           .          *          .
  8   -.115   .263           .        **|          .
  9   -.251   .265           .     *****|          .
 10    .057   .277           .          |*         .
 11   -.221   .277           .      ****|          .
 12   -.116   .286           .        **|          .
 13   -.134   .288          .        ***|           .
 14   -.025   .292          .           *           .
 15   -.073   .292          .          *|           .
 16   -.002   .293          .           *           .
 17    .062   .293          .           |*          .
 18   -.041   .293          .          *|           .

Plot Symbols:          Autocorrelations *     Two Standard Error Limits .
Total cases:   21        Computable first lags after differencing:   19
```

Table(2.3b)
Sample Partial Autocorrelations for First Difference Series of Live-Births

```
Lag   Corr.   Err.  -1  -.75 -.5 -.25   0   .25  .5   .75   1   (1-Rx^2)
                    +----+----+----+----+----+----+----+----+
  1    .179   .224             .        |****    .             0.968
  2    .247   .224             .        |*****   .             0.909
  3    .126   .224             .        |***     .             0.895
  4   -.146   .224             .     ***|        .             0.876
  5   -.220   .224             .    ****|        .             0.834
  6   -.138   .224             .     ***|        .             0.818
  7    .208   .224             .        |****    .             0.783
  8    .027   .224             .        |*       .             0.782
  9   -.332   .224             . *******|        .             0.696
 10    .038   .224             .        |*       .             0.695
 11   -.112   .224             .      **|        .             0.686
 12    .010   .224             .        *        .             0.686
 13   -.093   .224             .      **|        .             0.686
 14   -.019   .224             .        *        .             0.667
 15   -.057   .224             .       *|        .             0.665
 16    .091   .224             .        |**      .             0.659
 17   -.051   .224             .       *|        .             0.657
 18   -.168   .224             .     ***|        .             0.638

Plot Symbols:          Autocorrelations *     Two Standard Error Limits .
Total cases:   21        Computable first lags after differencing:   19
```

Figure (2.3a)

Sample Autocorrelations for First Difference Series of Live-Births

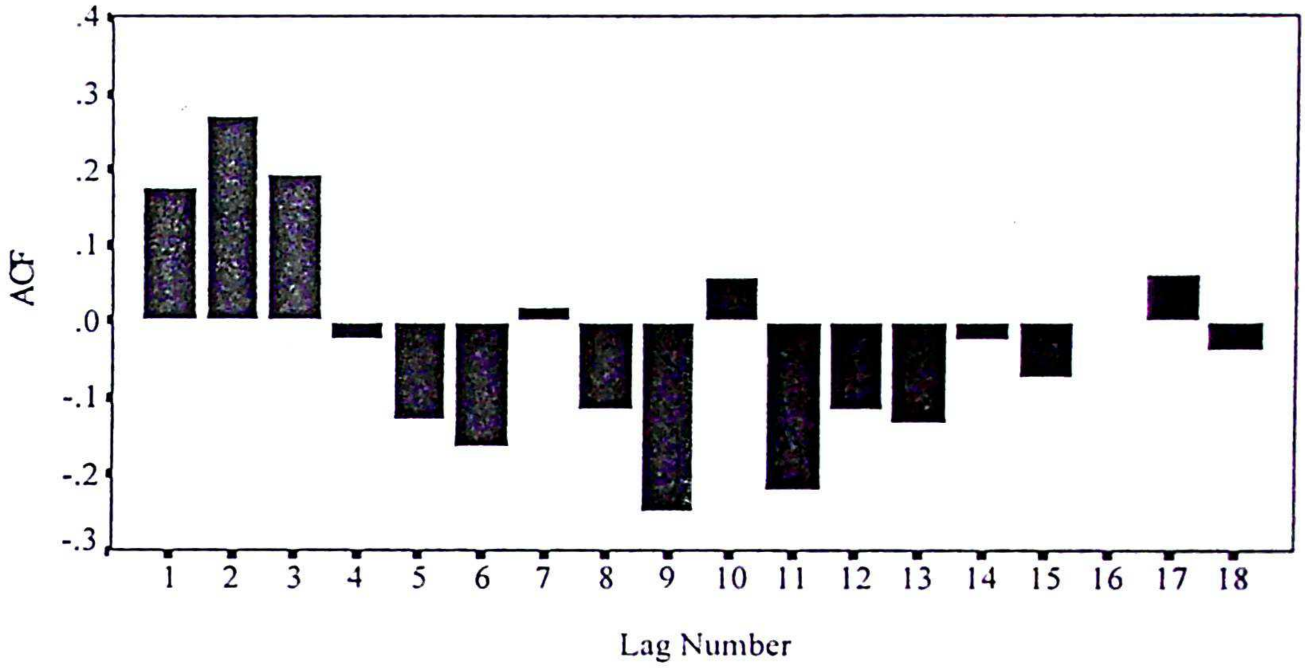
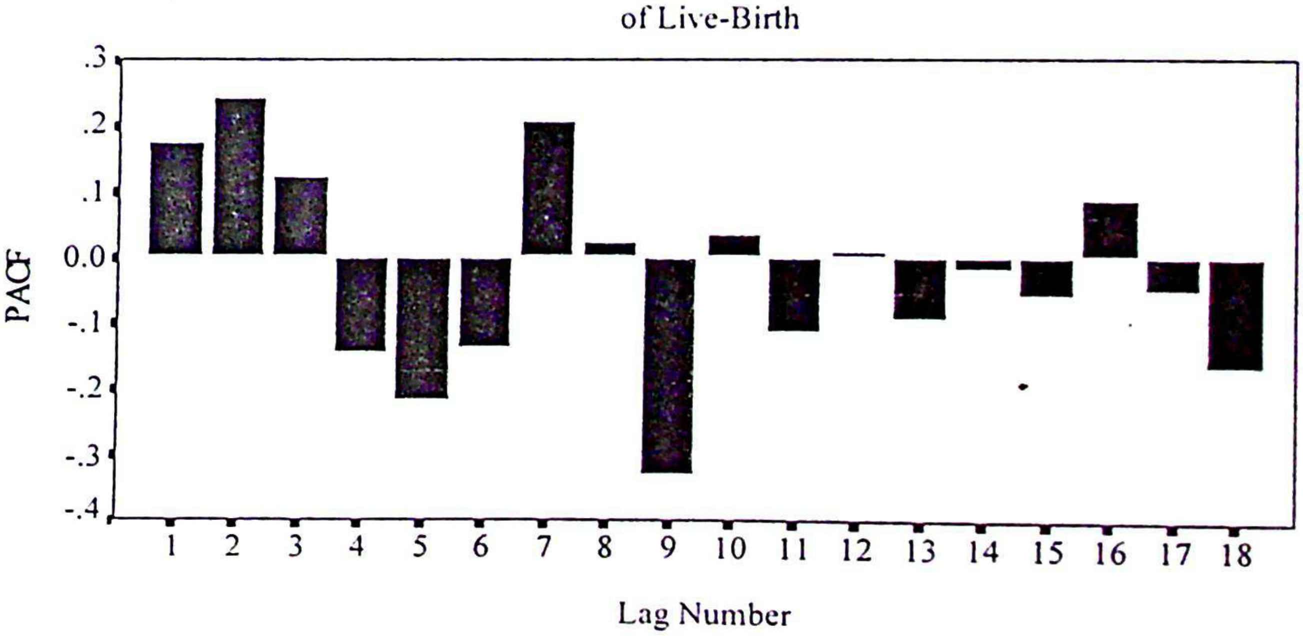


Figure (2.3b)

Sample Partial Autocorrelations for First Difference Series

of Live-Birth



2.6.2 Inter- Dependence Structure of First Difference Series for Mid-Year Estimated Population

At the Table (2.4a) and Table (2.4b), the sample autocorrelations and the partial autocorrelations are described and their correlograms are also presented in Figures (2.4a) and (2.4b) respectively.

From Table (2.4a), the autocorrelation for lag(1) is out of 95 percent confidence limit of a random series, i.e (-0.4383, +0.4383). Since these values are gradually decreasing from lag (1) through lag (6) and can be assumed to tail off, an AR process may be taken as the underlying process.

After studying the sample partial autocorrelation function , it is apparent that the underlying process can be taken as AR(1). Since it cut off after lag(1) and the partial autocorrelation for lag(1) is only out of the confidence interval. The same result is obtained when the complement of the coefficient of the determination is 0.62 in lag(1), 0.58 in lag(2) and 0.56 in lag(3). Thus, the series obey AR(1) scheme and we have to use AR(1) process as the underlying process of the first difference series for mid-year estimated population.

Table(2.4a)
Sample Autocorrelations for First Difference
Series of Mid-Year Estimated Population

```
Lag  Corr.  Err. -1   -.75  -.5  -.25    0   .25   .5   .75    1
                   +----+----+----+----+----+----+----+----+----+
  1  -.616  .224           ***.********|        .
  2   .227  .297           .           |*****      .
  3   .071  .305           .           |*          .
  4  -.166  .306           .        ***|           .
  5   .132  .310           .           |***        .
  6  -.104  .313          .          **|             .
  7   .017  .315          .            *             .
  8  -.117  .315          .          **|             .
  9   .093  .317          .            |**           .
 10  -.079  .319          .          **|             .
 11  -.002  .320          .            *             .
 12  -.008  .320          .            *             .
 13   .032  .320          .            |*            .
 14   .009  .320          .            *             .
 15   .006  .320          .            *             .
 16   .005  .320          .            *             .
 17  -.002  .320          .            *             .
 18   .005  .320          .            *             .

Plot Symbols:        Autocorrelations *     Two Standard Error Limits .
Total cases:   21      Computable first lags after differencing:  19
```

Table(2.4b)
Sample Partial Autocorrelations for First Difference
Series of Mid-Year Estimated Population

```
Lag  Corr.  Err. -1   -.75  -.5  -.25    0   .25   .5   .75    1  (1-R_k^2)
                   +----+----+----+----+----+----+----+----+----+
  1  -.616  .224           ***.********|        .             0.620
  2  -.245  .224              .   *****|        .             0.580
  3   .162  .224              .        |***     .             0.560
  4   .023  .224              .        *        .             0.560
  5  -.013  .224              .        *        .             0.560
  6  -.089  .224              .      **|        .             0.556
  7  -.086  .224              .      **|        .             0.552
  8  -.271  .224              .   *****|        .             0.511
  9  -.143  .224              .     ***|        .             0.501
 10  -.062  .224              .       *|        .             0.499
 11  -.058  .224              .       *|        .             0.497
 12  -.150  .224              .     ***|        .             0.486
 13  -.051  .224              .       *|        .             0.485
 14   .014  .224              .        *        .             0.458
 15  -.001  .224              .        *        .             0.458
 16  -.082  .224              .      **|        .             0.455
 17  -.084  .224              .      **|        .             0.452
 18  -.101  .224              .      **|        .             0.447

Plot Symbols:        Autocorrelations *     Two Standard Error Limits .
Total cases:   21      Computable first lags after differencing:  19
```

Figure (2.4a)

Sample Autocorrelations for First Difference Series

of Mid-Year Estimated Population

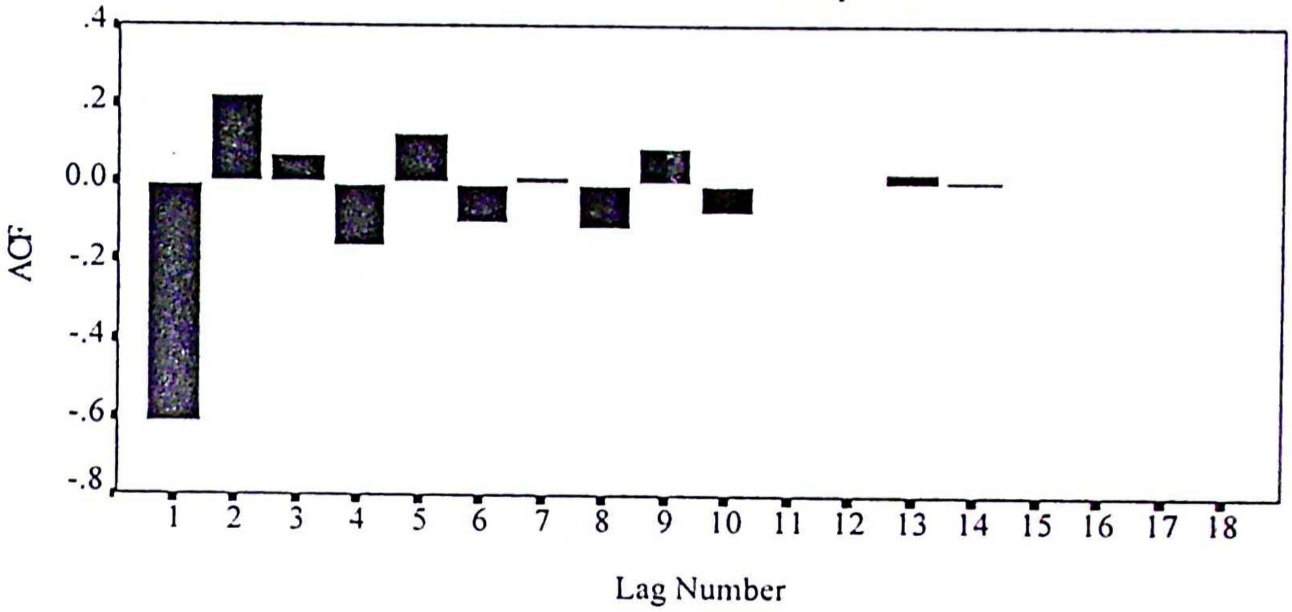
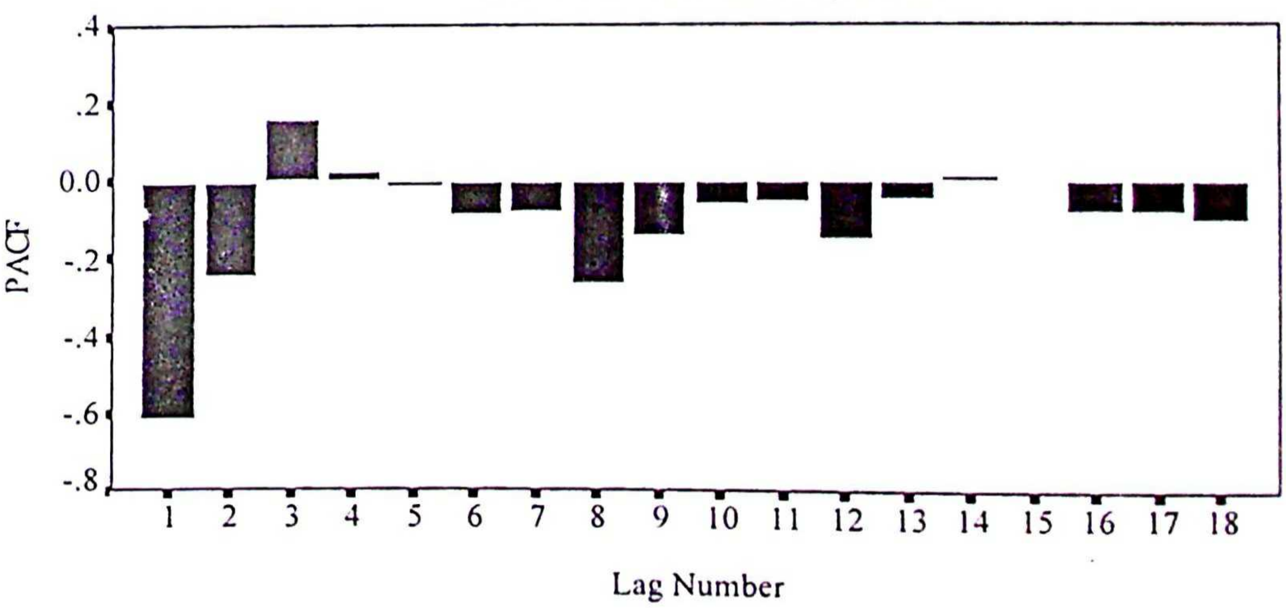


Figure (2.4b)

Sample Partial Autocorrelations for First Difference Series

of Mid-Year Estimated Population



2.6.3 Inter- Dependence Structure of Log-Transformed Series for Live-Births of Urban Myanmar

The sample autocorrelations and the partial autocorrelations of log-transformed series for live-births for k=1,2,...,20 are presented in Table (2.5a) and (2.5b).

From Table (2.5a), the sample autocorrelation for lag(1) is highest and it can be seen that the sample autocorrelations are gradually decreasing and the autocorrelation function can be assumed to tail off. The correlations for lag(1) and lag(2) are out of the confidence limits. Therefore, this process may be autoregressive process.

When the partial autocorrelations are studied, the partial autocorrelation for lag(1) is out of the confidence limits (-0.4277,+0.4277) and it cuts off after lag(1). The complement of the coefficient of determination is 0.355 in lag(1) and 0.352 in lag(2). Thus we have to use AR(1) process as the underlying process of log-transformed series for live-births. The sample correlogram and the partial correlogram are also presented in Figures (2.5a) and(2.5b).

Table(2.5a)
Sample Autocorrelations for Log-Transfformed Series of Live-Births

Lag	Auto-Corr.	Stand. Err.	Plot (-1 … 1)
1	.803	.218	****************
2	.614	.330	************
3	.391	.381	********
4	.181	.399	****
5	.005	.403	*
6	-.141	.403	***
7	-.236	.406	*****
8	-.330	.412	*******
9	-.357	.424	*******
10	-.381	.438	********
11	-.350	.454	*******
12	-.235	.467	*****
13	-.132	.472	***
14	-.052	.474	*
15	-.025	.474	*
16	-.009	.474	*
17	-.043	.474	*
18	-.098	.475	**
19	-.057	.475	*

Plot Symbols: Autocorrelations * Two Standard Error Limits .
Total cases: 21 Computable first lags: 20

Table(2.5b)
Sample Partial Autocorrelations for Log-Transfformed Series of Live-Births

Lag	Pr-Aut-Corr.	Stand. Err.	Plot (-1 … 1)	$(1-R_k^2)$
1	.803	.218	****************	0.355
2	-.085	.218	**	0.352
3	-.212	.218	****	0.336
4	-.131	.218	***	0.330
5	-.064	.218	*	0.329
6	-.085	.218	**	0.327
7	-.037	.218	*	0.327
8	-.157	.218	***	0.319
9	.009	.218	*	0.319
10	-.108	.218	**	0.315
11	.016	.218	*	0.315
12	.176	.218	****	0.305
13	-.030	.218	*	0.276
14	-.113	.218	**	0.272
15	-.136	.218	***	0.267
16	-.038	.218	*	0.267
17	-.115	.218	**	0.263
18	-.137	.218	***	0.258
19	.230	.218	*****	0.244

Plot Symbols: Autocorrelations * Two Standard Error Limits .
Total cases: 21 Computable first lags: 20

Figure (2.5a)

Sample Autocorrelations for Log-transformed Series of Live-Births

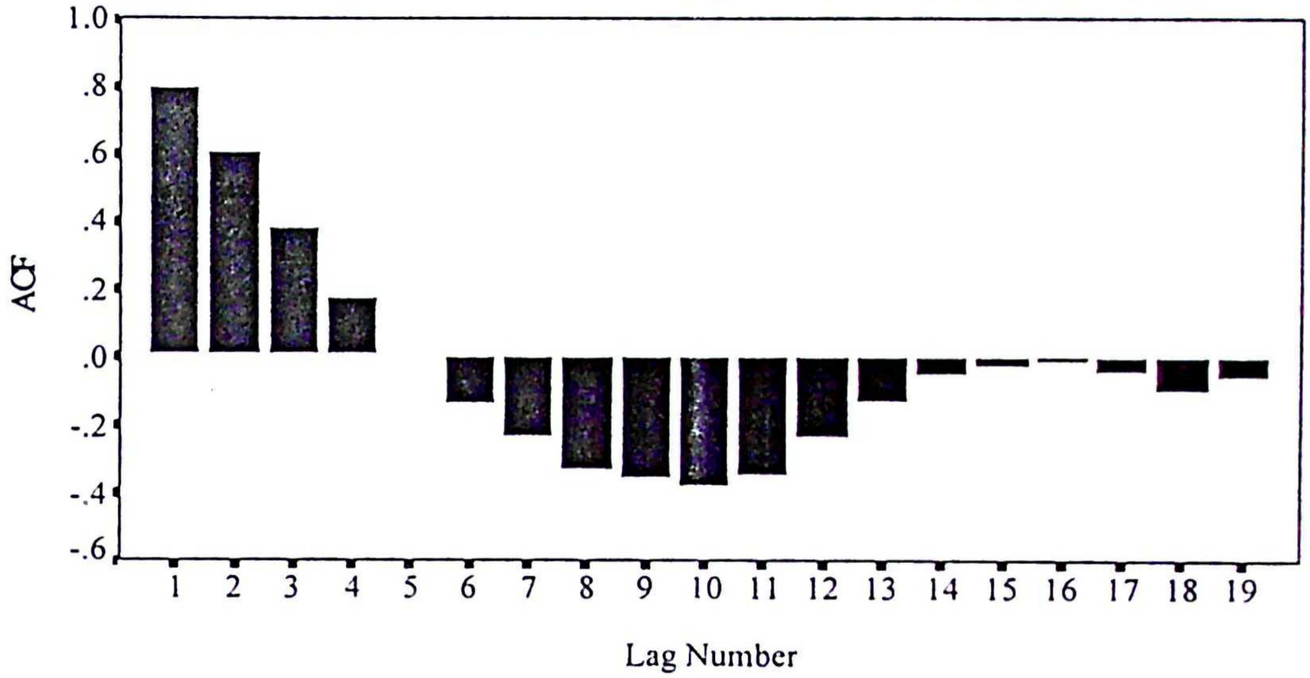
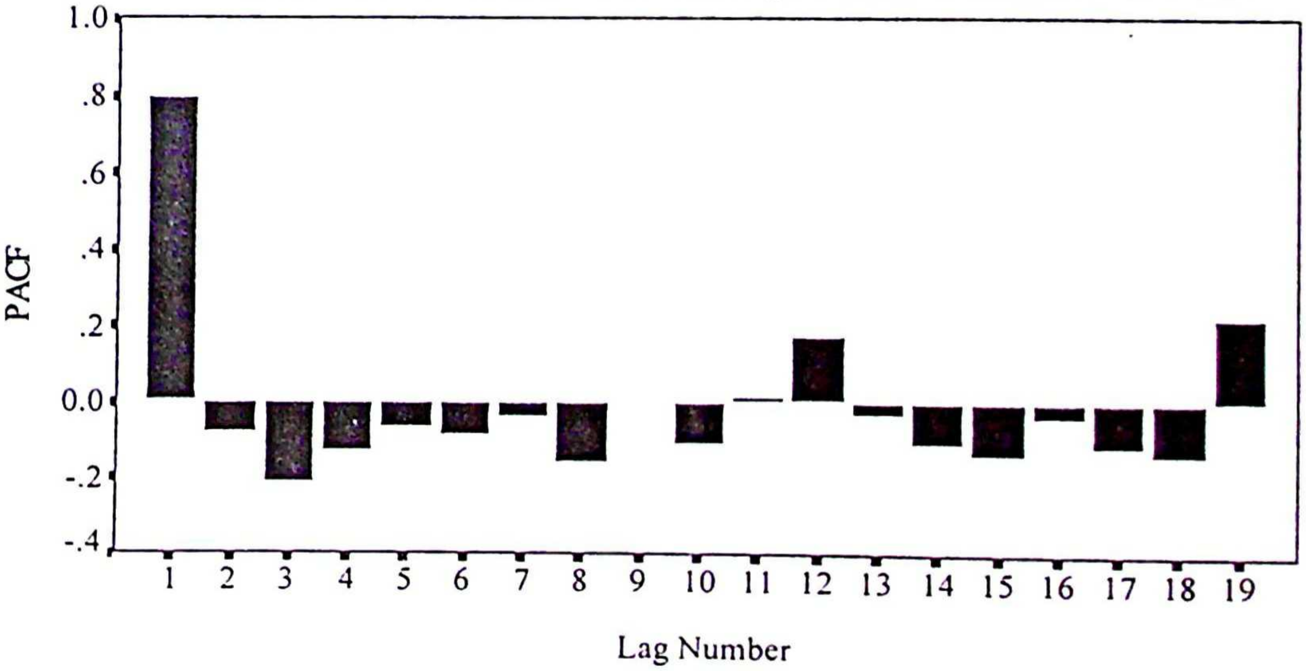


Figure (2.5b)

Sample Partial Autocorrelations for log-transformed Series of Live-Birth



2.6.4 Inter- Dependence Structure of Log-Transformed Series for Mid-Year Estimated Population

At the Table (2.6a) and Table (2.6b)the sample autocorrelations and the partial autocorrelations are described.

In this Table (2.6a), the sample autocorrelation for lag(1), 0.855 is the highest and it gradually declines up to lag(14). Therefore, the sample autocorrelation function can be assumed to tail off and the lag(1) to the lag(3) are out of the 95 percent confidence interval (-0.4277,+0.4277). Hence, the process may be taken as an AR process.

The figure of $\hat{\gamma}_k$ and $\hat{\phi}_{kk}$ for k=1,2,...,20 are also presented in Figure (2.6a) and (2.6b). When the partial autocorrelations are examined, the lag one value is significantly out of the 95 percent limits and it cuts off after lag(1). Hence, underlying process can be an AR(1) process.The complement of determination $(1-R^2_k)$, k=1,2,...,20 is 0.268 in lag(1) and 0.267 in lag(2). Thus, we have to use the AR(1) process as underlying process of log transformed series for mid-year estimated population.

Table(2.6a)
Sample Autocorrelations for Log-Transformed Series of Mid-Year Estimated Population

```
Lag  Corr.  Err.  -1  -.75 -.5  -.25  0   .25  .5   .75   1
                  +----+----+----+----+----+----+----+----+
  1   .855  .218             .        |********.********
  2   .718  .343        .             |**************
  3   .572  .408      .               |***********    .
  4   .420  .445    .                 |********         .
  5   .284  .463   .                  |******            .
  6   .186  .471   .                  |****              .
  7   .046  .475   .                  |*                 .
  8  -.050  .475   .                 *|                  .
  9  -.138  .475   .               ***|                  .
 10  -.220  .477   .              ****|                  .
 11  -.286  .482   .            ******|                  .
 12  -.336  .490  .            *******|                   .
 13  -.375  .501  .           ********|                   .
 14  -.405  .514              ********|                   .
 15  -.387  .529              ********|
 16  -.398  .542              ********|
 17  -.361  .556               *******|
 18  -.297  .567                ******|
 19  -.215  .574                  ****|

Plot Symbols:      Autocorrelations *     Two Standard Error Limits .
Total cases:  21     Computable first lags:  20
```

Table(2.6b)
Sample Partial Autocorrelations for Log-Transformed Series of Mid-Year Estimated Population

```
Lag  Corr.  Err.  -1  -.75 -.5  -.25  0   .25  .5   .75   1  (1-R_k^2)
                  +----+----+----+----+----+----+----+----+
  1   .855  .218             .        |********.********     0.268
  2  -.050  .218             .       *|        .             0.267
  3  -.114  .218             .      **|        .             0.264
  4  -.116  .218             .      **|        .             0.260
  5  -.047  .218             .       *|        .             0.259
  6   .047  .218             .        |*       .             0.258
  7  -.246  .218             .   *****|        .             0.242
  8   .025  .218             .        |*       .             0.242
  9  -.069  .218             .       *|        .             0.241
 10  -.076  .218             .      **|        .             0.240
 11  -.059  .218             .       *|        .             0.239
 12  -.089  .218             .      **|        .             0.237
 13  -.005  .218             .        *        .             0.237
 14  -.117  .218             .      **|        .             0.234
 15   .100  .218             .        |**      .             0.232
 16  -.193  .218             .    ****|        .             0.223
 17   .105  .218             .        |**      .             0.221
 18   .055  .218             .        |*       .             0.220
 19   .044  .218             .        |*       .             0.220

Plot Symbols:      Autocorrelations *     Two Standard Error Limits .
Total cases:  21     Computable first lags:  20
```

Figure (2.6a)

Sample Autocorrelations for Log-transformed Series

of Mid-Year Estimated Population

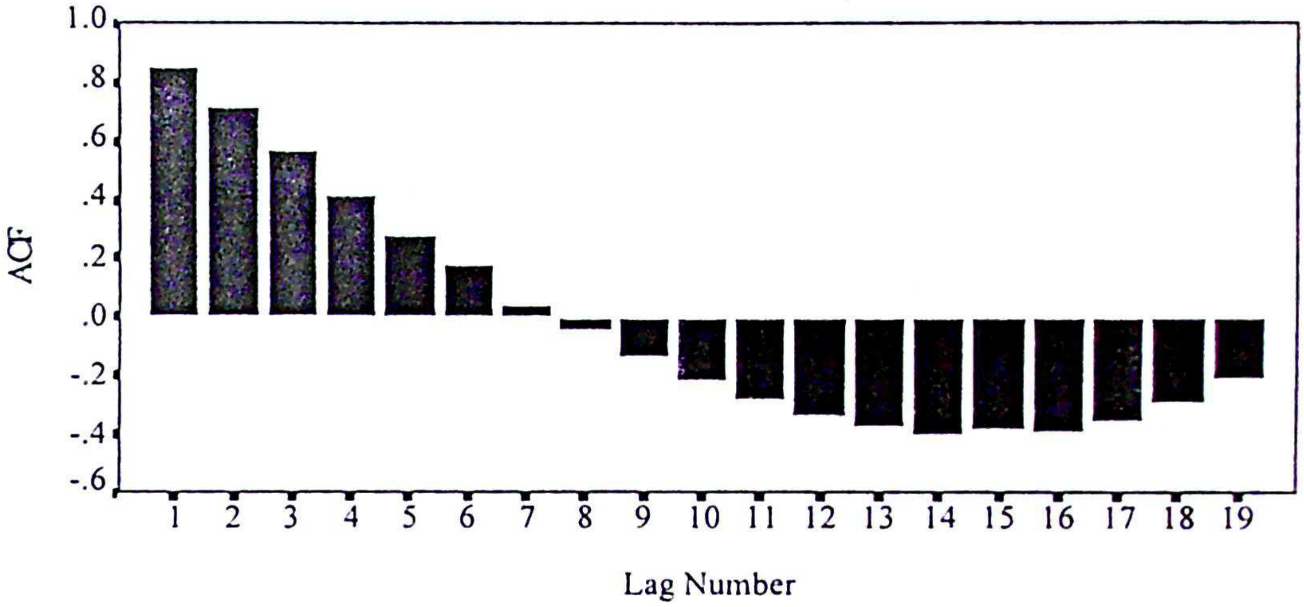
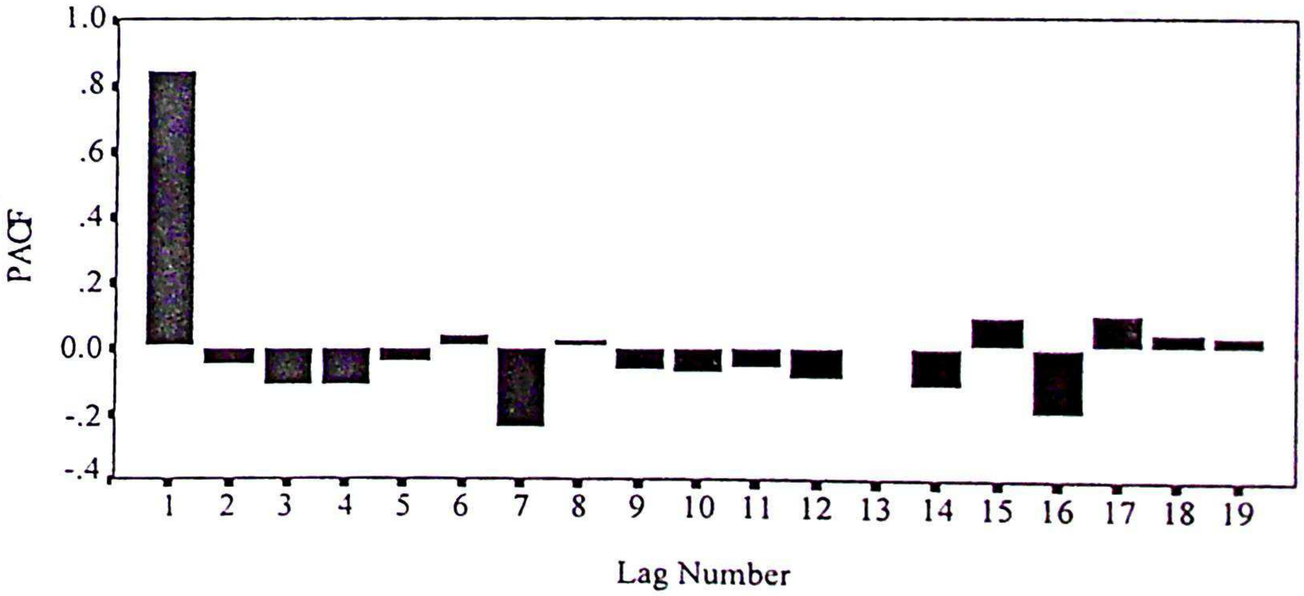


Figure (2.6b)

Sample Partial Autocorrelations for Log-transformed Series

of Mid-Year Estimated Population



2.6.5 Inter- Dependence Structure of Standardized Series for Live-Births of Urban Myanmar

The sample auto and partial correlations together with $(1-R^2_k)$, k=1,2,...,20 are computed to study the internal dependent structure of the standardized series. The values are presented in Table (2.7a) and (2.7b).

For standardized series, the lag(1) and lag(2) sample autocorrelations are out of the 95 percent confidence limits of a random series, i.e, -0.4277 and+0.4277. Since, the values are gradually decreasing from lag(1) through lag(11) and can be assumed to tail off, an AR process may be taken as the underlying process.

After studying the sample partial autocorrelation function, it is apparent that the underlying process can be taken as AR(1) since it cuts off after lag(1). The same result is obtained when the complement of the coefficient of determination is examined since the difference between lag(1) value and lag(2) value is quite small. Its value for the lag(1) is 0.365 and for lag(2) is 0.364.

Table(2.7a)
Sample Autocorrelations for Standardized Series of Live-Births

```
     Auto- Stand.
Lag  Corr.   Err. -1   -.75   -.5 -.25    0    .25  .5    .75   1
                  +----+----+----+----+----+----+----+----+----+
  1   .797   .218                .        |******** .*******
  2   .611   .329            .            |************.
  3   .393   .379          .              |********         .
  4   .189   .398         .               |****               .
  5   .019   .402         .               *                   .
  6  -.125   .402         .           ***|                    .
  7  -.223   .404         .          ****|                    .
  8  -.319   .410         .        ******|                    .
  9  -.347   .422        .        *******|                     .
 10  -.375   .435        .        *******|                     .
 11  -.347   .450       .         *******|                      .
 12  -.241   .463      .            *****|                       .
 13  -.142   .469      .              ***|                       .
 14  -.064   .471      .                *|                       .
 15  -.036   .471      .                *|                       .
 16  -.018   .471      .                 *                       .
 17  -.051   .471      .                *|                       .
 18  -.106   .472      .               **|                       .
 19  -.063   .473      .                *|                       .
```

Plot Symbols: Autocorrelations * Two Standard Error Limits .
Total cases: 21 Computable first lags: 20

Table(2.7b)
Sample Partial Autocorrelations for Standardized Series of Live-Births

```
   Pr-Aut- Stand.
Lag  Corr.   Err. -1   -.75   -.5 -.25    0    .25  .5    .75   1 (1-Rk^2)
                  +----+----+----+----+----+----+----+----+----+
  1   .797   .218              .      |******** .*******      0.364
  2  -.069   .218              .     *|         .             0.358
  3  -.199   .218              .  ****|         .             0.343
  4  -.125   .218              .    **|         .             0.337
  5  -.064   .218              .     *|         .             0.336
  6  -.092   .218              .    **|         .             0.336
  7  -.043   .218              .     *|         .             0.333
  8  -.149   .218              .   ***|         .             0.332
  9   .011   .218              .      *         .             0.325
 10  -.108   .218              .    **|         .             0.321
 11   .012   .218              .      *         .             0.321
 12   .158   .218              .      |***      .             0.313
 13  -.019   .218              .      *         .             0.313
 14  -.094   .218              .    **|         .             0.313
 15  -.129   .218              .   ***|         .             0.305
 16  -.044   .218              .     *|         .             0.304
 17  -.128   .218              .   ***|         .             0.299
 18  -.147   .218              .   ***|         .             0.278
 19   .229   .218              .      |*****    .             0.293
```

Plot Symbols: Autocorrelations * Two Standard Error Limits .
Total cases: 21 Computable first lags: 20

Figure (2.7a)

Sample Autocorrelations for Standardized Series of

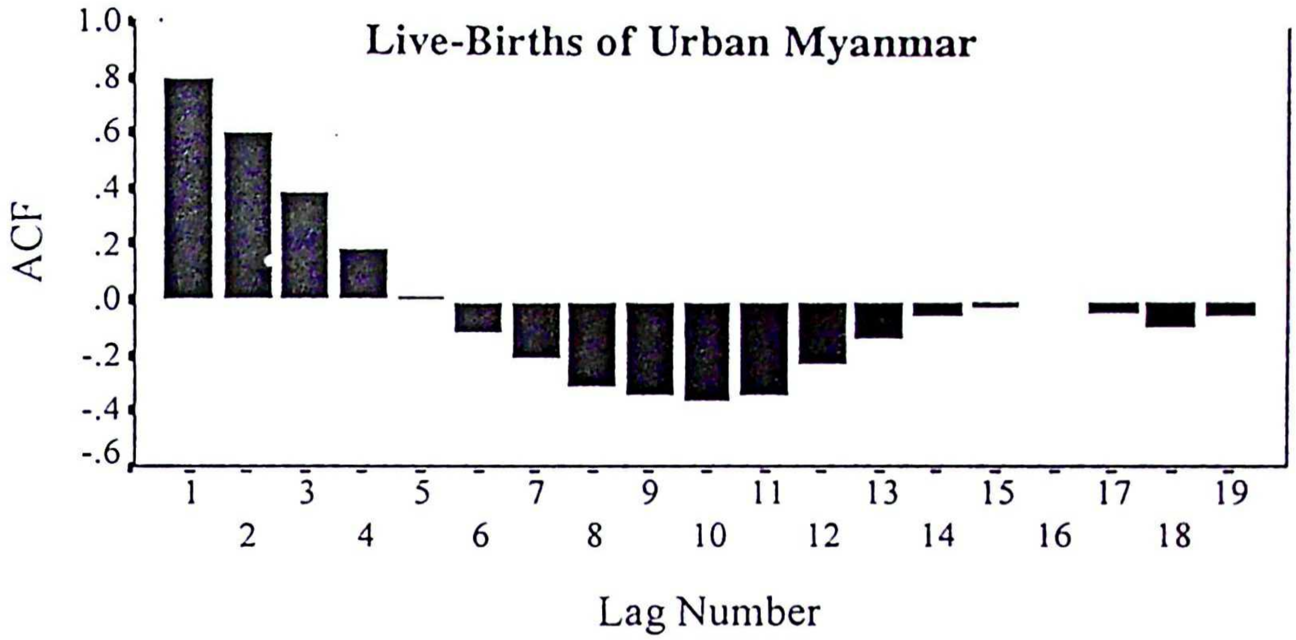
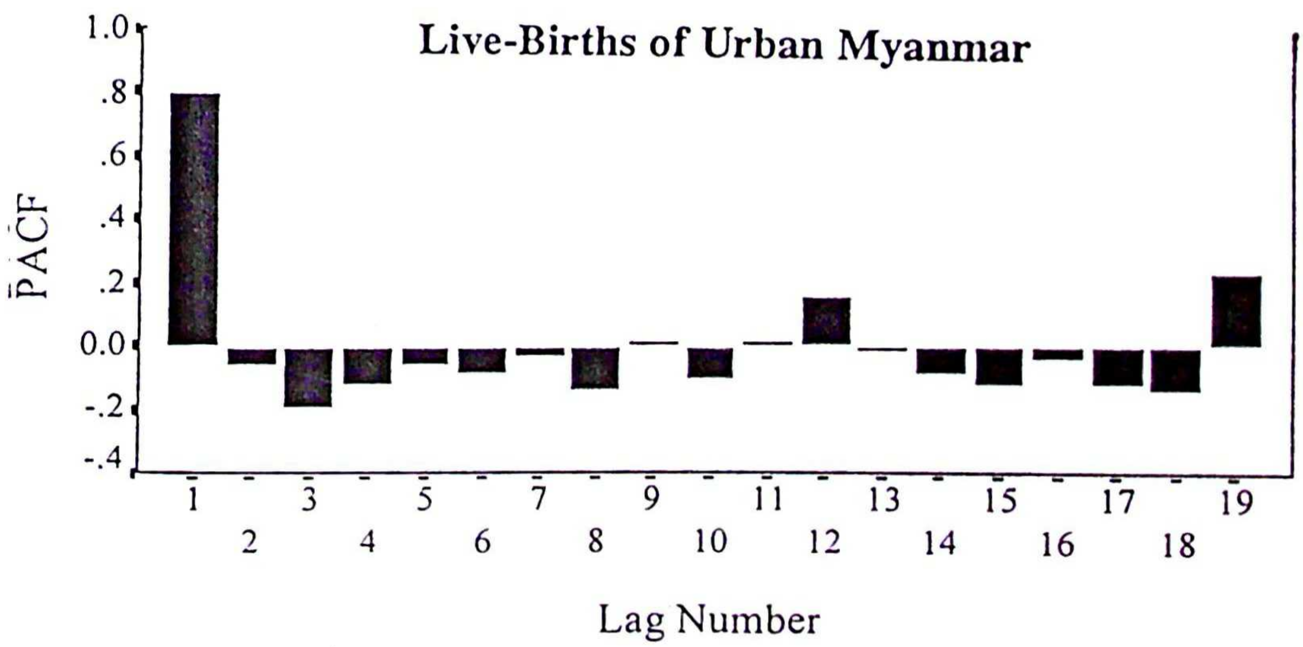


Figure (2.7b)

Sample Partial Autocorrelations for Standardized Series of



2.6.6 Inter - Dependence Structure of Standardized Series for Mid-Year Estimated Population

The sample auto and the partial correlations together with $(1-R^2_k)$, are computed for lag(1) to(20) and presented in Table (2.8a) and (2.8b). Sample correlogram and partial correlogram are also presented in Figure (2.8a) and (2.8b)

The sample correlation for lag(1) is the highest and the lag(1) to lag(4) are out of the 95 percent confidence limits (-.4277, +.4277) of a random series. The lag one sample autocorrelation is 0.857 and the sample autocorrelations pattern can be assumed to tail off. Thus an AR process may be taken to represent the series.

To determine the order of the process, the partial autocorrelations and the complement of coefficient of determination are used. The lag one sample partial atuocorrelation is the only value which is significantly out of the 95 percent limits of a random series and the AR(1) process can be used to represent the series.

Similar result is also obtained by examining the complement of the coefficient of determination $(1-R^2_k)$, k=1,2,...,20. Its value for the lag(1) is 0.265 and for lag(2) is 0.255.Therefore, the series obey AR(1) scheme.

From the above results, the standardized series, the logarithmic transformed series, the first difference series and the original series are found to be the series, which can be represented by an AR(1) process properly. Thus, all of the series will be used in the live-births and population for selected towns in Urban Myanmar.

Table(2.8a)
Sample Autocorrelations for Standardized Series
of Mid-year Estimated Population

```
      Auto-  Stand.
Lag   Corr.   Err.  -1   -.75 -.5  -.25 0    .25  .5   .75  1
                    +----+----+----+----+----+----+----+----+
  1    .857   .218              .       |********.********
  2    .722   .343         .            |**************
  3    .579   .409       .              |************  .
  4    .431   .446     .                |*********      .
  5    .296   .466    .                 |******           .
  6    .195   .475    .                 |****             .
  7    .056   .478    .                 |*                .
  8   -.044   .479    .                *|                 .
  9   -.136   .479    .              ***|                 .
 10   -.222   .481    .             ****|                 .
 11   -.290   .486    .           ******|                 .
 12   -.343   .494   .           *******|                  .
 13   -.383   .505   .          ********|                  .
 14   -.413   .519              ********|
 15   -.394   .534              ********|
 16   -.405   .548              ********|
 17   -.366   .562               *******|
 18   -.302   .573                ******|
 19   -.219   .581                  ****|
```

Plot Symbols: Autocorrelations * Two Standard Error Limits .
Total cases: 21 Computable first lags: 20

Table(2.8b)
Sample Partial Autocorrelations for Standardized Series
of Mid-year Estimated Population

```
      Pr-Aut- Stand.
Lag   Corr.   Err.  -1   -.75 -.5  -.25 0    .25  .5   .75  1 (1-R_k^2)
                    +----+----+----+----+----+----+----+----+
  1    .857   .218              .       |********.********     0.266
  2   -.046   .218              .      *|       .              0.265
  3   -.111   .218              .     **|       .              0.262
  4   -.108   .218              .     **|       .              0.258
  5   -.054   .218              .      *|       .              0.257
  6    .032   .218              .       |*      .              0.257
  7   -.235   .218              .  *****|       .              0.241
  8    .013   .218              .       *       .              0.241
  9   -.068   .218              .      *|       .              0.240
 10   -.082   .218              .     **|       .              0.238
 11   -.058   .218              .      *|       .              0.237
 12   -.081   .218              .     **|       .              0.235
 13   -.013   .218              .       *       .              0.235
 14   -.109   .218              .     **|       .              0.232
 15    .102   .218              .       |**     .              0.230
 16   -.189   .218              .   ****|       .              0.222
 17    .102   .218              .       |**     .              0.220
 18    .056   .218              .       |*      .              0.219
 19    .050   .218              .       |*      .              0.218
```

Plot Symbols: Autocorrelations * Two Standard Error Limits .
Total cases: 21 Computable first lags: 20

Figure (2.8a)

Sample Autocorrelations for Standardized Seies

of Mid-Year Estimated Population

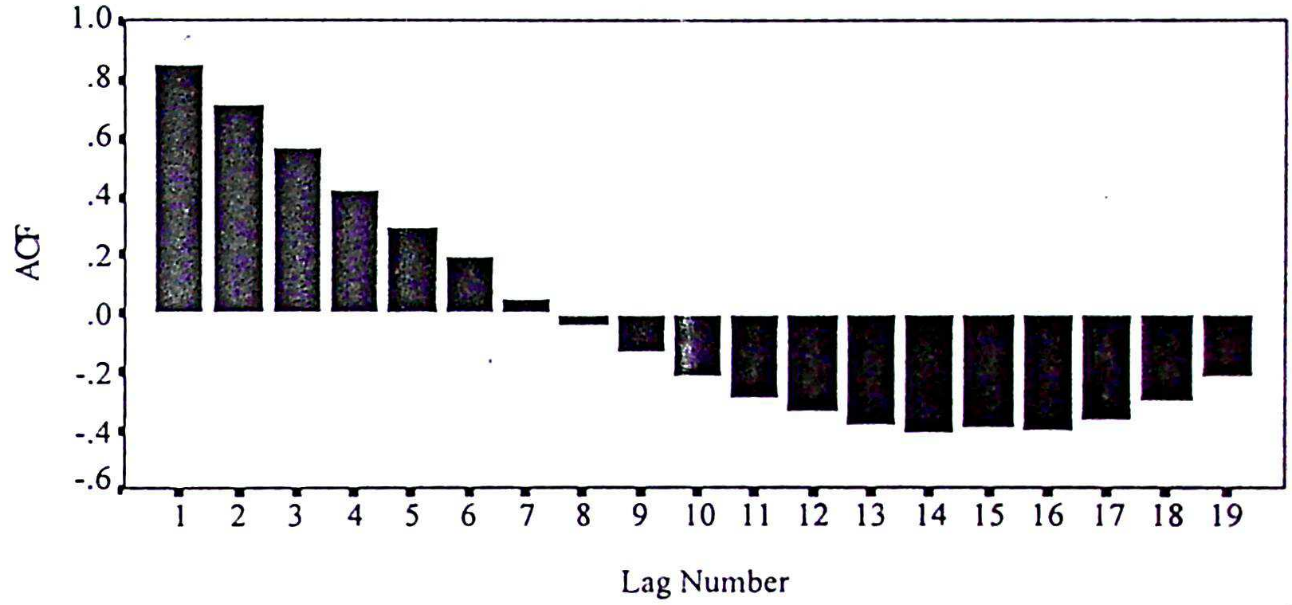
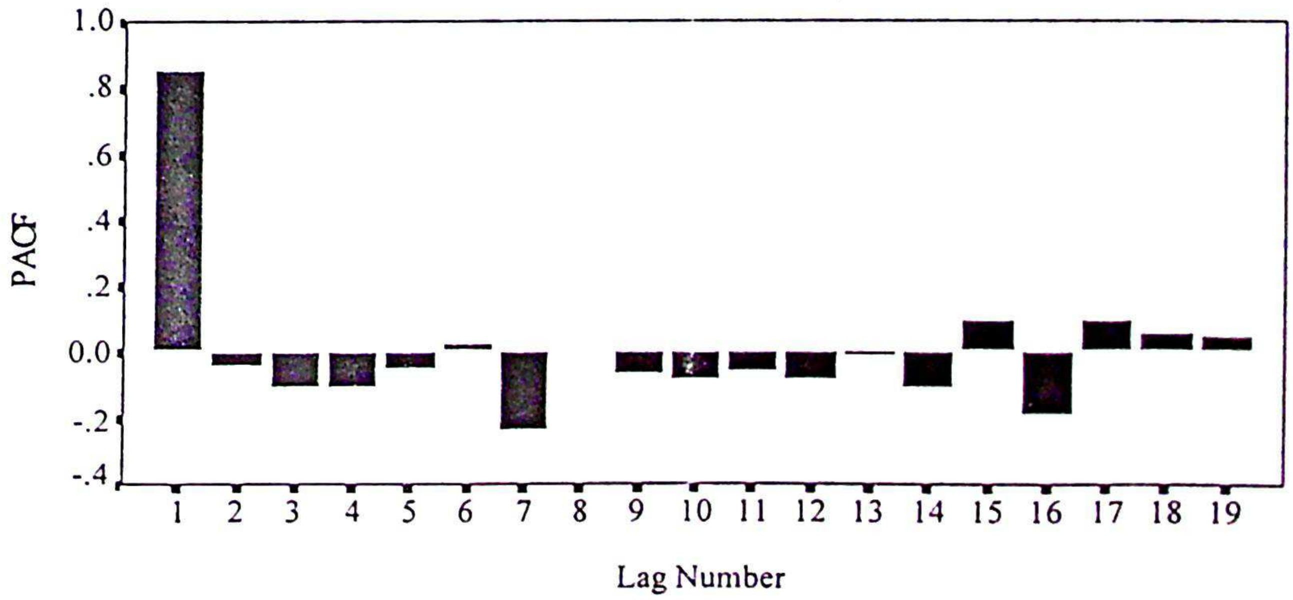


Figure (2.8b)

Sample Partial Autocorrelations for Standardized Series

of Mid-Year Estimated Population



CHAPTER III

AUTOREGRESSIVE MODEL FITTING IN FERTILITY DATA

3.1 Introduction

From the theoretical point of view, it may be seemed preferable to use behavioural models for calculating predictions. After studying the autocorrelation and the partial autocorrelation of fertility data in the previous chapter, it is apparent that the series can be represented by autoregressive (AR) model. This model generates a new predictor variable by using the y variable lagged one or more periods.

The application of this model has been attractive in fertility mainly because the autoregressive form has an intuitive type of time dependence, i.e., the value of a variable at the present time depends on the value at previous time and the autoregressive models are the simplest model to use.

In this chapter, the general discrete autoregressive models presented with its properties, such as acvf, partial acf and the spectrum. Then, the estimation of parameters by using the method of moments and the maximum likelihood method is described in detail.

The test associated with the model fitting and the estimation of its parameters are also described in section (3.3). Then the autoregressive models are fitted to the selected series.

3.2 Discrete Autoregressive Model and Its Properties

A stochastic model which can be extremely useful in the representation of certain practically occurring series is called autoregressive model. In this model, the current value of the process is expressed as finite, linear aggregate of previous values of the process and a shock a_t.

Let us denote the values of a process at equally spaced times t, t-1, t-2, ... by x_t, x_{t-1}, x_{t-2}, ... Also let $\tilde{x}_t$, $\tilde{x}_{t-1}$, ... be deviation from μ : for example,

$$\tilde{x}_t = x_t - \mu$$

Then

$$\tilde{x}_t = \phi_1 \tilde{x}_{t-1} + \phi_2 \tilde{x}_{t-2} + \ldots + \phi_p \tilde{x}_{t-p} + a_t \qquad 3.2.1$$

If we define an autoregressive operator of order p by

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p \qquad 3.2.2$$

then the autoregressive model may be written economically as,

$$\phi(B) \tilde{z}_t = a_t \qquad 3.2.3$$

The model contains p+2 unknown parameters μ, ϕ_1, ϕ_2, ..., ϕ_p, σ_a^2 which in practice have to be estimated from the data.

The white noise process a_t is assumed to be independently and identically distributed random error term with mean zero and variance σ_a^2. That is

$E(a_t) = 0$

$E(a_t, a_{t+k}) = 0 \qquad k \neq 0$

$= \sigma_a^2 \qquad k = 0$

The most widely used AR models in fertility data are only the lower order autoregressive models, especially of order 1 and 2. [Ronald, JASA, Vol 69, p-608]

For p=1, the first order autoregressive process is

$$x_t - \mu = \phi_1 (x_{t-1} - \mu) + a_t \; \ldots \qquad 3.2.4$$

Since $\tilde{x}_t$ depends only on the previous observations, it is referred to as a Markov process and there are three parameters (μ, ϕ, σ_a^2) to be estimated.

Similarly, the AR(2) models is

$$X_t - \mu = \phi_1 (X_{t-1} - \mu) + \phi_2 (X_{t-2} - \mu) + a_t \qquad 3.2.5$$

and there are four parameters $(\mu, \phi_1, \phi_2, \sigma_a^2)$ to be estimated. The AR(2) model is also known as Yule model. The properties of the discrete general AR(p) model, AR(1) and AR(2) models are discussed in term of the autocovariance function, autocorrelation function, partial autocorrelation function, the stationarity condition to be met and the power spectrum in the following subsections.

3.2.1 Autocovariance and Autocorrelation Functions

The autocovariance function of the autoregressive AR(p) process is found by multiplying throughout in

$\tilde{x}_t = \phi_1 \tilde{x}_{t-1} + \phi_2 \tilde{x}_{t-2} + \ldots + \phi_p \tilde{x}_{t-p} + a_t$ .. by $\tilde{x}_{t-k}$ to obtain

$$\tilde{x}_{t-k} \tilde{x}_t = \phi_1 \tilde{x}_{t-k} \tilde{x}_{t-1} + \phi_2 \tilde{x}_{t-k} \tilde{x}_{t-2} + \ldots + \phi_p \tilde{x}_{t-k} \tilde{x}_{t-p} + \tilde{x}_{t-k} a_t \qquad (3.2.6)$$

on taking expected values in (3.2.6),

We obtain the different equation

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \ldots + \phi_p \gamma_{k-p} + \ldots k>0 \qquad 3.2.7$$

$E(\tilde{x}_{t-k}\, a_t)$ vanishes when k>0, since $\tilde{x}_{t-k}$ can only involve the shock a_j up to time t-k, which are uncorrelated with a_t. On dividing γ_k by γ_o, it can be obtained the autocorrelation function

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \ldots + \phi_p \rho_{k-p} \ldots k>0 \qquad 3.2.8$$

where $\gamma_o = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \ldots + \phi_p \gamma_p + \sigma_a^2$ since $\gamma_j = \gamma_{-j}$ and $\rho_0 = 1$

By replacing $\rho_j = \dfrac{\gamma_j}{\sigma_x^2}$ or $\gamma_j = \sigma_x^2 \rho_j$ in this eq^n gives

$$\sigma_x^2 = \phi_1 \rho_1 \sigma_x^2 + \phi_2 \rho_2 \sigma_x^2 + \ldots + \phi_p \rho_p \sigma_x^2 + \sigma_a^2$$

$$\sigma_x^2 (1 - \phi_1\rho_1 - \phi_2\rho_2 - \ldots - \phi_p\rho_p) = \sigma_a^2$$

$$\sigma_x^2 = \frac{\sigma_a^2}{1 - \phi_1 \rho_1 - \phi_2 \rho_2 - \ldots - \phi_p \rho_p} \qquad 3.2.9$$

equation (3.2.9) is the relation between σ_x^2 and σ_a^2. The eq^n (3.2.8) is known as Yule-Walker equation. This equation is due to Yule (1927) and Walker (1931) and commonly used for estimating the parameters of the AR(p) model by the method of moment [Box-Jenkin, 1976]. The other important use of this equation is for dertermining the correlogram ρ_k for a given set of parameters ϕ_j, j= 1, 2, ..., p. It is important to know the shape of ρ_k for a given model because it will serve for identifying the order of the model.

The acvf and acf of the AR(1) process is found as

$$(1 - \phi_1 B)\, \tilde{x}_t = a_t$$

$$\tilde{x}_t - \phi_1 \tilde{x}_{t-1} = a_t$$

where $B\tilde{x}_t = \tilde{x}_{t-1}$ which is the backward shift operator

$$\therefore\ \tilde{x}_t = a_t\ (1-\phi_1 B)^{-1}$$

$$= \sum_{j=0}^{\alpha} (\phi_1 B)^j a_t$$

$$= \sum_{j=0}^{\alpha} \phi_1^j B^j a_t$$

$$= \sum_{j=0}^{\alpha} \phi_1^j a_{t-j} \quad \text{since } B^j a_t = a_{t-j} \qquad 3.2.10$$

Then $\tilde{x}_t$ is liner process $\sum_{j=0}^{\alpha} h_j a_{t-j}$ which $h_j = \phi_1^j$.

The autocovariance function is

$$\gamma_k = E(\tilde{x}_t\, \tilde{x}_{t+k})$$

$$= E\left(\sum_{j=0}^{\alpha} h_j a_{t-j} \quad \sum_{j=0}^{\alpha} h_j a_{t+k-j}\right)$$

$$= E\left(\sum_{j=0}^{\alpha} h_j h_{j+k} a_{t-j}^2 + C.P.T\right)$$

$$= \sigma_a^2 \sum_{J=0}^{\alpha} h_j h_{j+k} \qquad 3.2.11$$

But $h_j = \phi_1^j$

The autocovariance function is

$$\gamma_k = \sigma_a^2 \sum_{j=0}^{\alpha} \phi_1^j \phi_1^{j+k}$$

$$= \sigma_a^2 \phi_1^k \sum_{j=0}^{\alpha} \phi_1^{2j}$$

$$= \frac{\sigma_a^2\ \phi_1^k}{1-\phi_1^2}$$

and acf of the AR(1) process becomes

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi_1^k\ ,\ k = 0, 1, 2, \ldots$$

Using the Yule-Walker equation, the autocorrelation function satisfies the first-order difference equation

$$\rho_k = \phi_1 \rho_{k-1} \quad \ldots \quad k>0$$

which with $\rho_0 = 1$ and $\rho_1 = \rho_{-1}$

$$\rho_1 = \phi_1 \rho_0$$

$$\rho_1 = \phi_1$$

$$\rho_2 = \phi_1 \rho_1 = \phi_1^2$$

$$\rho_k = \phi_1^k \qquad k \geq 0$$

When ϕ_1 is positive, the autocorrelation function decays exponentially to zero and, ϕ_1 is negative, it oscillates in sign

The variance of the process is

$$\sigma_x^2 = \frac{\sigma_a^2}{1-\phi_1\rho_1} = \frac{\sigma_a^2}{1-\phi_1^2}$$

The acvf and acf of the AR(2) process is found as

$$\tilde{x}_t = \phi_1 \tilde{x}_{t-1} + \phi_2 \tilde{x}_{t-2} + a_t$$

$$(1-\phi_1 B - \phi_2 B^2)\ \tilde{x}_t = a_t$$

Let G_1^{-1} and G_2^{-1} be the root of the characteristic equation $(1-\phi_1 B - \phi_2 B^2) = 0$. Then

$$(1-G_1 B)(1-G_2 B) = (1-\phi_1 B - \phi_2 B^2)$$

$$\tilde{x}_t = \frac{1}{B(G_1 - G_2)}\left[(1-G_1B)^{-1} - (1-G_2B)^{-1}\right] a_t$$

$$= \sum_{j=0}^{\alpha} \frac{G_1^{j+1} - G_2^{j+1}}{G_1 - G_2} B^j a_t$$

$$= \sum_{j=0}^{\alpha} \frac{G_1^{j+1} - G_2^{j+1}}{G_1 - G_2} a_{t-j}$$

Thus, it is a linear process with $h_j = \dfrac{G_1^{j+1} - G_2^{j+1}}{G_1 - G_2}$. The acvf of the AR(2) process becomes

$$\gamma_k = \sigma_a^2 \sum_{j=0}^{\alpha} h_j h_{j+k}$$

$$= \frac{\sigma_a^2}{G_1 - G_2} \left[\frac{G_1^{k+1}\left(1-G_2^2\right) - G_2^{k+1}\left(1-G_1^2\right)}{\left(1-G_1^2\right)\left(1-G_2^2\right)\left(1-G_1 G_2\right)} \right]$$

$$\gamma_0 = \frac{\sigma_a^2}{G_1 - G_2} \left[\frac{G_1\left(1-G_2^2\right) - G_2\left(1-G_1^2\right)}{\left(1-G_1^2\right)\left(1-G_2^2\right)\left(1-G_1 G_2\right)} \right]$$

The acf of the AR(2) process is

$$\rho_k = \frac{G_1^{k+1}\left(1-G_2^2\right) - G_2^{k+1}\left(1-G_1^2\right)}{G_1\left(1-G_2^2\right) - G_2\left(1-G_1^2\right)}$$

$$= \frac{G_1\left(1-G_2^2\right)G_1^k - G_2\left(1-G_1^2\right)G_2^k}{G_1\left(1-G_2^2\right) - G_2\left(1-G_1^2\right)}$$

In this process, G_1^{-1} and G_2^{-1} are the roots of the characteristic equation $\phi(B) = 0$. When the roots are real, the autocorrelation function consists of a mixture of damped exponentials. It occurs when $\phi_1^2 + 4\phi_2 \geq 0$. If the roots are complex ($\phi_1^2 + 4\phi_2 < 0$) the autocorrelation function consists of a damped sine waves.

On the other hand, substituting p = 2 in Yuke-Walker equation,

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} \qquad (3.2.12)$$

and for k = 1, 2, ...

$$\rho_1 = \phi_1 \rho_0 + \phi_2 \rho_1$$

$$\rho_2 = \phi_1 \rho_1 + \phi_2 \rho_0$$

$$\rho_3 = \phi_1 \rho_2 + \phi_2 \rho_1 \quad \text{and etc.}$$

Thus the acf can be obtained by solving (3.2.12) recursively, beginning with the initial values $\rho_0 = 1$ and

$\rho_1 = \dfrac{\phi_1}{1-\phi_2}$ when ϕ_1 and ϕ_2 are given.

For AR(1) model,

$$\sigma^2 = \frac{\sigma_a^{\,2}}{(1-\phi_1^{\,2})} \quad \text{since} \quad \rho_1 = \phi_1$$

Similarly, for AR(2) model

$$\sigma^2 = \frac{\sigma_a^{\,2}(1-\phi_2)}{(1+\phi_2)\,[(1-\phi_2)^2 - \phi_1^{\,2}]}$$

since $\rho_1 = \dfrac{\phi_1}{1-\phi_2}$ and $\rho_2 = \phi_2 + \dfrac{\phi_1^{\,2}}{(1-\phi_2)}$ are obtained from two equations of (3.2.12).

3.2.2 The Partial Autocorrelation Function

The partial autocorrelation function is useful to decide the order of autoregressive process to fit to an observed time series. It is described in terms of p non-zero function of the autocorrelation.

To decide the order of autoregressive process to be fitted to an observed time series is the same as deciding on the number of independent variables to be included in a multiple regression. Denote by ϕ_{kj} .,the j^{th} coefficient in an autoregressive process of order k, so that the partial correlation is ϕ_{kk} which is the last coefficient of the K^{th} order AR process.

From Yule-Walker equation, the ϕ_{kj} satisfy the set of equations.

$$\rho_j = \phi_{k1}\rho_{j-1} + \ldots + \phi_{kk}\rho_{j-k}, \quad j = 1, 2, \ldots, k$$

which may be written as

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 \cdots\cdots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 \cdots\cdots & \rho_{k-2} \\ \cdot & \cdot & \cdots\cdots\cdots & \cdot \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \cdot \\ \phi_{kk} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \cdot \\ \rho_k \end{bmatrix} \qquad (3.2.13)$$

or

$$P_k \phi_k = p_k$$

solving these equations for k=1, 2, 3,....., successively, we obtained $\phi_{11} = \rho_1$

$$\phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

$$\phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} \qquad \text{and so on}$$

In general, for ϕ_{kk}, the determinant in the numerator has the same elements as that in the denominator, but with the last column replaced by $[\rho_1, \rho_2, \ldots, \rho_k]'$.

On the otherhand the partial autocorrelation function can be obtained by solving Durbin's recursive formulae [Box-Jenkin,1976].

$$\phi_{p+1,j} = \phi_{pj} - \phi_{p+1,p+1}\ \phi_{p,p-j+1}\ ,\ j = 1,\ 2,\ldots,p$$

$$\phi_{p+1,p+1} = \frac{\rho_{p+1} - \sum_{j=1}^{p} \phi_{pj}\rho_{p+1-j}}{1 - \sum_{j=1}^{p} \phi_{pj}\rho_j} \qquad 3.2.14$$

with starting value $\phi_{11} = \rho_1$

The estimated partial autocorrelation can be obtined by substitution estimates r_j for the theoretical autocorrelations.

$$r_j = \hat{\phi}_{k1}\ r_{j-1} + \hat{\phi}_{k2}\ r_{j-2} + \ldots + \hat{\phi}_{k,k-1}\ r_{j-k-1} + \hat{\phi}_{kk}\ r_{j-k}\ \ j=1,2,\ldots,k.$$

and solving the resultant equations for k=1,2,.... . Under the assumption that the process is autoregressive of order p, the estimated partial autocorrelations of order p+1 and higher are approximately independently distributed. [Box-Jenkin,1976]

Also if n is the number of observations used in fitting,

$$\mathrm{Var}[\hat{\phi}_{kk}] \approx 1/n;\ k=p+1,\ p+2,\ldots$$

Thus the standard error of the estimated partial autocorrelation $\hat{\phi}_{kk}$ is

$$\mathrm{SE}[\hat{\phi}_{kk}] \approx 1/\sqrt{n}\ \ ;\ k=p+1,\ p+2,\ldots \qquad (3.2.15)$$

The partial autocorrelation of AR(1) process are

$\phi_1 = \rho_1 = \phi_{11}$

and the partial autocorrelations of order higher than one are, from (3.2.14),

$$\phi_{22} = \frac{\rho_2 - \phi_{11}\rho_1}{1 - \phi_{11}\rho_1} = 0,$$

since $\rho_k = \phi_1^k = \rho_1^k$

similarly, $\phi_{21} = \phi_{11} - \phi_{22}\,\phi_{11} = \rho_1$

$$\phi_{33} = \frac{\rho_3 - \phi_{11}\rho_2 - \phi_{22}\rho_1}{1 - \phi_{11}\rho_1 - \phi_{22}\rho_2} = 0$$

since $\rho_3 = \rho_1^3$, $\phi_{11}\rho_2 = \rho_1^3$ and $\phi_{22}\,\rho_1 = 0$

Generally, for the AR(1) process,

$\phi_{kk} = 0$, $k = 2, 3, 4, \ldots$ can be obtained by calculating recursively.

For the AR(2) process, the partial autocorrelations are obtained as,

$\phi_{11} = \rho_1$

$$\phi_{22} = \frac{\rho_2 - \phi_{11}\rho_1}{1 - \phi_{11}\rho_1} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} = \phi_2$$

and $\phi_{21} = \phi_{11} - \phi_{22}\,\phi_{11}$

$$= \rho_1 - \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}\rho_1 = \frac{\rho_1(1 - \rho_2)}{1 - \rho_1^2}$$

then $\phi_{33} = \dfrac{\rho_3 - \phi_{21}\rho_2 - \phi_{22}\rho_1}{1 - \phi_{21}\rho_1 - \phi_{22}\rho_2} = 0$

by substituting $\rho_3 = \dfrac{\rho_1}{1-\rho_1^2}[2\rho_2 - \rho_1^2 - \rho_2^2]$

The expression for ρ_3 can be obtained by solving the Yule-Walker equation for p=2, i.e

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} \quad , \; k=1,\ 2,\ 3, \ldots$$

with $\phi_1 = \dfrac{\rho_1(1-\rho_2)}{1-\rho_1^2}$ and $\phi_2 = \dfrac{\rho_2 - \rho_1^2}{1-\rho_1^2}$

It can be easily shown that $\phi_{kk} = 0$, k=3, 4, 5,... by the same way. Therefore, for AR(2) process,

$$\phi_{11} = \rho_1 = \frac{\phi_1}{1-\phi_2}$$

$$\phi_{22} = \frac{\rho_2 - \rho_1^2}{1-\rho_1^2} = \phi_2$$

and $\phi_{kk} = 0$; k = 3, 4, 5,...

Generally, it can be concluded that, for an AR(p) process, the partial autocorrelation function ϕ_{kk} will be nonzero for k = 1, 2, 3, ...,p and zero for k = p+1, p+2,.. In other words the partial autocorrelation function of a p^{th} order autoregressive process has a cut off after lap p. This can be used to determine the order of an AR process. By computing the sample partial autocorrelation function of the given series and if the value of ϕ_{kk} is within the interval $(2/\sqrt{n}, -2/\sqrt{n})$, after lag(m), the underlying process can be taken as an AR(m).

3.2.3 Stationary Conditions for Autoregressive Process

The set of adjustable parameters $\phi_1, \phi_2, \ldots \phi_p$ of an AR(p) process

$$\tilde{x}_t = \phi_1 \tilde{x}_{t-1} + \ldots + \phi_p \tilde{x}_{t-p} + a_t$$

$[1-\phi_1 B \ldots \phi_p B_p^{\ p}]\ \tilde{x}_t = a_t$ must satisfy certain conditions for the process to be stationary.

The first order autoregressive process is

$$(1-\phi_1 B)\tilde{x}_t = a_t$$

$$\tilde{x}_t = (1-\phi_1 B)^{-1} a_t = \sum_{j=0}^{\alpha} \phi_1^j a_{t-j}$$

Hence $\psi(B) = (1-\phi_1 B)^{-1} = \sum_{j=0}^{\alpha} \phi_1^j B^j$

For stationarity, $\psi(B)$ must converge for $|B| \leq 1$. So, the parameter ϕ_1, of an AR(1) process, must satisfy the condition $|\phi_1| < 1$ to ensure stationarity. And the same way the root of $1-\phi_1 B = 0$ must lie outside the unit circle.

For the general AR(p) process

$$\tilde{x}_t = \phi_1^{-1}(B)\ a_t$$

$$\phi(B) = (1-G_1 B)(1-G_2 B)\ldots(1-G_p B)$$

and expanding in partial fractions

$$\tilde{x}_t = \phi^{-1}(B) a_t = \sum_{i=0}^{p} \frac{K_i}{1-G_i B} a_t$$

$$\psi(B) = \phi_1^{-1}(B)$$

If $\psi(B)$ converge for $|B| \leq 1$, then the root G_i must satisfy the condition $|G_i| < 1$ where $i=1, 2, \ldots, p$. Equivalently, the root of $\phi(B) = 0$ must lie outside the unit circle. [Box-jenkin, 1976]

3.2.4 The Spectrum of the AR(p) Process

The power spectrum of the AR(p) process is also useful as acvf and acf to analyze a stationary stochastic process. It shows how the variance of the stochastic process varies with frequency. It is the Fourier transform of the acvf and hence, acvf and acf belongs to the time domain study where as the power spectrum belongs to the frequency domain study of the process. Since the spectrum is the Fourier transform of the acvf of the process, one half of the spectrum of the process can be obtained by substituting $B = e^{-i2\pi f}$,where $i=\sqrt{-1}$ in the autocovariance generating function $r(B)=\sum_{-\infty}^{\infty} r_k B^k$ [Box-Jenkin, 1976] The autocovariancee generating function of a linear process is

$$r(B) = \sigma_a^2 \, \phi^{-1}(B)\phi^{-1}(f)$$

and the power spectrum of the linear process is

$$X(f) = 2\sigma_a^2 \, \phi^{-1}(e^{-i2\pi f})\phi^{-1}(e^{i2\pi f}) \qquad (3.2.16)$$

$$= 2\sigma_a^2 \, |\phi(e^{-i2\pi f})|^{-2} \; ; \; 0 \le f \le 1/2$$

$$= \frac{2\sigma_a^2}{\left|\phi(e^{-i2\pi f})\right|^2} ; \quad 0 \le f \le 1/2$$

For the AR(p) process

$$\phi(B) = (1 - \phi_1 B - \phi_2 B^2 \ldots - \phi_p B^p)$$

and the power spectrum becomes

$$X(f) = \frac{2\sigma_e^2}{\left|1 - \phi_1 e^{-i2\pi f} - \phi_2 e^{-i4\pi f} - ... \phi_p e^{-i2\pi p f}\right|^2}, 0 \le f \le 1/2 \quad (3.2.17)$$

Fir the AR(1) process

$$\phi(B) = (1 - \phi_1 B)$$

$$X(f) = \frac{2\sigma_e^2}{\left|1 - \phi_1 e^{-i2\pi f}\right|^2};$$

$$= \frac{2\sigma_e^2}{1 + \phi_1^2 - 2\phi_1 \text{Cos} 2\pi f}; 0 \le f \le 1/2 \quad (3.2.18)$$

Similarly, for AR(2) process

$$\phi(B) = (1 - \phi_1 B - \phi_2 B^2)$$

and its powerspectrum becomes.

$$X(f) = \frac{2\sigma_e^2}{\left|1 - \phi_1 e^{-i2\pi f} - \phi_2 e^{-i4\pi f}\right|^2}$$

$$= \frac{2\sigma_e^2}{1 + \phi_1^2 + \phi_2^2 - 2\phi_1(1 - \phi_2)\text{Cos} 2\pi f - 2\phi_2 \text{Cos} 4\pi f}, 0 \le f \le 1/2 \quad (3.2.19)$$

3.3 Estimation of the Parameters

After the model identification, the efficient estimates of the parameters are needed. Thus the parameters are needed to estimate in a model.

The identification process having lead to a tentative formulation for the model, then it need to obtain efficient estimates of the parameters. After the parameters have been estimated, the fitted model will be subjected to diagonstic checks and tests of goodness of fit. For testing of goodness of fit to be relevant, it is necessary that efficient use of the data should have been made in fitting process. If this is not so, inadequacy of fit may simply arise because of the inefficient fitting and not because the form of the model is inadequate.

Therefore, the estimation of the parameters is important in the fitting process. Among the estimation methods for the estimation of the parameters of the autogressive process, the method of moments and the maximum likelihood methods are widely used.

The method of moment is the simplest method and is based on the Yule-Walker equations. The moment estimates of the autoregressive parameters are sometimes called as Yule-Walker estimates. The maximum likelihood method is the most powerful and relevant method when the underlying assumptions are fulfilled.

Let the given set of data be $X_1, X_2, \ldots, X_n$ where n is the number of observations and also let the underlying process of X be the p^{th} order autoregressive process. Denote μ be means of X, σ^2 be the variance of X and ρ_j be the lag "j" autocorrelation coefficient.

3.3.1 Method of Moments

The method of moments rests on the fact that the moments of the sample and its parent population may have the same properties. In this method, the population moments are equated by their sample counterparts and the equations obtained are solved to get the estimates.

The moment estimate of the population mean is

$$\hat{\mu} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_t \tag{3.3.1}$$

and the sample mean is used to estimate the population mean. Similarly, the population variance or the variance of the process is estimated by

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_t - \bar{x})^2 \tag{3.3.2}$$

but this estimator is a biased estimator and the commonly used one is an unbiased estimator.

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_t - \bar{x})^2 \tag{3.3.3}$$

To estimate the autoregressive coefficients, the sample autocorrelations r_s are substituted in place of the theoretical autocorrelation ρ_s and $\hat{\phi}_s$ are replaced in $\phi_,$ in the Yule-Walker equations. Then

$$r_k = \hat{\phi}_1 r_{k-1} + \hat{\phi}_2 r_{k-2} + \ldots + \hat{\phi}_p r_{k-p}, \qquad k=1,\ 2,\ 3, \ldots, p$$

and by setting k = 1, 2, ..., p a system of equations is obtained as

$$\begin{aligned} r_1 &= \hat{\phi}_1 + \hat{\phi}_2 r_1 + \ldots + \hat{\phi}_p r_{p-1} \\ r_2 &= \hat{\phi}_1 r_1 + \hat{\phi}_2 + \ldots + \hat{\phi}_p r_{p-2} \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \qquad (3.3.4) \\ r_p &= \hat{\phi}_1 r_{p-1} + \hat{\phi}_2 r_{p-2} + \ldots + \hat{\phi}_p \end{aligned}$$

or $\underline{\hat{\phi}} = \underline{R}\underline{\hat{\phi}}$, where $r' = [r_1, r_2, \ldots, r_p]$,

$\underline{\hat{\phi}}' = [\hat{\phi}_1, \ldots, \hat{\phi}_p]$ and

$$\underline{R} = \begin{bmatrix} 1 & r_1 & . & r_p \\ r_1 & 1 & . & r_{p-2} \\ . & . & . & . \\ r_{p-1} & r_{p-2} & . & 1 \end{bmatrix}$$

By solving these equations simultaneously, the moment estimates or the Yule-Walker estimates of the parameters are obtained easily as

$$\underline{\hat{\phi}} = \underline{R}^{-1}\underline{r}$$

The error variance σ_a^2 is estimated from the relation of σ^2 and σ_a^2, i.e

$$\sigma_a^2 = \sigma^2 \left(1 - \sum_{j=1}^{p} \phi_j \rho_j\right)$$

replacing $\hat{\phi}_j$ in ϕ_j and r_j in ρ_j. In order to get an unbiased estimate of error variance [Salas, 1980], it is necessary to multiply with n/(n-p), then $\hat{\sigma}_a^2$ becomes,

$$\hat{\sigma}_a^2 = \frac{n}{n-p}\hat{\sigma}^2\left(1-\sum_{j=1}^{p}\hat{\phi}_j r_j\right) \qquad (3.3.5)$$

For a given set of data $\hat{\mu}$ and $\hat{\sigma}^2$ are constant for whatever the order of process, but $\hat{\phi}_1$, $\hat{\phi}_2$, ..., $\hat{\phi}_p$ and $\hat{\sigma}_a^2$ aare changed according to the order of autoregression. For example, the estimates of the parameters of AR(1) model are obtained from equations (3.3.4) and (3.3.5) as

$$r_1 = \hat{\phi}_1$$

and

$$\hat{\sigma}_a^2 = \frac{n}{n-1}\hat{\sigma}^2\,(1-\hat{\phi}_1 r_1)$$

$$= \frac{n}{n-1}\hat{\sigma}^2\,(1-\hat{\phi}_1^2)$$

For AR(2) process, equation (3.3.5) becomes

$$r_1 = \hat{\phi}_1 - \hat{\phi}_2\, r_1$$

$$r_2 = \hat{\phi}_1\, r_1 - \hat{\phi}_2$$

and $\hat{\phi}_1 = \dfrac{r_1(1-r_2)}{1-r_1^2}$, $\quad \hat{\phi}_2 = \dfrac{r_2-r_1^2}{1-r_1^2}$

Then the estimate of error variance is

$$\hat{\sigma}_a^2 = \frac{n}{n-2}\hat{\sigma}^2\left(1-\sum_{j=1}^{2}\hat{\phi}_j r_j\right)$$

or $$\hat{\sigma}_a^2 = \frac{n}{n-2}\hat{\sigma}^2 \frac{1+\hat{\phi}_2}{1-\hat{\phi}_2}\left\{(1-\hat{\phi}_2)^2 - \hat{\phi}_1^2\right\}$$

replacing $r_1 = \dfrac{\hat{\phi}_1}{1-\hat{\phi}_2}$ and $r_2 = \dfrac{\hat{\phi}_1^2}{1-\hat{\phi}_2} + \hat{\phi}_2$

3.3.2 Maximum Likelihood method of Estimation

To use the maximum likelihood method of estimation, it is important to know the exact distribution of the variable under study. But in practice, it is impossible to know the exact distribution and it needs to make assumptions about the distribution. Since the autoregressive process of order p can be written as

$$x_t - \mu = \phi_1(x_{t-1} - \mu) + \phi_2(x_{t-2} - \mu) + \ldots + \phi_p(x_{t-p} - \mu) + a_t$$

or

$$w_t = \phi_1 w_{t-1} + \phi_2 w_{t-2} + \ldots + \phi_p w_{t-p} + a_t$$

where w_t is the mean deviation of x_t and μ is estimated by $\bar{x}$ if it is unknown.

Let us assume a_t is independently and normally distributed with mean zero and variance $\hat{\sigma}_a^2$ and hence w is normally distributed random variable with mean zero. Then, the joint probability density function of the

$\underline{w}_n' = (w_1, w_2, \ldots w_n)$ is

$$f(\underline{w}_n' / \phi, \hat{\sigma}_a^2) = (2\pi\sigma_a^2)^{-\frac{n}{2}} |\underline{M}|_n^{\frac{1}{2}} \exp-\left\{\frac{\underline{w}_n' \underline{M}_n \underline{w}_n}{2\sigma_a^2}\right\}$$

$$-\infty \le w_t \le \infty \; ; \; t = 1, 2, \ldots, n \qquad (3.3.6)$$

where $\hat{\underline{\phi}}' = (\phi_1, \phi_2, \ldots \phi_p)$ is the vector of autoregressive parameters.

$\hat{\sigma}_a^2 = V(a_t) = E(a_t^2)$ is the variance of random error term and $\hat{\sigma}_a^2 \underline{M}_n^{-1}$ is the variance-covariance matrix of. $\underline{w}_n$.

The variance-covariance matrix of $\underline{w}_n$ can be expressed as $\hat{\sigma}_a^2 \underline{M}_n^{-1}$ since all the variances and covariances of w_n can be expressed as the product of $\hat{\sigma}_a^2$ and autocovariances.

i.e
$$\sigma_a^2 \underline{M}_n^{-1} = \sigma_a^2 \begin{bmatrix} E(w_1w_1) & E(w_1w_2) & . & E(w_1w_n) \\ E(w_2w_1) & E(w_2w_2) & . & E(w_2w_n) \\ . & . & . & . \\ E(w_nw_1) & E(w_nw_2) & . & E(w_nw_n) \end{bmatrix}$$

$$= \sigma_a^2 \begin{bmatrix} r_0 & r_1 & . & r_{n-1} \\ r_1 & r_0 & . & r_{n-2} \\ . & . & . & . \\ r_{n-1} & r_{n-2} & . & r_0 \end{bmatrix}$$

where $\underline{r}_j = r_j$ is the lag j autocovariance of AR(p) processs. But actually $r_j = 0$ for $j \geq p$. The matrix $\underline{M}_n$ is symmetric about both of its principal diagonals or a doubly symmetric matrix(say).

Now,

$$f(\underline{w}_n' / \phi, \hat{\sigma}_a^2) = f(w_{p+1}, w_{p+2}, \ldots, w_n / \underline{w}_p, \phi, \hat{\sigma}_a^2) f(\underline{w}_p / \phi, \hat{\sigma}_a^2)$$

where $\underline{w}_p' = (w_1, w_2, \ldots w_p)$

To get the conditional distribution of

$\underline{w}_{n-p}' = (w_{p+1}, \ldots w_n)$ given $(w_1, w_2, \ldots w_p)$, the distribution of $(a_{p+1}, \ldots, a_n)$ can be used, i.e,

$$f(a_{p+1}, \ldots, a_n) = (2\pi\sigma_a^2)^{-\frac{n-p}{2}} \exp -\left\{ \frac{1}{2\sigma_a^2} \sum_{t=p+1}^{n} a_t^2 \right\}$$

$$-\infty \le a_t \le \infty \; ; \; t = p+1, \ldots, n \qquad (3.3.7)$$

Since a_t are independently and normally distributed with zero mean and variance $\hat{\sigma}_a^2$.

For fixed $\underline{w}_p$, $(a_{p+1}, a_{p+2}, \ldots, a_n)$ and $(w_{p+1}, w_{p+2}, \ldots, w_n)$ are related by the transformation.

$$a_{p+1} = w_{p+1} - \phi_1 w_p - \ldots\ldots - \phi_p w_1$$

$$a_{p+2} = w_{p+2} - \phi_1 w_{p+1} - \ldots\ldots - \phi_p w_2$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$a_n = w_n - \phi_1 w_{n-1} - \ldots\ldots - \phi_p w_{n-p}$$

The Jacobian of the transformation is

$$|J| = \left| \frac{\partial \underline{a}_{n-p}}{\partial \underline{w}_{n-p}} \right|$$

and $|J|$ is obviously unity. Then,

$$f(\underline{w}_{n-p}/\underline{w}_p, \underline{\phi}, \hat{\sigma}_a^2) = (2\pi\sigma_a^2)^{-\frac{n-p}{2}} \exp\left\{ -\frac{1}{2\sigma_a^2} \sum_{t=p+1}^{n} (w_t - \phi_1 w_{t-1} \ldots - \phi_p w_{t-p})^2 \right\}$$

$$-\infty \le w_t \le \infty \; ; \qquad (3.3.8)$$

Also, the distribution of $\underline{w}_p$ is

$$f(\underline{w}_p / \underline{\phi}, \hat{\sigma}_a^2) = (2\pi\sigma_a^2)^{-\frac{p}{2}} |\underline{M}_p|^{\frac{1}{2}} \exp\left\{ -\frac{\underline{w}_p' \underline{M}_p w_p}{2\sigma_a^2} \right\}; -\infty \leq w_t \leq \infty ;$$

where $\sigma_a^2 \ \underline{M}_p^{-1}$ is the variance-covariance matrix of $\underline{w}_p$.

Thus

$$f(\underline{w}_n / \underline{\phi}, \hat{\sigma}_a^2) = (2\pi\sigma_a^2)^{-\frac{n}{2}} |\underline{M}_p|^{\frac{1}{2}} \exp\left\{ -\frac{S(\phi)}{2\sigma_a^2} \right\} \qquad (3.3.9)$$

where

$$S(\phi) = \underline{w}_p' \underline{M}_p \underline{w}_p + \sum_{t=p+1}^{n} (w_t - \phi_1 w_{t-1} \cdots - \phi_p w_{t-p})^2$$

$$= \sum_{i=1}^{p} \sum_{j=1}^{p} w_i m_{ij} w_j + \sum_{t=p+1}^{n} (w_t - \phi_1 w_{t-1} \cdots - \phi_p w_{t-p})^2 \qquad (3.3.10)$$

where m_{ij} is the i^{th} row and j^{th} column element of $\underline{M}_p$

Also $\underline{M}_p = \{m_{ij}\} = \hat{\sigma}_a^2 \{\gamma_{|i-j|}\}^{-1} =$

$$= \sigma_a^2 \begin{bmatrix} r_0 & r_1 & \cdot & r_{p-1} \\ r_1 & r_0 & \cdot & r_{p-2} \\ \cdot & \cdot & \cdot & \cdot \\ r_{p-1} & r_{p-2} & \cdot & r_0 \end{bmatrix}$$

From the equation (3.3.6) and (3.3.9), it can be seen that $|\underline{M}_n| = |\underline{M}_p|$.

Now, Let n = p+1. so that from equation(3.3.10)

$$\underline{w}_{p+1}'\,\underline{M}_{p+1}\,\underline{w}_{p+1} = \sum_{i=1}^{p}\sum_{j=1}^{p} w_i m_{ij} w_j + (w_{p+1} - \phi_1 w_p \cdots \phi_p w_1)^2$$

$$= \sum_{i=1}^{p}\sum_{j=1}^{p} w_i m_{ij} w_j + (\phi_p^2 w_1^2 + \phi_p\phi_{p-1} w_1 w_2 + \ldots + \phi_p\phi_1 w_1 w_{p-}$$

$$-\phi_p w_1 w_{p+1} + \phi_p\phi_{p-1} w_1 w_2 + \phi_{p-1}^2 w_2^2 +$$

$$\phi_{p-1}\phi_1 w_2 w_p - \phi_{p-1} w_2^2 + \phi_{p-1} w_2 w_{p+1} + \ldots$$

$$\phi_p\phi_1 w_1 w_p + \phi_{p-1}\phi_1 w_1 w_p + \phi_1^2 w_p^2 -$$

$$\phi_1 w_p w_{p+1} - \phi_p w_1 w_{p+1} - \phi_{p-1} w_2 w_{p+1}$$

$$\ldots - w_{p+1}^2)$$

$$= [w_1\ w_2\ \ldots\ w_{p+1}] \left[\begin{matrix} \cdot & \cdot & \cdot & 0 \\ \cdot & \underline{M}_p & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & 0 \end{matrix}\right] + \left[\begin{matrix} \phi_p^2 & \phi_p\phi_{p-1} & \cdot & -\phi_p \\ \phi_p\phi_{p-1} & \phi_{p-1}^2 & \cdot & -\phi_{p-1} \\ \cdot & \cdot & \cdot & \cdot \\ -\phi_p & -\phi_{p-1} & \cdot & 1 \end{matrix}\right] \left[\begin{matrix} w_1 \\ w_2 \\ \cdot \\ w_{p+1} \end{matrix}\right]$$

Then the element of $\underline{M}_p$ can be deduced from the consideration that both $\underline{M}_p$ and $\underline{M}_{p+1}$ are doubly symmetric. For example, For AR(1) process,

$$\underline{M}_2 = \begin{bmatrix} m_{11} + \phi_1^2 & -\phi_1 \\ -\phi_1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -\phi_1 \\ -\phi_1 & m_{11} + \phi_1^2 \end{bmatrix}$$

and after equating elements in the two matrices,

$$m_{11} + \phi_1^2 = 1 \quad \text{or} \quad m_{11} = \underline{M}_1 = 1 - \phi_1^2 = |\underline{M}_1|$$

Preceeding in this way, the elements of $\underline{M}_2$ of AR(2) process are, from $\underline{M}_3$ as

$$\underline{M}_3 = \begin{bmatrix} m_{11}+\phi_2^2 & m_{12}+\phi_1\phi_2 & -\phi_2 \\ m_{12}+\phi_1\phi_2 & m_{22}+\phi_1^2 & -\phi_1 \\ -\phi_2 & -\phi_1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -\phi_1 & -\phi_2 \\ -\phi_1 & m_{22}+\phi_1^2 & m_{12}+\phi_1\phi_2 \\ -\phi_2 & m_{12}+\phi_1\phi_2 & m_{11}+\phi_2^2 \end{bmatrix}$$

$m_{11} + \phi_2^2 = 1$ or $m_{11} = 1 - \phi_2^2$

$m_{12} + \phi_1\phi_2 = -\phi_1$ or $m_{12} = -\phi_1(1 + \phi_2)$

and

$$\underline{M}_2 = \begin{bmatrix} 1-\phi_2^2 & -\phi_1(1+\phi_2) \\ -\phi_1(1+\phi_2) & 1-\phi_2^2 \end{bmatrix}$$

Since $m_{11} = m_{22}$, the matrix $\underline{M}_2$ is doubly symmetric.

Therefore, $|\underline{M}_2| = (1 - \phi_2^2)^4 - \phi_1^2(1+\phi_2)^2$

$$= (1 + \phi_2)^2[(1 - \phi_2)^2 - \phi_1^2]$$

Thus, it is clear that the elements of $S(\phi) = \underline{w}_n'\underline{M}_n\underline{w}_n$ are not only a quadratic form in the w's, but also a quadratic form in the parameters ϕ_s. Writing

$\underline{\phi}_u' = (1, \phi_1, \ldots, \phi_p)$, the (P+1)x(p+1) matrix $\underline{D}$ can be found, whose elements are quadratic forms of w's.

$$\underline{w}_n'\underline{M}_n\underline{w}_n = \underline{\phi}_u'\underline{D}\,\underline{\phi}_u$$

and write

$$\underline{D} = \begin{bmatrix} D_{11} & -D_{12} & -D_{13}\ldots & \ldots -D_{1,p+1} \\ -D_{12} & D_{22} & D_{23}\ldots & \ldots D_{2,p+1} \\ . & . & . & . \\ -D_{1,p+1} & D_{2,p+2} & D_{3,p+1} & D_{p+1,p+1} \end{bmatrix} \quad (3.3.11)$$

From equation (3.3.10) it can be seen that the elements of $\underline{D}$, say D_{ij} are symmetric sums of squares and lagged products, defined by

$$D_{ij} = D_{ji} = W_i W_j + W_{i+1} W_{j+1} + \ldots + W_{n+1-j} W_{n+1-i} \qquad (3.3.12)$$

where the sum D_{ij} contains n-(i-1)-(j-1) terms.

Then, the likelihood function or the joint probability function of $\underline{w}_n$ is

$$f(\underline{w}_n / \underline{\phi}, \hat{\sigma}_a^2) = L(\underline{\phi}, \hat{\sigma}_a^2 / \underline{w}_n)$$

$$= (2\pi\sigma_a^2)^{-\frac{n}{2}} |\underline{M}_p|^{\frac{1}{2}} \exp\left\{-\frac{S(\phi)}{2\sigma_a^2}\right\}$$

where

$$S(\phi) = \underline{w}_p' \underline{M}_p \underline{w}_p + \sum_{t=p+1}^{n} (w_t - \phi_1 w_{t-1} \cdots \phi_p w_{t-p})^2$$

$$= \underline{\phi}_u' \underline{D} \underline{\phi}_u \qquad (3.3.13)$$

and the log likelihood is

$$L(\underline{\phi}, \sigma_a^2 / \underline{w}_n) = -(n/2) \mathrm{Ln}\sigma_a^2 + (1/2) \ln |\underline{M}_p| - S(\phi)/(2\sigma_a^2) \quad (3.3.14)$$

To get the maximum likelihood estimates of the parameters. the log likelihood function must be differentiated by the parameters σ_a^2 , $\phi_1, \ldots,$ ϕ_p and equating with zero, the equations must be solved simultaneously.

That is

$$\frac{\partial L}{\partial \sigma_a^2} = -\frac{n}{\sigma_a^2} + \frac{S(\underline{\phi})}{\left[\sigma_a^2\right]^2} = 0 \qquad (3.3.15)$$

$$\frac{\partial L}{\partial \phi_j} = M_j + \sigma_a^{-2}[D_{1,j+1} - \phi_1 D_{2,J+1} - \dots - \phi_p D_{p+1,j+1}]$$

$$= 0 \quad , \; j = 1, 2, \dots, p \qquad (3.3.16)$$

where $M_j = \dfrac{\partial\left\{\frac{1}{2}\operatorname{Ln}\left|\underline{M}_p\right|\right\}}{\partial \phi_j}$

By solving (3.3.15) and (3.3.16)

$\sigma_a^2 = S(\phi)/n$ can be obtained easily, but it is not easy to find the estimates of ϕ_s since the M_j' are the complicated functions of the ϕ_s'.

Thus three alternative procedures may be used rather than using the exact likelihood function.

(1) Least-square Estimates

Whereas the expected value of $s(\phi)$ is proportion to n, the value of $|M_p|$ is independent of n and for moderate to large samples , (3.3.12) is dominated by the term in $s(\phi)$ and the term in $|M_p|$ is comparatively small. Thus ,by ignoring that term , the log likelihood becomes,

$$L(\phi, \sigma_a^2 / W_n) \simeq -n/2 \operatorname{Ln} \sigma_a^2 - S(\phi)/2\sigma_a^2 \qquad (3.3.17)$$

and the estimates $\hat{\phi}$ of ϕ obtained by maximization of (3.3.17) are the least-square estimates obtained by minimizing $S(\phi)$. The normal equations obtained by minimizing $S(\phi)$ are the same as second term of equation (3.3.16). That is ,

$$\sigma_a^2[\ D_{i\ ,j+i} - \phi_1 D_{2,j+i} - \ldots\ldots - \phi_p D_{p+1,J+1}] = 0 \qquad j=1,2, \quad ,p$$

or $\quad D_{1,j+1} = \phi_1 D_{2,j+1} + \phi_2 D_{3,j+1} + \ldots\ldots + \phi_p D_{p+1,j+1} \qquad j=1,2, \quad ,p$

substituting $j=1,2, \quad ,p$ in the above equations **gives**

$$D_{12} = \hat{\phi}_1 D_{22} + \hat{\phi}_2 D_{23} + \ldots\ldots + \hat{\phi}_p D_{2,p+1}$$
$$D_{13} = \hat{\phi}_1 D_{23} + \hat{\phi}_2 D_{33} + \ldots\ldots + \hat{\phi}_p D_{3,p+1} \qquad (3.3.18)$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$D_{1,p+1} = \hat{\phi}_1 D_{2,p+1} + \hat{\phi}_2 D_{3,p+1} + \qquad + \hat{\phi}_p D_{p+1,p+1}$$

The matrix notation

$$\underline{\hat{\phi}} = \underline{D}^{-1}\underline{d} \qquad \text{where} \quad \underline{d} = [D_{12}, D_{13}, \qquad D_{1,p+1}]$$

(2) Approximate Maximum Likelihood Estimates

Since the autocovariance of lag j of AR(p) process can be expressed as,

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \ldots\ldots + \phi_p \gamma_{j-p}$$

$$\gamma_j - \phi_1 \gamma_{j-1} - \phi_2 \gamma_{j-2} - \ldots\ldots - \phi_p \gamma_{j-p} = 0\ ,\ j = 1,\ 2,\ 3,\ \ldots (3.3.19)$$

Since $\quad D_{ij} = \sum_{i=0}^{n+1-(i+j)} w_{i+1} w_{j+1}$,

the expectation of D_{ij} becomes

$E(D_{ij}) = [n+2-(i+j)] \quad \gamma_{|i-j|}$

Taking expectation of (3.3.16) gives ,

$$\sigma_a^2 M_j + (n-j)\gamma_j - (n-j-1)\phi_1 \gamma_{j-1} - \ldots\ldots - (n-j-p) \phi_p \gamma_{j-p} = 0 \quad (3.3.20)$$

or

$$\sigma_a^2 M_j = -(n-j)\gamma_j + (n-j-1)\phi_1 \gamma_{j-1} + \ldots\ldots + (n-j-p) \phi_p \gamma_{j-p} \quad (3.3.20)$$

and adding n times the equation(3.3.19) gives

$$\sigma_a^2 M_j = j \gamma_j - (j+1)\phi_1 \gamma_{j-1} \ldots\ldots - (j+p) \phi_p \gamma_{j-p} = 0 \qquad (3.3.21)$$

Therefore, by using $\dfrac{D_{i+1,j+1}}{n-i-j}$ as an estimate of $\gamma_{|j-i|}$, a natural estimate of M_j is

$$\sigma_a^{-2}\left[\frac{j}{n-j}D_{1,j+1} - \frac{j+1}{n-(j+1)}\phi_1 D_{2,j+1\ldots\ldots} - \frac{j+p}{n-(j+p)}\phi_p D_{p+1,j+1}\right]$$

Substituting this estimate in (3.3.16) and solving the resulting equations give

$$n\sigma_a^2 \left[\frac{D_{1,j+1}}{n-j} - \phi_1 \frac{D_{2,j+1}}{n-(j+1)} - \ldots\ldots - \phi_p \frac{D_{p+1,j+1}}{n-(j+p)}\right] = 0; \quad ,j=1,\ 2,\ldots,p$$

Then a set of linear equations which is of the same form as (3.3.18) is obtained. Thus the least square estimate and maximum likelihood estimate(approximate) are the same except that D_{ij}^{*} is used instead of D_{ij} , where

$$D_{ij}^{*} = \frac{nD_{ij}}{n+2-i-j} = \frac{n}{n+2-i-j}\sum_{i=0}^{n+1-(i+j)} w_{i+1}w_{j+1} \qquad (3.3.22)$$

(3) Yule-Walker Estimates

If n is moderate to large, as an approximation the symmetric sum of squares and product in (3.3.18) can be replaced by

$$nC_k = \sum_{t=1}^{n-k} w_t w_{t+k}$$

on dividing by nC_0 throughout the resultant equation, we obtain the Yule-Walker equations expressed in terms of estimated autocorrelations $\gamma_k = (C_k/C_0)$.

For the parameter estimates obtained by the above three methods, the differences are small for moderate to large sample sizes. Box and Jenkins [Box-Jenkin, 1976] normally used the first approximation which uses the least-square estimates. However Salas, et . al [Box-Delleur-Yevjevich-Lane, 1980] used the second approximation. We choose the moment estimates since they can be computed without rigid assumptions, such as, that the distribution of $\bar{e}$ is normal.

For example, the maximum likelihood estimates of AR(1) process are

$$\hat{\mu} = \frac{1}{n}\sum_{t=1}^{n} X_t = \overline{X}$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{t=1}^{n}(X_t - \overline{X})^2$$

$$\hat{\phi} = \frac{D_{12}^*}{D_{22}^*}$$

and $\hat{\sigma}^2 = \frac{1}{n-1}(D^*_{11} - \hat{\phi}_1 D^*_{12})$ where

$$D^*_{ij} = D^*_{ji} = \frac{n}{(n+2-i-j)} \sum_{l=0}^{n+1-(i+j)} (X_{i+l} - \overline{X})(X_{j+l} - \overline{X})$$

Similarly for the AR(2) process, the system of linear equations for the parameter ϕ_k is

$$D^*_{12} = \hat{\phi}_1 D^*_{22} + \hat{\phi}_2 D^*_{23}$$

$$D^*_{13} = \hat{\phi}_1 D^*_{23} + \hat{\phi}_2 D^*_{33}$$

and it gives

$$\hat{\phi}_1 = \frac{D^*_{12} D^*_{33} - D^*_{13} D^*_{23}}{D^*_{22} D^*_{33} - D^{*2}_{23}} \quad , \quad \hat{\phi}_2 = \frac{D^*_{13} D^*_{22} - D^*_{12} D^*_{23}}{D^*_{22} D^*_{33} - D^{*2}_{23}}$$

Then, the estimate of the error variance becomes

$$\hat{\sigma}^2_a = \frac{1}{n-2}(D^*_{11} - \hat{\phi}_1 D^*_{12} - \hat{\phi}_2 D^*_{13})$$

3.4 Test for Model fitting and Parameter Estimation

The goodness of fit tests of the AR models fitted to a certain series can be accomplished by

(1) testing on the assumptions made for the AR model

(2) comparing the model correlogram with the fitted correlogram and

(3) checking whether the fitted model is adequate so as to obtain a parsimonious model.

In these three portions, the second portion does not need any test but just compare the two correlograms.

Tests on the assumption of the model

Two assumptions (a) independence and (b) the normality of the residuals of the fitted model have to be tested. The residuals of the fitted model a_t may be determined from the equation

$$\hat{a}_t = (X_t - \overline{X}) - \hat{\phi}_1(X_{t-1} - \overline{X}) - \ldots - \hat{\phi}_p(X_{t-p} - \overline{X}) \quad (3.4.1)$$

$$= w_t - \hat{\phi}_1 w_{t-1} - \ldots\ldots - \hat{\phi}_p w_{t-p}, \quad t = p+1, \ldots, n$$

where $w_t = X_t - \overline{X}$ are used to test the validity of the above assumptions.

To test whether a_t is an independent series or not, the Porte Manteau lack of fit test [Box-Pierce,1970] is appropriate to use. The test statistic is

$$Q = n \sum_{k=1}^{l} r_k^2 \quad (3.4.2)$$

where n is the number of observations, r_k is the lag k sample autocorrelation of the estimated residual series and 'l' is the maximum lag considered. It is appropriate to use 10 to 30 percent of n as l. Then the statistic Q is approximately χ^2_{l-p}.

Davies, Trigges and Newbold [Davies-Newbold, 1979] show that the χ^2 tables tend to overestimate the critical values of Q for finite samples and give briefed tables of exact percentage points. Ljung and Box [Ljung-Box, 1978] suggest the revised statistic

$$Q' = n(n+2)\sum_{j=1}^{l}\frac{r_j^2}{n-j}$$

whose distribution appears closer to that of χ^2_{1-p} for finite samples. Since they compute r_j as,

$$r_j = \frac{\sum_{t=1}^{n-j} a_t a_{t+j}}{\sum_{t=1}^{n} a_t^2}$$

and we use the definition,

$$r_j = \frac{n}{n-j}\frac{\sum_{t=1}^{n-j} a_t a_{t+j}}{\sum_{t=1}^{n} a_t^2},$$

it seems better for us to use Q.

If $Q < \chi^2_{1-p}$, then $\hat{a}_t$ can be taken to be independent and it in turn implies that the selected model is adequate.

To test the normality of the residuals of the fitted model, the skewness test of the normality is used. The skewness test of normality is based on the fact that the skewness coefficient for a normal variable has zero mean and the variance $\sigma/\sqrt{n}$ [Snedecor-Cochran, 1980]. Then the 95% limits of the skewness coefficient become $-1.96\sqrt{\sigma/n}$ and $+1.96\sqrt{\sigma/n}$.

If the skewness of the residual series is between the limits, the residual series is regarded as normal.

Test of the Adequacy of the model

To check whether the order of the fitted model is adequate, the Akaike Information Criterion (AIC) can be used [Kendall-Ord, 1983]. The AIC for an AR(p) model is

$$AIC(p) = n \ln \hat{\sigma}_a^2 + 2p$$

where $\hat{\sigma}_a^2$ is the maximum likelihood estimate of the error variance.

The comparison between the AICs of the AR(p-1), AR(p) and AR(p+1) models points out the model to be used. If the AIC(p) is less than both AIC(p-1) and AIC(p+1), then the AR(p) model is the best for the given data set. Otherwise, the model with less AIC have to be used in the fitting.

If the methods of moment is used in fitting, the estimates of error variance can be used to choose a model instead of using the AICs. If $\hat{\sigma}_p^2$, the estimate of the error variance of AR(p) is less than both $\hat{\sigma}_{p-1}^2$ and $\hat{\sigma}_{p+1}^2$, then AR(p) model is the best for the given data set. Otherwise, the model with less $\hat{\sigma}_a^2$ will be used in the fitting.

Similarly for comparing two autoregressive models of order p and p+r, MC Clave [Mc Clave,1978] proposed the criterion,

$$MC = (n-p-r)\ \frac{\hat{\sigma}_p^{\,2} - \hat{\sigma}_{p+r}^{\,2}}{\sigma_{p+r}^{\,2}}$$

which is asymptotically distributed as χ^2_r. In this criterion, $\hat{\sigma}_p^2$ and $\hat{\sigma}_{p+r}^2$ are the estimates of the error variance of AR(p) and AR(p+r) models.

3.5 AR Model Fitting to the Live-births and Population

In section 4,5,6 of Chapter 2, the primary investigations for the underlying process of various transformed series were done by using the sample auto and partial autocorrelations. From those investigations, it was found that the autoregressive models (AR models) can be used to represent the following selected transformed series.

(1) The standardized series
(2) The logarithmic transformed series
(3) The difference series

The next steps in AR model fitting after the primary investigations are

(i) finding the AR model of best fit or finding the most appropriate order of the AR model.
(ii) estimating the model parameters and
(iii) the diagnostic checking of the model

To be efficient in model fitting, the above steps have to be done in the following detail procedure.

(a) First of all, it needs to decide on the highest order of the AR model to be fitted. This can be done by examining the sample auto and partial correlograms of the series as described in the previous chapter.

Let the highest order chosen be p. Then for each of the order 1 to p.

(b) The model parameters have to be estimated. These estimates of the model parameters can be obtained by using the method of moments (Yule-Walker equations) or the maximum likelihood method of estimation.

(c) The estimates of the residuals are computed by using,

$$\hat{a}_t = w_{t+p} - \hat{\phi}_1 w_{t+p-1} - \hat{\phi}_2 w_{t+p-2} \ldots\ldots - \hat{\phi}_p w_t \ , \ t=1,\ 2,\ldots,21$$

where $w_t = X_t - \overline{X}$

(d) Then, the Porte Monteau lack of fit test is used to decide whether the residuals of a dependence models are uncorrelated.

(e) The AIC (Akaike Information Criterion) is used to determine which order is adequate among the fitted orders of the dependence model if the ML method is used in (b), or the estimated error variances, $\hat{\sigma}_a^2$ are simply used to determine the most appropriate order if the method of moment is used in (b).

The order whose AIC or $\hat{\sigma}_a^2$ is less than respective neighbouring value is the adequate order to use in further analysis.

(f) If the order can not be determined sharply, the Mc Clave criterion is used to choose the order of the model.

(g) After choosing the order of the model, it is tested whether the parameter estimates of the chosen model meet the stationary conditions or not.

3.5.1 AR Model Fitting to the Live-Births of Urban Myanmar

Since the sample autocorrelations for the lags 1 and 2 are consecutively out of the 95 percent confidence interval of a random series (-0.4277, +0.4277) and the partial correlation for lag(1) is out of the above interval, the AR(2) model is chosen as the highest order to be fitted to the observed series.

Method of moment is used to estimate the parameters of the model. The estimates of the autoregressive parameters and the estimate of the error variance, $\hat{\sigma}_a^2$ are as shown in Table (3.1).

Table (3.1)

Method of Moments Estimates of the Parameters for Live-Births of Urban Myanmar

Order	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\sigma}_a^2$
1	0.797	-	-	6968
2	0.849	-0.066	-	7314
3	0.836	0.105	-0.202	7368

After finding the estimates of the parameters for each order of the model, the estimated residual series is computed and tested to see whether the series is random or still contains some systematic parts. In so doing, the Porte Manteau lack of fit test is applied to the residual series. The value of the test statistics and the corresponding decisions are shown in the following Table (3.2).

Table (3.2)

Porte Manteau Lack of Fit Test

Order	Q	Deg.Fr.	C.V	Sig-Level	Decision
1	20.51	17	27.59	5%	random
2	19.38	15	25	5%	random
3	19.65	13	22.36	5%	random

From above tests, it can be seen that the residual series are random. Thus, the AR(1) to AR(3) explain the systematic component of the observed series and AR(1) to AR(3) can be used to represent the underlying process of live-births.

Since the method of moment is used to estimate the parameters, the estimated error variance is used to choose one of models, AR(1) to AR(3). By examining the error variance describe in Table (3.1), the underlying process of live-births series is AR(1) since the $\hat{\sigma}_a^2$ for the AR(1) is smaller than their respective neighbouring values.

The stationary condition of the model is met.

3.5.2 AR Model Fitting to the Standardized Series of Live-births

The sample autocorrelations for lag(1) and lag(2) are out of 95 percent confidence interval of a random series and the partial autocorrelation for lag(1) is out of the above interval for the standardized series of live-births. Thus, the AR(1) to AR(3) are chosen to be fitted to the standardized series.

Then the method of moment is used to estimate the parameters of the model. The method of moments estimates of the mean μ and variance σ^2 are zero and 1.05. The method of moments estimates of the autoregressive parameters and the estimated error variance, σ_a^2 , for each order of the AR model are as shown in Table (3.3).

Table (3.3)

Method of Moments Estimates of the Parameters for Standardized Series of Live-Births

Order	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\sigma}_a^2$
1	0.797	-	-	0.042
2	0.849	-0.066	-	0.422
3	0.836	0.105	-0.202	0.425

After finding the estimates of the parameters for each order of the model, the estimated residual series is computed and tested to see whether the series is simply random or still contains some systematic parts. Thus, the same test, as before is applied to the residual series.

The value of Q with its level of significance for the Porte Manteau lack of fit test is as shown in Table (3.4).

Table (3.4)

Porte Manteau Lack of Fit Test

Order	Q	Deg.Fr.	C.V	Sig-Level	Decision
1	20.62	17	27.59	5%	random
2	19.52	15	25	5%	random
3	19.85	13	22.36	5%	random

From Table (3.4), it is clear that the residual series are random for all orders. Therefore AR(1) to AR(3) models can be used as the underlying process of the standardized series.

Since the method of moments is used to estimate the parameters, the error variance is used to choose one of the models AR(1) to AR(3). From Table (3.3), it can be seen that the estimates of the error variance of the AR(1) is less than their respective neighbouring values. Thus, the AR(1) model is chosen to represent the underlying process of standardized series. The stationary condition for the AR(1) model is $|\hat{\phi}_1| < 1$. Since $\hat{\phi}_1 = 0.797$, the stationary condition for model is met.

3.5.3 AR Model Fitting to the Logarithmic Transformed Series of Live-Birth

Since the sample autocorrelations for lag(1) and lag (2) are out of the 95 percent confidence interval of a random series (-0.4277, +0.4277) and the partial autocorrelation for lag one is out of the above interval, the AR(3) model is chosen as the highest order to be fitted to the logarithmic transformed series.

Then, the method of moments is used to estimate the parameters of the model. The method of moments estimates of the autoregressive parameters and the estimated error variance, $\hat{\sigma}_a^2$, for each order of the AR model are as shown in Table (3.5).

Table (3.5)

Method of Moments Estimates of the Parameters for Log-transformed Series of Live-Births

Order	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\sigma}_a^2$
1	0.803	-	-	0.0018
2	0.873	-0.087	-	0.0019
3	0.859	0.109	-0.208	0.0019

After finding the estimates of the parameters for each order of the model, the estimated residual series is computed and tested to see whether the series is simply random. Thus, the Porte Monteau lack of fit test is applied to the residual series.

The value of Q with its level of significance for the Porte Manteau lack of fit test with the decision for each order are as shown in Table (3.6).

Table (3.6)

Porte Manteau Lack of Fit Test

Order	Q	Deg.Fr.	C.V	Sig-Level	Decision
1	27.59	17	27.59	5%	random
2	25	15	25	5%	random
3	22.36	13	22.36	5%	random

From Table (3.6) it is clear that the residual series are random for all orders. Therefore AR(1) to AR(3) model can be used as the underlying process of the log-transformed series.

Since the method of moments is used to estimate the parameters, the error variance is used to choose one of the models AR(1) to AR(3). Thus, the AR(1) model is chosen to represent the underlying process of log-tranformed series. The stationary condition for AR(1) model is $|\hat{\phi}_1| < 1$. Since $\hat{\phi}_1$ the stationary condition for model is met.

3.5.4 AR Model Fitting to the First Difference Series of Live-Birth

The sample auto and partial autocorrelations of the first difference series of live-births are in the 95 percent confidence interval of a random series. The AR(3) model is chosen as the highest order to be fitted to the first difference series of live-births.

Then the method of moment is used to estimate the parameters of the model. The method of moments estimates of the autoregressive parameters and the estimated error variance, $\hat{\sigma}_a^2$, for each order of the AR models are as shown in the Table (3.7).

Table (3.7)

Method of Moments Estimates of the Parameters for First Difference Series of Live-Births

Order	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\sigma}_a^2$
1	0.179	-	-	2835
2	0.135	0.247	-	2809
3	0.104	0.2297	0.127	2764

After finding the estimates of the parameters for each order of the model, the estimated residual series is computed and tested to see whether the series is simply random or still contains some systematic parts. Thus, the same test, as before is applied to the residual series.

The value of Q with its level of significance for the Porte Manteau lack of fit test with the decision for each order is as shown in Table (3.8).

Table (3.8)

Porte Manteau Lack of Fit Test

Order	Q	Deg.Fr.	C.V	Sig-Level	Decision
1	5.637	16	26.03	5%	random
2	5.389	14	23.68	5%	random
3	7.473	12	21.03	5%	random

From Table (3.8) it is clear that the residual series are random for all orders. Therefore AR(1) to AR(3) models can be used as the underlying process of the first difference series.

Since the method of moments is used to estimate the parameters, the error variance is used to choose one of the models AR(1) to AR(3). From Table(3.7), it can be seen that estimates of the error variances are different slightly small. The Mc Clave Criterion has to be used in determining the order of the best fit sharply.

From Table(3.7 $\hat{\sigma}_a^2$ (1) = 2835 , $\hat{\sigma}_a^2$ (2) = 2809 ,

$\hat{\sigma}_a^2$ (3) = 2764 the Mc Clave criterion is

MC = (21-2)(2835-2809)/2809

= 0.176 $< \chi^2_{(0.05,1)}$ = 3.84,

MC = (21-3)(2835-2764)/2764

= 0.4624 $< \chi^2_{(0.05,2)}$ = 5.99,

and since Mc = 0.176 is less than $\chi^2_{(0.05,1)}$ = 3.84, Mc = 0.462 is less than $\chi^2_{(0.05,2)}$ = 5.99, the AR(1), AR(2) and AR(3) models are not significantly different. Thus AR(1) model is chosen to represent the underlying process of the first difference series.

The stationary condition for AR(1) model is $|\hat{\phi}_1| < 1$. Since ϕ = 0.179, the stationarity condition is met.

3.5.5 AR Model Fitting to the Mid-Year Estimated Population

Since the sample autocorrelations for the lag 1 to 4 were consecutively out of the 95 percent confidence interval of a random series (-0.4277, +0.4277) and the partial autocorrelations for the lag(1) was out of the above interval, the AR(3) model is chosen as the highest order to be fitted to the mid-year estimated population.

Then the method of moments is used to estimate the parameters of the model.

The estimates of the autoregressive parameters and the estimates of the error variance, σ_a^2 are as shown in Table (3.9).

Table (3.9)

Method of Moments Estimates of the Parameters for Series of Mid-Year Estimated Population

Order	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\sigma}_a^2$
1	0.857	-	-	2504
2	0.897	-0.047	-	2630
3	0.892	0.049	-0.108	2747

After finding the estimates of the parameters for each order of the model, the estimated residual series is computed and tested to see whether the series is simply random or still contains some systematic parts. In so doing, the Porte Manteau Lack of fit test is applied to the residual series. The values of test statistics and the corresponding decisions are shown in the following table, Table (3.10).

Table (3.10)

Porte Manteau Lack of Fit Test

Order	Q	Deg.Fr.	C.V	Sig-Level	Decision
1	7.84	17	27.59	5%	random
2	6.54	15	25	5%	random
3	7.79	13	22.36	5%	random

From the Porte Manteau Lack of fit test, it can be seen that the residual series are random. Thus, the AR(1) to AR(3) can explain the systematic component of the observed series and all these models, AR(1) to AR(3) can be used to represent the underlying process of the mid-year estimated population.

Since the method of moments is used to estimate the parameters, the estimated error variance is used to choose one of the models, AR(1) to AR(3). By examining the estimates of the error variance described in Table (3.11), the underlying process of the observed series is AR(1). Since $\hat{\sigma}_a^2$ for AR(1) is smaller than AR(2) and AR(3).

The stationary condition for the AR(1) model $|\hat{\phi}_1| < 1$. Since $\hat{\phi}_1 = 0.857$, the stationary condition of the model is met.

3.5.6 AR Model Fitting to the Log-transformed Series of Mid-Year Estimated Population

The sample autocorrelation for Lag(1) to Lag(4) were out of 95 percent confidence interval of a random series and the partial autocorrelation for lag(1) was out of the above interval. Thus AR(3) model is chosen as the highest order to be fitted to the log-transformed series.

Then the method of moment is used to estimate the parameters of the model. The method of moments estimates of the autoregressive parameters and the estimated error variance, $\hat{\sigma}_a^2$, for each order of the AR model are as shown in Table (3.11).

Table (3.11)

Method of Moments Estimates of the Parameters for Log-Transformed Series of Mid-Year Estimated Population

Order	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\sigma}_a^2$
1	0.855	-	-	0.007
2	0.896	-0.048	-	0.007
3	0.895	0.05	-0.112	0.008

After finding the estimates of the parameters for each order of the model, the estimated residual series is computed and tested to see whether the series is simply random or still contains some systematic parts. Thus, the same test as before is applied to the residual series.

The value of Q with its level of significance for the Porte Manteau Lack of fit test with the decision for each order are as shown in Table (3.12).

Table (3.12)

Porte Manteau Lack of Fit Test

Order	Q	Deg.Fr.	C.V	Sig-Level	Decision
1	7.22	17	27.59	5%	random
2	6.21	15	25.00	5%	random
3	7.79	13	22.36	5%	random

From Table (3.12) it is clear that the residual series are random for all orders. Therefore AR(1) to AR(3) models can be used as the underlying process of the log-transformed series.

Since the method of moments is used to estimate the parameters, the error variance is used to choose one of the models AR(1) to AR(3). It can be seen that the estimates of AR(1) and AR(2) are equal but they are less than the value of AR(3). Thus, the AR(1) model is chosen to represent the underlying process of log-transformed series.

The stationary condition for the AR(1) model is $|\hat{\phi}_1|<1$. Since $\hat{\phi}_1 = 0.855$, the stationary condition for AR(1)model is met.

3.5.7 AR Model Fitting to the Standardized Series of Mid-Year Estimated Population

The sample autocorrelations for Lag(1) to Lag(4) were out of the confidence interval and the partial autocorrelation for lag(1) is also out of the above interval for the standardized series. Thus AR(3) model is chosen as the highest order to be fitted to the standardized series.

Then the method of moment is used to estimate the parameters of the model. The method of moments estimates of the autoregressive parameters and the estimated error variance, $\hat{\sigma}_a^2$, for each order of the AR model are as shown in Table (3.13).

Table (3.13)

Method of Moments Estimates of the Parameters for the Standardized Series of Mid-Year Estimated Population

Order	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\sigma}_a^2$
1	0.857	-	-	0.293
2	0.897	-0.047	-	0.308
3	0.892	0.049	-0.108	0.321

After finding the estimates of the parameters for each order of the model, the estimated residual series is computed and tested to see whether the series is simply random or still contains some systematic parts. Thus, the same test as before is applied to the residual series.

The value of Q with its level of significance for the Porte Manteau Lack of fit test with the decision for each order are as shown in Table (3.14).

Table (3.14)

Porte Manteau Lack of Fit Test

Order	Q	Deg.Fr.	C.V	Sig-Level	Decision
1	7.899	17	27.59	5%	random
2	6.540	15	25.00	5%	random
3	7.796	13	22.36	5%	random

From Table (3.14) it is clear that the residual series are random for all orders. Therefore, the AR(1) to AR(3) models can be used as the underlying process of the standardized series.

Since the method of moments is used to estimate the parameters, the error variance is used to choose one of the models AR(1) to AR(3). From Table (3.13), it can be seen that the estimates of the error variance of AR(1) is less than the value of AR(2) and AR(3). Therefore the underlying process of standardized series is AR(1).

The stationary condition for AR(1) is $|\hat{\phi}_1| < 1$. Since $\hat{\phi}_1 = 0.857$, the stationary condition for model is met.

3.5.8 AR Model Fitting to the First Difference Series of Mid-Year Estimated Population

Since the sample autocorrelations for lag(1) was out of the 95 percent confidence interval of a random series (-0.4293, +0.4293) and the partial autocorrelation for the lag(1) was out of the above interval, the AR(2) model is chosen as the highest order to be fitted to the first difference series of mid-year estimated population.

Then, the method of moments is used to estimate the parameters of the model.

The estimates of the autoregressive parameters and the estimates of the error variance, $\hat{\sigma}_a^2$, are as shown in Table (3.15).

Table (3.15)

Method of Moments Estimates of the Parameters for First-difference Series of Mid-Year Estimated Population

Order	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\sigma}_a^2$
1	-0.616	-	-	1792
2	-0.767	-0.246	-	1779
3	-0.728	-0.122	0.161	1559

After finding the estimates of the parameters for each order of the model, the estimated residual series is computed and tested to see whether the series is simply random or still contains some systematic parts. Thus, the same tests, as before is applied to the residual series.

The value of Q with its level of significance for the Porte Manteau lack of fit test with the decision for each order are as shown in Table (3.16).

Table (3.16)

Porte Marteau Lack of Fit Test

Order	Q	Deg.Fr.	C.V	Sig-Level	Decision
1	3.57	16	26.30	5%	random
2	4.41	14	23.68	5%	random
3	4.32	12	21.03	5%	random

From Table (3.16) it is clear that the residual series are random for all orders. Therefore, the AR(1) to AR(3) model can be used as the underlying process of difference series.

Since the method of moments is used to estimate the parameters, the error variance is used to choose one of the models AR(1) to AR(3). From Table (3.18), it can be seen that the estimates of the error variance of AR(1), AR(2), AR(3) are not different very much. So Mc Clave criterion has to be used in determining the order of the best fit sharply.

From Table(3.15) $\hat{\sigma}_a^2$ (1) = 1792 and $\hat{\sigma}_a^2$ (2) = 1779 and $\hat{\sigma}_a^2$ (3) = 1559 the Mc Clave criterion is

$$MC = (21-2)(1792-1779)/1779$$

$$= 0.138$$

and since Mc is less than $\chi_{(0.05,1)}^2$ = 3.84, the AR(1) and AR(2) models are not significantly different.

$$MC = (21-3)(1792-1559)/1559$$
$$= 2.69$$

and since Mc is less than $\chi_{(0.05,2)}^2$ = 5.99, the AR(3) and AR(1) models are not significantly different. Thus, the AR(1) model is chosen to represent the underlying process of first different series.

The stationary condition for AR(1) model is $|\hat{\phi}_1|<1$. Since $\hat{\phi}_1=-0.616$, the stationary condition for model is met.

Comments on the Model Fitting

When the AR model is fitted to the selected transformed series, it is found that the AR(1) model can be used to represent all the series. It is also found that the moments method and the maximum likelihood method are not significantly different in their results as to the model fitting but the moments method is simple to use and the results obtained can be checked manually.

As to various transformations, the first difference transformation can be said to be the best since the forecast values of the first difference transformed series are the nearest values of the actual values.

In all the computation of this thesis, the method of moments estimates were obtained by solving the Yule-Walker equations. The rigid assumptions are not needed when the moments method is used.

CHAPTER IV

EX-ANTE PROJECTION OF FERTILITY DATA

4.1 Projection of Live-births Data for Urban Area

Demographic predictions have had a poor record in recent decades . The predictions have been misleading with respect to the future levels. They have also failed to anticipate the likely range of error. Nor have recent refinements of technique improved the situation, either with respect to the accuracy of predictions, or the derivation of appropriate confidence interval. Indeed the continued efforts of demographers to predict population are sometimes justified on the grounds that they facilitate demographic analysis, rather than the reverse.

Demographic analysis is concerned with precisely defined rates and corresponding subtleties of population structure, it proceeds by successive refinements of population . But demographic forecasting techniques failed with post-transition populations dominated by fluctuating fertility; time series analysis of fertility can improve the forecasts[Ronald,JASA,1974].

In this chapter, an attempt is made to predict live-births for selected towns in urban Myanmar by using the first-order autoregressive model. In the previous chapter,various transformed methods for live-births and model fitting have been mentioned. In this chapter, the best transformed live-birth data are used to predict the future live-births. Moreover, three forms for forecast are described in this chapter.

4.1.1 Three Basic Forms for the Forecast

The expression for the forecast in any one of three different ways can be written down corresponding to the three ways of expressing the model. [Box-Jenkin,1970]. For simplicity in notation, squared brackets imply that the conditional expectation, at time t. Thus

$$[a_{t+l}] = E_t[a_{t+l}]$$

$$[Z_{t+l}] = E_t[Z_{t+l}] \quad \text{for} \quad l>0$$

The three explicit forms for the general ARIMA model are

(i) Directly in terms of the difference equation by

$$z_{t+l} = \varphi_1 z_{t+l-1} + \ldots + \varphi_{p+d} z_{t+l-p-d} - \theta_1 a_{t+l-1} - \ldots - \theta_q a_{t+l-q} + a_{t+l} \qquad (4.1.1)$$

(ii) As an infinite weighted sum of current and previous shocks a_j,

$$Z_{t+l} = \sum_{j=-\infty}^{t+l} \psi_{t+l-j} a_j = \sum_{j=0}^{\infty} \psi_j a_{t+l-j} \qquad (4.1.2)$$

where $\psi_0 = 1$ and the ψ weights may be obtained by equating coefficients in

$$\varphi(B) \cdot (1 + \psi_1 B + \psi_2 B^2 + \ldots) = \theta(B) \qquad (4.1.3)$$

Equivalently , for positive $l>q$, the model may be written in the truncated form

$$Z_{t+l} = C_t(l) + a_{t+l} + \psi_1 a_{t+l-1} + \ldots + \psi_{l-1} a_{t+1} \quad (4.1.4)$$

where the 'complementary function $c_t(l)$ is equal to the truncated infinite sum

$$C_t(l) = \sum_{j=-\infty}^{t} \psi_{t+l-j}\, a_j = \sum_{j=0}^{\infty} \psi_{l+j}\, a_{t-j} \quad (4.1.5)$$

(iii) As an infinite weighted sum of previous observations, plus a random shock

$$Z_{t+l} = \sum_{j=1}^{\infty} \pi_j\, Z_{t+l-j} + a_{t+l} \quad (4.1.6)$$

Also, if $d \geq 1$

$$\bar{Z}_{t+l-1}(\pi) = \sum_{j=1}^{\infty} \pi_j\, Z_{t+l-j} \quad (4.1.7)$$

will be a weighted average, since then $\sum_{j=1}^{\infty} \pi_j = 1$.

The π weight may be obtained from

$$\varphi(B) = (1 + \pi_1 B + \pi_2 B^2 + \ldots).\theta(B) \quad (4.1.8)$$

Forecast from Difference Equation

Taking conditional expectations at time t in (4.1.1), it can be obtained the following equation;

$$[Z_{t+1}] = \hat{Z}_t(1) = \varphi_1 [Z_{t+1-1}] + \ldots + \varphi_{p+d} [Z_{t+1-p-d}]$$

$$-\theta_1 [a_{t+1-1}] - \ldots - \theta_q [a_{t+1-q}] + [a_{t+1}] \qquad (4.1.9)$$

By using the equation (4.1.2), the forecast in integrated form

$$[Z_{t+1}] = \hat{Z}_t(1) = \psi_1 [a_{t+1-1}] + \ldots + \psi_{1-1} [a_{t+1}] + \psi_1 [a_t] +$$

$$\psi_{1+1} [a_{t-1}] + \ldots + [a_{t+1}]$$

Alternatively, using the truncated form of the model (4.1.4), for positive l>q.

$$[Z_{t+1}] = \hat{Z}_t(1) = C_t(1) + [a_{t+1}] + \psi_1 [a_{t+1-1}] + \ldots + \psi_{1-1} [a_{t+1}] \qquad (4.1.11)$$

where $c_t(l)$ is the complementary function at origin t.

Forecasts as a weighted average of previous observations and forecast made at previous lead times from the same origin. Finally, taking conditional expectation in 4.1.6.

$$[Z_{t+1}] = \hat{Z}_t(1) = \sum_{j=1}^{\infty} \pi_j [Z_{t+1-j}] + [a_{t+1}] \qquad (4.1.12)$$

To calculate the conditional expectations which occur in the expression (4.1.10) to (4.1.12), if j is a nonnegative integer,

$$[Z_{t-j}] = E_t [Z_{t-j}] = Z_{t-j} \qquad j = 0, 1, 2, \ldots$$

$$[Z_{t+j}] = E_t\ [Z_{t+j}] = \hat{Z}_t\ (j) \qquad j = 1, 2, \ldots.$$

$$[a_{t-j}] = E_t[a_{t-j}] = a_{t-j} = Z_{t-j} - \hat{Z}_{t-j-1}\ (1) \qquad j = 0, 1, 2, \ldots$$

$$[a_{t+j}] = E_t[a_{t+j}] = 0$$

Therefore, to obtain the forecast Z_t, one writes down the model for Z_{t+1} in any one of the above three explicit forms and treats the terms on the right according to the following rules.

The $Z_{t-j}(j=0,1,2,\ldots.)$, which have already happened at origin t, are left unchange.

The $Z_{t+j}(J=1,2,\ldots)$, which have not yet happened, are replaced by their forecasts $\hat{Z}_t(j)$ at origin t.

The $a_{t-j}(j=0,1,2,\ldots)$, which have happened, are available from $Z_{t-j}-\hat{Z}_{t-j-1}(1)$.

The $a_{t+j}(j=1,2,\ldots)$, which have not yet happened are replaced by zeros.

The variance of the 1 steps ahead forecast error for any origin t is the expected value of

$$e_t^2\ (1) = \{Z_{t+1} - \hat{Z}_t(1)\}^2$$

and is given by

$$V(1) = \{\ 1 + \sum_{j=1}^{l-1} \psi_j^2\ \}\ \sigma_a^2$$

Assuming that the a's are normal. So the conditional probability distribution $P[Z_{t+1} / Z_t, Z_{t-1}, \ldots]$ of a future value of Z_{t+1} of the process will be normal with mean $\hat{Z}_t(1)$ and standard deviation $\{ 1 + \sum_{j=1}^{l-1} \psi_j^2 \}^{1/2} \sigma_a$

The approximate 1-ε probability limits $Z_{t+1}(-)$ and $Z_{t+1}(+)$ for Z_{t+1} will be given by

$$Z_{t+1} \ (\pm) = \hat{Z}_t(1) \pm \mu_{\varepsilon/2}\{ 1 + \sum_{j=1}^{l-1} \psi_j^2 \}^{1/2} s_a$$

where $\mu_{\varepsilon/2}$ is the deviate exceeded by a proportion ε/2 of the unit normal distribution. The 50% and 95% limits for $\hat{Z}(1)$ are given by

$$50\% \text{ limits} \quad \hat{Z}_t(1) \pm (0.674) \{ 1 + \sum_{j=1}^{l-1} \psi_j^2 \}^{1/2} s_a$$

$$95\% \text{ limits} \quad \hat{Z}_t(1) \pm (1.96) \{ 1 + \sum_{j=1}^{l-1} \psi_j^2 \}^{1/2} s_a$$

The probability of the actual value z_{t+1} will occur within these limits $z_{t+1}(-)$ and $z_{t+1}(+)$ is

$$p_r \ (Z_{t+1}(-) < Z_{t+1} < Z_{t+1}(+)) = 1-\varepsilon$$

4.1.2 AR Model For Live-Births Forecasting

Data for the live-births time series for selected towns in urban Myanmar for 1968-1988 were taken from various issues of the Vital Statistics Reports of urban Myanmar.Various transformed for live-births data fitted in AR model were tried,and the best encountered were an AR(1) model for first difference live-births time series for

1968-2000, together with 95% confidence interval were obtained using AR(1) model are shown in the Table(4.1)

Forecasts for the AR(1) model of live-births series is written as

$$(1-B)\phi(B)z_{t+1} = a_{t+1} \quad .... \qquad 4.1.2(1)$$

Where $\phi(B)=1-\phi B$ and a_{t+1} is the white noise with mean zero and variance σ_a^2.

$$Z_{t+1} = (1+\phi_1)Z_{t+1-1}-\phi_1 Z_{t+1-2}+a_{t+1} \qquad 4.1.2(2)$$

Taking conditional expectation in 4.1.2(2)

$$Z_t(1)=(1+\phi_1)Z_t(1-1)-\phi_1 Z_{t+1-2}+a_{t+1}$$

And then z_{t+1} can be written as an infinite weighted sum of current and previous shocks a_j,

$$Z_{t+1}=\psi(B)\ a_{t+1} \quad .. \qquad 4.1.2(3)$$

By equating eq^n 4.1.2(1) and 4.1.2(3), we get

$$(1-B)\phi(B) = \psi^{-1}(B)$$

$$(1-B)\psi(B) = 1/\phi(B)$$

$$(1-B)(1-\psi_1 B-\psi_2 B^2-....) = 1/(1-\phi_1 B)$$

$$(1-\psi_1 B-\psi_2 B^2- B+\psi_1 B+\psi_2 B^3+\psi_3 B^4+\ ) = \sum_{j=0}^{\infty} (\phi_1 B)^j$$

By equating

$$\phi_1 = -1\ -\psi_1$$

$\psi_1 = -\phi_1 - 1$

$\phi_1^2 = \Psi_1 - \Psi_2$

$\Psi_2 = \Psi_1 - \phi_1^2$

$\Psi_3 = \Psi_2 - \phi_1^3$

....

In general $\Psi_i = \Psi_{i-1} - \phi_1^i$

The interval forecast of the live-births time series can be obtained by using this equation

$$\hat{Z}_t(1) \pm (1.96) \{ 1 + \sum_{j=1}^{l-1} \psi_j^2 \}^{1/2} s_a$$

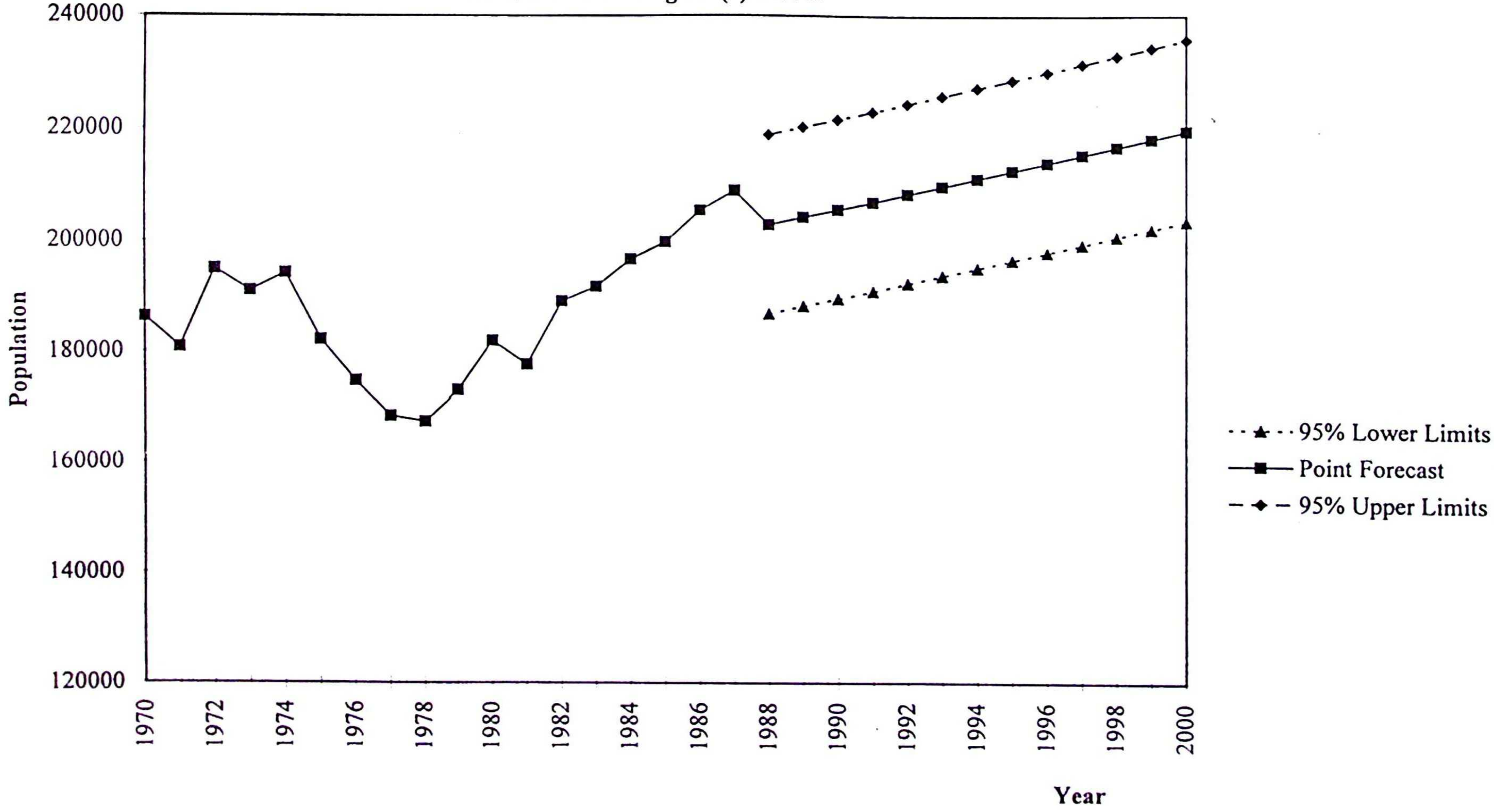
The forecat of live-births for the year 1968 was 186272 and its actual value was 181879. So, it was very reasonable. The forecasts overestimated the actual number of births in 1968 due to the fertility decrease during that year. And also 1971 to 1988 forecasts were very reasonable.

TABLE(4.1)

Forecast for The Total Live-Births Series for 1968-2000 Using AR(1) Model

Year	95% lower Limits	Point Forecasts	95% Upper Limits
1970		186272	
1971	164644	180779	196914
1972	178813	194948	211083
1973	174822	190957	207092
1974	177986	194121	210256
1975	165922	182058	198193
1976	159364	174599	190734
1977	152054	168189	184324
1978	150993	167128	183263
1979	156752	172887	189022
1980	165656	181791	197854
1981	161430	177565	193700
1982	172825	188960	205095
1983	175445	191580	207715
1984	180354	196489	212624
1985	183412	199547	215682
1986	189283	205148	221283
1987	192579	208714	224849
1988	186522	202657	218792
1989	187935	204070	220205
1990	189348	205483	221618
1991	190761	206896	223031
1992	192174	208309	224444
1993	193587	209722	225857
1994	195000	211135	227270
1995	196413	212548	228683
1996	197826	213961	230096
1997	199239	215374	231509
1998	200652	216787	232922
1999	202065	218200	234335
2000	203478	219613	235748

Table(4.1)
Forecasts for the Total Live-Births Series
for 1968-2000 Using AR(1) Model



4.2 Projection of Population Data for Selected Towns in Urban Myanmar

In this thesis, various transformed for population data fitted in AR model were tried, and the best encountered were an AR(1) model for first difference population series, together with 95 per cent confidence intervals were obtained by using AR(1) model are shown in Table (4.2).

The total population for that period were estimated by the C.S.O and they are based on the 1973 and 1973 national census figures (VSR ,1988). So, the population data seem to be a mathematical curve. Therefore, some of the mathematical curves are described and fit to the population data in this thesis.

Linear Change: The pattern of growth may be assumed arithmetic progression i.e. that there is a constant amount of increase per unit of time. This concept of a straight line is often used not only to describe population growth but also to project it into the future.

In symbol, $P_t = a + b\,t$ (4.2.1)

The estimated population for the selected towns in urban Myanmar by the linear equation is shown in Table(4.2).

The geometric change is a compound interest type of change, i.e.

$$P_{t2} = P_{t1}(1+r)^{t2 - t1} \qquad (4.2.2)$$

If the compounding is assumed to take place continuously, the above expression can be represented by the exponential curve, as

$$P_{t2} = P_{t1} \, e^{r(t2-t1)} \qquad (4.2.3)$$

The population of three selected towns in Urban Myanmar are estimated by the linear curve, geometric curve and exponential curves. They are shown in Table (4.2).

Application of the model to selected towns in Urban Myanmar population yielded unsatisfactory results. But, the population estimated by the mathematical curve seem to give satisfactory results. Moreover, the estimated population of the exponential method and geometric method has the same results.

4.3 Future Trend OF fertility Data In Urban Myanmar

The most basic form of measures of fertility is the crude birth rate which is a ratio of birth occurred during a year to the mid-year population,usually multiplied by 1000.In other words,the crude birth rate shows the number of births per 1000 population per year,indicating an overall effect of fertility upon the growth of population during a year. It can be expressed as follows;

$$CBR = B/P \cdot K$$

Where B = the number of births during a year

P = population at the middle of year

K =1000

TABLE (4.2)

The Estimated Population of The Selected Towns In Urban Myanmar

Year	Observed Value	Linear	Geometric	Exponential
1968	4700950	4664573		
1969	4714048	4819551	4817349	4817349
1970	4842986	4974530	4936629	4936629
1971	4938418	5129509	5058863	5058863
1972	5149647	5284487	5184124	5184124
1973	5680534	5439466	5312486	5312486
1974	5376960	5594445	5444027	5444027
1975	5911850	5749424	5578825	5578825
1976	6024437	5904402	5716960	5716960
1977	6180754	6059381	5858515	5858515
1978	6338052	6214360	6063576	6063576
1979	6512218	6369338	6152229	6152229
1980	6637718	6524317	6304562	6304562
1981	6763526	6679296	6460667	6460667
1982	6921701	6834274	6620637	6620637
1983	6941515	6989253	6784568	6784568
1984	7084613	7144232	6952559	6952559
1985	7226444	7299211	7124708	7124708
1986	7380948	7454189	7301121	7301121
1987	7507803	7609168	7481901	7481901
1988	7667158	7764147	7667158	7667158
1989		7919125	7857002	7857002
1990		8074104	8051546	8051546
1991		8229083	8250908	8250908
1992		8384061	8455205	8455205
1993		8539040	8664562	8664562
1994		8694041	8696884	8696884
1995		8849019	8881477	8881475
1996		9003997	9069989	9069987
1997		9158975	9262502	9262499
1998		9313953	9459101	9459098
1999		9468931	9659873	9659869
2000		9623909	9864906	9864902

The CBR for the selected towns in urban Myanmar are shown in the Table(4.3). CBR* is estimated by C.S.O.It shows that CBR* in the urban areas fluctuated arround 40 per thousand. Until the early 1970s,except for the years 1952 to 1954,thereafter,a feritility decline has been in progress.Between 1971 and 1976,the reported CBR declined by more than 29 per cent.Comparing the reported CBR for five year averages,there was an 11.7 per cent decline between 1971-1975 and 1976-1980.

But,the level of fertility could hardly be considered reliable because of the average area, which vary from year to year.Nevertheless,the fertility decline in urban areas started around 1970. It accelerated in the late 1970's and become stable again at around 28 per thousand population in 1980's.

In this thesis,CBR(1) is calculated by using the data of AR(1)model estimates of live-birth and estimated population from (VSR) and CBR(2) is calculated by using the data of AR(1) model estimates of live-births and exponential curve of population.

It reveals the trend in crude birth rates from.1970 to2000.The rates were obtained from various forecasts.It shows that the CBR for the selected towns in Urban Myanmar have declined slowly since 1972.As may be seen,the crude birth rates fluctuated around 27 per thousand population until the mid 1990.Since1981,fertility declined slightly to about 27 per thousand.

TABLE(4.3)
The CBR for Selected Towns in Urban Myanmar

Year	No.of Reporting Towns	CBR*	No.of Towns	CBR(1)	CBR(2)
1970	139	37.6	125	38.46	37.73
1971	138	39.2	125	36.61	35.74
1972	126	38.0	125	37.86	37.60
1973	170	32.5	125	33.62	35.94
1974	178	34.1	125	36.10	35.65
1975	164	29.7	125	30.79	32.63
1976	145	28.2	125	28.98	30.54
1977	115	27.2	125	27.21	28.71
1978	159	27.0	125	26.37	27.56
1979	158	27.8	125	26.55	28.10
1980	146	26.9	125	27.39	28.83
1981	145	27.7	125	26.25	27.48
1982	151	27.6	125	27.30	28.54
1983	167	28.3	125	27.59	28.24
1984	167	28.3	125	27.73	28.26
1985	168	28.6	125	27.61	28.01
1986	169	28.6	125	27.79	28.09
1987	169	28.6	125	27.79	27.89
1988	169	28.6	125	26.43	26.43
1989	169	28.5	125		25.97
1990					25.52
1991					25.08
1992					24.64
1993					24.20
1994					24.28
1995					23.93
1996					23.59
1997					23.25
1998					22.92
1999					22.59
2000					22.26

Sources: Nyan Myint (1988, pp-35, 69); CSO (1990); Ministry of Planning and Finance, MPF (1990, p-207)

CBR* is estimated by CSO

CBR(1) is calculated by using the estimated live-births from AR(1) model and the estimated population from VSR.

CBR(2) is calculated by using the estimated live-birth by AR(1) model and the estimated population by using exponential curve.

Crude birth rate remained stable between the 1982 and mid-1990 accorrding to various method of data. Since 1973, a tendency towards declining fertility has been observed. Total fertility fell by about 17.9 per cent between 1973 and 1983.

The demographic data in this thesis, although not representative of the whole country, are comparatively reliable. The information provided in Table (4.3) suggest a declining trend in crude birth rates. It can be seeen that fertility levels in Urban Myanmar have been relatively low and may decline in the future.

CHAPTER V

CONCLUSION

Model fitting for yearly fertility data in Urban Myanmar Fertility for selected towns in urban Myanmar, this study demonstade the importance of the meethodological issues in fertility studies. Because time series analysis enhancess confidence in the results.

Time series analysis is a useful method for examining changes in fertility level. The AR model is a simple time series model for projecting the live births series. Analysis of selected towns in urban live biths data for the period 1968-1988 resulted in the AR model(1) with lag 1 and AR parameter coefficient of 0.179. And ,forecasts of the live births data is clears indication of the precision of the forecasting from a stochastic rather than the traditional deterministic births-forecasting models. In contrast, by presenting point rather than interval forecasts, deterministic models tend to create the illusion that the future is more certain than it is- this , in turn, may result in serious errors being made in planing for the future, for the properties of optional decisions made under considerable uncertainty are usually rather different from those made in situations characterized by certain or almost certain outcomes. For instance, in a certain world specialisation is usually on optional strategy, but when the future is uncertain, policy should be more flexible to allow for the different outcomes possible. Also, ass new information becomes available and the degree of uncertainty is reduced, decisions should be revised. The advantage of the stochastic model is that the model and reflect the uncertainty regarding the future. In contrast, deterministic models tend to ignore it.(1980, JASA,p39).

Although Myanmar is still considered as under-populated, population growth must be controlled in order to bring about significant economic development. Since Myanmar is one of the developing countries, economic development is certainly hamped by the currently moderately high population growth rate. The higher level fertility, the grater the problems for socio-economic development. For instance, even though the magnitude of food production has declined. Therefore, it is necessary to slow population growth by controlling fertility in order to achieve economic and social development in Myanmar.

Application of the model to selected towns in urban Myanmar fertility rates yielded satisfactory results. Crude birth rates fluctuated within the range of 26.25 to 38.46 per thousand population over twenty years period, 1968-1988, it is apparent that the crude birth rate started to decline during the decade 1970-1980. This may be due to many reasons. Of course, Myanmar is a Buddhist country and Buddhism does not oppose any kind of contraception, and because the people are familiar with modern methods of contraception, especially pills, injections and condoms. If the Government were to set up a family planning programme, fertility decline could be faster than in some neighbouring countries such as Bangladesh and India. Currently Myanmar still seems to maintain a pro-natalist altitude, although the establishment of a child-spacing programme has recently been accepted by the Government. The birth spacing project was funded by United Nations Fund for Population Activities (UNFPA) in 1991 and was implemented in (20) pilot townships. The project made a significant ontribution by supporting the initiative to introduce birth spacing in Union of Myanmar. The expansion of birth spacing programme to another (52) townships, making a total of (72) townships were implemented in 1996. The objective of this project is to improve the health status of mothers and children by lowering the high fertility, morbidity and mortality rates through birth spacing services. Therefore, fertility level of the country seems to have a decline in the future.

Appendix Table(1)

Selected Towns in States and Divisions(Urban Area)(1968-1988)

No	Kachin	Kayah	Kayin	Chin	Mon	Rakhine	Shan
1	Bhamo	Loikaw	Kawkareik	Tiddin	Kyaikto	Kyaukpyu	Kalaw
2	Myitkyina		Kya-in-Seikkyi	Paletwa	Belin	Sittwe	Kyaukme
3	Momauk		Pa-an	Falam	Mawlamyaing	Myohaung	Hse-Hsang
4	Shwegu		Thandaung	Matupi	Thaton	Kyauktaw	Taunggyi
5				Mindat		Maungdaw	Lashio
6				Haka			Kutkai
7						Thandwe	Liolem

Appendix Table(1)
(Concluded)
Selected Towns in States and Divisions(Urban Area)(1968-1988)

No	Sagaing	Tanintharyi	Bago	Magway	Mandalay	Yangon	Ayeyarwaddy
1	Katha	Kawthaung	Kyauktaga	Chauk	Kyaukpadaung	Yangon	Kyangin
2	Kawlin	Tanintharyi	Gyobingauk	Gangaw	Kyaukse	Hlegu	Kyonpyaw
3	Monywa	Dawai	Zigon	Minbu	Mahlaing	Kayan	Kyaiklat
4	Mawlaik	Myeik	Nyaunglebin	Salin	Tada-U	Kyauktan	Kyaunggon
5	Myinmu	Thayetchaung	Taungoo	Taungdwingyi	Pyinmana	Thanlwin	Ngaputaw
6	Sagaing		Htantabin	Natmauk	Pyawbwe	Taikkyi	Yandoon
7	Shwebo		Daik-U	Pakokku	Mandalay	Thongwa	Dedaye
8	Wuntho		Natalin	Pauk	Meiktila	Twante	Danubyu
9	Ye-U		Bago	Magway	Myingyan	Kungyangon	Pathein
10			Paungde	Thayet	Pyin-Oo-Lwin	Hmawbi	Pantanaw
11			Pyay	Aunglan	Yamethin		Pyapon
12			Pyu		Thazi		Bogale
13			Yedashe		Myitnge		Maubin
14			Minhla		Pagan		Myanaung
15			Shwegin				Moulyaingyun
16			Shwedaung				Myaungmya
17			Letpadan				Yegyi
18			Waw				Ngathainggyaung
19			Thanatpin				Labutta
20			Thayawaddy				Wakema
21			Thonze				Enime
22							Hinthada

REFERENCES

Box, George and Gwillyn Jenkin (1970) *Time Series Analysis; Forecasting and Control* Sanfrancisco, Holden-Day.

Box, G.E.P and Piere, D.A(1970) Distribution of Residual Autocorrelations in ARIMA Time Series Models Journal of American Statistical Association, 65;1509-1526

CENTRAL STATISTICAL ORGANIZATION (CSO)(1968-1988) Vital Statistis Reports. Central Statistical Organization and Department of Health, Yangon, Myanmar.

Davies, N, Triggs, C.M.and Newbold,D.P(1977) Significance Levels of the Box -Pierce Portmanteam Statistics in Finite Samples Biometrika, 64;517-522.

IMMIGRATION AND MANPOWER DEPARTMENT (IMD)(1976) 1973: Population Census. Immigration and Manpower Department, Yangon, Myanmar.

John MC Donald (1980) Births Time Series Models and Structural Interpretation, Journal of American Statistical Association,vol75, No 359

Joop Dee Beer(1985) A time Series Model for Chort Data Journal of American Statistical Association, Vol 80. No 391

Kendall, M.G, Stuart,A and Ord ,J.K.(1983) *The Advanced Theory of Statistics*, Vol (iii). Fourth Edition, London, Holden-Day.

Liung,G.M.and Box,G.E.P(1978) On a Masure of Lack of Fit in Time Series Models Biometrika,65;297-303.

MC Clave,J.T(1978) Estimating The Order of Autoregressive Model; THE Max χ^2 Method. The journal of American Statistics Association,73,122-128.

MYINT,NYAN *Levels and Trends of Fertility and Mortality in Burma*. M.A Thesis in Demography, ANU, Canberra, Australia.

NYUNT,TIN TIN(1978) *Estimation of Vital Rates for Burma and Their Effects on the Future Size of Population*. M.A Thesis in Demography, ANU, Canberra,Australia.

Ronald Demos Lee(1974) Forecasting Births in Post-Transition Populations:Stochastic Renewal with Serially Correlated Fertility Journal of American Statistical Association.Vol69,No.347.

Salas, J.D., Delleur,J.W., Yevjevich, V.and Lane.W.L(1980) Applied Modeling of Hydrologic Time Series, Water Resources Publications,Colorado,Fort Collins.

Sendecar ,G.W. and Cohran,W.G(1980) *Statistical Methods* 7th Edition, Ames, Jowa, Jowa Stat University Press.

Sundrum, R.M (1957) *Population Statistics of Burma Statistical Paper No.3, Economics Research Project*, Economics, Statistics and Commerce Department, University of Rangoon.