SUDAN UNIVERSITY OF TECHNOLEGY & SCIENCES

COLLEGE OF GRADULATE STUDIES

A Comparative Study to Estimate and Forecasting Mortality Using Demographic and Statistical Models

دراسة مقارنة للتقدير والتنبؤ بالوفاة باستخدام النماذج الديموغرافية والاحصائية

Case Study: Sudan

Time Period: (1973 -2008)

Thesis Submitted in Fulfillment of the Requirements for the phD. Degree in Statistics

BY

Shaza Mustafa Abd Elgader Abd Ellatif

Supervised by

Dr. Adeil Musa Younis

March 2017

بسم الله الرحمن الرحيم

Sudan University of Science & Technology

College of Graduate Studies

جامعة السودان للعلوم والتكنولوجيا

كلية الدراسات العليا

كلية الدراسات العليا

Ref: SUST/ CGS/A11

Approval Page

(To be completed after the college council approval)

Name of Candidate: Shaza Mustafa Abd elgader Abd Ellatif

Thesis title: A comparative Study to Estimate and forecasting Mortality Using Demographic and Statistical Models (Case study: Sudan Time Pero 1973-2008)

Degree Examined for: PHD

Approved by:

1. External Examiner

Name: Dr. Hunza Ibrahim Hamza

Signature: .... Date: 12-3-2017

2. Internal Examiner

Name: Mohd Arafa Ahmed Elkhatib

Signature: MDHA Elkhatieb Date: 12-3-2017

3. Supervisor

Co-Supervisor:

Name: Adil Musa Younis / Khalid Rahamtalla Khider

Signature: .... Date: 11.5.2017

HOLLYVERIOUS

الاية القرانية

قال تعالي:

(هُوَ الَّذِي خَلَقَكُم مِّن تُرَابٍ ثُمَّ مِن نُّطْفَةٍ ثُمَّ مِنْ عَلَقَةٍ
ثُمَّ يُخْرِجُكُمْ طِفْلا ثُمَّ لِتَبْلُغُوا أَشُدَّكُمْ ثُمَّ لِتَكُونُوا
شُيُوخًا وَمِنكُم مَّن يُتوَفَّى مِن قَبْلُ وَلِتَبلُغُوا أَجَلا
مُّسَمًّى وَلَعَلَّكُمْ تَعْقِلُونَ)

صدق الله العظيم

سورة غافر الآية (67)

DEDICATION

I dedicate this research to:

My mother's Soul

My Dearest Father

My Brothers and Sisters who always wishing me all the success.

ACKNOWLEDGEMENT

I wish to express my deep thanks and gratitude to Allah for giving me the ability and mercifulness to do and complete this work.

I would like to express my sincere profound gratitude, appreciation and debt to Dr. Adeil Musa Younis**,** for his continuous supervision, kindness, encouragement and skillful guidance throughout this research. To him I should express my deep and sincere gratitude.

My deep thanks to Dr. Kkalid Ram Allah. Thanks are also due to Dr Mustafa Mohammed, Dr Arbab, and Dr. Mohammed El Amin for helping me during the stages of this research.

Appreciation and many thanks are due to my colleagues in University of Al Zaiem Al Azhari, Al Gazira University and University of El Neelain for their help and cooperation during the preparation of this research.

Finally, I cannot forget to express my thanks and love to my family for their constant help to me morally during my study.

ABSTRACT

Estimation and forecasting for indicators of mortality are of great importance for the decision-making, it is considered to be the main economic and social planning factor which contributes to the development of policies and Strategies in the long term. The problem of this study is that, there is no studies on mortality and life expectation at birth (e_0) as an indicator of mortality levels, the studies on life expectation at birth (e_0) are scarce in Sudan due to the problem of incomplete vital records of mortality beside, the high rates of illiteracy along together with some other reasons have contributed to the problem of the incompletion of the records .The importance of this study is that, there is no study on estimating and forecasting life expectation at birth (e_0) using demographic and statistical models in Sudan because such study needs accurate data therefore, this study will be of great benefit in the areas of statistics and its applications. The study aimed to apply statistical and demographic models used in the field of mortality and adapted to be use in Sudan for estimating and forecasting life expectancy at birth (e_0), also the attempt to correct defective data in mortality to study the phenomenon understudy in Sudan using demographic and statistical models(and compare between them). The use of population censuses through indirect estimation techniques, to determine levels, trends and patterns of mortality in Sudan .The hypotheses of the study are: It is mathematically plausible to derive an adjusted model life table to estimate the mortality in Sudan. The time series of life expectation at birth (e_0) in Sudan for period (1973 – 2008) is stationary. There is no significant difference between statistical and demographic models to forecast the phenomenon understudy. The research methodology comprises of descriptive and inferential statistical methods also used demographic techniques. The data were obtained from Sudan Population Censuses (1973 - 1983- 1993- 2008)., using programs (Statistical Package for Social Sciences, Excel, Eviews and Spectrum) in order to reach results. The study found that, the levels of mortality in Sudan are high compared to developed countries. Life expectation at birth (e_o) increased by about 14 years for both sexes, 13 years for males and 12 years for females during the period (1973-2008). When we look to the pattern of deaths in Sudan at that time, we found that, the central death rate (nM_x) by age groups took the form of the Latin letter (U). The study also used demographic models and (Box - Jenkes time series - ARIMA) as statistical model to forecast the life expectation at birth (e_o) during the period (2009-2020) for each of the (both sexes - males - females). The study found that, the time series is not stationary for both sexes and males, but the series is stationary for females, in order to make time series stationary for both sexes by taking the second difference of the series, the best model to forecast (both sexes) is ARIMA (0,2,1), as for males it stabilized the series after taking the first difference and best model to (males) is ARIMA (1,1,0). For females the best forecasting model is ARMA (1, 0), .The study made comparison between demographic and statistical models, which is better in forecasting the life expectation at birth (e_o) for each of the (both sexes- males - females), as one of the objectives of the study. The study concluded that, the Statistical model is better than demographic model when data are homogeneous, used in several statistical criteria's to compare between statistical and demographic models. This study affirmed the importance of demographic indicators reached. It is necessary to develop demographic package programs such as (MORTPAK Package and CHADMOR Package) in order to apply the modern demographic models in a country like Sudan

suffering from deficient and incomplete data problems. The necessity to increase the volume of data and remove all the discrepancies of time series data before the application in order to raise the efficiency of each of the models Box –Jankins and other demographic models in life expectation at birth (e_0) analysis. Separate data studies should be performed when the data are homogeneous and another study when the data are heterogeneous focuses on the comparison of statistical and demographic models to forecast life at birth (e_0) using other models not used in this study.

المستخلص

يعتبرالتقدير والتنبؤ بمؤشرات الوفاة ذات اهمية عظيمة في اتخاذ القرارات ، فهي مركزالتخطيط الاقتصادي و الاجتماعي مما يساهم في وضع السياسات والاستراتيجات في المدى الطويل.ومن الملاحظ ان الدراسات عن الوفيات في السودان قليلة. ومن هنا تكمن مشكلة الدراسة في عدم اكتمال السجلات الحيوية الخاصة بالوفيات ولارتفاع معدلات الامية والى جانب اسباب اخرى التي جعلت هذه السجلات غير مكتملة مما ادى ان الدراسات عن الوفيات وتوقع الحياة عند الميلاد(e_0)كمؤشر لمستويات الوفاة في السودانانادرة. وترجع اهمية هذه الدراسة في انه لا توجد دراسة عن تقديرات والتنبؤ بتوقع الحياة عندالميلاد(e_0) مستخدمة النماذج الديموغرافية والاحصائية في السودان لانهاتحتاج لبيانات دقيقة لذا فان هذه الدراسة ستكون ذات فائدةعظيمة في مجالات الاحصاءوتطبيقاتها. تهدف هذه الدراسة الى تطبيق النماذج الاحصائية والديموغرافية المستخدمة في مجال الوفيات في تقدير والتنبؤ بتوقع الحياة عند الميلاد (e_0) في السودان ، وايضا اصلاح البيانات المعيبة باستخدام تلك النماذج لدراسة الظاهرة قيدالدراسة والمقارنة بين تلك النماذج، اضف الى ذلك استخدام تعدادات السكان والمساكن في السودان من خلال طرق التقدير الغير مباشرة في تحديد مستويات واتجاهات وانماط الوفيات في السودان وقد افترضت الدراسة عدة افتراضات: من الممكن رياضيا استخلاص نموذج جدول الحياة المعدل لتقدير الوفاة في السودان،السلسلة الزمنية لتوقع الحياة عند الميلاد (e_0) في السودان خلال الفترة(1973 - 2008)هي ساكنة .لا يوجد فرق معنوي بين النماذج الإحصائية والديموغرافية للتنبؤ بهذه الظاهرة قيد الدراسة .اعتمدت الدراسة على بيانات تعداد السكان والمساكن في السودان خلال الفترة (1973- 1983- 1993 -2008). استخدمت الدراسة في تحليل البيانات الاحصاء الوصفي والاحصاء الاستنتاجي مستخدمة برنامج(الحزم الاحصائية للعلوم الاجتماعية ،الاكسل ،Eviews, Spectrum) من اجل التوصل للنتائج. وقد توصلت الدراسة الى عدة نتائج هي : ان مستويات الوفيات في السودان مرتفعة مقارنة بالدول المتقدمة، وان توقع الحياة عند الميلاد (e_0) زاد بحوالي 14 سنة للجنسين، 13 سنة للذكور و 12 سنة للاناث خلال الفترة (1973- 2008).

وبالنظرلنمط الوفيات في السودان في تلك الفترة وجدنا ان معدل الوفاة المركزي حسب العمر (nM_x) اخذ شكل الحرف اللاتيني (U). كذلك استخدمت الدراسة النماذج الديموغرافية ونموذج (Box – Jenkes time series -ARIMA) كنموذج احصائي للتنبؤ بتوقع الحياة عند الميلاد (e_0) خلال الفترة (2009- 2020) لكل من(الجنسين– الذكور–الاناث). وقد توصلت الدراسة ان السلسلة الزمنية غير مستقرة لكل من الجنسين والذكور بينما السلسلة مستقرة عند الاناث ، ومن أجل جعل السلسلة الزمنية مستقرة بالنسبة للجنسين اخذ الفرق الثاني للسلسلة وكان افضل نموذج للتنبؤ(الجنسين) هو(ARIMA (0,2,1 .اما بالنسبة للذكور فقد استقرت السلسلة بعد اخذ الفرق الاول وكان افضل نموذج للتبؤ (الذكور) هو(ARIMA (1,1,0 .اما الاناث فافضل نموذج للتنبؤهو (ARMA (1,0، كما قامت الدراسة بالمقارنة بين النماذج الديموغرافية والاحصائية ايهما افضل في التنبؤ بتوقع الحياة عند الميلاد(e_0) لكل من(الجنسين – الذكور – الاناث)كاحد اهداف الدراسة،وقد توصلت الدراسة الى ان النماذج الاحصائية افضل من النماذج الديموغرافية عندما تكون البيانات متجانسة ، مستخدمة في ذلك عدة معايير احصائية للمقارنة بين النماذج الاحصائية والديموغرافية.وقد اوصت هذه الدراسة باهمية المؤشرات الديموغرافية التي توصلت اليها، فلا بد من الاهتمام بها. ولا بد من تطوير برامج الحزم الديموغرافية مثل (MORTPAK Package and CHADMOR Package)حتى نتمكن من تطبيق النماذج الديموغرافية الحديثة في دولة مثل السودان لانها تعاني من مشاكل نقص وعدم اكتمال البيانات. وزيادة حجم البيانات وازالة كافة التغيرات من البيانات السلسلة الزمنية قبل التطبيق لرفع كفاءة كل من نماذج Box –Jankins والنماذج الديموغرافية في تحليل توقع الحياة عند الميلاد(e_0). لا بد من دراسات منفصلة للبيانات عندما تكون البيانات متجانسة و دراسة اخرى عندما تكون البيانات غير متجانسة تركز على المقارنة بين النماذج الاحصائية والديموغرافية لتبوء بتوقع الحياة عند الميلاد باستخدام نماذج اخرى غير مستخدمة في الدراسة.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter One

Introduction

1.1 Background of the Study :

Estimating and forecasting methods of mortality are rapidly gaining recognition. Models for estimates and forecasts have been produced for the US, Australia and several European countries. In the Netherlands, official statistical agencies now use forecasting methods in mortality, and other countries, such as Germany, the US, are also adopting them in official forecasts.

Estimates and Forecasts of the size and structure of the population are central to social and economic planning, from the provision of services in the short term to policy development in the long term.

Not least of the demographic challenges facing developed countries is the rapid ageing of the population. Already developed-country populations are experiencing unprecedented large elderly proportions.

Estimating and forecasting of population must take proper account of all the three components of demographic change, which are mortality, fertility and migration. Mortality estimating and forecasting have received considerable attention in recent years. Methods for estimating and forecasting fertility and migration are less well developed: as with human behavior in general, these demographic behaviors are difficult to forecast. A further problem for demographic forecasting is the estimation of uncertainty:

Estimates may vary considerably depending on the method of estimation (Keilman., 2001).

Any forecasting and estimating exercise presupposes that the data on which the forecast and estimate are based exist in suitable form. For mortality, this is generally the case for developed countries: vital registration provides lengthy series of data with the necessary detail for that study of estimating in length of life and forecasting of mortality in developing countries are rare or almost non-existent.

There is no method that gives complete satisfactory results in estimating mortality. What direct methods for measuring mortality are available? Survey data on deaths that occurred in the last year and registration data are often available, but in developing countries these data are rather poor. Although many surveys have been undertaken, none has given satisfactory results, since they are affected by problems both reference period and omission. Civil registration is also generally deficient, expect in a few developing countries such as Chile. It is possible to attempt to correct omissions in civil registration by comparing the estimation of childhood mortality with direct estimation obtained from registration or from survey data, and to use this comparison to estimate the overall incompleteness of the effective data. It can be seen therefore, that estimation of adult mortality presents a very difficult problem. (Brass, W.,1975).

One approach would be to select a model life table with a childhood mortality level, such a model life table could easily be selected from one of the available sets reference tables; I,e; the United Nations Model Life Tables, the Regional Model Life Tables, and those based on the Logic System. However, whichever of the model systems is used, this approach is not satisfactory because the relationship between childhood mortality and adult mortality is not very strong. Populations with the same mortality level for the adult population many have very different levels of childhood mortality. For example Turkey has a relatively low adult in comparison with its childhood mortality, (Brass,W.,1975).

It is significant that two seemingly opposed designations have been applied to the same thing; "Life Table" .Quite in accord with this dual character of the life table, applications may be broadly classified into two categories-applications relating primarily to mortality and death rates, and applications relating to survivals (Deevey,S,E,L.,2008).

Mortality is general cause of the physical notion of a human population. Mortality is object of great attention and analysis by demographers. A demographer constructs a life table in order to describe the survivorship of a population subjected to the risk of death. Such a table describes the effect of mortality. An abridged life table presents the mortality pattern by age groups. The problem of estimating a complete life table, when the data are provided in age intervals has been extensively discussed in demographic, biostatistical, as well as actuarial literature. The main reason for providing data in an abridged form is related to the phenomenon of age heaping, caused by age misstatements in data registration. Another typical reason is the unstable documentation of vital statistics based on insufficiently small samples. The most typical age-misstatements are the preference of ages heaps.

Life table can be extended to include other information in addition to mortality, for instance health information to calculate health expectancy, health expectancies, of which disability- free life expectancy (DFLE) and Health Life Years (HLY) are the best–known examples, are the remaining number of years a person can expect to live in a specific health state, such as free disability. Also there is a curvilinear relationship between period life expectancy at birth (e_0) and gross domestic product per capita (GDP).

Life tables can also be extended to show life expectancies in different labor force states or marital status states. Life tables are also used extensively in biology and epidemiology. The concept is also of importance in life product life cycle management (Deevey, S,E,J.,1988) .As well as life tables used in livestock and agriculture of various kinds also useful for insurance companies.

There is a close relationship between life tables and forecasting mortality, Where you can find all of forecasting of mortality through the parameter e_0 (life expectation at birth). It must be here to take advantage of the parameter e_0 and therefore this study will focus on the life tables to find forecasting of mortality and thus will benefit from these parameters of life tables and used in statistical models to find forecasting of mortality.

Actually in the fields of Demography and Actuarial Science, there have been many attempts to find an appropriate model that represents mortality. One of these successful attempts is Box-Jenkins models, also known as autoregressive moving average process (ARIMA) (Box y Jenkins, 1976), are usually used on forecasting.

1 - 2 Statement of the Problem:

There are problems in estimating and forecasting measures of mortality and mortality- related knowledge as well, especially in developing countries, including Sudan and that ascribed to lack of or incomplete data and records of deaths because of illiteracy and the lack of the required capabilities despite their importance in the policies and strategies of the country for their key role they play in financial and economic affairs in the areas of health, insurance companies and other. Vital registration provides lengthy series of data with the necessary detail for that study of estimating in length of life and forecasting of mortality in developing countries are rare or almost non-existent. Moreover, although statistical models to estimate and forecasting of the life expectation at birth (e_0) accurate, but most of the studies used demographic models, there is no study on life expectation at birth (e_0), using Univariate Time Series Modeling {Autoregressive Integrated Moring Average (ARIMA) models} in Sudan.

1-3 Research Important:

The concept of the life table originated in longevity studies of man, where it was always presented as a subject peculiar to public health, demography and actuarial science. As a result, its development has not received sufficient attention in the field of statistics.

Also among the few demographic studies on Sudan, mortality has not been a subject of intense attention. Studies conducted on mortality in Sudan have focused exclusively on child or maternal mortality, using mostly survey data. In fact, studies about mortality and life expectation are never taken into account due to lack of reliable detailed information.

There are a few studies of life expectancy that estimate and forecast mortality in developing countries, so this study will be great benefit in the areas of statistics and its applications. The study presents a straightforward and potentially useful way to think about the components of a difference in life expectancies in two complementary senses. All formulas present the discrete case most directly applicable to observed empirical data. Conversion to the more abstract continuous case would be a simple matter for those who prefer more precise mathematical expressions. On the one hand, we explore the ages at which a difference in life expectancy originates. Important policy decisions are made today based on forecasts of the elderly population far into the future. So the life expectancy is very important to estimating and forecasting mortality so we will use it for this purpose using (1973 – 1983-1993-2008) Sudan population census data.

1-4 Research Objectives:

This section includes search objectives, which are as follows:

The main objective of this research is to apply statistical and demographic models used in the field of mortality and adapted for use in Sudan for estimating and forecasting life expectation at birth e_0.

The sub objectives of this research are as follows:

1- An attempt to correct defective data in mortality to study the phenomenon understudy in Sudan using demographic and statistical models (and compare between them).

2- To use of population censuses to appraise through indirect estimation techniques, to determine levels, trends and patterns of mortality in Sudan.

1-5 Research Hypotheses:

1-It is mathematically plausible to derive an adjusted model life table to estimate the mortality in Sudan.

2-The time series of life expectation at birth (e_0) in Sudan for period (1973 – 1983-1993- 2008) is stationary.

3- There is no significant difference between statistical and demographic models to forecast the phenomenon understudy.

1-6 Research Methodology:

The scientific approach attempts to answer the problem of estimating and forecasting (e_0) in Sudan will be as follows:

The methodology of this study is descriptive and inferential statistics. Demographic models such as Brass technique methods are used to find the value of life expectation. Also Statistical models such as Box –Jankins (ARIMA) time series model is adopted for forecasting mortality. All demographic and statistical measures are calculated by (Mortpak Package, SPSS, Excel and E-Views program system).

1-6 -1 Data Used:

We used longitudinal settings study. The study will used secondary data, which it will be gathered from Censes Data (1973 – 1983 – 1993 – 2008) for Sudan. All the population of Sudan (both sexes, males and females) will be included in the study.

The data used to estimate mortality indicators presented in this study are as follows:

1-The number of children ever born classified by sex and five –year age groups of mother.

2-The number of children surviving classified by sex and five –year age groups of mother.

3- The number of births during the year preceding the census classified by five year age group of mother.

4- Population by sex according to survival status of mother.

5-Deaths during the year preceding the census classified by five year age group of mother.

6-Population distribution by sex and age group.

1-6 -2 Data Evaluation:

Experience in data collection over all censuses shows that, all the mortality data which mentioned above except deaths during the 12 months preceding the census are expected to be reliable .The data on deaths during the 12 months preceding the census is suspected to be under enumerated. However, for the sake of comparison of mortality estimates obtained from different types of data and for the sake of further evaluation of data on deaths during the 12 months preceding the census mortality indicators based on these data have also been presented.

There are some indicators may be used to evaluate data on children ever born ,children surviving and female population, such as sex ratios of children ever born is expected to range between (102–107) males per 100 females for all age groups of mothers. Also the average number of children ever born is expected to increase by age of mother. The proportion of children dead is expected to be higher for mother in younger and older ages than those in middle ages. In general, the quality of mortality data obtained from Sudan Population Censuses (1973 -1983–1993 – 2008) is acceptable.

1 – 7 Previous of the Studies:

This session will discuss previous related studies that touched on modeling, estimating and forecasting mortality. The study tried to discuss previous studies according to the models used .

There are plenty of studies that focused on the events that directly affect population dynamics, and the most important demographic events directly affect population dynamics are fertility and mortality. Studies of mortality are facing many problems, especially in developing countries, due to incomplete data from censuses or lack of vital records. So the studies of mortality with respect to contrast and predict deaths are rare and few, despite the fact that there are a lot of good and useful studies in this area.

1 – 7- 1 Life Tables Models:

Andersson and Philipov (2002) studied Life-table representations of family dynamics so they presented a system of descriptions of family-demographic behavior in developed countries. It used life-table techniques in order to describe the experience of men, women, and of children in processes related to family formation and family dissolution. It was developed a large number of descriptive measures, and applied them to survey data from Sweden, Norway, Finland, France, the USA, Austria, Germany, Flanders, Italy, Spain, the Czech Republic, Hungary, Slovenia, Latvia, Lithuania, and Poland, in order to describe patterns in the family demographic behavior during the late 1980s and early 1990s. It used Sweden and Hungary as examples when presenting the outline of our system of tabulations and provide results for the whole set of countries in an appendix to the paper.[1]

Ediev (2003) had a study on Monotonic Convergence to Stability, the paper introduced age-distribution population vectors, a class of distances to the stable equivalent; each of these distances converges monotonically to zero as the population approaches stability. The Kullback distance, considered earlier to be unique as a measure with this property, turns out to be a mere specimen of this class. It was shown that the very feature of monotonic convergence agrees with demographic potential of age specific stabilization measures. The paper also introduced a class of monotonic measures of the convergence to each other of two non-stable populations with similar reproduction regimes. Numerical illustrations and demographic applications were used as tools. This paper used the classical population projection model in a form more convenient for the purposes of the study.[2]

Vladimir, et al (2003) This paper presented a toolkit for measuring and analyzing inter-individual inequality in length of life by Gini coefficient. Gini coefficient and four other inequality measures were defined on the length-of-

---

1- Andersson, G., Philipov, D., 2002. " study, Life-table representations of family dynamics in Sweden, Hungary, and 14 other FFS countries", www.demographic-research.org .,2002.
2- Ediev D., 2003. "Study on Monotonic Convergence to Stability", www.demographic-research.org/Volumes/Vol8/2/, 22 January 2003.

life distribution. Properties of these measures and their empirical testing on mortality data suggest a possibility for different judgments about the direction of changes in the degree of inequality by using different measures. A new computational procedure for the estimation of Gini coefficient from life tables was developed and tested on about four hundred real life tables. New formulae had been developed for the decomposition of differences between Gini coefficients by age and cause of death. A new method for decomposition of age components into effects of mortality and composition of population by group was developed. Temporal changes in the effects of elimination of causes of death on Gini coefficient were analyzed. [1]

Merrill (2004) studied Life Expectancy among LDS and Non-LDS in Utah. This study showed that, LDS males and females experience higher life and age-conditional life expectancy compared with non-LDS in Utah. Only some of the higher life expectancy in LDS is explained by historically lower use of tobacco among LDS. Tobacco-related deaths had a larger impact on the difference in life expectancy when conditioned on older ages because of the long latency period often involved with these tobacco-related diseases. Other explanations for the higher life expectancy in LDS may be related to their comparatively high level of church activity, with religiously active people often displaying better physical health, better social support, and healthier lifestyle behaviors, each of which promote longer life. Religious activity may also have an independent protective effect against mortality.[2]

1- Vladimir M. Shkolnikov, Evgueni E. Andreev, Alexander Z. Begun., 2003. “Gini coefficient as a life table function: computation from discrete data, decomposition of differences and empirical examples”, www.demographic-research.org/Volumes/Vol8/11/, 11-8-2003.

1- Merrill, R. M., 2004. Life Expectancy among LDSand Non-LDS in Utah. www.demographic-research.org/Volumes/Vol10/3/, 12 MARCH 2004.

Schoen and Canudas- Romo (2005), highlighted that Period life expectancy varies with changes in mortality, and should not be confused with the life expectancy of those alive during that period. Given past and likely future mortality changes, a recent debate has arisen on the usefulness of the period life expectancy as the leading measure of survivorship. An alternative aggregate measure of period mortality which has been seen as less sensitive to period changes, the cross-sectional average length of life (CAL) has been proposed as an alternative, but has received only limited empirical or analytical examination. Here, the study introduced a new measure, the average cohort life expectancy (ACLE), to provide a precise measure of the average length of life of cohorts alive at a given time. To compare the performance of ACLE with CAL with period and cohort life expectancy, first it used population models with changing mortality. Then the four aggregate measures of mortality are calculated for England and Wales, Norway, and Switzerland for the years 1880 to 2000. CAL is found to be sensitive to past and present changes in death rates. ACLE requires the most data, but gives the best representation of the survivorship of cohorts present at a given time.[1]

Joshua R. Goldstein.,(2006), the study compared the conventional life expectancy with two other possible measures .Bongaarts and Feeney (2002) considered a specific, closed to migration and with constant number of births annually population with frozen mortality conditions. It considers different measures of life expectancy and compares them for specific populations using stochastic ordering of the corresponding random variables. The study found that stochastic comparisons of populations should be explored to a greater

1-Schoen R and Vladimir Canudas-Romo.,2005, “Changing mortality and average cohort life expectancy” http://www.demographic-research.org/Volumes/Vol13/5/.,05 OCTOBER 2005.

extent in the future studies. This gives possibility to look at the problem in a more general way. It showed that, when mortality rates are declining in time, the conventional observed life expectancy substantially overestimates the life expectancy for this model, obtained using numerical methods. In the current paper the study proves analytically that conventional life expectancy also majors the average age at death. It considers stochastic comparisons of age distributions for specific populations and obtains inequalities for the corresponding mean values as a simple consequence. [1]

Congdon (2006),[1] developed a joint approach to life and health expectancy based on 2001 UK Census data for limiting long term illness and general health status, and on registered death occurrences in 2001. The model takes account of the interdependence of different outcomes as well as spatial correlation in their patterns. A particular focus is on the proportionality assumption or ’multiplicative model’ whereby separate age and area effects multiply to produce age-area mortality rates. Alternative non-proportional models are developed and shown to be more parsimonious as well as more appropriate to actual area-age interdependence. The application involves mortality and health status in the 33 London Boroughs.[2]

Carlson (2006), stated that decomposition of a difference in life expectancies may identify ages at which difference originates in mortality differences, or may identify ages at which difference results in different values of e_0. This study shows that, the two approaches are orthogonally related to each other,

[2-]Joshua R. Goldstein., 2006"A cohort perspective on tempo-adjusted life expectancy". http://www.demographic-research.org/Volumes/Vol14/5/.,2006.
2- Congdon P. , 2006 "A model for geographical variation in health and total life expectancy"
,http://www.demographic-research.org/Volumes/Vol14/9/,Research Article. , 2006.

and derives an origin-destination decomposition matrix in which summing in one direction produces Andreev's origin-decomposition results, while summing in the other direction produces destination-decomposition corresponding to directly-observed differences in nLx values.[1]

Scand (2007), examined trends in age-specific mortality in a rural South African population from 1992 to 2003, a decade spanning major sociopolitical change and emergence of the HIV/AIDS pandemic. Changing mortality patterns are discussed within a health-transition framework. Data on population size, structure, and deaths, obtained from the Agincourt health and demographic surveillance system, were used to calculate person-years at risk and death rates. Life tables were computed by age, sex and calendar year. Mortality rates for the early period 1992–93 and a decade later, 2002–03, were compared. The study findings that, demonstrate significant increases in mortality for both sexes since the mid-1990s, with a rapid decline in life expectancy of 12 years in females and 14 years in males. The increases are most prominent in children (0–4) and young adult (20–49) age groups, in which increases of two- and fivefold respectively have been observed in the past decade. Sex differences in mortality patterns are evident with increases more marked in females in most adult age groups.[2]

Canudas-Romo (2008)., described that the closure of Mongolia to the international community during the 20th century resulted in a dearth of available data and analytic demographic studies. In the absence of mortality analysis during the socialist period, this paper proposes the use of indirect

---

[1] : Carlson E.,2006 “Ages of origin and destination for a difference in life expectancy”, http://www.demographic-research.org/Volumes/Vol14/11/.,2006

2 - Scand J.,2007, “Public Health Suppl ,Mortality trends in a new South Africa: Hard to make a fresh start”., August, 2007.

census-based techniques to estimate mortality levels and trends of the last two socialist decades (1969-1989). Due to census data quality and choice of model life table, results are not homogeneous. The respective effects of these two components are discussed in order to understand the results. However, despite these shortcomings, it is shown that during the last socialist decades in Mongolia, the health conditions of the population deteriorated. The Mongolian pattern is relatively similar to the situation documented for the ex-socialist republics. Causes of this similarity are discussed.[1]

Tymicki (2009), his study has two main goals. The first is to review the context for studying infant mortality, which includes a review of the theoretical framework, the covariates used to examine mortality over the first 60 months of life, and the major findings of empirical studies. Second, to add some new empirical evidence that comes from the longitudinal reconstitution of church registers of Bejsce prish, located in the south of Poland. This rich database allows for an analysis of mortality trends of cohorts born between them 18^{th} and 20^{th} centuries in the parish. The analysis includes a reconstruction of descriptive measures of infant and childhood mortality, and a hazard model of mortality over the first 60 months of life. The hazard model has been calculated for each cohort separately in order to demonstrate the change in the relative importance of analyzed factors during the process of mortality decline in the parish.[2]

Cai *et al* (2010), highlighted that the multistate life table (MSLT) model is an important demographic method to document life cycle processes. In this

1-Canudas- mortality hypothesis, Vladimir".,July 2008. "The modal age at death and the shifting Romo.,2008,

[2] -: Tymicki K., 2009 "Correlates of infant and childhood mortality: A theoretical overview and new evidence from the analysis of longitudinal data of the Bejsce (Poland) parish register reconstitution study of the 18th-20th centuries", VOLUME 20, ARTICLE 23, PAGES 559-594 , PUBLISHED., 26 MAY 2009.

paper, they present the SPACE (Stochastic Population Analysis for Complex Events) program to estimate MSLT functions and their sampling variability. It has several advantages over other programs, including the use of micro-simulation and the bootstrap method to estimate the variance of MSLT functions. Simulation enables researchers to analyze a broader array of statistics than the deterministic approach, and may be especially advantageous in investigating distributions of MSLT functions. The bootstrap method takes sample design into account to correct the potential bias in variance estimates.[1]

Waddad., (2010), concerned that with asocial study, targeted towards maternal deaths in Sudan. The aim was to combine M.M Ratio with the Orphanhood Method to estimate adult mortality; with view of improving maternal mortality estimation indirectly .The paper used data on mortality collected in Sudan2008 census. The MM Ratio was augmented with the Orphanhood Method, using Logit system, in one model in order to estimate MM Ratio and to build a life - table for women in aged 15 - 49 in Sudan. Value of R^2 and r, suggested that, the whole variation on the dependant variable MM Ratio explained by the variation on the probability of dying, with a very strong positive relationship. The census data suggested a MM Ratio of 508 per 100,000 live births, the ratio raised to 541 per 100,000 live births using the combined method.[2]

Abu El-Yamen et al (2010),presented the analysis of mortality data obtained from the 2008 population census .The analysis discussed the estimates of

1- Cai L , Mark D.Hayward, Saito A , Lubitz J, Hagedorn A , Crimmins E.,2010, “Estimation of multi-state life table functions and their variability from complex survey data using the SPACE Program”, http://www.demographic-research.org/Volumes/Vol22/6/ .,2010.

2 - Waddad A.,2010, “Combined Maternal Mortality Ratio and Orphandhood Technique to estimate Maternal Mortality logit System From Sudan 2008 Census data”, Ministry of the Cabinet Central Bureau of Statistics.,2010

mortality indicators for national level, Northern Sudan ,Southern Sudan and states .The discussion included mortality levels, trends and differentials derived from the final results of the 2008 population census. Different methods of estimation were used to calculate mortality indicators listed Brass methods to estimate child mortality , Brass methods to estimate female adult mortality indicators from data on maternal Orphanhood , estimation of male adult mortality for male and female child mortality and female adult mortality as used in Chadmor and life table.[1]

Ediev (2013), had introduced, for age-distribution population vectors, a class of distances to the stable equivalent; each of these distances converges monotonically to zero as the population approaches stability. The Kullback distance, considered earlier to be unique as a measure with this property, turns out to be a mere specimen of this class. It is shown that, the very feature of monotonic convergence agrees with demographic potential of age specific stabilization measures. The paper also introduces a class of monotonic measures of the convergence to each other of two non-stable populations with similar reproduction regimes. Numerical illustrations and demographic applications conclude the paper.[2]

1 – 7- 2Time Series and Box-Jenkins Models:

Lee *et al* (1992) [2] developed time series methods and used to make long-run forecasts, with confidence intervals, of age-specific mortality in the United States from 1990 to 2065. First, the logs of the age-specific death rates are modeled as a linear function of an unobserved period-specific intensity index,

---

1 - Abu El-Yamen S, Musafa I, Ibrahim S.,2010, “Analysis of Mortality Levels Trends and differentials Census Data” ,Ministry of the Cabinet Central Bureau of Statistics.,2010
1-Ediev .,2013, “Modeling and Forecasting U.S. Mortality”., Sep1992.

with parameters depending on age. This model is fit to the matrix of U.S. death rates, 1933 to 1987, using the singular value decomposition (SVD) method; it accounts for almost all the variance over time in age-specific death rates as a group. Whereas e_0 has risen at a decreasing rate over the century and has decreasing variability, k (t) declines at a roughly constant rate and has roughly constant variability, facilitating forecasting K (t), which indexes the intensity of mortality, is next modeled as a time series (specifically, a random walk with drift) and forecast. The method performs very well on within-sample forecasts, and the forecasts are insensitive to reductions in the length of the base period from 90 to 30 years; some instability appears for base periods of 10 or 20 years, however. Forecasts of age-specific rates are derived from the forecasts of k, and other life table variables are derived and presented.[1]

TuljapurkarS et al.,(2007), discusses three widely used classes of methods for national population forecasts. They are demographic accounting equations, statistical time series, and structural modeling methods. This paper combines the first two of these techniques. An advantage of the time series approach is that forecast variances can be derived. In this paper he assumes that time series methods have been used -to model and to forecast fertility, survival, and migration rates. Either univariate (Box and Jenkins, 1970) or multivariate (Tiao and Box, 1981) autoregressive-integrated-moving average (ARIMA) time series models could be employed. This paper shows how to use the accounting equations to calculate the mean and the variance of future populations. [2]

---

[1] - LEE R, LAWRENCE R. CARTER., 1992, “Modeling and Forecasting U.S. Mortality”., Sep1992

2 - TuljapurkarS , Ryan D. Edwards. , 2007 ,”Variance in Death and Mortality Decline”., September 20, 2007

Chen and Cox (2007), stated in this paper, it incorporated a jump-diffusion process into the original Lee-Carter model, and uses it to forecast mortality rates and analyze mortality securitization. The outlier-adjusted Lee-Carter model is examined to provide further evidence of mortality jumps. It also explore alternative models with transitory versus permanent jump effects. The paper uses the Swiss Re mortality bond as an example to show how to apply the model and the distortion measure approach to value mortality-linked securities. Pricing the Swiss Re mortality bond is difficult because the mortality index is correlated across countries and over time. In this paper, it has a deep discussion in mortality modeling and mortality-linked security pricing.[1]

Carter .Rlawrence., (2008) stated that, this paper compares two methodologies for modeling and forecasting statistical time series models of demographic processes: Box-Jenkins ARIMA and structural time series analysis. The Lee-Carter method is used to construct nonlinear demographic models of U.S. mortality rates for the total population, gender, and race and gender combined. Single time varying parameters of k, the index of mortality, are derived from these models and fitted and forecasted using the two methodologies. Forecasts of life expectancy at birth, e_0, are generated from these indexes of k. Results showed marginal differences in fit and forecasts between the two statistical approaches with a slight advantage to structural models. Stability across models for both methodologies offers support for the robustness of this approach to demographic forecasting**.** [2]

1 - Chen and Cox .,2007, " a jump-diffusion process into the original Lee-Carter model"

[2] - Carter .Rlawrence (2008).,Forecasting U.S. Mortality: A Comparison of Box-Jenkins ARIMA and Structural Time Series Models. 28, June, 2008.

Adhistya Erna Permanasari at el (2009),This paper analyzed and presented the use of Seasonal Autoregressive Integrated Moving Average (SARIMA) method for developing a forecasting model that able to support and provide prediction number of zoon sis human incidence. The dataset for model development was collected on a time series data of human Salmonellosis occurrences in United States which comprises of fourteen years of monthly data obtained from a study published by Centers for Disease Control and Prevention (CDC). The result showed that, the SARIMA(9,0,14)(12,1,24) is the fittest model. While in the measure of accuracy, the selected model achieved 0.062 of Theil's U value. It implied that the model was highly accurate and a close fit. It was also indicated the capability of final model to closely represent and made prediction based on the tuberculosis historical dataset.[1]

Nan Li and Patrick Gerland.,(2012),this paper addressed that, the application of the Lee-Carter model to age-specific death rates by gender in Argentina. These rates are available for the period that goes from 1979 to 2006. The index of the level of mortality for each gender, and the shape and sensitivity coefficients for nine age groups were obtained through the Lee-Carter method. The autoregressive moving average (ARIMA) and the space-state (SSM) models are used to forecast the general index for the time period that goes from 2007 to 2011 in order to project life expectancy at birth using life tables.[2]

[1] - Adhistya Erna Permanasari at el .,2009, Prediction of Zoonosis Incidence in Human using Seasonal Auto Regressive Integrated Moving Average (SARIMA), (IJCSIS) International Journal of Computer Science and Information Security, Vol. 5, No. 1, 2009

2- Nan Li and Patrick Gerland., 2012"The Lee Carter method for estimating and forecasting of mortality: an application for Argentina" , National University of Rosario, Argentina , Faculty of Economics and Statistics., 2012

Smart A. Sarpong (2013), The study covered that, the feasibility for application of Box-Jenkins Approach to time series (ARIMA) in modeling and forecasting Maternal Mortality ratios (MMR). Analyses were based on data available at the Bio-Statistics Department of the Obstetrics& Gynecology directorate of the facility. The result showed that ,the hospitals (MMR) was relatively stable but had a very alarming average quarterly MMR of 967.7 per 100,000 live births which is about twice the National ratio of 451 per 100,000 live births. With AIC (581.41), it concluded that, the ARIMA (1,0,2) model is adequate for forecasting quarterly maternal mortality ratios at the hospital. [1]

Smart A. Sarpong (2013), examined maternal mortality ratios at the Okomfo Anokye Teaching Hospital in Kumasi from the year 2000 to 2010. The study explores the feasibility for application of Box-Jenkins Approach to time series autoregressive integrated moving average (ARIMA) in modeling and forecasting Maternal Mortality ratios (MMR). Analyses were based on data available at the Bio-Statistics Department of the Obstetrics and Gynecology directorate of the facility. The result shows that the hospitals Maternal Mortality Ratio (MMR) was relatively stable but had a very alarming average quarterly MMR of 967.7 per 100,000 live births which is about twice the National ratio of 451 per 100,000 live births With AIC (581.41), we conclude that the ARIMA (1,0,2) model is adequate for forecasting quarterly maternal mortality ratios at the hospital. [2]

Musa Mohammed (2013), Sudan is a sub-Saharan African country that is highly affected by malaria with 7.5 million cases and 35,000 deaths every

---

[1] - Smart A. Sarpong. ,2013, (Modeling and Forecasting Maternal Mortality; an Application of ARIMA Models) ,International Journal of Applied Science and Technology Vol. 3 No. 1., January 2013

[2] - Smart A. Sarpong.,2013 Modeling and Forecasting Maternal Mortality; an Application of ARIMA Models, School of Graduate Studies, Research and Innovations Kumasi Polytechnic –Ghana

year. (ARIMA) model was used to predict the spread of malaria in the Sudan. The ARIMA model used malaria cases from 2006 to 2011 as a training set, and data from 2012 as a testing set, and created the best model fitted to forecast the malaria cases in Sudan for years 2013 and 2014. The ARIMAX model was carried out to examine the relationship between malaria cases and climate factors with diagnostics of previous malaria cases using the least Bayesian Information Criteria (BIC) values. The results indicated that there were four different models, the ARIMA model of the average for the overall states is $(1,0,1)(0,1,1)^{12}$. The ARIMAX model showed that there is a significant variation between the states in Sudan**.** [1]

Ahmed I. Abdelghani at el (2013) this study aimed to developed a forecasting model on the number cerebrospinal meningitis case (CSM) in Sudan using time series analysis (SARIMA) model had been developed on the data collected between 2000 and 2010 and then the model validated using the data collected between January-December2011. The results showed that the regressive forecast curves were consistent with the pattern of actual values. The most suitable model was the SARIMA (1,0,0) ×(0,1,1)12 with adequate fitting for the data the Q test of white noise for the residuals (Q =33.46, p>0.758). Bartlett test for white noise based on the Cumulative Period gram of the residuals (B=36, p>0.9). In-sample forecasting was drawn for the year 2011. It showed no significance differences between the observed and model-predicted monthly CSM cases using paired t test. This indicated that there was no significant autocorrelation between residuals at different lag times in the SARIMA (1,0,0)×(0,1,1)12 model**.** [2]

---

[1] - Musa Mohammed , Malaria Disease Distribution in Sudan Using Time Series ARIMA Model (2013).

[2] - Ahmed I. Abdelghani at el.,April2013Predicting Cerebrospinal Meningitis in Sudan Using TimeSeries Modling SUDANESE JOURNAL OF PUBLIC HEALTH VOL. 8 No. 2,

Ekezie Dan Dan et al (2014), this paper examined the modeling and forecasting malaria mortality rate using SARIMA Models. Among the most effective approaches for analyzing time series data is the method propounded by Box and Jenkins, the Autoregressive Integrated Moving Average (ARIMA). This paper employed Box-Jenkins methodology to build ARIMA model for malaria mortality rate for the period January 1996 to December 2013 with a total of 216 data points. The model obtained in this paper was used to forecast monthly malaria mortality rate for the upcoming year 2014. The forecasted results will help Government and medical professionals to see how to maintain steady decrease of malaria mortality in other to combat the predicted rise in mortality rate envisaged in some months. [1]

1 – 7 - 3 Summaries of the above findings:

From the previous studies, we found that, the studies concerted on life table of population are important because there are many areas in need of such as live stock, Insurance, Minerals, oil fields and other. Given the limitations and disadvantages of previous studies, we noted that, it focused on Life tables studies mostly in Africa only for HIV/AIDS virus and its impact on the level of mortality, but Africa in need of in-depth studies in this area due to the high of mortality rates in that region. Also

these studies did not address any comparisons between statistical and demographic models to forecasting the life expectation at birth. As there are no studies in developing countries, including Sudan used statistical models to forecasting the life expectation at birth, so this study is a real addition at the

---

[1] - Ekezie Dan Dan, Opara Jude, Okenwe Idochi, Modelling and Forecasting Malaria Mortality Rate using SARIMA Models (A Case Study of Aboh Mbaise General Hospital, Imo State Nigeria), Science Journal of Applied Mathematics and Statistics. Vol. 2, No. 1, 2014, pp. 31-41. Doi.

regional and local level because it is trying to cover these aspects. Also previous studies did not include studies that show historical and economic analyses can benefit from an examination of forecasting in age at death in addition to the traditionally important study of life.

1-8 Research Organization:

The study consists of five chapters. Chapter one includes background, statement of the problem, important of the study, objectives of the study, hypotheses of the study, research methodology, pervious of the studies and research organization. Chapter two and chapter three that contains the definition of demographic and statistical models that will be used in this study namely, Brass technique methods, to find life expectation and Univariate Time Series Model (ARIMA) time series model is used to forecast mortality. Chapter four includes the statistical analysis of the data (Application of the study) to find the value of (e_0) of the population of Sudan, the following demographic models which it will be used Brass technique methods. Also includes statistical models and demographic technique methods for forecasting the life expectation at birth (e_0) in Sudan, where we will use statistical models Box – Jankins (ARIMA) model, this is followed by comparison of statistical and demographic models, whichever is best to forecast the life expectation at birth (e_0). Chapter five presents conclusion of the study and recommendations.

Chapter Two

The Demographic Models

2 -1 Preface:

This chapter discusses the demographic models that used in demographic analysis currently, which will be compared to statistic model, which will present in next chapter, it will describe genera1 demographic models, with a focus on life tables form, with interpretations of its various functions and present a method of constructing a current life tables.

2 -2 Demographic Models:

Demographic models are an attempt to represent demographic processes in the form of a mathematical function or set of functions relating two or more measurable demographic variables. The primary purpose of modeling is simplification, to reduce a confusing mass of numbers to a few, intelligible basic parameters, or to make possible an approximate representation of reality without its complexity.

Because all demographic models attempt to represent reality, they are based to a greater or lesser extent on actual data. Yet, according to their degree of dependence upon observed data, two broad categories can be distinguished. On the one hand, there are models that can be derived solely from a set of simple assumptions or postulates. An example is the stable population model which arises from assuming that fertility and mortality have remained constant for a sufficiently long period. The proof of the convergence of almost any initial age distribution to a stable state after being subject for a long time to constant fertility and mortality can be carried out mathematically without recourse to any type of observed data. Yet, a model of this sort is powerful only to the extent to which it reflects an actual process. Therefore, even though the stable population model is not derived from actual observations, only the good fit it provides to the age structures of populations that might reasonably be considered "stable".

On the other hand, there are models that could not be derived at all if suitable data were not available. In this category most of the model life tables that have been proposed. These models arise from the systematic analysis of observed mortality patterns and from the discovery and exploitation of common patterns present in them. In general, most models fall between these two extremes: that is, their basis is neither purely empirical nor purely theoretical. In fact, some have evolved from a purely empirical to a purely theoretical foundation. An example of this type of evolution is the nuptiality model first presented by Coale. Ansley J: Coale.

Unfortunately, the modeling of other demographic phenomena has not been equally successful. Yet, undoubtedly, the aim of researchers in this field is to arrive at models that, while being as economical as possible in the number of parameters they incorporate, are also flexible enough to approximate all the relevant variations observable in real data, and whose form and parameters have meaningful interpretations in reality. The next sections describe several useful demographic models in the areas of mortality, attention is focused on those models which have been relevant in the development of indirect estimation techniques, mainly because of the important part such models have played in allowing the simulation of demographic data. Measured without difficulty and. on the other hand, more useful demographic parameters whose values are not easily determined directly. [1]

2 -2-1 Mortality Models (Life Tables Models):

A life table provides a summary description of the effects of age-specific mortality rates upon a birth cohort. The very earliest demographic models attempted to describe in statistics and stochastic processes the variations of mortality with age, particularly the increase in the risk of dying after childhood. Attempts to describe by a single mathematical function mortality experience throughout life have found it difficult to reproduce the letter U or J shape of mortality rates by age. This difficulty leads to a new approach in creating mortality models or life table models. Instead of trying to relate the risk of dying solely to age, risks at a given age were related to the risks observed at other ages or to risks observed in other populations at similar ages.

Life tables is a demographic measurement of mortality levels, at different ages .It is generated from age specific death rates (ASDRs).A commonly used index of mortality is life expectation at birth, which measures the average number of years a birth is expected to survive in a population given the current mortality conditions of the population .The life expectation at birth, denoted by (e_0), for a population reflects the general health conditions of the population. So many problems in population study must be answered. Incidentally, it should be remarked that life table, as generally constructed, represent a fixed mortality of a particular calendar year or period. Such a table tells us what would be the number – of survivors to ages: (0), (1–4), (5–9) ,etc;

[1] - Alfred J *et al* (1972)., having been subject to unchanging fertility and mortality for many years

if the mortality at each age remained constant as of the calendar year or period for which it is constructed.

Because the relations explored have, in general, shed no light on a plausible theoretical interpretation of how the process of mortality occurs, most of the life table models existing to date depend heavily upon empirical data. At least four systems of life table models have been developed on the principle of narrowing the choice of a life table to those deemed feasible on the basis of examination of mortality risks calculated for actual populations. These systems vary in the range of human experience they encompass, so that one may be more appropriate than another for a particular case. Each of them is described in detail blew.

There are two forms of the life table in general: the cohort (or generation) life table and the current life table. In its strictest form:

A cohort life table: Records the actual mortality experience of a particular group of individuals (the cohort) from birth to the death of the last member of the group. The difficulties involved in constructing a cohort life table for a human population are apparent. Statistics covering a period of 100 years are available for only a few populations and even those are likely to be less reliable than current statistics. Individuals in a given cohort may have emigrated or died unrecorded, and the life expectancy of a group of people already dead is of little more than historical interest.

The current life table: as the name implies, gives a cross-section view of the mortality and survival experience of all ages in a population during one short period of time, for example, the California population of 1970. It is dependent entirely on the age-specific death rates prevailing in the year for which it is constructed. Such tables project the life span of each individual in a hypothetical cohort on the basis of the actual death rates in a given population. When we speak of the life expectancy of an infant born in a current year, for example, .We mean the life expectancy that would be obtained if he were subjected throughout his life to the same age-specific mortalities prevailing in the current year. The current life table is then a fictitious pattern reflecting the mortality experience of a real population during a calendar year. However, it is the most effective means of summarizing mortality and survival experience of a population, and is a sound basis for making statistical inference about the population understudy. The reader can no doubt confirm from his own experience that the current life table is a standard and useful tool for comparing international mortality data, and for assessing mortality trends on the national level.

The advantage of such a table is to reduce the possible abnorma1ities in mortality pattern which may exist in a single calendar year. Data for constructing, life tables are sometimes refined by graduation or other methods for smoothing or reducing, the effect of extreme values. Techniques for refinement of life tab1e data were developed by acuaria1scientists. While refinement of data has its merit in smoothing, data, it is difficult to make proper statistical inference of life table functions which is based on such information.[1]

[1] - Coale A and Demeny P .,1966,"Regional Life Tables Model and Stable Populations", Princeton: Princeton University Press.

2 -2-1 -1 Types of Life Tables:

2-2-1-1-1 United Nations Model Life Tables:

The first set of life tables models were developed by the Population Division of the United Nations Secretariat during the 1950s. This set, subsequently published in a revised form is based on a collection of 158 observed life tables for each sex. The life tables models were constructed by assuming that the value of each $5q_4$, (the probability of dying between age (x) and (x + 5) in a life table) is a quadratic function of the previous q value, namely, $5q_{x-5}$(except for the first two age groups, $1q_0$ and $4q_1$. all the other groups considered are five years in length). This assumption implies that the knowledge of only one mortality parameter (Iq_o or an equivalent "level" that indices the $1q_o$values used) would determine a complete life table. The United Nations life table models are thus a one-parameter system**.**[1]

Since the coefficients of the quadratic equations relating each $5q_x$value to its predecessor were not known priori, they had to be estimated on the basis of observed data. Regression was used to estimate these coefficients from the 158 mortality schedules available for each sex. Once they were estimated, calculation of the actual life table model was straightforward: a convenient value of $1q_0$ would be chosen arbitrarily; it would then be substituted in the equation relating $1q_o$to $4q_1$ so that a value of $4q_1$would be obtained which, in turn would be used to generate a value of $5q_5$through the equation relating $5q_5$ to $4q_1$ and so on.

[1] - United Nations. ,1981,"Life Tables Model for Developing Countries. New York: United Nations", (Sales No. E.81.XIII.7.

There are inherent disadvantages in the "chaining" technique used in calculating the United Nations tables model, especially when, as in this case, the quadratic equations relating one parameter to the next are not exact. Although the regression method of fitting equations to data does make some allowance for the existence of errors, it assumes that the distribution of these errors is known and that their mean is zero. The presence of systematic errors with non-zero mean (biases) may severely affect the estimates yielded by fitted regression equations, and their adverse effects are very likely to be augmented by the chaining technique described above whereby errors in one estimate would only contribute to accentuated errors in the next.

Furthermore, the 158 tables used as the data base from which the regression coefficients of the fitted equations were estimated were not all of the same quality. Because many tables from developing countries were used in the data base, mortality data with numerous deficiencies were included; and since large areas of the world did not possess any mortality statistics, life tables for those areas could not be included in the data base.

Owing to these shortcomings and to the fact that a one-parameter system lacks the flexibility required' to approximate adequately the variety of cases encountered in reality, the United Nations life tables were soon superseded by other model sets which, though based on a similar approach, tried to avoid the pitfalls encountered by the creators of the first system. However, the United Nations set established the usefulness of a life table model system.[1]

[1] - C.J.L.Murrary et al .,2010, GPE Discussion Paper Series:No8,EIP / GPE / EBD World Health Organization.

2-2-1-1- 2The Coale and Demeny Regional Model Life Tables:

These were first published in 1966. They were derived from a set of 192 life tables, by sex, from actual populations. This set included life tables from several time periods (39 from before 1900 and 69 from after the Second World War) and mostly from Western countries. Europe, North America, Australia and New Zealand contributed a total of 176 tables. Three were from Israel; 6 from Japan, 3 from Taiwan; and 4 from the white population of South Africa. All of the 192 selected life tables were derived from registration data, and were subjected to very stringent standards of accuracy.

Further analysis of the underlying relationships identified four typical age patterns of mortality, determined largely by the geographical location of the population, but also on the basis of their patterns of deviations from previously estimated regression equations. Those patterns were called: North, South, East, and West. Each had a characteristic pattern of child mortality. The East model comes mainly from the Eastern European countries, and is characterized by high child mortality in relation to infant mortality. The North model is based largely on the Nordic countries, and is characterized by comparatively low infant mortality, high child mortality and low old age mortality beyond age 50. The South model is based on life tables from the countries of Southern Europe (Spain, Portugal, and southern Italy), and has a mortality pattern characterized by (a) high child mortality in relation to infant mortality at high overall mortality, and (b) low child relative to infant mortality at low overall mortality. The West model is based on the residual tables not used in the other regional sets (i.e., countries of Western Europe and most of the non-European populations). It is characterized by a pattern intermediate between North and the East patterns. Because this model is derived from the largest number and broadest variety of cases, it is believed to represent the most general mortality pattern. In this system, any survivorship probability, whether from birth or conditional on having attained a certain age, uniquely determines a life table, once a family has been selected. Although technically a one parameter system, it could be argued that the choice of a family constitutes a separate dimension. [1]

2-2-1-1- 3 Coale and Demeny (1959 - 1969):

This system is based on a factor analysis of some 157 empirical tables. The method of selection was less rigid than in the Coale- Demeny tables, but they represent more developing country experiences. Analysis of the tables disclosed five factors that apparently explained a large proportion of the variability among the life tables. The extracted factors related to (a) general level of mortality, (b) relation between childhood and adult mortality, (c) mortality at older ages, (d) mortality under age five, and (e) male-female difference in mortality in the age range 5-70 years. Later, Lederman developed a series of one- and two-parameter model life tables based on these result. . [2]

2-2-1-1-4 Brass Logit System (1971):

This system provides a greater degree of flexibility than the empirical models discussed above. It rests on the assumption that two distinct age-patterns of mortality can be related to each other by a linear transformation of the logit of their respective survivorship probabilities. Thus for any two observed series of survivorship values, lx and l_x^s, where the latter is the standard, it is possible to

1 - C.J.L. Murray et al.,2011, GPE Discussion Paper Series: No. 8EIP/GPE/EBD, World Health Organization.
2 - C.J.L. Murray et al.,2011, GPE Discussion Paper Series: No. 8EIP/GPE/EBD, World Health Organization.

find constants β and α. So model will explain more clearly as one models that will be used in the analysis.[1]

2-2-1-1-5The UN Model Life Table for Developing Countries (1981):

These were designed to address the needs of developing countries. The underlying data consisted of 36 life tables covering a wide range of mortality levels from developing countries, by sex. Sixteen pairs of life tables came from 10 countries in Latin America, 19 pairs from 11 countries in Asia, and one pair from Africa. Five families of models were identified, each with a set of tables ranging from a life expectancy of 35 to 75 years for each sex. Each family of models covers a geographical area: Latin American, Chilean, South Asian, Far Eastern and a General. The general model was constructed as an average of all the observations.

2 -2 -2 Complete and Abridged Life Tables:

The major visual difference between the complete and abridged life tables lies in the age groupings for which the estimates have been produced.

In complete life tables, there is only one age value per row, which indicates the exact age for the number of survivors, the cumulative number of life years lived and the life expectancy. For the number of deaths, death and survival probabilities, as well as the number of life years lived, the interval in the life table represent the interval between two exact ages. That means, the presentation is the same in abridged life tables, but the age intervals are of the form (x, x+ (n-1)); that is, both ages (x) and (x+ (n-1) are included in the interval. Most age intervals in abridged life tables span five years. The

1 - Brass W.,1971 On the Scale of Mortality, in W. Brass (ed.), Demography London :Taylor & Fracis

exceptions occur in the first two rows of these tables and for the last row: the first row (age 0) represents a one-year interval and the second row, a four-year interval (ages 1 to 4). The last row is an open age interval, 75 or 85 years and over.

A bridged life table is a method of deriving measures which are representative of average life expectancy prevailing at a given time. It is a purely hypothetical calculation. It is a statistical tool typically used to portray expectation of life at various ages. A life table describes the course of mortality throughout the life cycle. It preferable to construct abridged life table, this might be because the data being used are not sufficiently reliable for a complete table or because a more concise –picture of mortality is preferred, in demography abridged life tables are far more common than complete ones.[1]

2-2-3 Life Table Used for Analysis:

In this study will be used functions of bridged life table(demographic and statistical functions), Brass Relational two-Parameter Logit system Life Table, UN model life table for developing countries and Coale and Demeny model (the West region) to find life expectation at birth e_0 .So now we must be addressed to life table components in detail.

2-2-3 -1 Abridged Life Table:

Is a method of deriving measures which are representative of average life expectancy prevailing at a given time. It is a purely hypothetical calculation. It is a statistical tool typically used to portray expectation of life at various ages. A life table describes the course of mortality throughout the life cycle. It

[1] - Bongaarts, J., & Feeney, G. .,2003, "Estimating mean lifetime", Proceedings of the National Academy of Sciences, 100(23), 13127-13133.

preferable to construct abridged life table, this might be because the data being used are not sufficiently reliable for a complete table or because a more concise –picture of mortality is preferred, in demography abridged life tables are far more common than complete ones.

2-2-3 -1-2Abridged Life Table Functions:

The components of a life table are described in the following:

There are two points to be clarified before starting to explain life table components:

i**- Moving Average:**

The moving average is a type of finite impulse response filter used to analyze asset of data points by creating a series of average of different of the full data set given a series of numbers and affixed subset size, the moving average can be obtained by first taking that average of the first subset. The affixed subset size is then shifted forward, creating a new subset of numbers which is averaged. The process is repeated over the entire data series. A moving average may also use unequal weights for each data value in the subset to emphasize particular values in the subset.

ii**- Age Intervals, ($_na_x$)**: As with the cohort table, each interval in this column is defined by the two exact ages stated except for the final age interval, which is open-ended such as 75 and over.

$_na_x$ The proportion of the interval lived by those who die , but sometimes it is defined as the average number of years lived by those who die. Thus, the normal 5 years intervals ${}_na_x$ maybe 0.5 or 2.5.The starting point for the final age interval is denoted by ω.

${}_n\boldsymbol{a}_x$, is included containing the average number of years that a person dying in the interval live inside this interval. For a five-year interval, the value of this function is about (*n/2*).

$$ {}_na_x = \frac{a_x d_x (1+a_{x+1})d_{x+1} + (1+a_{x+2})d_{x+2} + \cdots (4+a_{x+n-1})d_{x+n-1}}{d_x + d_{x+1} + d_{x+2} + \ldots\ldots + d_{x+n-1}} \quad \ldots\ldots\ldots\ldots \quad (2.1) $$

obtained from the abridged life tables published by the INE (1998) and then derived:

$$ a_1 = \frac{4a_1(d_1+d_2+d_3+d_4) - 1.5d_2 + 2.5d_3 + 3.5d_4}{d_1} $$

$$ 4a_1 = \frac{4L_1 + 4L_5}{4d_1} $$

The length of the typical interval (x_i, x_{i+n}) in the abridged table is greater than one year (commonly, n = 5 years). The essential element here is the average fraction of the interval lived by each person who dies at an age included in the interval. This fraction, called the fraction of last age interval of life, denoted by ${}_na_x$, is conceptually

logical extension of the fraction of the last year of life, a_x. Starting with the values of a_i we can construct the abridged life. It will start from this point in the formation of life table as we see in the equation (2-4) (2-6). The components of a life table are described in the following:

At first it is necessary to explain (Crude Death Rate, probability of death and Age Specific Death Rate) in some detail because the method presented here is based on an observed relationship between these indicators and then the rest of the equations of life table can be composition of these indicators.

1 .Death Rates and Adjustment of Rates:

A. Age Specific Death Rate:

nM_x : Central death rate for the age interval (x, x+n). Usual notation is nm_x

$$nM_x = \frac{deaths\ during year\ of persons\ aged(x)}{Population\ aged(x) at\ mid-year} \quad \ldots\ldots\ldots\ldots\ldots \quad (2\text{- }2)$$

$$nM_x = \frac{ndx}{nLx}$$

For age a specific interval (x_i , x_{i+n}) the death rate nM_x, , is defined as follows:

$$M_i = \frac{Number of dying\ aged(\ x_i ,x_{i+n}\ )}{\text{Number of years lived in aged } (\ x_i ,x_{i+n}\ ) \text{by those alive at x,}} \quad \ldots\ldots\ldots\ldots \quad (2\text{- }3)$$

Suppose that ℓ_i of people living at exact age x_i, d_i die between age x_i, and x_{i+n} and each of d_i people lives on the average a fraction, a_i of the interval formula m_i. Then the death rate m_i, defined in (2- 3) may be expressed in the formula:

$$M_i = \frac{di}{ni\ (\ell_i - di) + ai\ ni\ di} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{- }4)$$

where $ni = x_{i+n} - x_i$ is the length of the interval $(x_i, x_{i+n}), n_i$ (ℓ_i-d_i) is the number of years lived in (x_i, x_{i+n}) by the (ℓ_i-d_i) survivors, and (a_i n_i d_i) is the number of years lived by the d_i people who die in the interval. The unit of a death rate is the number of deaths per person-years. The corresponding estimate of probability of dying, given by:

$$q_i = \frac{d_i}{\ell_i} \quad \text{..........................} \quad (2\text{-}5)$$

is a pure number. From (2-4) and (2-5), we find a relationship between $q_i$and M_i

$$q_i = \frac{ni\ Mi}{1+(1-ai)\ ni\ Mi} \quad \text{......................} \quad (2\text{-}6)$$

Remark that:x_1 = 1, x_5= 5, n_i = 5 – 1

We see then that the age-specific death rate and the probability of dying are two different concepts and they are related by formula (2-6).

Formula (2-4) of the age-specific death rate is expressed in terms of a life table framework where $\ell_i$people are followed for n_i years to determine the number of deaths (d_i) and the number of survivors (ℓ_i - d_i) at the end of $n_i$years. Instead of d_i defined in a life table, we haveD_i, the observed number of deaths occurring to people in the age group (x_i ,x_{i+n}) during a calendar year. To derive a formula for the death rate as in (2- 4), we let $N_i$be the (hypothetical) number of people alive at exact age x_i among them D_i deaths occur. Then we have the death rate:

$$M_i = \frac{D_i}{n_i(N_i - D_i) + a_i n_i D_i} \quad \text{........................} \quad (2\text{-}7)$$

and an estimate of the probability q_i

$$q_i = \frac{D_i}{N_i} \quad \text{.............................} \quad (2\text{-}8)$$

They also have the relationship in (2-6).

Since N_i is a hypothetical number, the denominator of (2-7) and the death rate for a current population cannot be computed from (2-7). Customarily, the denominator of (2-7) is estimated by the mid-year (calendar year) population P_i for age group (x_i, x_{i+n}), and hence the age-specific death rate is given by:

$$M_i = \frac{D_i}{P_i} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{-}9)$$

Although it is a well known and accepted definition of age-specific death rates, formula (2-9) is much more meaningful when P_i is interpreted as an estimate of the denominator in (2-3).

A death rate usually is a small number; its significance is not easily appreciated. To remedy this, the numerical value of a death rate is multiplied by a number, such as 1000, which is called the base. The formula of a death rate often appears as:

$$M_i = \frac{D_i}{P_i} * \text{k} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{-}10)$$

where k is radix (k =100,1000,10000,……)

Note that, the result is written so per 1,000 person - year, but often delete the words "person -years"**.**

It should be clear that in formula (2-9) and (2-10) the number of deaths D_i in the numerator and the mid-year population P_i in the denominator refer to the same population. The population and the base must be clearly stated in a death when a death rate is for an entire life; it is called the crude death rate. In formula:

$$M = \frac{D}{P} * \text{k} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{-}11)$$

Where :

D = $\sum_i D_i$: is the total number of deaths occurring during a calendar year, and

P = $\sum_i p_i$: is the total mid-year population of a community, or a country in question.

Death rates may be computed for any specific category of people in a population. Sex-specific death rates, occupation-specific death rates, age-sex-specific death rates, are examples. In each case, the specific rate is defined as the number of deaths occurring to people in the stated category during a calendar year divided by the mid-year population of the same category.

B. Adjustment of Rates:

Specific death rates presented above are essential in mortality analysis. Individually, these rates describe mortality experience within respective categories of people. Collectively, they represent a mortality pattern of the population in question. When a collective measure of mortality of an entire population is required, specific rates provide the fundamental components. One of the central tasks in statistical analysis of mortality data is making comparisons of experiences of various communities or countries; summarization of specific rates in a single number is extremely important. Since age-sex distribution varies from one community to another, adjustment for such variation will have to be made in summarizing specific rates. The resulting single figure is called the adjusted rate. Adjustment can be made with respect to age, sex, occupation and possibly others. For simplicity, we shall consider only age-adjusted rates. Adjusted rates for other variables, such as sex-adjusted rates, age-sex-adjusted rates, etc., can be computed similarly. Various methods of adjustment have been proposed; some of these are listed in Table (2-1). It is the purpose of this section to review them. But first, let us introduce some notations. In the adjustment of rates, two populations are usually involved:

There are several methods of adjustment of age-specific death rates were presented, Although each method was developed on the basis of a specific philosophic argument and designed to serve a definite purpose, they all assume a general form of a weighted mean of the age-specific death rates. These methods of adjustment are reproduced in Table (2-1) for easy reference.

With the exception of the indirect method of adjustment, the weights add to unity. The sum of the weights in the indirect method can be greater or less than unity, depending upon the difference between community and standard populations in age composition. For this reason, the indirect method is not strictly comparable with any other adjusted rate, and neither is its standard error.

The inclusion of the crude death rate in the list of adjusted rates is of significance. Since it is usually expressed as the ratio of all deaths to the total mid-year population, the crude death rate is occasionally treated as a binomial proportion, which leads to an incorrect formula for the standard deviation. Individuals differing in age and sex obviously do not have the same probability of dying, and the notion of an average probability is incomprehensible; therefore, a direct application of the binomial theory is inappropriate. If, however, it is visualized as the weighted mean of specific death rates, with the actual population size employed as weights, then the crude death rate is perhaps the most meaningful measure of mortality for a single community. This way of viewing the crude rate is also essential in the derivation of its standard deviation.

In all the adjusted rates, the choice of weights applied to specific rates is based on: (1) The proportion of those in a specific age group to the total population, i.e., population proportion.

(2) The relative interval length of a specific age group.

For the crude rate, the weights used are the community population proportions in specific age groups (P_{ui}/P_u), for the direct method of adjustment, the standard population proportions in specific age groups (P_{si}/P_s), for the comparative mortality rate, the average of the two population proportions, for the life table death rate, the life table population proportions for specific age groups(L_i/T_0), and for the equivalent average death rate, the relative interval lengths of the age groups ($n_i \sum_i n_i$). The weights used in the indirect method of adjustment are functions of the age-specific rate for the standard population, community population proportions, and standard population proportions.

The methods of adjustment listed in Table (2-1) also include two indices, the relative mortality index, the mortality index, and the standardized mortality ratio.

As seen from the second panel of Table (2 -1), the two indices are weighted means of the ratios of a community's specific death rates to the corresponding specific rates for the standard population. The difference is that the relative mortality index uses the population proportions of the community for specific age groups as weights, while the mortality index uses the relative lengths of the age intervals. In the derivation of their standard deviations, however, we shall consider them as linear functions of age-specific death rates of a community with coefficients as listed in the weight column[1].

Table (2-1): Formulas and Weights Used to Compute the Crude Death Rate, Age-Adjusted Rates and Mortality Indices

Title	Crude Death Rate (C.D.R.)	Direct Method of Adjustment (D.M.D.R.)	Weight (W_i)
Crude death rate (C.D.R.)	$\frac{\sum P_{ui} M_{ui}}{P_u}$	Linder, F. E. and Grove, R. D. (1943)	$\frac{P_{ui}}{P_u}$
Direct method of adjustment (D.M.D.R.)	$\frac{\sum P_{si} M_{ui}}{P_s}$	"The Registrar General's Statistical Revies of England & Wales for the Year 1934"	$\frac{P_{si}}{P_s}$
Indirect method of adjustment (I.B.D.R.)	$\frac{(D_s/P_s)(D_u/P_u)}{\sum P_{ui} M_{si} / P_u}$	'~he Resistrar General's Decennial Supplement, England and Wales, 1921, Part III."	$\frac{(D_s/P_s)}{\sum P_{ui} M_{si}} P_u$
Life table death rate (L.T.D.R.)	$\frac{\sum \mathcal{L}_i M_{ui}}{\sum \mathcal{L}_i}$	Brownlee, J. (1913) (1922)	$\frac{\mathcal{L}_i}{\sum \mathcal{L}_i}$

Source:" Statistical Methods for Rates and Proportions", Fleiss, J. 1, Wiley, New York.,1983.

C. Crude death rate: Crude death rate is the ratio of the total number of deaths occurring in a community during a calendar year to the community's total midyear-population:

1- Jack Johnston& John Dinardo., " Econometric Methods, Four Edition, Univariate Time Series Modeling", four edition http :// www.mh college.com.

$$\text{C.D.R.} = D_u / P_u \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.. \quad (2\text{-}12)$$

The crude death rate, which is the most commonly used and conveniently computed single value, bears a close relationship to age-specific death rates. The numerator in (2-12) is the sum of the number of deaths occurring in all age categories :

$$D_u = \sum_i D_{ui}$$

By definition, the age-specific death rate for age interval (x_i, x_{i+n}) is given by:

$$M_{ui} = \frac{D_{ui}}{P_{ui}}$$

So that the number of deaths (D_{ui}) is the product of the age-specific death rate (M_{ui}.) and the corresponding mid-year population

(P_{ui}):

$$D_{ui} = P_{ui} M_{ui} \ \ldots\ldots\ldots\ldots\ldots\ldots.. \quad (2\text{-}13)$$

Therefore, the total number of deaths in (2-13) may be rewritten as:

$$D_u = \sum_i P_{ui} M_{ui} \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{-}14)$$

Substituting (2-10a) in (2- 9) yields:

$$\text{C.D.R} = \sum_i \frac{D_{ui}}{P_u} M_{ui} \ \ldots\ldots\ldots\ldots\ldots\ldots.. \quad (2\text{-}15)$$

Where the summation is taken over the entire life span. Thus the C.D.R. is a weighted mean of age-specific death rates with the actual population proportions $\frac{D_{ui}}{P_u}$ experiencing the mortality used as weights. From this view point, the C.D.R. is the most meaningful single figure summarizing the mortality experience of a given population.

The C.D.R., however is not without deficiencies. The quantity on the right-hand side of (2- 15) is a function of both the age-specific death rates and the age-specific population proportions. As a weighted mean of age-specific death rates, the C.D.R. is affected by the population composition of the community in question. This disadvantage becomes apparent when the C.D.R. is used as a common measure to compare the mortality experience of several communities.

D. Life Table Death Rate (L.T.D.R.): Most of the methods of adjustment rely on a standard population or its rates. One exception is the L.T.D.R. which is defined as:

$$\text{L.T.D.R} = \sum_i \frac{L_i}{T_0} M_{ui} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{-}14)$$

Where L_i is the number of years spent in (x_i , x_{i+n}) by a life table population and

$$T_0 = L_0 + 4L_1 + \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2 - 15)$$

A full appreciation of this method of adjustment requires the knowledge of the life table discussed later on, a brief discussion of formula (2-15) follows. Given $\ell_0$people alive at age (0) who are subject to the age specific death rates of the community, (L_i/T_0) is the proportion of their life time spent in the age interval(x_i , x_{i+n}). In other words, the L.T.D.R. shown in formula (2-15) is a weighted mean of the age specific death rates (M_{ui}) with the proportion of life time spent in (x_i , x_{i+n}) being used as weights. Since the weights (L_i/ T_0) depend solely on the age-specific death rates, the

L.T.D.R. is independent of the population composition either of a community or a standard population.

As we will see blow, the age specific death rate (M_{ui}) is equal to the ratio d_i/ℓ_0,

$$M_{ui} = d_i/\ell_i,$$

Hence :

$$M_{ui}\ell_i = d_i \;\ldots\ldots\ldots\ldots\ldots \qquad (2\text{-}16)$$

Where d_i is the life table death in age interval$(x_i\;, x_{i+n}\;)$ the sum:

$$d_0 + 4d_1 + 9d_5 + \ldots\ldots = \;\ell_0 \;\ldots\ldots\ldots\ldots\ldots \;(2\text{- }17)$$

is equal to the total number of individuals $\ell_0$at age (0). Substituting (2- 16) in (2-14) and recognizing (2-14), we have:

L. T .D.R. =

$$\textstyle\sum_i d\;_i/\,T_0 = \ell_0\;/\;T_0 \;\ldots\ldots\ldots\ldots\ldots\ldots \;\; (2\text{-}18)$$

The inverse:

$$T_0\;/\;\ell_0 = e_0$$

is known as the (observed) expectation of life at age (0), therefore :

L.T.D.R=

$$\frac{\sum L_x M_{ux}}{\sum L_x} = \frac{\sum d_x}{\sum L_x} = \frac{\ell_0}{T_0} = \frac{1}{e_0} \;\ldots\ldots\ldots\ldots\ldots \qquad (2\text{-}19)$$

Probability of Death and the Age-Specific Death Rate:

As we know, the probability of death and the age-specific death rate are two measures of the risk of mortality acting on individuals in the population. While the probability of death is an established concept in the field of statistics, analytic meaning of the age-specific death rate is not fully appreciated. The age-specific death rate either is regarded as an ill-defined statistical quantity, or else it is treated as if it were another name for the probability of death. These misconceptions need be corrected. The age-specific death rate is just as meaningful analytically as the probability. For easy reference, we state again the estimate of the probability and the age-specific death rate below.[1]

Let P_i be the number of individuals alive at the exact age X_i among them a number D_i dying during the interval (X_i, X_{i+n}). Then the estimate of the probability of dying in (X_i, X_{i+n}) is:

$$q_i = \frac{D_i}{P_i} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{-}20)$$

On the other hand, the age-specific death rate, M_i , is the ratio of the number of deaths, D_i to the total number of years lived in the interval (X_i, X_{i+n}) by the P_i people. In formula:

$$M_i = \frac{D_i}{n_i(P_i - D_i) + a_i n_i D_i} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{-}21)$$

Solving equations (2-20) and (2-21) yields the basic relationship between q_i and M_i:

$$q_i = \frac{n_i M_i}{1 + (1 - a_i) n_i M_i} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{-}22)$$

1- Colin Newell. (July 1987),"Methods and Models in Demography".

Here $n_i = (X_{i+n} - X_i)$ and a_i is the average fraction of the age interval (X_i, X_{i+n}) lived by individuals dying at any age included in the interval. For a current population, the age-specific death rates are determined from the vital and population statistics:

$$M_i = \frac{D_i}{P_i} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{-}23)$$

WhereD_i is the number of deaths occurring in age group (X_i, X_{i+n}) during a calendar year and P_i is the corresponding mid-year population. The probability of death is computed from formula (2-22) or (2- 23). After we explained in detail (C.D.R, A.S.D.R and P.O. D) we can find the rest of the components of life tables[1].

nq_x: This is probability of dying between exact age (x) and (x +n) as mentioned above.

$$nq_x = \frac{ndx}{\ell_x} \quad \ldots\ldots\ldots\ldots\ldots\ldots.. \ (2\text{ -}24)$$

From the following equation can be found $_nq_x$:

$$np_x + nq_x = 1$$

$$nq_x = 1 - np_x \quad \ldots\ldots\ldots\ldots\ldots. \ (2\text{- }25)$$

np_x: This is probability of surviving between exact age (x) and (x +n).It is just the component of np_x Thus:

Also can be calculated by this equation:

$$np_x = \frac{l_{x+n}}{lx} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots..(2\text{- }26)$$

1- Chin Long Chiang ., 2007 Life Table and Mortality Analysis, World Health Organization.

or

$$np_x = 1- \frac{ndx}{\ell_x} \text{ ………………….. } (2\text{- }27)$$

or

$$np_x = 1 - nq_x \text{ …………………… } (2\text{-}28)$$

Often np_x like nq_x is expressed per1000 or 10,000 populations in order to avoid lots of zeros.

$\boldsymbol{\ell_x}$**:** This is the number of persons alive at exact age x. Given a survival model, with survival probabilities p_x, we can construct the life table for the model from some initial age $x_0$to a maximum age ω. We define a function { ℓ_x } for $x_0 \leq x \leq \omega$ as follows.

ℓ_0 is an arbitrary number called the radix. Usually it will be around number such as 100 or 1000 or 10,000.To calculate ℓ_x**,** first choose a suitable radix, and then work down the table using the formula:

$$\ell_x = \ell_{x-n}.\, np_x \text{ …………….. } (2\text{- }29)$$

From this definition we see that for $0 \leq x \leq x + n \leq \omega$,

$$\ell_{x+n} = \ell_{x0x+n-x}\ 0p_{x0}$$

$$= \ell_{x0\ x-x}\ 0p_{x0}\ np_x$$

$$= \ell_x np_x$$

Also can be calculated as follows:

$$\ell_{x+n} = \ell_x\text{- } nd_x \text{………………. } (2\text{-}30)$$

nd_x: This is the number of persons dying between exact ages (x)and (x+ n). It is just the difference between two ℓ_x algebraically:

$$nd_x = \ell_x - (\ell_{x+n}) \quad \ldots\ldots\ldots\ldots\ldots (2\text{- }31)$$

in addition to ℓ_x as these are used to compute ${}_nq_x$. From (2- 30) we have:

$$nd_x = \ell_x + \left(1 + \frac{\ell x+n}{\ell x}\right)$$

$$= \ell_x(1- \ np_x) = \ell_x \ nq_x$$

$$d_0 = (L_0 - {}_1\ell) / 0.3$$

${}_nL_x$: This is the number of persons years lived between age (x) and (x + n)

When assuming ${}_na_x = 1/2$, we have

$${}_nL_x = \ell_{x+n} + \left(1 + \frac{1}{2}\right) ndx \quad = \quad \frac{n\,(\ell_x + \ell_{x+n})}{2} \quad \ldots\ldots\ldots\ldots (2\text{- }31)$$

Note that:

$$L_0 = 0.3\ell_0 + 0.7\ell_1$$

or

$$L_0 = \ell_1 + 0.1d_0$$

$$L_0 = a_0\,\ell_0 + (1- a_0)\,\ell_1$$

$$4L_1 = 1.3\,\ell_1 + 2.7\,\ell_5$$

$$\infty L_{85} = \ell_{85}\ \log_{10}\ \ell_{85}$$

or

$$\infty L_{85} = \frac{5d_{85}}{5M_{85}}$$

***T_x* :** This is the number of persons-years lived after exact age x. It is thus simply the $_nL_x$ column cumulated from the bottom .That is :

$$T_x = T_{(x+n)} + nL_x$$

or

$$T_x = \sum_{y=x}^{w} nL_y \quad \dots\dots\dots\dots\dots\dots \quad (2\text{- }32)$$

Note that: T0= Total population

$$T_0 = L_0 + L_1 + L_2 + \dots\dots + L_w$$

$$T_\infty = \infty L_{85}$$

***e_x*:** This is the expectation of life age (x), or the average of years a person aged **x** has to live.

The value of $\mathbf{e_x}$ at age zero has a special significance in demographic analysis because represents a measure for life expectation at birth which represents the average length of the period, which is expected to live a new born if it was exposure to current death rates in the table of life. It thus represents an important indicator brief and to measure the level of mortality.

$$e_x = \frac{T_x}{\ell_x} \quad \dots\dots\dots\dots\dots\dots\dots\dots \quad (2\text{- }33)$$

$$e_0 = \frac{T_0}{\ell_0}$$

Data Smoothing:

To correct defective data, the strong method for smoothing data applied to the Abridged Life Table:

$$M_0 = \frac{q_0}{(1-0.7q_0)} \quad \dots\dots\dots\dots\dots \quad (2\text{-}34)$$

$$4m_1 = \frac{q_1}{(4-2.4q_1)} \quad \ldots\ldots\ldots\ldots\ldots \quad (2\text{-}35)$$

nm_x: For Age Specific Death Rate m_0 the Moving Average is uses to estimate the value of m_x to correct the defective data.

nq_x : The probability of dying between exact age (x) and(x +n)

$$nq_x = \frac{n.\ nMx}{1+n\ (\ 1-nax)\ nMx} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{- }36)$$

$$q_0 = \frac{m_0}{1+0.7m_0} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{- }37)$$

$$4q_1 = \frac{4\ (4m_1\ )}{1+2.7\ (4m_1\ )} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{- }38)$$

2 -2 -3-2 Statistical Theories of Life Table Functions:

The purpose of this section to give a brief presentation of the theoretical aspects of the life table. A typical abridged life table is reproduced below. As we mentioned above, the radix ℓ_0 is conventionally set equal to a convenient number, such $\ell_0 =$ 100,000, so that the value of ℓ_x clearly indicates the proportion of survivors to age (x) we adopt the convention and consider ℓ_0a constant in deriving the probability distributions of other life table functions. The distributions of the quantities in columns L_x and $T_x$are not discussed because of their limited use. One remark should be made regarding the final age interval ($\mathcal{W}$ and over). In a conventional table the last interval is usually ($\mathcal{W}$) an open interval, e.g., 75 and over; statistically speaking, ($\mathcal{W}$) is a random variable and it's treated accordingly. Throughout this appendix we shall assume a homogeneous population in which all individuals are subjected to the same force of mortality, and in which one individual's survival is independent of the survival of any other individual in the group. [1]

1 - Aa1en, O., 2003, Nonparametric inference in connection with multiple decrement models, Scand. J. Stat., 3, 15 -27.

A. Probability Distribution of (ℓ_x) the Number of Survivors at Age (x):

The various functions of the life table are usually given for integral ages or for other discrete intervals. In the derivation of the distribution of survivors, however, age is more conveniently treated as a continuous variable with formulas derived for ℓ_x the number of individuals surviving the age interval (o, x), for all possible values of x.

The probability distribution of ℓ_x depends on the force of mortality, or intensity of risk of death, μ(x) defined as follows:

μ(x) Δ + o (Δ) = Pr {an individual alive at age (x) will die in interval (x, x+ Δ)}

Let the continuous random variable (x) be the life span of a person so that the distribution function:

$$F_x\ (x) = \Pr\ (X < x)$$

is the probability that the individual will die prior to (or at) age (x).Consider now the interval (o, x+Δ) and the corresponding distribution function F_x(x+Δ) =Pr (x<x+Δ).For an individual to die prior to (x+Δ)he must die prior to (x) or else he must survive to (x) and die during the interval (x, x+Δ).Therefore, the corresponding probabilities have the relation:

$$F_x(x+\Delta) = F_x\ (x)\ [1-\ F_x\ (x)][\ \mu(x)\ \Delta + (\Delta)]\ \ldots\ldots\ldots\ (2\text{-}39)$$

Or

$$\frac{F_x(\ x+\Delta) - F_x\ (x)}{\Delta} = [1-\ F_x(x)][\ \mu(x) + \frac{o(\Delta)}{(\Delta)}]$$

Taking the limits of both sides of (2-39) as $\Delta \rightarrow o$ we have the differential equation:

$$\frac{d}{d_t} F_x(\text{x}) = [1- F_x(\text{x})]\, \mu(\text{x}) \ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{-}40)$$

with the initial condition:

$$F_x(0) = 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{-}41)$$

Integrating (2-23) and using (2-23a) yields the solution:

$$1- F_x(\text{X}) = e^{-\int_0^x \mu(t)dt} = P_{0x} \ldots\ldots\ldots\ldots \quad (2\text{-}42)$$

Equation (2-42) gives the probability that one individual alive at age (0) will survive to age (x). If there are (ℓ_0) individuals alive at age (0) who are subject to the same force of mortality, the number (ℓ_x) of survivors at age(x) is clearly a binomial random variable with the probability (P_{0x})to x and the probability distribution given by:

$$\Pr\{\ell_x = k\} =$$

$$\frac{\ell_0!}{k!(\ell_0-k)!} {P_{0x}}^k (1- P_{0x})^{\ell_0-k}, \text{k=0,1}\ldots,\ell_0 \ldots\ldots\ldots\ldots \quad (2\text{-}43)$$

For x = x_i the probability that an individual will survive the age interval (0,x_i) is:

$$P_{0i} = \exp\{-\int_0^{x_i} \mu(\text{t})\text{dt}\} \ldots\ldots\ldots\ldots \quad (2\text{-}44)$$

and the probability distribution of the number of survivors, ℓ_i is:

$$\Pr\{\ell_i = k_i \mid \ell_0\} =$$

$$\binom{\ell_0}{k_i} {P_{0i}}^{k_i} (1- P_{0i})^{\ell_0-k_i} \quad , \text{k=0,1}\ldots\ldots,\ell_0 \ldots\ldots\ldots \quad (2\text{-}45)$$

respectively.

In general, the probability of surviving the age interval (x_i, x_j) is:

$$P_{ij} = \exp\{-\int_{x_i}^{x_j} \mu(t)dt\} \qquad \text{for } i \le j \; \ldots\ldots\ldots\ldots \; (2\text{-}46)$$

with the obvious relation:

$$P_{\alpha j} = P_{\alpha i} P_{ij} \qquad \text{for } \alpha \le i \le j \; \ldots\ldots\ldots\ldots \; (2\text{-}47)$$

If we start with $\ell_i$individuals at x_i the number of survivors $\ell_j$at $x_j$for $i \le j$is also a binomial random variable with the probabilityP_{ij} and:

$$\Pr\{\ell_j = k_j \mid \ell_i\} =$$

$$\frac{\ell_i!}{k_j\,(\ell_i - k_j)!} P_{ij}^{\,k_j} (1 - P_{ij})^{\ell_i - k_j} \;, k_j = 0,1 \ldots., \ell_i \; \ldots\ldots(2\text{-}48)$$

When j=i+l,(2-48)becomes:

$$\Pr\{\ell_{i+1} = k_{i+1} \mid \ell_i\} =$$

$$\frac{\ell_i!}{k_{i+1}\,(\ell_i - k_{i+1})!} P_i^{\,k_{i+1}} (1 - P_i)^{\ell_i - k_{i+1}} \; \ldots\ldots\ldots\ldots\ldots \quad (2\text{-}49)$$

It is intuitively clear that given (ℓ_i)people alive at age $(x_i)$the properties distribution of the number of people alive at $x_j$for $x_j > x_i$is independent of $\ell_0, \ell_1, \ldots\ldots \ell_{i-1}$.This means that for each k_j

$$\Pr\{\ell_j = k_j \mid \ell_0, \ell_1, \ldots\ldots \ell_i\} =$$

$$\Pr\{\ell_j = k_j \mid \ell_i\} \ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{-}50)$$

In other words, for each μ the sequence $\ell_0, \ell_1, \ldots\ldots \ell_\mu$ is a Markov process.

B. **Mortality Laws:**

The survival probability in (2-42) has been known to life table researchers for more than two hundred years. Unfortunately, it has not been given due recognition by investigators in statistics although differing forms of this function have appeared in various areas of research. We shall mention a few below in terms of the probability density function of (x):

$$F_x(\mathrm{x}) = \frac{d\,F_x\,(\mathrm{x})}{d_x} = \mu(x)e^{-\int_0^x \mu(\mathrm{t})\mathrm{dt}} \qquad \mathrm{x} \geq 0$$

$$F_x(\mathrm{x}) \quad = 0 \qquad \mathrm{x} < 0 \ldots\ldots\ldots\ldots\ldots \quad (2\text{-}51)$$

(i) Gompertz Distribution:

In a celebrated paper on the law of human mortality, Benjamin Gompertz attributed death to two causes: chance, or the deterioration of the power to withstand destruction. In deriving his law of mortality, however, he considered only deterioration and assumed that man's power to resist death decreases at a rate proportional to the power itself. Since the force of mortality μ(t) is a measure of man's susceptibility to death, Gompertz used the reciprocal [l / μ (t)] as a measure of man's resistance to death and thus arrived at the formula:

$$\frac{d}{d_t}\left(\frac{1}{\mu\,(\mathrm{t})}\right) = -\,\mathrm{h}\left(\frac{1}{\mu\,(\mathrm{t})}\right) \ldots\ldots\ldots\ldots \qquad (2\text{-}52)$$

Where (h) is a positive constant. Integrating (2-51) gives:

$$ln\left(\frac{1}{\mu\,(\mathrm{t})}\right) = -\,h_{t+k} \ldots\ldots\ldots\ldots \quad (2\text{-}53)$$

Which when rearranged becomes the Gompertz law of mortality:

$$\mu\,(\mathrm{t}) = Bc^t \ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{-}54)$$

The corresponding density function and distributions are given, respectively, by:

$$f(x) = Bc^{x}e^{-B[c^{x}-1]/lnc} \quad \ldots\ldots\ldots\ldots\ldots \quad (2\text{-}55)$$

and

$$F_{x}(x) = 1 - \exp\{-\frac{B}{lnc}(c^{x}-1)\} \ldots\ldots\ldots\ldots\ldots \quad (2\text{-}56)$$

Bongaarts and Feeney (2002, 2003) have stimulated a new debate about how to interpret period summary measures of mortality when rates of death vary over time. The parallel shift in adult mortality analyzed by Bongaarts and Feeney can be characterized by a Gompertz mortality change model (Vaupel 1986, Vaupel and Canudas- Romo 2003). Their formulation is an extension of the Gompertz (1825) model of mortality, which has a changing force of mortality component,

$$\mu(\alpha, t) = \mu(0,t)\, e^{\beta\alpha} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{-}57)$$

Where μ (α,t) reflects the value of the rate of mortality decrease over time and parameter β>0 is the fixed rate of mortality increase over age.

Force of mortality [Bongaarts and Feeney (2002, 2003)]:

$$M(t) = \frac{ln\beta - \ln[\mu(0,t)]}{\beta} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{- }58)$$

The survival function for this model is:

$$\ell(M,t) = e^{\frac{\{\mu(0,t)\}}{\beta}-1} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{- }59)$$

Bongaarts and Feeney (2002, 2003) showed that the value of μ (0, *t*) declines over time. When the reduction in mortality is almost negligible, and the value of μ (0, *t*) approaches zero so number of survivors:

$$\lim_{\mu(0,t)\to 0} \ell(M, t) = e^{-1} \approx 0.37$$

and thus the number of deaths is $d(M; t) = \beta e^{-1}$(Pollard 1991, Pollard and Valkovics1992). However, the modal age at death increases to infinity. Therefore, under this model the rectangularization process of the survival curve has stopped completely and no more concentration is observed in *l* (*M, t*). Instead, a shift occurs in the modal age towards advanced ages.

So the force of mortality at age (*a)* and time (*t)* is defined as:

$$\mu\ (a,t) = e^{\alpha - pt + \beta a}$$

where α is a constant that reflects the value of the force of mortality at age zero and time zero, μ(*0*,0), and parameter *p* is the rate of mortality decrease over time.

(ii) **Makeham's Distribution:** In 1860 W. M. Makeham suggested the modification:

$$\mu(t) = A + Bc^{t} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{-}60)$$

which is a restoration of the missing component, "chance" to the Gompertz formula. In this case, we have:

$$f(x) = \{A + Bc^{x}\}\ exp\{-[Ax + B(c^{x} - 1)\ /\ ln\ c]\} \ \ldots\ldots\ (2\text{-}61)$$

and

$$f_x(x) = 1 - exp\{-[Ax + B(c^{x} - 1)\ /\ ln\ c]\} \ \ldots\ldots\ (2\text{-}62)$$

(iii) Weibull Distribution: When the force of mortality (or failure rate) is assumed to be a power function of t, $\mu(t) = \mu a t^{a-1}$we have

$$f\ (x) = \mu a x^{a-1} e^{-\mu x^{a}} \ \ldots\ldots\ldots\ldots\ldots \ (2\text{-}63)$$

and

$$f_x(\mathrm{x}) = 1- e^{-\mu x^a} \text{ ……………………..} \quad (2\text{-}64)$$

(iii) Exponential Distribution: $\mu(\mathrm{t}) = \mu$ If is a constant, then :

$$f\ (\mathrm{x}) = \mu e^{-\mu x} \text{ ……………………….} \quad (2\text{-}65)$$

and

$$f_x(\mathrm{x}) = 1- e^{-\mu x} \text{ ………………….} \quad (2\text{- }66)$$

a formula that plays a central role in the problem of life testing.

C. Joint Probability Distribution of the Number of Survivors:

Let us consider, for a given u, the joint probability distribution of

$$\ell_1, \ell_2, \dots\dots, \ell_u \qquad \text{given} \quad \ell_0$$

$$Pr\ \{\ell_1 = k_1, \dots\dots, \ell_u = k_u \mid \ell_0\} \text{……………} \quad (2\text{-}67)$$

It follows from the Markovian property in (2-50) that:

$$Pr\ \{\ell_1 = k_1, \ell_2 = k_2 \dots\dots\dots, \ell_u = k_u \mid \ell_0\}$$

$$= Pr\ \{\ell_1 = k_1 \mid \ell_0\}\ Pr\ \{\ell_2 = k_2 \mid k_1\}$$

$$\text{………}\ Pr\ \{\ell_u = k_u \mid k_{u-1}\} \text{………………} \quad (2\text{-}68)$$

Substituting (2-48) in (2-68) yields a chain of binomial distributions:

$$Pr\ \{\ell_1 = k_1, \ell_2 = k_2 \dots\dots\dots, \ell_u \qquad = \prod_{i=0}^{u-1} \frac{k_i!}{k_{i+1!}(k_i - k_{i+1})!} {p_i}^{k_{i+1}} (1 - p_i)^{k_i - k_{i+1}}$$

$$k_{i+1} = 0,1,\text{…………},k_i \qquad \text{with } k_0 = \ell_0 \text{ ………..} \quad (2\text{-}69)$$

Formula (2-69) shows that when a cohort of people is observed at regular points in time, the number of survivors, ℓ_{i+1} to the end of the interval (x_i, x_{i+1}) has a binomial distribution depending solely on the number of individuals alive at the beginning of the interval k_i=ℓ_i.

D. Joint Probability Distribution of the Numbers of Deaths:

In a life table covering the entire life span of each individual in a given population, the sum of the deaths at all ages is equal to the size of the original cohort. Symbolically:

$$d_0 + d_1 + \ldots\ldots\ldots\ldots\ldots + d_w = \ell_0 \; \ldots\ldots\ldots\ldots \quad (2\text{-}70)$$

Where d_w is the number of deaths in the age interval (x_w and over). Each individual in the original cohort has a probability $p_{0i}\, q_i$ of dying in the interval (x_i, x_{i+1}), ί = 1,…….., $\mathcal{W}$. Since an individual dies once and only once in the span covered by the life table:

$$p_{00}\, q_0 + \ldots\ldots\ldots\ldots\ldots + p_{0w}\, q_w = 1 \; \ldots\ldots\ldots\ldots\ldots \quad (2\text{- }71)$$

Where $p_{00} = 1$, $q_0 = 1$, thus we have the well-known results: The numbers of deaths, d_0 ,…… , d_w in a life table have a multinomial distribution with the joint probability distribution.

In the discussion above, age (0) was chosen only for simplicity. For any given age, say (x_α) the probability that an individual alive at age(x_α) will die in the interval (x_i, x_{i+1}) subsequent to (x_α) is $(p_\alpha q_i)$and the sum:

$$\sum_{i=\alpha}^{w} p_\alpha q_i = 1$$

and thus the numbers of deaths in intervals beyond (x_α) also have a multinomial distribution.

E. Distribution of $e_\alpha$the Observed Expectation of Life at Age (x_α):

The observed expectation of life summarizes the mortality experience of a population from a given age to the end of the life span. At age$(x_i)$the expectation expresses the average number of years remaining to each individual living at that age if all individuals are subjected to the estimated probabilities of death $\widehat{q_J}$for j ≥ i this is certainly the most useful column in the life table.

To avoid confusion in notation, let denote a fixed number and x_α a particular age. We are interested in the distribution of $\boldsymbol{e_\alpha}$ the observed a expectation of life at age x_α.Consider ℓ_α , the number of survivors to age x_α , and let $y_\alpha$denote the future lifetime beyond age x_α of a particular individual. Clearly, y_α is a continuous random variable that can take on any non-negative real value. Let y_α be the value that the random variable$y_\alpha$assumes, then $(x_\alpha + y_\alpha)$ is the entire life span of the individual. Let $f\,(y_\alpha)$ be the probability density function of the random variable y_α and let dy_α be an infinitesimal time interval. Since y_α can assume values between y_α and (y_α +dy_α) if and only if the individual survives the age interval (x_α , x_α +y_α), and dies in the interval (x_α +y_α , x_α +$y_\alpha + dy_\alpha$) we have:

$$f\,(y_\alpha)\;dy_\alpha =$$

$$e^{-\int_{x_\alpha}^{x_\alpha+y_\alpha} u(t)dt} \quad u(x_\alpha{+}y_\alpha)dy_\alpha y_\alpha \geq 0 \;\ldots\ldots\ldots\; (2\text{-}73)$$

Function $f\,(y_\alpha)$ in (2-73) is a proper probability density function since it is never negative and since the integral of the function from y_α= 0 to y_α = α is equal to unity. Clearly, $f\,(y_\alpha)$ can never be negative whatever the value of y_α ,to evaluate the integral:

$$\int_0^{\alpha} f\ (y_\alpha)\ dy_\alpha =$$

$$\int_0^{\alpha} e^{-\int_{x_\alpha}^{x_\alpha+y_\alpha} u(t)dt} u(x_\alpha+y_\alpha)dy_\alpha \text{ ………… } (2\text{- }74)$$

We define a quantityϕ:

$$\phi=\int_{x_\alpha}^{x_\alpha+y_\alpha} u(t)dt \quad =\int_0^{y_\alpha} u(x_\alpha+t)dt \text{ ……… } (2\text{- }75)$$

and substitute the differential:

$$d\,\phi=u(x_\alpha+y_\alpha)dy_\alpha \quad \text{…………………..} \quad (2\text{- }76)$$

in the integral to give the solution:

$$\int_0^{\infty} f\ (y_\alpha)\ dy_\alpha =\int_0^{\infty} e^{-\phi} d\,\phi = 1 \text{ …………… } (2\text{- }77)$$

The mathematical expectation of the random variable y_α is the expected length of life beyond age x_α and thus is the true expectation of life at age x_α. In accordance with the definition given the symbol e_α ,we may write:

$$e_\alpha = \int_0^{\infty} y_\alpha\ f\ (y_\alpha)\ dy_\alpha$$

$$=\int_0^{\infty} y_\alpha\ e^{-\int_{x_\alpha}^{x_\alpha+y_\alpha} u(t)dt}\ u(x_\alpha\ +y_\alpha)dy_\alpha \text{ ……… } (2\text{- }78)$$

The expectation of life at age $x_\alpha$is conventionally defined as:

$$e_\alpha =\int_0^{\infty} e^{-\int_{x_\alpha}^{x_\alpha+y_\alpha} u(t)dt} \quad \text{…………… } (2\text{-}79)$$

It is instructive to prove that the two alternative definitions (2-78) and (2-49) are identical. Let $u = y_\alpha$, $du = dy_\alpha$:

$$\text{v} = -\ e^{-\int_{x_\alpha}^{x_\alpha+y_\alpha} u(t)dt} \quad \text{………………..} \quad (2\text{- }80)$$

and

$$\text{dv} = e^{-\int_{x_\alpha}^{x_\alpha+y_\alpha} u(t)dt}\ u(x_\alpha\ +y_\alpha)dy_\alpha \text{ ………… } (2\text{-}81)$$

Integrating (2-78) by parts gives:

$$\int_0^{\infty} y_\alpha \, e^{-\int_{x_\alpha}^{x_\alpha + y_\alpha} u(t)dt} \, u(x_\alpha + y_\alpha) dy_\alpha$$

$$= -y_\alpha \, e^{-\int_{x_\alpha}^{x_\alpha + y_\alpha} u(t)dt} \Big|_0^{\infty} + \int_0^{\infty} e^{-\int_{x_\alpha}^{x_\alpha + y_\alpha} u(t)dt} \, dy_\alpha \quad \ldots\ldots\ldots (2\text{-}82)$$

The first term on the right vanishes and the second term is the same as (2-79) proving the identity.

2 -2 -4Brass Relational two-Parameter Logit System Life Table:

The main shortcoming of the model life-table systems described above such as United Nations model life tables, Coale and Demeny are their dependence upon the type of data that generated them. The rather restricted data base used for this purpose and the fact that the model systems themselves consist of only a finite number of cases which cannot be expected to represent all possible human experience make them less than ideal. Another type of model is needed. Naturally, this model should adequately reflect the patterns found in empirical mortality data. However, it should not be constrained to represent exclusively the patterns these data embody for, as pointed out earlier, the true mortality experience of many populations has not yet been ascertained with any degree of accuracy, and it may or may not strictly conform to patterns observed in countries where accurate measurement has been possible. A model that provides a greater degree of flexibility is that proposed by Brass and colleagues, better known as the "logit system".Brass attempted to relate mathematically two different life tables. He discovered that a certain transformation of the probabilities of survival to age x (ℓ_x) values in life-table terms made the relationship between corresponding probabilities for different life tables approximately linear. Also the system has the advantage that it does not require large books of tables or computer programs for it to be used easily.

So Brass model discovers which one of the finite family of life table that can be generated by varying $(\alpha)$and (β), the raw data are most like .The raw data may be a complete life table , or they may be just a handful of values.

The logit system is used to fit the adjusted mortality rates of a population and to synthesize independent estimates of child and adult mortality into coherent mortality schedules.

The logits of observed (ℓ_{xs}) are fitted plotted on a graph against the logist of standard life table. Then a straight line is fitted to the points. The standards are plotted on the X-axis, the observed values on the Y-axis.

2 -2 -4- 1 The Logit Transformation Formula:

The formula is:

$$\text{Logit (p)} = 0.5 \text{ loge } \left(\frac{1-p}{p}\right) \ldots\ldots\ldots\ldots\ldots \quad (2\text{-}83)$$

The transformation has some interesting:

When p = 0 then logit (p) = 0.5loge (1 / 0) = + infinity.

When p = 1 then logit (p) = 0.5loge (0 / 1) = - infinity.

When p = 0.5 then logit (p) = 0.5loge (0.5 / 0.5) = 0.0.

Whereas, of course, the range of the original proportions is from 0.0 to 1.0.

Occasionally, the transformation is expressed as:

$$\text{Logit (p)} = 0.5 \text{ loge } \left(\frac{p}{1-p}\right) \ldots\ldots\ldots\ldots\ldots \quad (2\text{-}84)$$

P = A set of (ℓ_x) value are just proportion if the radix is 1.0.

Rather than as above and this changes the sign of all the values. Also , the 0.5 in the transformation is omitted in most books .These differences have no fundamental effect, but, of course, it is essential to be consistent, and it is best to follow demographic convention and use the first form given above. To return to model life table, a set of ($\ell_{\mathbf{x}}$) values are just proportion, thus it is possible to take logits of ($\ell_{\mathbf{xs}}$). This is very useful because it has been found by Brass (1971) that if one takes two life tables ,and takes logits of their ($\ell_{\mathbf{xs}}$,) then the relationship between the two sets of logits is remarkably liner ,that is drawing a graph of one against the other, produces a straight line .This important discovery is essentially an empirical rather than a theoretical finding though (Brass 1975) gives some justification .Sometimes it does not give exactly a straight line, especially at the extremes of the age range.

2 -2 -5 The Linear Regression:

The equation of a straight line can be written as:

$$Y = a + b\,x \ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{-}85)$$

So it is possible to describe one set of logit ℓ_{xs} by using another set of logit (ℓ_{xs}) and an(a) and(b) it is possible to produce another set of (ℓ_{xs}). Alternatively, from any one standard life table it is possible to generate a whole series of other related life table just by varying (**a**) and (**b**), It is this aspect of the models that makes the system "relational".

The two parameters (**a**) and (**b**) are in fact known by the Greek letters(α)and(β) . If logit ($\ell_{\mathbf{x}}$) is denoted by $Y_{\mathbf{(x)}}$ and the standard as $Y_{\mathbf{s(x)}}$ as has become conventional in the literature, then the equation of the straight line can be rewritten as:

$$Y_{(x)} = \alpha + \beta Y_{s(x)} \ldots\ldots\ldots\ldots\ldots \quad (2\text{-}86)$$

Altering $\boldsymbol{\alpha}$ will affect the level of mortality, while altering (β) will affect the relationship between childhood and adult mortality.

2 -2 -5- 3 Fitting A logit Model Life Table:

The intercept and the slope of the line,α and (β), are calculated using the Group Average Methods, this simply involves taking the average of the childhood mortality points and the average of the adulthood mortality points, Thus having started out of the three child and five adult mortality estimates .The fitted logits are then computed by putting α and β and the standard logits on the straight line equation. Lastly anti logits are taken to produced asset of fitted ℓ_{xs} - the fitted of life table.

Having fitted the line, the values of α and (β) can be calculated, for (β) using the group average point such as:

$$\beta = \frac{y_a - y_c}{x_a - x_c} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{-}87)$$

For $\boldsymbol{\alpha}$ we can also conveniently use the group average points:

$$\alpha = y_c - \beta . X_c \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{-}88)$$

If the logit ℓ_x is denoted by $Y_{(x)}$ and the logit of the standard as $Y_{s(x)}$ as has become conventional in literature.

Having now calculated,α and β , it is possible to calculate the logits of the fitted model through the equation:

$$Y_{fit\,(x)} = \alpha + \beta . y_{s(x)} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{-}89)$$

Then the fitted ℓ_{xs} computed by taking the anti-logits of $Y_{fit\,(x)}$ using the equation:

$$Fitted\ \ell_x = \frac{1}{1 + e^{2Yfit(x)}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2\text{-}90)$$

Chapter Three

The Statistic Models

Univariate Time Series Modeling

3 -1 Preface:

In this chapter we will discuss in details a Time Series and Box-Jenkins Methodology .In the univariate case series model which it uses in data analysis**.**

3 -2 Definition of Time Series:

Time series is a sequential set of data points, measured typically over successive times. It is mathematically defined as a set of vectors **x** (***t***), **t** = 0,1,2,... where **t** represents the time elapsed. The variable **x** (***t***) is treated as a random variable. The measurements taken during an event in a time series are arranged in a proper chronological order. A time series containing records of a single variable is termed as univariate. But if records of more than one variable are considered, it is termed as multivariate. Time series can be continuous or discrete. In a continuous time series observations are measured at every instance of time, whereas a discrete time series contains observations measured at discrete points of time. For example temperature readings can be recorded as a continuous time series. On the other hand a series of life expectation at birth over many years may represent discrete time series. Usually in a discrete time series the consecutive observations are recorded at equally spaced time intervals such as hourly, daily, weekly, monthly or yearly time separations. The variable being observed in a discrete time series is assumed to be measured as a continuous variable using the real number scale. Furthermore a continuous time series can be easily transformed to a discrete one by merging data together over a specified time interval. In practice a suitable model is fitted to a given time series and the corresponding parameters are estimated using the known data values. The procedure of fitting a time series to a proper model is termed as Time Series Analysis. It comprises methods that attempt to understand the nature of the series and is often useful for future forecasting and simulation. In time series forecasting, past observations are collected and analyzed to develop a suitable mathematical model which captures the underlying data generating process for the series.[1]

There are many techeques used in time series for instance, Box-Jenkins ARIMA models, Box-Jenkins Multivariate Models, Holt-Winters Exponential Smoothing (single, double, triple).

3 -3Components of Time Series:

Time series in general is supposed to be affected by four main components, which can be separated from the observed data. These components are: Trend, Cyclical, Seasonal and Irregular components. A brief description of these four components is given here:

1. Secular Trend: The general tendency of a time series to increase, decrease or stagnate over a long period of time is termed as simply Trend. Thus, it can be said that trend is a long term movement in a time series. For example, series relating to population growth. Show upward trend, whereas downward trend can be observed in series relating to mortality rates, epidemics, etc.
2. Seasonal Variations: Seasonal variations in a time series are fluctuations within a year during the season. The important factors causing seasonal

[1] - Orlaith Burke., 2011, Statistical Methods Autocorrelation Box-Jenkins Modeling Introduction, University of Oxford, Department of Statistics

variations are: climate and weather conditions, customs, traditional habits, etc. Seasonal variation is an important factor for businessmen, shopkeeper and producers for making proper future plans.

3. Cyclical Variation: The cyclical variation in a time series describes the medium-term changes in the series, caused by circumstances, which repeat in cycles. The duration of a cycle extends over longer period of time, usually two or more years. Most of the economic and financial time series show some kind of cyclical variation.
4. Irregular Variations: Irregular or random variations in a time series are caused by unpredictable influences, which are not regular and also do not repeat in a particular pattern. These variations are caused by incidences such as war, strike, earthquake, flood, revolution, etc. There is no defined statistical technique for measuring random fluctuations in a time series.[1]

3 – 4 Fragmentation of the Time Series:

Considering the effects of these four components, two different types of models are generally used for a time series viz. Multiplicative and Additive models.

Multiplicative Model:

$$Y(t) = T(t) \times S(t) \times C(t) \times I(t).$$

Additive Model:

$$Y(t) = T(t) + S(t) + C(t) + I(t).$$

Where $Y(t)$ is the observation and $T(t)$, $S(t)$, $C(t)$ and $I(t)$ are respectively the trends, seasonal, cyclical and irregular variation at time t. Multiplicative

1 - Box G, Jenkins G, (1970), Time Series Analysis, Forecasting and Control, San Francisco: Holden-Day.

model is based on the assumption that the four components of a time series are not necessarily independent and they can affect one another; whereas in the additive model it is assumed that the four components are independent of each other. Time series has many applications such as, economic forecasting, stock market analysis, sales forecasting, census analysis, etc.

3 – 5 Techniques of Time Series:

There are many methods used to model and forecast time series which includes the following fitting models:

• Box-Jenkins ARIMA models

• Box-Jenkins Multivariate Models

• Holt-Winters Exponential Smoothing (single, double, triple).

3 – 6 Objectives of Time Series Analysis:

The usage of time series models is to:

1- Obtain an understanding of the underlying forces and structure that produced the observed data.

2- Fit a model and proceed to forecasting, monitoring or even feedback and forward control.

3-To access precise description - summary statistics and graphs of time series from which to describe the process to generate time series.

4-Generation model explains and interprets the behavior of this series with other variables.

5-Analysis and interpretation - find a model to describe the time dependence in the data, can we interpret the model.

6 - Control - adjust various control parameters to make the series fit closer to a target. [1]

3- 7Analysis of Time Series:

There are two directions for the analysis of time series:
1- Time Domain Analysis: Where the function uses is Auto Covariance Generating Function (ACGF) , Auto Correlation Functions (ACF) and Partial Auto Correlation Function(PACF) .In the analysis of time toward time-series would be Linear vehicle in the follow-up to the limits of random errors independent and identical.

2- Frequency Domain Analysis: It depends on the spectrum, so it called spectrum analysis. In this analysis described the function of the time series behavior in the limits of sine, cosine and different frequencies and this is done through Fourier Transform.[2]

3 – 8 A stationary:

A special class of stochastic processes is a stationary stochastic process. A statistical process is stationary if the probability distribution is the same for all starting values of *t*. This implies that the mean and variance are constant for all values of *t*. A series that exhibits a simple trend is not stationary because the values of the series depend on *t*. A stationary stochastic process is completely defined by its mean, variance, and autocorrelation function. One of the steps in the Box - Jenkins method is to transform a non-stationary series into a stationary one.

---

[1] - John Frain.(January 1999), " Lecture Notes on Univariate Time Series Analysis and Box Jankins Forecasting ", Economic Analysis, Research and Publications.

2 - Orlaith Burke., 2011)," Statistical Methods Autocorrelation- Box-Jenkins Modeling Introduction", Department of Statistics, 1 South Parks Road, Oxford OX1 3TG.

To analyze the time series according to the stages of analysis, it should be stationary and achieved if the following conditions are available:

The elements of our time series are denoted by:

$$X_1, X_2, X_3 \ldots\ldots, X_t$$

The mean and variance of the observation at time t are given by:

$$\mu_t = E\{X_t\} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3 - 1)$$

$$\sigma_t^2 = E\{(X_t - \mu_t)^2\} \ldots\ldots\ldots\ldots\ldots\ldots (3 - 2)$$

Respectively and covariance of X_t, X_{t-s} by

$$\text{Cov}(X_t, X_{t-s}) =$$

$$E\{(X_t - \mu_t)(X_{t-s} - \mu_{t-s})\} = \lambda_{t,t-s} \ldots\ldots\ldots\ldots (3 - 3)$$

In this system there is obviously too little information to estimate μ , σ_t^2 and $\lambda_{t,t-s}$ as we only have one observation for each time period. To proceed we need two properties, stationarity and periodicity. A series is second order stationary if:

$$\mu_t = \mu \qquad t = 1, 2, 3, \ldots\ldots$$

$$\sigma_t^2 = \sigma^2 \qquad t = 1, 2, 3, \ldots\ldots$$

$$\lambda_{t,t-s} = \tau_{t-s} \qquad t \neq s$$

That mean, variance and covariance are independent of time. A series is strictly stationary if the joint distribution of $(X_1, X_2, X_3 \ldots\ldots, X_t)$ is the same as that of $(X_{1+\tau}, X_{2+\tau}, X_{3+\tau} \ldots\ldots, X_{t+\tau})$ for all t and τ. If a series has a multivariate normal distribution then second order stationary implies strict stationary. Strict stationary implies second order stationary if the mean and variance exist and are finite.

For a non stationary series we may try to proceed in the following ways:

1-Find a transformation or some operation that makes the series stationary.
2- Estimate parameters.
3- Reverse the transformation or operation.

We can know the series whether stationary or non-stationary in several methods, including drawing the series and see the graph. If the form takes an upward trend or waived, it refers to a trend in other words, the time series is non-stationary in the middle and in order to be the series static in the middle we take differences. It is calculated as follows:

Time series after taking the first difference will be

$$y_t^{/} = y_t - y_{t-1} \ldots\ldots\ldots\ldots \quad (3-4)$$

$$y_t^{/} = y_t - \beta\, y_t = (1-\beta)\, y_t \ldots\ldots\ldots \quad (3-5)$$

After that test the stationary of the series ($\boldsymbol{y_t^{/}}$). If shows that, is still non-stationary can then find differences of second-class:

$$y_t^{//} = \nabla^2\, y_t = (1-1\beta)^2\, y_t \ldots\ldots\ldots \quad (3-6)$$

$$y_t^{//} = y_t^{/} - y_{t-1}^{/} = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2})$$

$$= y_t - 2y_{t-1} + y_{t-2} \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3-7)$$

$$y_t^{//} = y_t - 2\,\beta y_t + \beta^2 y_t \ldots\ldots\ldots\ldots\ldots \quad (3-8)$$

$$y_t^{//} = (1-\beta)^2\, y_t \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3-9)$$

As in the case of the fact that the series is non-stationary in variance, it addresses the transfer of data series using the conversion function:

$$y_t = \lambda \ (y_t) = \frac{y^{\lambda} - 1}{\lambda} \ \ ; \ \lambda \in (-\alpha \, , \alpha) \ \ldots\ldots\ldots\ldots\ldots \ (3 - 10)$$

Where $\boldsymbol{\lambda}$ is the Parameter of conversion.

The following table shows the most commonly used values for the parameter λ Transfers with the corresponding.

Table (3-1): The Most Commonly Used Values for the Parameter $\boldsymbol{\lambda}$ Transfers with the Corresponding

$\boldsymbol{\lambda}$	-1.0	-0.5	0.0	0.5	1.0
y_t	$\frac{1}{y_t}$	$\frac{1}{\sqrt{y_t}}$	$ln y_t$	$\sqrt{y_t}$	y_t

Source : an Application of ARIMA Models" ,International Journal of Applied Science and Technology.,1998

Note that, d is the amount of the difference, which is calculated Data, it makes data stationary and smooth most often be worth equal to 1.2 Can be made more than a difference but usually is the first or second teams enough. (Spyros M.Steren.,1983 &Anderson .,1971).

3-9 Time Series Analysis Models:

In general models for time series data can have many forms and represent different stochastic processes. There are two widely used linear time series models in literature, Autoregressive (AR) and Moving Average (MA) models. Combining these two, the Autoregressive Moving Average (ARMA) and Autoregressive Integrated Moving Average (ARIMA) models have been proposed in literature. The Autoregressive Fractionally Integrated Moving Average (ARFIMA) model generalizes ARMA and

ARIMA models. For seasonal time series forecasting, a variation of ARIMA, viz, the Seasonal Autoregressive Integrated Moving Average (SARIMA) model is used. ARIMA model and its different variations are based on the famous Box-Jenkins principle and so these are also broadly known as the Box-Jenkins models.

Linear models have drawn much attention due to their relative simplicity in understanding and implementation. However many practical time series show non-linear patterns. Non-linear models are appropriate for predicting volatility changes in economic and financial time series. Considering these facts, various nonlinear models have been suggested in literature. Some of them are the famous Autoregressive Conditional Heteroskedasticity (ARCH) model and its variations like Generalized ARCH (GARCH), Exponential Generalized ARCH (EGARCH) etc., the Threshold Autoregressive (TAR) model, the Non-linear Autoregressive (NAR) model, the Nonlinear Moving Average (NMA) model, etc.

In this section we shall discuss about the important linear stochastic time series models with their different properties, which it should be used in our analysis. Before beginning the models it is necessary to explain the basic equations.

The general equation is:

$$x_t = f\,(x_{t-1}, x_{t-2}\,, \dots \dots .\,.\,\mu_i) \dots\dots\dots\dots \quad (3 - 11)$$

Define the backshift or lag operator β by:

$$\beta y_t \quad = \quad y_t - 1$$

Now suppose a time series y_t satisfies the following:

$$\alpha\,(\beta)\,(y_t - \mu\,) = \theta\,(\,\beta\,) \in_t \dots\dots\dots\dots\dots \quad (3\text{– } 12)$$

Where

$$\alpha\ (\beta) = (\ 1\text{-}\ \alpha_1\ \beta\ \text{-}\ \ldots\ldots\text{-}\ \alpha_{p+d}\ \beta^{p+d}\ )\ldots\ldots\ldots\ldots\ (3 – 13)$$

and $$\theta\ (\beta\ ) =\ (\ 1\ \text{-}\ \theta_1\beta\ \text{-}\ \ldots\ldots\theta_q\beta^q\ )\ \ldots\ldots\ldots\ (3 – 14)$$

and where $\in_t$ is IID white noise with mean 0 and variance σ^2 and the α_i and $\theta_i$are parameters.

Let (z) be a complex variable and define:

$$\alpha\ (\ z) = 1\text{-}\ \alpha_1 z_1\text{-}\ \ldots\ldots\ \text{-}\ \alpha_{p+d} z^{p+q}\ \ldots\ldots\ (3 – 15)$$

We will discover that the behavior of the time-series y_t is affected by the location in the complex plane of the roots of α (z).

To help us visualize how the roots do, in fact, influence the time series we begin by re-expressing (3 –1 2):

$$\alpha\ (\beta)\ (y_t -\ \mu\ ) = \theta\ (\ \beta\ )\ \in_t$$

Suppose (z) = +1 has d unit roots, z = +1, where d ≥ 0 then we can be rewritten (3-12) as:

$$\alpha\ (\beta) = \emptyset\ (\beta)\ (1-\beta)^d\ \ \ldots\ldots\ldots\ \ (3\text{-}16)$$

So that (3-12) can be rewritten as

$$\emptyset\ (\beta)\ \ \Delta^d(y_t -\ \mu\ ) = \theta\ (\ \beta\ )\ \in_t\ \ldots\ldots\ (3\text{-}17)$$

Where the difference operator

$$\Delta = (\beta\ \text{-}\ 1)$$

and

$$\emptyset\ (\beta)\ \ = (1\ \text{-}\ \emptyset_1\ \ \beta - \cdots \emptyset_p\ \beta^p\ \ )$$

$$\theta\ (\beta\ ) = (1\ \text{-}\ \theta_1\ \ \beta - \cdots \theta_q\ \beta^q\ \ )$$

The model (3-17):

$$\emptyset\,(\beta) = (\,1 - \emptyset_1\ \beta - \cdots \emptyset_p\ \beta^p\ )\ \Delta^d(y_t - \mu\,)$$

$$= (1 - \theta_1\ \beta - \cdots \theta_q\ \beta^q\ )\ \in_t\ \ldots\ldots\ldots\ldots\ (3\text{-}18)$$

Now that we have seen the definition of basic equations, let us investigate the models.

3-10 ARMA Modeling:

There are three steps in ARMA modeling:

1- Check the series for stationary and if necessary transform the series to include stationary.

2-From the autocorrelation properties of the transformed series choose a few ARMA specifications for estimation and testing in order to arrive at preferred specification with white noise residuals.

3-Calculate forecasting over a relevant time horizon from the preferred specification.

We will concentrate on stage 2 first before looking at stage 1 .The basic idea is to derive the autocorrelation patterns associated with various low-order AR, MA and ARMA schemes .Comparing these with the empirical patterns computed from the series under analysis then suggests one or more ARMA specifications for statistical estimation and testing.

3 – 10-1 Autoregressive Model:

In these models the current value of the series (x_t) is expressed as a liner combination of the previous values ($x_{1,}\ x_2\ ,x_3\ ,$, ...).If we take the series value as the deviation from its mean ($z_t = x_t - \mu$), then the autoregressive model of the order (p) may be written as:

$$Z_t = \emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \cdots + \emptyset_p Z_{t-p} + \in_t \ldots\ldots\ldots \quad (3-19)$$

Where :

$\emptyset_1 \emptyset_2 \emptyset_3 \ldots\ldots, \emptyset_p$ are parameters of the model.

$\in_t$ is the error variable that satisfies the assumption which were mentioned above.

If we use the actual value of the series x_t equation (3 - 19) becomes:

$$x_t - \mu = \emptyset_1 (x_{t-1} - \mu) + \emptyset_2 (x_{t-2} - \mu) + \cdots + \emptyset_p (x_{t-p} - \mu) + \in_t$$

$$x_t = \mu(1 - \emptyset_1 - \emptyset_2 - \cdots - \emptyset_p) + \emptyset_1 x_{t-1} + \emptyset_2 x_{t-2} + \ldots. + \emptyset_p x_{t-p} + \in_t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots. \quad (3-20)$$

AR (1) Model:

Consider the autoregressive model of order one (1):

$$Z_t = \emptyset_1 Z_{t-1} + \in_t \ldots\ldots\ldots\ldots \quad (3-21)$$

by $a$ctual values: $x_t = \mu (1 - \emptyset_1) + \emptyset_2 x_{t-2} + \in_t$

With $\mu (1 - \emptyset_1)$ as the intercept.

Properties of AR (1) Model :

i- Mean :

$$\text{E} (Z_t) = E (\emptyset_1 Z_{t-1} + \in_t)$$

$$= \text{E} (\emptyset_1 Z_{t-1}) + E (\in_t)$$

$E\ (\emptyset_1 Z_{t-1})$ since $E\ (Z_{t-1})$

$= E\ (x_t - \mu) = 0$ ………….. (3 -22)

ii- Variance :

$V\ (Z_t) = E\ (z_t^2) - (E\ (z_t\ ))2 = E(z_t^2)$

$= E\ \{\emptyset_1 Z_{t-1} + \in_t\}^2$

$= \emptyset_1^2\ E\ (z_{t-1}^2\ ) + 2\ \emptyset_1 E\ (Z_{t-1} \in_t\ ) + E\ (\in_t^2)$

$= \emptyset_1^2\ v\ (z_{t-1}\ ) + 0 + \sigma_\epsilon^2$

$$\lambda_0 = \emptyset_1^2\ \lambda_0 + \sigma_\epsilon^2 = \frac{\sigma_\epsilon^2}{(1-\emptyset_1^2)} \text{ ……….. } (3-23)$$

Since the series is stationary $V\ (Z_t) = 0$

Since the series is positive then :

$\emptyset_1 < 1$ or $-1 < \emptyset_1$. If $\emptyset_1 \neq 1$ then $V\ (Z_t) = \infty$ .Becomes infinite the series which is not stationary.

iii − Convariance:

$\lambda_s = Cov\ (Z_t, Z_{t-s}\ )$

$\lambda_1 = Cov\ (Z_1, Z_{t-1}\ )$

$= E\ (Z_t, Z_{t-1}) - E\ (Z_t\ )\ E(Z_{t-s}\ ) = E\ (Z_t, Z_{t-1})$

$E\ (Z_t, Z_{t-1}) = E(\emptyset_1 Z_{t-1} + \in_t\ )\ Z_{t-1}$

$= \emptyset_1\ E\ (z_{t-1}^2\ ) + (Z_{t-1} \in_t\ )$

$\lambda_1 = \emptyset_1 V\ (Z_{t-1}) = \emptyset_1\ \lambda_0 + 0$

$\therefore\ \lambda_2 = \emptyset_1\ \lambda_1 = \emptyset_1^2\ \lambda_0$

By the same method we find that:

$$\lambda_3 = \emptyset_1 \ \lambda_2 = \emptyset_1^3 \ \lambda_0 \ , \qquad \lambda_4 = \emptyset_1 \ \lambda_3 = \emptyset_1^4 \ \lambda_0$$

In general :

$$\lambda_k = \emptyset_1 \ \lambda_{k-1} = \emptyset_1^k \ \lambda_0 \ , \qquad k=1,2,3,\ldots. \qquad (3-24)$$

***iv*) Autocorrelation:**

$$p_s = \frac{\lambda_s}{\lambda_0} \ , \qquad s=1,2,3,\ldots.$$

$$p_1 = \frac{\lambda_1}{\lambda_0} = \ p_1 = \frac{\emptyset_1 \lambda_0}{\lambda_0} = \emptyset_1$$

$$p_2 = \frac{\lambda_2}{\lambda_0} = \ p_2 = \frac{\emptyset_1^2 \lambda_0}{\lambda_0} = \emptyset_1^2$$

$$p_3 = \frac{\lambda_3}{\lambda_0} = \ p_3 = \frac{\emptyset_1^3 \lambda_0}{\lambda_0} = \emptyset_1^3$$

$$p_k = \frac{\lambda_k}{\lambda_0} = \ p_k = \frac{\emptyset_1^k \lambda_0}{\lambda_0} = \emptyset_1^k \ , \qquad k=1,2,3,\ldots. \qquad (3-25)$$

Memory Function:

The autoregressive model describe by having long memory (strong memory), AR(1) model as equation (3 –21) can be written as:

$$Z_t = \emptyset_1 \ (Z_{t-1} + \ \in_{t-1}) + \in_t$$

$$= \ \in_t + \emptyset_1 \ \in_{t-1} + \emptyset_1^2 \ Z_{t-2}$$

$$= \ \in_t + \emptyset_1 \ \in_{t-1} + \emptyset_1^2 (\emptyset_1 \ Z_{t-3} \in_{t-2})$$

$$\vdots$$

$$= \ \in_t + \emptyset_1 \ \in_{t-1} + \emptyset_1^2 \in_{t-2} + \emptyset_1^3 \in_{t-3} + \ldots\ldots\ldots\ldots$$

$$\rightarrow \sum_{j=0}^{\infty} \emptyset_1^j \in_{t-j} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \qquad (3-26)$$

So AR (1) model can be written as a sum of current random error and infinite number of past random, then for the current observation (z_t) affected by the random errors that happened in the long past. That is mean AR (1) model is said to have infinite memory. If the series is stationary then $\emptyset_1 < 1$ this make that the effect past random errors will be gradually disappeared.

The memory coefficient defined as coefficient of $\in_{t-1}$ in equation (3 - 26) that is $\emptyset_1$.

AR (2) Model:

$$Z_t = \emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \in_t \ldots\ldots\ldots\ldots \quad (3\text{-}27)$$

By actual values:

$$x_t = \mu (1 - \emptyset_1 - \emptyset_2) + \emptyset_1 x_{t-1} + \emptyset_2 x_{t-2} + \in_t$$

With intercept: $\mu (1 - \emptyset_1 - \emptyset_2)$

Properties oF AR (2) MODEL:

i) Mean:

$$E(Z_1) = E(\emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \in_t)$$
$$= \emptyset_1 E(Z_{t-1}) + \emptyset_2 E(Z_{t-2}) + E(\in_t)$$
$$E(Z_1) = 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3\text{- }28)$$

ii) Variance :

$$V(Z_t) = E(Z^2{}_t) - \left(E(Z_t)\right)^2 = E(Z^2{}_t)$$

$$= E[(\emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \in_t)^2]$$

$$= \varnothing_1^2 E(Z^2{}_{t-1}) \varnothing_1^2 E(Z_{t-2})^2 + E(Z_{t-2})^2 + E(\in_t)^2$$
$$+ 2\emptyset_1\emptyset_2 E(Z_{t-1} Z_{t-2}) + 2\emptyset_1 E(Z_{t-1} \in_t) + 2\emptyset_2 E(Z_{t-2} \in_t)$$

$$= Ø_1^2 \text{ V } (Z_{t-1}) + Ø_2^2 \text{ V } (Z_{t-2}) + \sigma^2 + 2\emptyset_1\emptyset_2 + 0$$

$$\lambda_0 = \frac{2\emptyset_1\emptyset_2\lambda_1+\sigma^2}{(1- Ø_1^2- Ø_2^2)} \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3\text{-}29)$$

iii) Covariance :

$$\text{Cov (x, y)} = E[(x - E(x))(y - E(y))]$$

$$\lambda_s = Cov\ (Z_t, Z_{t-s})$$

$$\lambda_1 = E(Z_t\ Z_{t-1}) = E(Z_t\ Z_{t-1}) - E(Z_t)E(Z_{t-1})$$

$$= E(\emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2})Z_{t-1}$$

$$= \emptyset_1\ \lambda_0 + \emptyset_2\lambda_1$$

$$\lambda_1 = \frac{\emptyset_1\lambda_0}{(1-\emptyset_2\ )}$$

$$\lambda_2 = E(\emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \in_t)Z_{t-2}$$

$$= \emptyset_1 E(Z_{t-1}\ Z_{t-2}) + \emptyset_2 Z_{t-2}^2 + E(\in_t\ Z_{t-2})$$

$$\lambda_2 = \emptyset_1\lambda_1 + \emptyset_2\lambda_0$$

$$\lambda_3 = E(Z_t, Z_{t-3}) = E(\emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \in_t)Z_{t-3}$$

$$= \emptyset_1\lambda_2 + \emptyset_2\lambda_1$$

$$\lambda_4 = \emptyset_1\lambda_3 + \emptyset_2\lambda_2$$

$$\lambda_k = \begin{cases} \dfrac{\emptyset_1}{(1-\emptyset_2\ )}\ \lambda_0 & ., \quad K = 1 \\ \emptyset_1\lambda_{k-1} + \emptyset_2\lambda_{k-2} & ., \quad K = 2, 3, \ldots\ldots \end{cases} \quad (3\ \text{-30})$$

iV) Autocorrelation:

$$\rho_s = \frac{\lambda_s}{\lambda_0}$$

$$\rho_1 = \frac{\lambda_1}{\lambda_0} = \frac{\emptyset_1\lambda_0}{(1-\emptyset_2\ )\lambda_0} = \frac{\emptyset_1}{1-\emptyset_2}$$

Ingernal :

$$\rho_k = \begin{cases} \frac{\emptyset_1}{(1-\emptyset_2)} & ., \quad K = 1 \\ \frac{\emptyset_1 \lambda_{k-1} + \emptyset_2 \lambda_{k-2}}{\lambda_0} & ., K = 2, 3, \ldots \end{cases} \quad (3\text{- }31)$$

$$\rho_s = \emptyset_1 \rho_1 + \emptyset_2 \rho_0 = \left(\frac{\emptyset_1^2}{1 + \emptyset_2} \right) + \emptyset_2$$

Memory Function:

$$Z_t = \emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \in_t$$

$$= \emptyset_1 (\emptyset_1 Z_{t-2} + \emptyset_2 Z_{t-3} + \in_{t-1}) + \in_t$$

$$= \emptyset_2 (\emptyset_1 Z_{t-3} + \emptyset_2 Z_{t-4} + \in_{t-2}) + \in_t$$

$$= \in_t + \emptyset_1 \in_{t-1} + \emptyset_2 \in_{t-2} + \emptyset_1^2 Z_{t-2} + 2\emptyset_1 \emptyset_2 Z_{t-4}$$

$$Z_t = \in_t + \emptyset_1 \in_{t-1} + \emptyset_1^2 \in_{t-2} + \emptyset_1^3 \in_{t-3} + \ldots \ldots \quad (3\text{-}32)$$

We see – that the Coefficient of $\in_{t-1}$ is $\emptyset_1$ which is the memory Coefficient.

3-10- 2 Moving Average Models:

In a pure MA process a variable is expressed solely in terms of the current and previous white noise disturbances. The MA of order q is MA (q) and its equations are as:

$$y_t = - \theta_1 \in_{t-1} - \theta_2 \in_{t-2} -$$

$$\theta_3 \in_{t-3} - \ldots\ldots\ldots - \theta_q \in_{t-q} + \in_t \ldots\ldots \quad (3 - 33)$$

Where :

$\theta_1, \theta_2, \theta_3,$ ……,θ_q are the parameters of the model.

$\in_1$, $\in_{t-1}, \in_{t-2}, \dots, \in_{t-q}$ are the errors which satisfy the following assumptions:

i- E($\in_t$) = 0

ii- Var ($\in_t$) = σ_t^2

iii- $\in_t \sim N$ (0, σ_t^2)

vi- E($\in_t, \in_{t-s}$) = 0

vii- E($\in_t, z_{t-s}$) =0

$z_{t=}\ x_t - \mu$ = deviations

MA (1) Models:

MA (1) models is a regression of z_t on $\in_{t-1}$as a predictor. The moving average model of order (1) as:

$z_{t=}$ μ - $\theta_1 \in_{t-1} + \in_t$, μ is the intercept or mean of time series.

Properties of MA (1):

(i) Mean :

E(z_t) = E (- $\theta_1 \in_{t-1} + \in_t$) = 0 ……… (3 -34)

(ii) Variance :

λ_o = V (z_t) = E (z_t^2)

= E $(- \theta_1 \in_{t-1} + \in_t)^2$

$$= E(\theta_1{}^2 \in_{t-1}{}^2 - 2\,\theta_1 \in_{t-1} \in_t + \in_t{}^2)$$

$$= \theta_1{}^2 E(\in_{t-1}{}^2) - 2\theta_1 E(\in_{t-1} \in_t) + E(\in_t{}^2)$$

$$= \theta_1{}^2\sigma_\in^2 + \sigma_\in^2$$

$$\lambda_o = (1 + \theta_1{}^2)\,\sigma_\in^2$$

(iii) $\lambda_s = \text{Cov}\ (z_{t,}\ z_{t-1})$ s=1,2,3,…..,$\frac{n}{2}$

$= E(z_{t,}\ z_{t-1})$, $\lambda_s = (z_{t,}\ z_{t-1})$

$= E(-\theta_1 \in_{t-1} + \in_t)\ z_{t-1}$

$= -\theta_1 E(\in_{t-1},\ z_{t-1})$

$= -\theta_1 E(\in_{t-1}\ (-\theta_1 \in_{t-2} + \in_{t-1}))$

$\lambda_1 = -\theta_1\sigma_\in^2$

$\lambda_2 = E(z_{t,}\ z_{t-2})$

$= E(-\theta_1 \in_{t-1} + \in_t)\ z_{t-2} = 0$

$\lambda_3 = 0$, $\lambda_4 = 0$

In general :

$$\lambda_s = \begin{cases} -\theta_1\ \sigma_\in^2 & ,\ s = 1 \\ 0 & .,\ s = 2,3,4,\ldots\ldots \end{cases} \quad (3-35)$$

Autocorrelation:

$$p_s = \frac{\lambda_s}{\lambda_0}, \qquad s = 1,2,3,\ldots..,\frac{n}{2}$$

$$p_1 = \frac{\lambda_1}{\lambda_0} = \frac{-\theta_1 \sigma_\in^2}{(1+{\theta_1}^2)\,\sigma_\in^2} = \frac{-\theta_1}{(1+{\theta_1}^2)}$$

$$p_2 = \frac{\lambda_2}{\lambda_0} = 0 \qquad , p_3 = \frac{\lambda_3}{\lambda_0} = 0$$

In general :

$$p_s = \begin{cases} \dfrac{-\theta_1}{(1+{\theta_1}^2)} & ., s = 1 \\ 0 & ., s = 2,3,\ldots \end{cases} \quad (3-36)$$

Memory Function:

Since $z_{t=}$ μ - $\theta_1 \in_{t-1}$ + $\in_t$ and the memory coefficienal is coefficient of the past error in the model .So (- θ_t) is the memory coefficient in MA (1) model.

MA (2) Model:

The moving Average of order (2) is given by:

$$Z_t = -\theta_1 \in_{t-1} - \theta_2 \in_{t-2} + \in_t \quad \ldots\ldots\ldots\ldots \quad (3-37)$$

by original data : $Z_t = \mu - \theta_1 \in_{t-1} - \theta_2 \in_{t-2} + \in_t$

with intercept of μ.

Properties of MA (2) Model :

i- Mean :

$$E(Z_t) = E(-\theta_1 \in_{t-1} - \theta_2 \in_{t-2} + \in_t) = 0 \ \ldots\ldots\ldots \ (3\text{- }38)$$

ii-Variance :

$$V(Z_t) = E(z_t)^2 = E(-\theta_1 \in_{t-1} - \theta_2 \in_{t-2} + \in_t)^2$$

$$= \theta_1^2\, E({\in_{t-1}}^2) + \theta_2^2\, E({\in_{t-2}}^2) + E({\in_t}^2)$$

$$\lambda_0 = (1 + \theta_1^2 + \theta_2^2)\, \sigma_\epsilon^2 \quad \ldots\ldots\ldots\ldots\ldots \quad (3\text{-}39)$$

iii – Covariance :

$$\lambda_s = \text{Cov}\,(Z_t, Z_{t-s})$$

$$\lambda_1 = E\,(Z_t, Z_{t-1}) = E\{(-\theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} + \epsilon_{t-1})\, Z_{t-1}\}$$

$$= -\theta_1\, E\,(\epsilon_{t-1} Z_{t-1}) - \theta_2\, E\,(\epsilon_{t-2} Z_{t-1})$$

$$= -\theta_1\, E\,\{\epsilon_{t-1}(-\theta_2 \epsilon_{t-2} - \theta_3 \epsilon_{t-3} + \epsilon_{t-1})\}$$

$$= -\theta_2\, E\,\{\epsilon_{t-2}(-\theta_1 \epsilon_{t-2} - \theta_3 \epsilon_{t-3} + \epsilon_{t-1})\}$$

$$= -\theta_1 \sigma_\epsilon^2 + \theta_1 \theta_2\, \sigma_\epsilon^2$$

$$\lambda_1 = \theta_1\,(\theta_2 - 1)\, \sigma_\epsilon^2$$

$$\lambda_2 = \text{Cov}\,(Z_t, Z_{t-s}) = E\,(Z_1, Z_{t-2})$$

$$= E\{(-\theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} + \epsilon_{t-1})\, Z_{t-2}\}$$

$$= -\theta_2\, E\,(\epsilon_{t-2} Z_{t-2})$$

$$= -\theta_2\, E\,\{\epsilon_{t-2}(-\theta_1 \epsilon_{t-3} - \theta_2 \epsilon_{t-4} + \epsilon_{t-2})\}$$

$$\lambda_2 = -\theta_2\, \sigma_\epsilon^2$$

$$\lambda_3 = \text{Cov}\,(Z_t, Z_{t-s}) = E\,(Z_1, Z_{t-3})$$

$$= E\,\{(-\theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} + \epsilon_t)\, Z_{t-3}\} = 0$$

$$\lambda_4 = 0, \qquad \lambda_5 = 0$$

In general :

$$\lambda_s = \begin{cases} \theta_1\,(\theta_2 - 1)\, \sigma_\epsilon^2 & , s = 1 \\ -\theta_2\, \sigma_\epsilon^2 & , s = 2 \\ 0 & , s = 3,4,5,\ldots \end{cases} \quad (3\text{-}40)$$

iv – Autocorrelation :

$$p_s = \frac{\lambda_s}{\lambda_0}$$

$$p_1 = \frac{\lambda_1}{\lambda_0} = \frac{\theta_1 (\theta_2 - 1) \sigma_\epsilon^2}{(1 + \theta_1^2 + \theta_2^2) \sigma_\epsilon^2} = \frac{\theta_1 (\theta_2 - 1)}{(1 + \theta_1^2 + \theta_2^2)}$$

$$p_2 = \frac{\lambda_2}{\lambda_0} = \frac{- \theta_2 \ \sigma_\epsilon^2}{(1 + \theta_1^2 + \theta_2^2) \sigma_\epsilon^2} = \frac{- \theta_2}{(1 + \theta_1^2 + \theta_2^2)}$$

$$p_3 = \frac{\lambda_3}{\lambda_0} = 0$$

In general :

$$p_s = \begin{cases} \frac{\theta_1 (\theta_2 - 1)}{(1 + \theta_1^2 + \theta_2^2)} & , s = 1 \\ \frac{- \theta_2}{(1 + \theta_1^2 + \theta_2^2)} & , s = 2 \\ 0 & , s = 3,4,5,\ldots \end{cases} \quad (3 – 41)$$

Memory Function: Memory coefficient of MA (2) is θ_2 .

3 -10-3 Autoregressive Moving Average Models (ARMA):

Actual extension of the autoregressive and moving average models is to combine the both models such as mixed process are referred to as autoregressive moving average model :

ARMA (p, q) is a symbol that denoted to a such model, where (P) is the order of autoregressive and (q) is the order of moving, this model is expressed by :

$$Z_t = \emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \ldots\ldots + \emptyset_p Z_{t-p} - \theta_1 \in_{t-1} - \theta_2 \in_{t-2} - \ldots\ldots - \theta_q \in_{t-q} + \in_t \ \ldots\ldots \ (3\text{-}42)$$

ARMA (1,1) Model :

$$Z_t = \emptyset_1 Z_{t-1} - \theta_1 \in_{t-1} + \in_t \quad \ldots\ldots\ldots \text{(3-43)}$$

by the Original date

$$\gamma_t = \mu\,(1-\emptyset_1) + \emptyset_1 \gamma_{t-1} - \theta_1 \in_{t-1} + \in_t$$

With intercept $\mu\,(1-\emptyset_1)$

Properties of ARMA (1, 1):

i) Mean:

$$E\,(Z_t) = \mathrm{E}\,(\emptyset_1 Z_{t-1} - \theta_1 \in_{t-1} + \in_t\,) = 0$$

ii) Variance:

$$\lambda_0 = V(Z_t) = E(Z_t^2)$$

$$= E\,(\emptyset_1 Z_{t-1} - \theta_1 \in_{t-1} + \in_t\,)^2 = E(\emptyset_1^2\, Z_{t-1}^2 + \theta_1^2\, \in_{t-1}^2 + \in_t^2 - 2\theta_1\emptyset_1 Z_{t-1} \in_{t-1} + 2\emptyset_1 Z_{t-1} \in_t - 2\theta_1 \in_{t-1}\in_t)$$

$$= \emptyset_1^2\, V(Z_{t-1}) + \theta_1^2\, V(\in_{t-1}) + V(\in_t) - 2\theta_1\emptyset_1 V(\in_t)$$

$$= \emptyset_1^2 \lambda_0 + \theta_1^2\, \sigma_\in^2 + \sigma_\in^2 - 2\theta_1\emptyset_1\sigma_\in^2$$

$$\lambda_0 = \frac{(\theta_1^2 + 1 - 2\emptyset_1\emptyset_1)\sigma_\in^2}{(1-\emptyset_1^2)} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad \text{(3- 44)}$$

iii) Covariance :

$$\lambda_s = Cov\,(Z_t\,, Z_{t-s})\,, \qquad S = 1, 2, 3, \ldots\ldots, \frac{n}{2}$$

$$\lambda_1 = Cov\,(Z_t\,, Z_{t-1}) = E(Z_t\,, Z_{t-1})$$

$$= \mathrm{E}\,(\emptyset_1 Z_{t-1} - \theta_1 \in_{t-1} + \in_t) Z_{t-1}$$

$$= \emptyset_1\, V(Z_{t-1}) - \theta_1 V(\in_t) = \emptyset_1 \lambda_0 - \theta_1 \sigma_\in^2$$

$$= \frac{(\emptyset_1 - \theta_1)(1-\theta_1\emptyset_1)\sigma_\in^2}{(1-\emptyset_1^2)}$$

$$\lambda_2 = Cov\,(Z_t\,, Z_{t-2}) = E(Z_t\,, Z_{t-2})$$

$$= \text{E}\,(\emptyset_1 Z_{t-1} - \theta_1 \in_{t-1} + \in_t) Z_{t-2}$$

$$\lambda_3 = \emptyset_1 \lambda_2 \quad \text{and} \quad \lambda_4 = \emptyset_1 \lambda_3$$

In general :

$$\lambda_k = \begin{cases} \frac{(\emptyset_1 - \theta_1)(1-\theta_1\emptyset_1)\sigma_\in^2}{(1-\emptyset_1^2)} & , \quad \text{K} = 1 \\ \emptyset_1 \lambda_{k-1} & , \quad \text{K} = 2, 3, \ldots\ldots, \frac{n}{z} \ldots\ldots\ldots \end{cases} \quad (3\text{-}45)$$

iv) Autocorrelation:

$$\rho_s = \frac{\lambda_s}{\lambda_0}$$

$$\rho_1 = \frac{\lambda_1}{\lambda_0} = \frac{\frac{(\emptyset_1 - \theta_1)(1-\theta_1\emptyset_1)\sigma_\in^2}{(1-\emptyset_1^2)}}{\left(\frac{(1+\theta_1^2 - 2\theta_1\emptyset_1)\sigma_\in^2}{(1-\emptyset_1^2)}\right)} = \frac{(\emptyset_1 - \theta_1)(1-\theta_1\emptyset_1)}{(1+\theta_1^2 - 2\theta_1\emptyset_1)}$$

$$\rho_2 = \frac{\lambda_2}{\lambda_0} = \frac{\emptyset_1\lambda_1}{\lambda_0} = \emptyset_1\rho_1$$

$$\rho_3 = \frac{\lambda_3}{\lambda_0} = \frac{\emptyset_1\lambda_2}{\lambda_0} = \emptyset_1\rho_2$$

In general :

$$\rho_k = \begin{cases} \frac{(\emptyset_1 - \theta_1)(1-\theta_1\emptyset_1)}{(1+\theta_1^2 - 2\theta_1\emptyset_1)} & , \text{K} = 1 \\ \emptyset_1\rho_{k-1} & , \text{K} = 2, 3, \ldots\ldots, \frac{h}{z} \ldots\ldots \end{cases} \quad (3\text{-}46)$$

V) Memory Function:

$$Z_t = \emptyset_t z_{t-1} - \theta_1 \in_{t-1} + \in_t$$

$$= \emptyset_1 (\emptyset_1 Z_{t-1}\ \theta_1 \in_{t-2} + \in_2) - \theta_1 \in_{t-1} + \in_t$$

$$= \in_t + (\emptyset_1 - \theta_1) \in_{t-1} - \emptyset_1\ \theta_1 \in_{t-2} + \emptyset_1^2\ (\emptyset_1 Z_{t-3} - \theta_1 \in_{t-3} + \in_{t-2})$$

$$= \in_t + (\emptyset_1 - \theta_1) \in_{t-1} + \emptyset_1 (\emptyset_1 - \theta_1) \in_{t-2}$$

$$- \emptyset_1^2 \theta_1 \in_{t-3} + \emptyset_1^3\ (\emptyset_1 Z_{t-4} - \theta_1 \in_{t-4} + \in_{t-3})$$

$$\vdots$$

$$= \in_t + (\emptyset_1 - \theta_1) \in_{t-1} + \emptyset_1 (\emptyset_1 - \theta_1) \in_{t-2}$$

$$+ \emptyset_1^2 (\emptyset_1 - \theta_1) \in_{t-3} + \emptyset_1^3\ (\emptyset_1 - \theta_1) \in_{t-4}$$

$$+ \emptyset_1^4 + (\emptyset_1 - \theta_1) \in_{t-5} + \ldots\ldots\ldots \qquad (3\text{-}47)$$

So there is a long memory (strong) for the model ARMA (1, 1), but according to stationary coalitions ($|\emptyset_1| < 1$ and ($|\emptyset_1 - \theta_1| < 1$), the memory becomes the terms of equation above (3 - 47) which consist of:

$\emptyset_1^2$, $\emptyset_1^3$, …., will be gradually disappeared, so the memory coefficient is ($\emptyset_1 - \theta_1$).

3–10-4 Autoregressive Integrated Moring Average Models (ARIMA):

If the time series is not stationary, we cannot use ARMA (p, q) models in analysis and forecasting from the certain time series. Because ARMA models assumed that the time series is stationary, so we must transform the non-stationary time series to stationary, by taking the difference. For such a series we use the ARIMA (Autoregressive Integrated Moving Average Models).

The ARIMA (p, d, q) Consist of Three Orders:

p = for autoregressive (p = 0, 1, 2, ……)

d = for differences (d = 0, 1, 2, ……) (at most d = 2 in many time series)

q = for moving average (q = 0, 1, 2, …….)

$$\text{let } W_t = \nabla^d Z_t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3\text{-}48)$$

where $\nabla = 1 - \beta$, as we mentioned above

Now ARIMA (p, d, q) for the time series Z_t is ARIMA (p, q) for the time series W_t

So: β is – the Back shift operator .

$$\beta Z_t = Z_{t-1}, \beta^2 Z_t = Z_{t-2} \ldots\ldots\ldots\ldots\ldots \quad (3\text{-}49)$$

$$W_2 = \emptyset_1 W_{t-1} + \emptyset_2 W_{t-2} + \ldots\ldots + \emptyset_p W_{t-p}$$

$$- \theta_1 \in_{t-1} - \theta_2 \in_{t-2} - \ldots.. - \theta_q \in_{t-q} + \in_t \quad \ldots\ldots. \quad (3\text{-}50)$$

$$= \emptyset_1 W_{t-1} - \emptyset_2 W_{t-2} - \ldots\ldots - \emptyset_p W_{t-p} - \in_t$$

$$- \theta_1 \in_{t-1} - \theta_2 \in_{t-2} - \ldots\ldots - \theta_q \in_{t-q}$$

$$= (1 - \emptyset_1 \beta - 1 - \emptyset_2 \beta^2 - \emptyset_p \beta^p) W_t$$

$$- (1 - \theta_1 \beta - \theta_2 \beta^2 - \ldots\ldots - \emptyset_q \beta^q) \in_t$$

If we define:

$$\emptyset(\beta) = 1 - \emptyset_1 \beta - \emptyset_2 \beta^2 - \dots\dots\dots - \emptyset_p \beta^q$$

$$\theta(\beta) = 1 - \theta_1 \beta - \theta_2 \beta^2 - \dots\dots\dots - \theta_q \beta^q$$

$$\Rightarrow \theta(\beta)\, W_t = \emptyset(\beta)\, \epsilon_t$$

Is the general form if ARIMA (p, d, q) model:

3 - 11The Three Stages of ARIMA Modeling:

The analysis performed by PROC ARIMA is divided into three stages, corresponding to the stages described by Box and Jenkins (1976). The Identify, Estimate, Diagnoses and Forecast statements perform these three stages, which are summarized below. [1]

3 –11–1 Identification of ARIMA Models:

In the identification stage, we use the Identify statement to specify the response series and identify candidate ARIMA models for it. The Identify statement reads time series that are to be used in later statements, possibly differencing them, and computes autocorrelations, inverse autocorrelations, partial autocorrelations, and cross correlations. Stationarity tests can be performed to determine if differencing is necessary. The analysis of the Identify statement output usually suggests one or more ARIMA models that could be fit. Options allow you to test for stationarity and tentative ARMA order identification.

---

1 - Jack Johnston& John Dinardo, Econometric Methods, Four Edition, Univariate Time Series Moddling,http :// www.mh college.com .

Auto correlation and Partial Auto Correlation Function when a series is assumed to be generated from a stationary process, the mean is estimated by:

$$y_t \quad = T^{-1} \sum_{i=1}^{t} y_t \text{..........} \qquad (3\text{ -}51)$$

and the variance is:

$$C_0 = (T - 1)^{-1} \sum_{i=1}^{y-k} (y_{t-} \bar{y})^2 \text{.............} \quad (3 - 52)$$

The covariance between y_t and y_{t-k} is estimated by:

$$C_0 = (T_{-k-1})^{-1} \sum_{i=1}^{y-k} (y_{t-} \bar{y})(y_{t-k-} \bar{y}) \text{........} \; (3 - 53)$$

and the covariance measure the degree of association between data (K) periods part if the data are completely random then there is no association between y_t and y_{t-k} so that $C_k = 0$ (for K = 0). Typically the degree of association declines as (k) increase and $C_k = C_{-k}$, that is the correlation between (y_t) and (y_{t-k}) is the same as that between (y_{t-k}) and (y_t).

Covariance is affected by scale of the data so that it is easier to analyze correlation. The correlation between y_t and y_{t-k} is:

$$r_k = {CK}/{C0}, \qquad K = 0, \pm 1 \pm 2, \text{........}$$

Since r_k measures the correlation between a variable and itself lagged (K) periods, it is known as an (AUTO) correlation and the sequence r_1, r_2, …. is the autocorrelation function (ACF). It can easily be shown that these (ACF) terms could be derived from simple two variables regression models. In particular, a regression of y_t on y_{t-k} and a constant produces a slope estimate, (r_k) which is the (K), the term in the (ACF).

It is also useful to define a Partial Autocorrelation Function (PACF) which can be computed from the regression coefficient on y_{t-k} in a regression of y_t on y_{t-k} and all of the intervening lagged y_{t-1} terms and a constant.

The General Liner Process:

The AR (1) Process can be written as an infinite weighted of past shocks with the weights declining geometrically. If the restriction that the weights decline geometrically is relaxed, the general model results:

$$Y_t = \in_t \psi_1 \in_{t-1} + \psi_2 \in_{t-2} + \ldots\ldots\ldots\ldots\ldots \quad (3\text{-}54)$$

That is a time series is composed of a weighted average of many (infinite number of) random shocks.

The AR(1) mode is a special case of (3 - 53) with ψ_1 - $\emptyset$ and higher AR(p) parameters can be written as the infantile number of ψ_1 parameters being a function of only (p) parameters, ($\emptyset_0$). The MA (1) process is also a restriction with $\psi_i = 0$, i > 1 and MA(q) processes can be defined by $\psi_i = 0$; i > q.

The development of the ARMA model is facilitated by introducing the lag operator, (L). Recall that (L) is defined by $y_t = y_{t-1}$ and, hence. $y = L^i y_t$
Equate (3 - 54) written as :

$$y_1 = \in_t + \psi_1 L \in_t + \psi_2 L^2 \in_t + \ldots\ldots\ldots\ldots$$

$$y_1 = (1 + \psi_1 L + \psi_2 L^2 + \ldots\ldots\ldots\ldots) \in_t$$

$$\text{or } y_1 = \psi (L) \in_t \ldots\ldots\ldots\ldots\ldots \quad (3\text{-}55)$$

Where $\psi(L)$ is as polynomial in the lag operator = $(1+\psi_1 L+\psi_2 L^2+ \ldots\ldots)$. $(\psi$ L) can be factored the same manner that a simple quadratic (az^2bz+c) can be factored into $(Z - T_1)(Z - T_2)$.

The AR(1) process:

$$Y_t = \emptyset y_{t-1} + \epsilon_t$$

Can be written as:

$$Y_t - \emptyset L y_t = \epsilon_t = (1 - \emptyset L)\, y_t = \epsilon_t$$

Hence be inverting $(1 - \emptyset L)$, providing the inverse exist:

$$Y_t - = (1 - \emptyset L)^{-1} \epsilon_t$$

Which can be compared to (3 - 54)(3 –55) thus, convenient representation of the general linear model is:

$$y_t = [\theta(L) / \emptyset(L)] \epsilon_t$$

Or the ARMA (p, q) model:

$$\emptyset(L)\, y_t = \theta\,(L)\, \epsilon_t$$

Where $\emptyset(L)$ and $\theta(L)$ are polynomial of degree (p) and (q) respectively. That is:

$$y_t = \emptyset_1\, y_{t-1} + \emptyset_2 y_{t-2} + \ldots + \emptyset_p\, y_{t-p}$$

$$- \theta_1\, \epsilon_{t-1} - \theta_2\, \epsilon_{t-2} - \ldots - \theta_q\, \epsilon_{t-q} \ldots\ldots\ldots \quad (3\text{- }56)$$

The general linear process given in (3 –54) is the function of the ARMA of Box- Jenkins class of models. As it stands. (3- 54) cannot be estimated because it has an infinite number of parameters and the ARMA class (3–56), which involves simplification, providing a feasible frame work for data analysis either by truncation or by representing ψ (L) as a ratio of two finite polynomial (3 -56) if an ARMA model contains a constant term that the mean is non zero.

$$\emptyset (L)\, y_i = \theta_0 + \theta(L) \in_t$$

The mean of y_i is given by:

$$\emptyset^{-1}(1)\, \theta_0 = \frac{\theta 0}{1-\sum_{i=1}^{p} \theta i} = \bar{y}$$

Since the lag of a constant is the same constant $\{L = \emptyset_0 = \theta_0\}$

Identifying MA (q) Process:

Unlike classical regression situation where the estimation of the parameters for a given specification is of a major concern, the faces of Box-Jenkins modeling is in identifying or choosing the appropriate model. The identification is based on examining the (ACF) so that it is of some importance to establish the theoretical (ACF_s) for a wide variety of models. Fortunately it will be shown that there are substantial differences between the (ACF_s) of AR and MA process so that this task is not two difficult.

For the MA(1) process :

$$y_t = \in_t - \theta_1 \in_{t-1}$$

The ACF is such that

$$r_1 = \hat{\theta}_1 / (1+\hat{\theta}_1^2)$$

and all higher – order autocorrelation are zero. A simple extension to this reasoning demonstrates that is the ACF has q non – zero autocorrelation at lags K = 1, 2, …… and r_1 is insignificant for $K > q$ and MA (q) should be identified.

$$y_t = \theta_0 + \in_t - \theta_1 \in_{t-1} - \dots\dots.. - \theta_q \in_{t-q} \dots\dots\dots \quad (3\text{- }57)$$

There is an important set of restrictions as the invariability restrictions. Just as the AR process was represented as an infinite AR process, an MA process can be back-substituted to from an infinite AR representation. Unless $|\theta_1|<1$ distant shocks take on increasing importance .The invert ability condition requires that $|-\theta_1|<1$ for an MA(1) process (Ronal Bewley, time series force sting, UNSW, July., 200 Draft, Box-Jenkins models-Internet).

Inevitability has importance for the size of leading term in theoretical ACF, P_1since $|\theta_1|<1$ implies, from $P_1 = [\theta_1 / (1+\theta_1^2)]$, that $|P_1|< 0.5$. Thus, if the empirical $(ACF)_s$ seemingly the single spike at lag one but $r_1 < -0.5$, say, the process is not likely to be an invertible MA(1) process. Of course, when the model is properly estimated, values of r_1 close to the invertible bounds may produce roots of $\theta(L)$ outside, but close to unit circle. A generalization of this rule is that $|\sum r_i<-0.5|$ for the invert ability of an MA (q) process.

Identifying AR (P) Process:

Because an AR (P) process can be written as infinite MA process, the (ACF) terms will not "cut-off" towards zero for an AR (P) process. For example, recall the AR (1) model.

$$y_t = \emptyset_1 y_{t-1} + \epsilon_t$$

since $\lambda_k = E(y_t\, y_{t-k})$ both sides of the first equation can be multiplied by y_{t-k} to give :

$$y_t\ y_{t-k} = \emptyset_1 y_{t-1}\ y_{t-k} + y_{t-k}\,\epsilon_t$$

Which, on taking expectations gives :

$$\lambda_k = \emptyset_1 \lambda_{k\text{-}1} + E\ (y_{t\text{-}k} \in_t) = \emptyset_1 \lambda_{k\text{-}1}$$

Since $\in_t$ occurs after $y_{t\text{-}k}$ and "unpredictable" on converting to correlations.

$$P_k = \emptyset_1 P_{k\text{-}1} \ldots\ldots\ldots\ldots\ (3\ \text{-}58)$$

Given that $P_0 = 1$, the theoretical (ACF) of AR(1) process declines geometrically. Higher order of AR process is characterized by the solutions to higher – order difference equation and the possible patterns are many and varied. Typically the decline towards zero quickly but can oscillate about zero. Therefore, a danger that an AR and MA process might be confused when the (ACF) oscillate. Thus emphasizes to look for patterns as well as significance.

The task becomes simplified if the partial auto correlation function is utilized. If the (PACF) terms are computed from regressions, it is clear that the K^{th} term $\emptyset_{kk}$ is used in place of $\emptyset_k$ because only the last coefficient in the sequence is considered and this distinguishes it from the direct regression coefficient is zero for all lags greater than the true length, P. Thus, the (PACF) cuts off at lag P for an AR (P) process.

3 – 11 -2 Estimation and Diagnoses of ARIMA Models:

In the estimation and diagnostic checking stage, we use the estimation statement to specify the ARIMA model to fit to the variable specified in the previous Identified statement, and to estimate the parameters of that model. The Estimate statement also produces diagnostic statistics to help you judge the adequacy of the model.

Significance tests for parameter estimates indicate whether some terms in the model may be unnecessary. Goodness-of-fit statistics aid in comparing this model to others. Tests for white noise residuals indicate whether the residual series contains additional information that might be utilized by a more complex model. If the diagnostic tests indicate problems with the model, you try another model, and then repeat the estimation and diagnostic checking stage.

Returning to the AR (1) specification in equation:

$$z_{t=} \mu - \emptyset\, y_{t-1} + \epsilon_t$$

Where $\in$ is white noise OLS is an obvious estimator the only qualification is that the value of y_t is taken as given and summations run over t = 2,3,……..n .The usual test statistics now only have an asymptotic justification ,because of the lagged regresses. OLS may also be seen to be a conditional ML estimator .If we take y_t as given, the conditional likelihood for the remaining (n-1) observations is :

$$L^* = \text{p} \left(y_{2\,,}\, y_{3\,,}\, y_{4\,,} \ldots\ldots\, y_n \,\middle|\, y_1 \right)$$

$$= \text{P} \left(y_2 \middle| y_1 \right) \text{P}\left(y_3 \middle| y_{2\,,}\right) \ldots\ldots \text{P}\left(y_n \middle| y_{n-1} \right)$$

If we assume Gaussian white noise,

$$\text{P}\left(y_t \middle| y_{t-1} \right) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[- \frac{1}{2\sigma^2} \left(y_t - \mu - \emptyset\; y_{t-1} \right)^2 \right]$$

$$\mathcal{L}^* = ln\ L^* =$$

$$\text{constant} - \frac{n-1}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{t=2}^{n} \left(y_t - \mu - \emptyset\; y_{t-1} \right)^2 \ldots\ldots\ldots \text{(3-59)}$$

Maximizing with respect to μ and ∅ gives the OLS estimates just described .To obtain full ML estimates one must maximize the unconditional likelihood :

$$L = p\,(y_1)\, L^*$$

Under the assumption of the AR (1) process:

$$y_1 \backsim N\left(\frac{\mu}{1-\emptyset}, \frac{\sigma^2}{1-\emptyset^2}\right)$$

Thus, ln P (y_1) =

$$\text{constant} + \frac{1}{2} ln\,(1-\emptyset^2) - \frac{1}{2}\, ln\,\sigma^2 - \frac{(1-\emptyset^2)}{2\sigma^2}\left(y_1 - \frac{\mu}{1-\emptyset}\right)^2$$

And so the unconditional log- likelihood is:

$$\mathcal{L} = ln\ \text{p}\,(y_1) + \mathcal{L}^*$$

$$= \text{constant} - \frac{n}{2} \ln \sigma^2 + \frac{1}{2}\, ln\,(1-\emptyset^2) - \frac{(1-\emptyset^2)}{2\sigma^2}\left(y_1 - \frac{\mu}{1-\emptyset}\right)^2 \quad \text{.... (3-60)}$$

$$- \frac{1}{2\sigma^2} \sum_{t=2}^{n} (y_t - \mu - \emptyset\, y_{t-1})^2$$

Taking first derivatives with respect to μ and ∅ no longer yields linear equations in these parameters .Thus iterative techniques are required to maximize equation(3- 60).In small samples the difference between maximizing equation (3-59) or (3-60) maybe important, but this difference diminishes with sample size.

Higher-order AR schemes are fitted in a similar fashion .Least Squares and conditional maximum likelihood take the first p observations as given. Full maximum likelihood procedures are also available.

The fitting of MA schemes is more complicated. Even the lowest-order processes involve nonlinear methods. For example the MA (1) scheme is:

$$y_t = \mu + \epsilon_t - \theta_1 \epsilon_{t-1}$$

If $\in_0$ is set at zero, then : $\in_1 = y_t - \mu$

and $\in_2 = y_2 - \mu - \in_t + \theta_1 \in_1$.

By proceeding in this fashion all n values of $\in_t$ can be expressed in terms of μ and θ. However, $\sum_{t=1}^{n} \in^2$ is a complicated nonlinear function of the parameters, and iterative techniques are required to obtain even conditional estimates, Full ML procedures are again available and are fully described in Hamilton's outstanding treatise. Full ARMA schemes share all estimation problems of the pure MA processes.

3 – 11 -3 Forecasting of ARIMA Models:

The main purpose of fitting ARMA schemes is to project the series forward beyond the sample period.Such projections are sometimes used as a benchmark to compare with forecasts yield by more complicated multivariate models.In projections or forcasts there are two inevitable of error , namely:

1- Error due to ignorance of future innovations .

2-Error due to differences between true and estimated parameter values.

In this section we will deal only with the first source of the error , illustrating principle involved with a few low – older processes.

Consider the frist AR(1) scheme:

$$y_t - \mu = \emptyset_1(y_{t-1} - \mu) + \in_t \quad |\emptyset| < 1 \quad \in_t \sim (0, \sigma^2)$$

which for some purposes is more conveniently written as:

$$y_t = (1 - \emptyset)\mu + \emptyset y_{t-1} + \in_t \quad \ldots\ldots\ldots\ldots \quad (3\text{- }61)$$

In all that follows we will assume that observations on y are available for 1 to n and that all forcasts are made conditional on information available at time n. Thus :

y_{n+s} = (unknown) value of y in future period (n+s).

$\hat{y}_{n+1}$= forcast of Y_{n+s} made on the basis of information available at time n

e_{n+s} = y_{n+s} -$\hat{y}_{n+1}$ = forecast error.

The Mean Square Error (MSE) of a forecast is simply the average or expected square forecast error .This treats positive and negative forecast errors symmetrically and is a widely used criterion for the choice of a forecasting rule. We wish to find a forecasting rule that will minimize MSE.It can be shown that the minimum MSE forecast of y_{n+s} is the conditional expectation of y_{n+s} given information available at time n. As an illustration, consider forecasting y_{n+s} for AR (1) process. The true value is:

$$y_{n+1} = (1 - \emptyset)\,\mu + \emptyset\, y_n + \epsilon_{n+1} \text{ .... (3-62)}$$

The minimum MSE forecast is then:

$$\hat{y}_{n+1} = E\,(y_{n+1} | y_n) = (1 - \emptyset)\,\mu + \emptyset\, y_n \text{ .....(3-63)}$$

The only unknown on right side of equation (3-62) is the innovation in period (n+1) and it is replaced by its zero expectation. The forecast in (3-63) may be rearranged as:

$$(\hat{y}_{n+1} - \mu) = \emptyset(y_n - \mu) \text{......... (3- 64)}$$

That is y_{n+1} is forecast to differ from the mean by a fraction θ of the deviation in period n. The forecast error is $e_{n+1} = y_{n+1} - \hat{y}_{n+1} = \epsilon_{n+1}$ and so var (e_{n+1}) = σ^2.

Turning to period (n+2), we need to express y_{n+2} in terms of y_n and innovations since that time.Substitution in equation (3- 58) gives :

$$y_{n+2} = (1 - \emptyset)\mu + \emptyset y_{n+1} + \epsilon_{n+2}$$

$$= (1 - \emptyset)\mu + \emptyset[(1 - \emptyset)\mu + \emptyset y_n + \epsilon_{n+1}] + \epsilon_{n+2}$$

$$= (1 - \emptyset^2)\mu + \emptyset^2 y_n + \emptyset \epsilon_{n+1} + \epsilon_{n+2}$$

The minimum MSE forecast is:

$$\hat{y}_{n+2} = (1 - \emptyset^2)\mu + \emptyset^2 y_n$$

This may be displayed as:

$$(\hat{y}_{n+2} - \mu) = \emptyset^2(y_n - \mu) = \emptyset(\hat{y}_{n+1} - \mu) \ldots\ldots \text{(3- 65)}$$

The forecast from the AR (1) model approach the unconditional μ exponentially as forecast horizon increase. The forecast error variance is:

$$\text{var}(e_{n+2}) = \sigma^2(1 + \emptyset^2).$$

Proceeding in this way we find:

$$y_{n+s} = (1 - \emptyset^2)\mu + \emptyset^s y_n + (\epsilon_{n+s} + \emptyset \epsilon_{n+s-1} + \ldots\ldots + \emptyset^{s-1}\epsilon_{n+1})$$

The forecast is:

$$(\hat{y}_{n+s} - \mu) = \emptyset^2(y_n - \mu) \ldots\ldots\ldots \text{(3-66)}$$

And the forecast error variance is :

$$\text{var}(e_{n+s}) = (1 + \emptyset^2 + \emptyset^4 + \cdots + \emptyset^{2(s-1)})\sigma^2 \ldots\ldots (3-67)$$

Clearly

$$\hat{y}_{n+s} \longrightarrow \mu \quad \text{as} \quad s \longrightarrow \infty$$

and var (e_{n+s}) $\longrightarrow \frac{\sigma^2}{1-\theta^2} = \sigma_y^2$ as $s \longrightarrow \infty$

Thus as forecast horizon increases, the forecast value tends to the unconditional mean of the process and the forecast error variance increases toward the unconditional variance of the process.

MA(1) Process:

The MA(1) process is:

$$y_t = \mu + \epsilon_t - \theta_1 \epsilon_{t-1} \ldots\ldots\ldots \quad (3\text{-}68)$$

Again forecasting from period n, we find:

$$\hat{y}_{n+1} = \mu - \theta \epsilon_n \quad \ldots\ldots\ldots\ldots \quad (3\text{-}69)$$

because ϵ_{n+1} is unknown at period n. Implementation of equation (3-69) however, does require knowledge of ϵ_n. From equation (3-68) it is clear that this in turn requires knowledge of previous values of ϵ. Thus, the strict implementation of equation (3-69) requires knowledge of ϵ_0. In practice this is often set at the expected value of zero in order to start the process off. This approximation is obviously of lesser importance as the sample size increases. Clearly var (e_{n+s}) = σ^2.Looking two periods a head, we see that:

$$y_{n+2} = \mu + \epsilon_{n+2} - \theta \epsilon_{n+1}$$

and the minimum MSE is:

$$\hat{y}_{n+2} = \mu$$

Thus for the MA (1) model:

$$\hat{y}_{n+s} = \mu \qquad s \geq 2 \ldots\ldots\ldots. \quad (3\text{- }70)$$

and

$$\operatorname{var}(e_{n+s}) = (1 + \theta^2)\ \sigma^2 = \sigma_y^2 \qquad s \geq 2 \ldots\ldots.(3-71)$$

From two periods out the forecast from the MA (1) model is simply unconditional mean of the series and the forecast error variance is the variance of series.

ARMA (1,1) Process :

As we know combine AR(1) and AM(1) processes to give the ARIMA(1,1)model:

$$(y_t - \mu) = \emptyset(y_{t-1} - \mu) + \epsilon_t - \theta\ \epsilon_{t-1} \ldots\ldots\ldots\ldots (3\text{-}72)$$

The minimum MSE forecast for period (n+1) is then

$$\hat{y}_{n+1} - \mu = \emptyset(y_n - \mu) - \theta\ \epsilon_n$$

This result differs from AR(1) forecast only by the term in $\theta\ \epsilon_n$.The forecast error variance is var $(e_{n+1}) = \sigma^2$.Repeated use of equation (3-72)gives:

$$(y_{n+2} - \mu) = \emptyset^2(y_n - \mu) + \epsilon_{n+2} + (\emptyset - \theta)\ \epsilon_{n+1} - \emptyset\ \theta\ \epsilon_n$$

The forecast for period (n+2)is then:

$$(\hat{y}_{n+2} - \mu) = \emptyset^2(y_n - \mu) - \emptyset\ \theta\ \epsilon_n = \emptyset(\hat{y}_{n+1} - \mu) \ldots\ldots \quad (3\text{ - }73)$$

Thus as in the AR(1) case successive forecast deviate from the mean in declining exponential fashion. The forecast error variance is var$(e_{n+2}) = \sigma^2$ $[1+(\emptyset - \theta)^2]$.By continuing this way it may be shown that :

$$(\hat{y}_{n+s} - \mu) = \emptyset^{s}(y_n - \mu) - \emptyset^{s-1}\,\theta\,\epsilon_n \ldots\ldots\ldots (3\text{-}74)$$

The forecast thus tends to the unconditional mean as the forecast horizon increases. Likewise it may be shown that:

$$\text{var}\,(e_{n+s}) \longrightarrow \sigma^2\left(\frac{1-2\emptyset\,\theta+\theta^2}{1-\emptyset^2}\right) \text{ as } s \longrightarrow \infty \ldots\ldots\ldots\ldots (3\text{-}75)$$

As we know this limiting variance is the variance of the y series.

ARIMA (1.1.0) Process:

As a final illustration consider a series z whose first differences follow an AR(1) model:

$$z_t - z_{t-1} = y_t$$

$$(y_t - \mu) = \emptyset(y_{t-1} - \mu) + \epsilon_t \ldots\ldots\ldots\ldots (3\text{-}76)$$

From equation (3- 76) we can write:

$$z_{n+s} = z_n + y_{n+1} + \ldots\ldots + y_{n+s}$$

$$= (z_n + \mu) + (y_{n+1} - \mu) + \ldots\ldots\ldots + (y_{n+s} - \mu)$$

Continuous substitution for the $(y_n - \mu)$ terms gives:

$$z_{n+s} = z_n + s\mu + \frac{\emptyset(1-\emptyset^{s})}{1-\emptyset}(y_n - \mu) + e_{n+s} \ldots\ldots\ldots (3\text{-}77)$$

Where:

$$e_{n+s} = \epsilon_{n+s} + (1+\emptyset)\epsilon_{n+s-1} + (1+\emptyset+\emptyset^2)\,\epsilon_{n+s-2} + \ldots$$

$$+ (1+\emptyset+\emptyset^2+\cdots+\emptyset^{s-1})\,\epsilon_{n+1} \ldots\ldots\ldots\ldots\ldots\ldots (3\text{-}78)$$

The forecasts are given by the three terms on the right side equation (3-77) two of which increase with the forecast horizon. Notice, however that the term in the initial value z_n does not fade away. From equation (3-75) the forecast error variance is:

$$\text{var}(e_{n+s}) = \sigma^2 [1+(1+\emptyset)^2+(1+\emptyset+\emptyset^2)^2$$

$$+\ldots\ldots+(1+\emptyset+\cdots+\emptyset^{s-1})^2 \ldots\ldots\ldots\ldots\ (3\text{-}79)$$

This variance increases monotonically with s. The forecast of a non stationary series become ever more imprecise as the forecast horizon increases.

All the formulas in this section are based on the assumption that the parameters of the process are known precisely. In practice they are replaced by sample estimates. The point forecasts will still be MSE asymptotically but the estimated forecast error variance will understate the true values because the formulae do not allow for coefficient error. [1]

Chapter Four

The Application of Demographic and Time series Models

4 -1 Preface:

[1] - Jack Johnston& John Dinardo, Econometric Methods, Four Edition, Univariate Time Series Moddling,http :// www.mh college.com.

This chapter consists of seven sections: section one is about the preface, section two is about calculate life expectation from life tables of Sudan Population Censuses (1973 -1983 – 1993 – 2008). Section three is over view about levels, trends and patterns of mortality in Sudan, section four is about comparison of population censuses using life tables, section five is about forecasting of life expectation at birth using demographic equations, section sex is about ARIMA model. Finally, a comparison of statistical and demographic models to forecast the life expectation at birth (e_0).

4-2 Calculation of Life Expectation:

Demographic information is scanty in Sudan and these related to mortality are even scantier. As a matter of fact, very little direct information on mortality is available in the country and whatever is paraded is based on indirect evidence. Even these are questionable, but since no alternative information were available, one had to make to with existing situation.

So as we know, deaths registration is incomplete and not reliable the only way to calculate mortality indicators is indirect methods using demographic and statistical techniques which mentioned in second chapter .We make use of the available data produced by the census and applied the relevant techniques.

The life expectation at birth (e_0) and probability of dying before age one (q_1) are calculated by Brass relational two – parameter logit system of estimation of child and adult mortality from data on children and mothers survivorship status, and the linkage of the two sets of information to produce abridged life table. The same methods were applied earlier in calculating mortality indicators from the censuses data which gives a reliable base for comparisons. The (q_1) obtained from Brass method is not considered because it is subject to errors resulting from the assumption that child mortality is a function only of age of child whereas children to mothers in age group (15 - 19) are at higher risk of death than those born to mothers at older ages. Instead the (q_1) is estimated from Coale and Demeny model (the West region) using the average level of child mortality levels obtained from Brass relational two – parameter logit system. Also we use linear regression method and the African model from regional model life tables was used in calculating.

In this section we will analysis and discuss life tables for population censuses (1973 – 1983 – 1993 - 2008).

4-2-1 Life Tables 1973:

As we know, the census of 1973 provided information on mortality through information on children and parent survival, however it should be clearly stated that raw data collected from the 1973 census are subject to several types of errors, defects, deficiencies, bias etc. Even though poor quality of data my affect some of the estimations, but fitting using in our study to improve the resolute of the data.

Table (4-1) (4-2) (4-3) show life tables for both sexes, males and females these tables indicate that, as life expectation at birth for both sexes **(e_0)**was 47.3 also the life expectation at birth **(e_0)** for males came out as 45. 5 and females life expectation at birth **(e_0)** about 49.1 so the males – females gap may be only around 3.6 years. The life expectation at birth **(e_0)** estimated may have been bloated up by increased value of ℓ_2 due to:

(1) Omission rate of dead children being higher than for living children.

(2) Higher child mortality among the not reported women.

(4-1) Life Table for Both Sexes -Census., 1973-Sudan

Agegroup	nM_x	nq_x	ℓ_x	$_nd_x$	nL_x	T_x	e_x
0	0.14496521	0.13161	100000	13161	90787.3	4731832.9	**47.3**
1 - 4	0.021400398	0.080926	86839	7027.5	328383.1	4641045.6	53.4
5 - 9	0.004004024	0.019822	79812	1582	395105	4312662.5	54.0
10 - 14	0.003030606	0.015039	78230	1176.5	388207.5	3917557.5	50.1
15 - 19	0.005107681	0.025216	77053	1943	380407.5	3529350	45.8
20 - 24	0.006985841	0.03433	75110	2578.5	369105	3148942.5	41.9
25 - 29	0.00720153	0.035371	72532	2565.5	356245	2779837.5	38.3
30 - 34	0.007555083	0.037075	69966	2594	343345	2423592.5	34.6
35 - 39	0.008522508	0.041724	67372	2811	329832.5	2080247.5	30.9
40 - 44	0.01014548	0.049473	64561	3194	314820	1750415	27.1
45 - 49	0.012853094	0.062265	61367	3821	297282.5	1435595	23.4
50 - 54	0.017113346	0.082056	57546	4722	275925	1138312.5	19.8
55 - 59	0.023447364	0.110745	52824	5850	249495	862387.5	16.3
60 - 64	0.034238305	0.157694	46974	7407.5	216352.5	612892.5	13.1
65 - 69	0.04905481	0.21848	39567	8644.5	176222.5	396540	10.0
70 - 74	0.076763705	0.32202	30922	9957.5	129717.5	220317.5	7.1
75 +	0.231402122	1.00000	20965	20964.5	90600	90600	4.3

Sources: The researcher from applied study(Sudan Census Data .,1973)

Table (4-1) above illustrates the life table for both sexes-Census (1973) - Sudan. From this table we observe that, the highest life expectation in the third age group (e_9), and the lowest probability of death and central death rate in the fourth age group ($10q_{14}$) ($10M_{14}$).

(4-2) Life Table for Males-Census., 1973-Sudan

Age group	nM_x	nq_x	ℓ_x	$_nd_x$	nL_x	T_x	e_x

0	0.15681	0.141300	100000	14130	90109	4550517	**45.5**
1 - 4	0.022589	0.085164	85870	7313	323734.9	4460408	51.9
05 - 09	0.004391	0.021717	78557	1706	388520	4136673	52.7
10 - 14	0.003327	0.016499	76851	1268	381085	3748153	48.8
15 - 19	0.0056	0.027612	75583	2087	372697.5	3367068	44.5
20 - 24	0.0077	0.037771	73496	2776	360540	2994370	40.7
25 - 29	0.007931	0.038886	70720	2750	346725	2633830	37.2
30 - 34	0.008326	0.040783	67970	2772	332920	2287105	33.6
35 - 39	0.009394	0.045891	65198	2992	318510	1954185	29.9
40 - 44	0.011191	0.054432	62206	3386	302565	1635675	26.3
45 - 49	0.014171	0.068429	58820	4025	284037.5	1333110	22.7
50 - 54	0.018746	0.089534	54795	4906	261710	1049073	19.1
55 - 59	0.025883	0.12155	49889	6064	234285	787362.9	15.8
60 - 64	0.037517	0.171500	43825	7516	200335	553077.9	12.6
65 - 69	0.053444	0.235727	36309	8559	160147.5	352742.9	9.7
70 - 74	0.083	0.343000	27750	9524	114940	192595.4	6.9
75 +	0.2347	1.000000	18226	18226	77655.4	77655.4	4.3

Sources: The researcher from applied study (Sudan Census Data .,1973)

Table (4-2) illustrates the life table for males-census (1973) -Sudan. From this table, we notice that, the highest value for (e_x) is the third age group (e_9) because they have reached the risk stage due to childhood illnesses. The lowest probability of death and the central death rate in the fourth age group ($10q_{14}$) ($10M_{14}$).

(4-3) Life Table for Females-Census., 1973-Sudan

Age group	nM_x	nq_x	ℓ_x	$_nd_x$	nL_x	T_x	e_x
0	0.133296	0.1219200	100000	12192	91465.6	4913277.7	**49.1**
1 - 4	0.020245	0.0767812	87808	6742	333028.6	4821812.1	54.9
05 - 09	0.00363	0.0179853	81066	1458	401685	4488783.5	55.4
10 - 14	0.002745	0.0136293	79608	1085	395327.5	4087098.5	51.3
15 - 19	0.004635	0.0229105	78523	1799	388117.5	3691771	47.0
20 - 24	0.006304	0.0310333	76724	2381	377667.5	3303653.5	43.1
25 - 29	0.00651	0.0320272	74343	2381	365762.5	2925986	39.4
30 - 34	0.006829	0.0335733	71962	2416	353770	2560223.5	35.6
35 - 39	0.007709	0.0378167	69546	2630	341155	2206453.5	31.7
40 - 44	0.009178	0.0448622	66916	3002	327075	1865298.5	27.9
45 - 49	0.011648	0.0565917	63914	3617	310527.5	1538223.5	24.1
50 - 54	0.015641	0.0752608	60297	4538	290140	1227696	20.4
55 - 59	0.021292	0.1010779	55759	5636	264705	937556	16.8
60 - 64	0.031411	0.1456218	50123	7299	232367.5	672851	13.4
65 - 69	0.045399	0.2038576	42824	8730	192295	440483.5	10.3
70 - 74	0.071914	0.305000	34094	10391	144492.5	248188.5	7.3
75 +	0.228582	1.000000	23703	23703	103696	103696	4.4

Sources:The researcher from applied study(Sudan Census Data .,1973)

Table(4-3) above explains the life table for females-Census (1973) -Sudan. From this table, we notice that, the highest value for me (e_x) is 55.4 in the third age group (e_9). The lowest probability of death and the central death rate in the fourth age group ($10q_{14}$) ($10M_{14}$).

4-2-2 Life Tables 1983:

It can also be constructed by using indirect data collected from 1983 census. Estimates of probabilities of dying of children before reaching age **(x)** for the whole population using Brass technique methods were obtained. From these probabilities the life tables have been generated. The African model from Regional Model Life tables was used in calculating these life tables.

Tables (4-4) through table (4- 6) show both sexes, males and females life tables .These tables indicate that, life expectation at birth (e_0) for both sexes was 53.5 years. Life expectation at birth (e_0) for males was 52.7 while (e_0) for females was 54.4 it is indicated the highest level of (e_0) for females compared to males. These findings are consistent with the pattern established worldwide. One of explanation for this pattern is that males are subject to risk of death more than females despite of increase maternal mortality among age group (15 – 19) up to (44 – 49) years for females. However the gap between e_0 for males and females is seemed to have narrowed since 1973, which this gap was 1.7 years in 1983 while it was 3.6 years in1973 as we mentioned above .This may be due to vagaries in the census data as well as in the application of techniques of estimation .Also increased medical care for males may not be ruled out.

These life tables also show that, in general, the mortality rates (nq_x) values for males were lower than those of females at younger ages, while they were higher at older ages. An exception is the infant mortality rates, which was higher for males than females. This suggests that the difference in (e_0) between males and females resulted mainly from mortality levels at older ages, which were higher for males than females.

(4-4)Life Table for Both Sexes -Census. 1983-Sudan

Age group	nM_x	nq_x	ℓ_x	$_nd_x$	nL_x	T_x	e_x
0	0.112398335	0.1042	100000	10417	92708	5353862	**53.5**
1 – 4	0.013499753	0.0521	89583	4669	348995	5261154	58.7
05 - 09	0.003274578	0.0182	84914	1379	421123	4912159	57.8
10 - 14	0.002471628	0.0123	83535	1026	415111	4491036	53.8
15 - 19	0.003627354	0.018	82509	1483	408838	4075925	49.4
20 - 24	0.004945238	0.0244	81026	1979	400183	3667088	45.3
25 - 29	0.004399708	0.0218	79047	1720	390935	3266905	41.3
30 - 34	0.004501347	0.0223	77327	1721	382330	2875971	37.2
35 - 39	0.005060178	0.025	75606	1889	373307	2493640	33.0
40 - 44	0.007712687	0.378	73717	2789	361612	2120333	28.8
45 - 49	0.009872926	0.0482	70928	3417	346098	1758721	24.8
50 - 54	0.035334794	0.0662	67511	4466	126391	1412623	20.9
55 - 59	0.019088169	0.0911	63045	5743	300867	1086232	17.2
60 - 64	0.028454251	0.1328	57302	7611	267482	785365	13.7
65 - 69	0.04228248	0.1912	49691	9501	224703	517883	10.4
70 - 74	0.065323986	0.2808	40190	11284	172739	293179	7.3
75 +	0.240003321	1.0000	28906	28906	120440	120440	4.2

Sources: The researcher from applied study (Sudan Census Data .,1983)

Table (4 -4) above illustrates the life table for Both Sexes-Census (1983) - Sudan. From this table we note that, there is no clear difference between e_4 and e_9. They have the highest value. The lowest probability of death and central death rate were in the fourth age group $(10q_{14})(10M_{14})$.

(4-5) Life Table for Males -Census., 1983-Sudan

Age group	nM_x	nq_x	ℓ_x	$_nd_x$	nL_x	T_x	e_x
0	0.118473	0.1094	100000	10941	92341	5268703	**52.7**
1 – 4	0.012883	0.0498	89059	4433	347370	5176362	58.1
05 - 09	0.003203	0.0156	84626	1316	410841	4828992	57.1
10 - 14	0.002306	0.0115	83310	955	414163	4409152	52.9
15 - 19	0.003498	0.0173	82355	1428	408207	3994988	48.5
20 - 24	0.004979	0.0246	80927	1990	399661	3586781	44.3
25 - 29	0.004266	0.0211	78937	1666	390520	3187120	40.4
30 - 34	0.004359	0.0216	77271	1666	382189	2796601	36.2
35 - 39	0.004456	0.022	75605	1666	373859	2414411	31.9
40 - 44	0.008779	0.0429	73939	3176	361755	2040553	27.6
45 - 49	0.011523	0.056	70763	3963	343908	1678797	23.7
50 - 54	0.01606	0.0772	66800	5157	321107	1334889	20.0
55 - 59	0.022527	0.1066	61643	6573	291782	1013782	16.4
60 - 64	0.03314	0.153	55070	8427	254283	722000	13.1
65 - 69	0.048555	0.2165	46643	10098	207971	467717	10.0
70 - 74	0.073059	0.3089	36545	11288	154506	259746	7.1
75 +	0.23999	1.0000	25257	25257	105240	105240	4.2

Sources: The researcher from applied study (Sudan Census Data .,1983)

Table (5-4) above shows life table for Males-Census (1983) -Sudan. From this table, we notice that, the highest value for (e_x) is the second age group (e_4). The lowest probability of death and the central death rate in the fourth age group ($10q_{14}$) ($10M_{14}$).

(4-6) Life Table for Females -Census., 1983-Sudan

Age group	$_nM_x$	$_nq_x$	ℓ_x	$_nd_x$	$_nL_x$	T_x	e_x
0	0.106256111	0.0989	100000	9892	93075	5439022	**54.4**
1 – 4	0.014118429	0.0544	90108	4906	350620	5345946	59.3
05 - 09	0.003413785	0.0169	85202	1442	422405	4995326	58.6
10 - 14	0.002636652	0.0131	83760	1097	416058	4572921	54.6
15 - 19	0.003758535	0.0186	82663	1539	409468	4156863	50.3
20 - 24	0.004906365	0.0242	81124	1966	400704	3747394	46.2
25 - 29	0.004538149	0.0224	79158	1776	391349	3346690	42.3
30 - 34	0.004640863	0.0229	77382	1775	382472	2955341	38.2
35 - 39	0.005663222	0.0279	75607	2111	372756	2572869	34.0
40 - 44	0.006650639	0.0327	73496	2404	361469	2200113	29.9
45 - 49	0.008237459	0.0404	71092	2869	348287	1838645	25.9
50 - 54	0.011384639	0.0553	68223	3776	331675	1490358	21.8
55 - 59	0.015850841	0.0762	64447	4913	309952	1158683	18.0
60 - 64	0.024208891	0.1141	59534	6795	280682	848730	14.3
65 - 69	0.036879339	0.1688	52739	8904	241436	568048	10.8
70 - 74	0.059071487	0.2574	43835	11281	190972	326613	7.5
75 +	0.24000118	1.0000	32554	32554	135641	135641	4.2

Sources : The researcher from applied study(Sudan Census Data .,1983)

Table (6 -4) above illustrates the life table for females-Census (1983) -Sudan. From this table, we notice that, the highest value for (e_x) is the second age group (e_4). The lowest probability of death and the central death rate in the fourth age group ($10q_{14)}$ ($10M_{14}$).

4 – 2 - 3 Life Tables 1993:

Tables (4 -7), (4 - 8) and (4 - 9) show that, life expectation at birth for Sudan census. , 1993 by sex and both sexes. It is shown that, the average number of years expected to be lived was around 55.3 years for both sexes .This figure contrasts with the life at birth of more than 70 years in some of developed countries during this period. Also we found (e_0) for the males about 55.34 years compared 55.31 years for the females. So the gap between males and females with respect to general health situation as indicated by e_0 reaches only 0.03 years .Note that, although the difference between males and females increased in census1983 but decreased very small in 1993.

(4-7)Life Table for Both Sexes-Census., 1993-Sudan

Age group	$_nM_x$	nq_x	ℓ_x	$_nd_x$	nL_x	T_x	e_x
0	0.105269989	0.098045141	100000	15049	86456.3248	5538664	**55.4**
1 - 4	0.028750414	0.106717584	84951	3290	330923.773	5452208	64.2
05 - 09	0.006362183	0.03131287	81662	1526	404493.857	5121284	62.7
10 - 14	0.003116667	0.015462854	80136	867	398511.161	4716790	58.9
15 - 19	0.004051506	0.020054404	79269	628	394773.463	4318279	54.5
20 - 24	0.005056043	0.024964661	78641	2182	387748.662	3923505	47.8
25 - 29	0.005298838	0.026147805	76459	827	380226.602	3535757	46.2
30 - 34	0.006527887	0.032115321	75632	1344	374800	3155530	41.7
35 - 39	0.006495036	0.031956288	74288	1740	367089.933	2780730	37.4
40 - 44	0.0096173	0.046957491	72548	1704	358478.523	2413640	33.3
45 - 49	0.011482139	0.055808686	70844	1697	349976.454	2055162	29.0
50 - 54	0.018318792	0.087582927	69147	1816	341195.286	1705185	24.7
55 - 59	0.017204649	0.082475827	67331	2055	331518.2	1363990	20.3
60 - 64	0.037299349	0.170589547	65276	2449	320257.747	1032472	15.8
65 - 69	0.040191234	0.182608061	62827	929	311811.126	712214	11.3
70 - 74	0.069088542	0.294565036	61898	5946	294622.879	400402.9	6.5
75 +	0.12690896	0.481711153	55952	55952	105780	105780	1.9

Sources :The researcher from applied study(Sudan Census Data .,1993)

Table (7-4) above explains the life table for both sexes-Census (1993) -Sudan. From this table, we found that, the 1993 census recorded the lowest life expectation for open group (e_{75+}). Compared to other censuses. The lowest probability of death and central death rate in the fourth age group ($10q_{14}$) ($10M_{14}$).

(4-8)Life Table for Males-Census., 1993-Sudan

Age group	nM_x	nq_x	ℓ_x	$_nd_x$	nL_x	T_x	e_x
0	0.117538	0.108602168	100000	15048.53	86456.32484	5534869.98	**55.3**
1 - 4	0.030467	0.112604184	84951.47204	3289.672	330923.7732	5448413.656	64.1
05 - 09	0.006886	0.033848001	81661.79981	1526.057	404493.8571	5117489.882	62.7
10 - 14	0.003547	0.017578471	80135.74303	867.0219	398511.1605	4712996.025	58.8
15 - 19	0.004413	0.021822471	79268.72117	628.057	394773.4633	4314484.865	54.4
20 - 24	0.005562	0.027427159	78640.66414	2181.863	387748.6623	3919711.402	46.8
25 - 29	0.006008	0.029593733	76458.80079	826.9607	380226.6021	3531962.739	46.2
30 - 34	0.006576	0.032347254	75631.84007	1343.68	374800.0002	3151736.137	41.7
35 - 39	0.007163	0.035185038	74288.16002	1740.347	367089.9327	2776936.137	37.4
40 - 44	0.011388	0.055365864	72547.81306	1704.217	358478.5229	2409846.204	33.2
45 - 49	0.014278	0.068929805	70843.59612	1696.611	349976.4539	2051367.681	29.0
50 - 54	0.00209	0.010394941	69146.98543	1815.857	341195.2859	1701391.227	24.6
55 - 59	0.0186	0.088866081	67331.12893	2054.978	331518.1995	1360195.941	20.2
60 - 64	0.039406	0.179360494	65276.15087	2449.203	320257.7472	1028677.742	15.8
65 - 69	0.041386	0.187527374	62826.94801	929.4455	311811.1263	708419.9947	11.3
70 - 74	0.075531	0.317670783	61897.50249	5945.853	294622.8793	396608.8685	6.4
75 +	0.130961	1.0000000	55951.64924	23347.2	101985.9891	101985.9891	1.8

Sources : The researcher from applied study(Sudan Census Data .,1993)

From table (8-4) the life table for males-Census (1993) -Sudan. We note that, the highest life expectation in the second age group (e_4) the lowest probability of death and the central death rate in the ($10q_{14}$) age group ($10M_{14}$).

(4-9)Life Table for Females -Census., 1993-Sudan

Age group	$_nM_x$	nq_x	ℓ_x	$_nd_x$	nL_x	T_x	e_x
0	0.091903	0.086347701	100000	15048.53	86456.32484	5531428.252	**55.3**
1 - 4	0.027036	0.100787233	84951.472	3289.672	330923.7732	5444971.928	64.1
05 - 09	0.005813	0.028650816	81661.7998	1526.057	404493.8571	5114048.154	62.6
10 - 14	0.002642	0.013124552	80135.743	867.0219	398511.1605	4709554.297	58.8
15 - 19	0.00368	0.018231362	79268.7212	628.057	394773.4633	4311043.137	54.4
20 - 24	0.004591	0.022692569	78640.6641	2181.863	387748.6623	3916269.674	49.2
25 - 29	0.004742	0.023433352	76458.8008	826.9607	380226.6021	3528521.011	46.1
30 - 34	0.00648	0.031884692	75631.8401	1343.68	374800.0002	3148294.409	41.6
35 - 39	0.005891	0.029026272	74288.16	1740.347	367089.9327	2773494.409	37.3
40 - 44	0.078696	0.328793038	72547.8131	1704.217	358478.5229	2406404.476	33.2
45 - 49	0.008744	0.042784731	70843.5961	1696.611	349976.4539	2047925.953	28.9
50 - 54	0.015469	0.074464175	69146.9854	1815.857	341195.2859	1697949.499	24.6
55 - 59	0.015537	0.074778606	67331.1289	2054.978	331518.1995	1356754.213	20.2
60 - 64	0.003479	0.017242743	65276.1509	2449.203	320257.7472	1025236.014	15.7
65 - 69	0.038589	0.175967505	62826.948	929.4455	311811.1263	704978.2667	11.2
70 - 74	0.061844	0.267812993	61897.5025	5945.853	294622.8793	393167.1405	6.4
75 +	0.122356	1.000000	55951.6492	55951.65	98544.26112	98544.26112	1.8

Sources: The researcher from applied study (Sudan Census Data., 1993)

Table (9 -4) above illustrates the life table for females-Census (1993) - Sudan. From this table we can observe that, the highest life expectation in the second age group (e_4) and the lowest (nq_x), (nM_x) in the fourth age group ($10q_{14}$) ($10M_{14}$).

4 - 2 – 4 Life Tables 2008:

Tables (4-10), (4 - 11) and (4 - 12) show the life expectation of both sexes , males and females. These tables indicate that, life expectation at birth (e_0)for both sexes was 59.8 years and Life expectation for males was 58.1 years while (e_0)for females was 61.4 years which the gap between them was 3.3 years. It has been well established that males health situation lag behind females health.

In view of the table (4-14), the life expectation at birth is low compared to each region of the world except in Africa. See table (14 - 4).

(4-10)Life Table for Both Sexes -Census., 2008 -Sudan

Age group	$_nM_x$	nq_x	ℓ_x	$_nd_x$	nL_x	T_x	e_x
0	0.08362	0.0790	100000	7900	94470	5979202	**59.8**
1 - 4	0.00883	0.0345	92100	3180	362045	5884732	63.9
05 - 09	0.00231	0.01150	88920	1020	442057	5522688	62.1
10 - 14	0.00176	0.00880	87900	770	437572	5080631	57.8
15 - 19	0.00268	0.01330	87130	1160	432756	4643059	53.3
20 - 24	0.00366	0.01810	85970	1560	425967	4210303	49.0
25 - 29	0.00157	0.00790	84410	660	420406	3784335	44.8
30 - 34	0.00161	0.00790	83750	670	417084	3363929	40.2
35 - 39	0.00194	0.00790	83080	800	413416	2946846	35.5
40 - 44	0.00482	0.02390	82280	1960	406502	2522430	30.8
45 - 49	0.00681	0.03340	80320	2690	394880	2126928	26.5
50 - 54	0.00993	0.04840	77630	3760	378770	1732048	22.3
55 - 59	0.01462	0.07060	73870	5210	356333	1353278	18.3
60 - 64	0.02236	0.10580	68660	7270	325127	996945	14.5
65 - 69	0.03469	0.15960	61390	9800	282467	671818	10.9
70 - 74	0.05533	0.24300	51590	12540	226622	389351	7.5
75 +	0.23997	1.00000	39050	39050	162729	162729	4.2

Sources :The researcher from applied study(Sudan Census Data .,2008)

Life table (4-10) for both sexes-Census (2008) -Sudan. This table shows that the life expectation is high in the first three age group and then gradually decreases with age. The lowest value (nq_x) was found at ($25q_{29}$) ($30q_{34}$) ($35q_{39}$) and the lowest value of (nM_x) at ($29M_{25}$).

(4-11) Life Table for Males-Census., 2008-Sudan

Age group	**$_nM_x$**	**nq_x**	**ℓ_x**	**$_nd_x$**	**nL_x**	**T_x**	**e_x**
0	0.09105608	0.0856	100000	8560	94006	5814465	**58.1**
1 - 4	0.008856852	0.0346	91438	3165	359422	5720459	62.6
05 - 09	0.002349654	0.0117	88273	1031	438788	5361037	60.7
10 - 14	0.001756842	0.0088	87242	763	434302	4922249	56.4
15 - 19	0.002757105	0.0137	86479	1184	429436	4489749	51.9
20 - 24	0.003909099	0.0194	85295	1651	422348	4058511	47.6
25 - 29	0.001266491	0.0063	83644	528	416900	3636164	43.5
30 - 34	0.001272139	0.0063	83116	527	414263	3219264	38.7
35 - 39	0.001280285	0.0064	82589	527	411627	2805001	34
40 - 44	0.006222372	0.0306	82062	2514	404026	2393374	29.2
45 - 49	0.001012185	0.0427	79548	394	389257	1989348	25
50 - 54	0.02085185	0.0617	79154	7695	369032	1600092	21
55 - 59	0.018816613	0.0899	71459	6421	341241	1231060	17.2
60 - 64	0.028392362	0.1326	65038	8621	303638	889818	13.7
65 - 69	0.042798838	0.1933	56417	10906	254820	586181	10.4
70 - 74	0.065890362	0.2829	45511	12873	195370	331361	7.3
75 +	0.240002941	1.0000	32638	32638	135990	135990	4.2

Sources :The researcher from applied study(Sudan Census Data .,2008)

Table (4-11) above shows life table for males-Census (2008) -Sudan. From this table we note that, the highest value for life expectation in the second age group (e_4), the lowest (nq_x) in age group ($25q_{29}$), ($30q_{34}$) and the lowest value of (nM_x) at ($25M_{29}$).

(4-12)Life Table for Females -Census., 2008-Sudan

Age group	${}_nM_x$	nq_x	ℓ_x	${}_nd_x$	nL_x	T_x	e_x
0	0.076265	0.0724	100000	7240	94933	6143940	**61.4**
1 - 4	0.008804	0.0344	92761	3189	364668	6049008	65.2
05 - 09	0.002277	0.0113	89572	1014	445327	5684340	63.5
10 - 14	0.001767	0.0088	88558	779	440843	5239013	59.2
15 - 19	0.002584	0.0128	87779	1127	436077	4798170	54.7
20 - 24	0.00342	0.017	86652	1469	429587	4362093	50.3
25 - 29	0.00189	0.0094	85183	801	423912	3932506	46.2
30 - 34	0.00191	0.0095	84382	802	419904	3508595	41.6
35 - 39	0.002596	0.0129	83580	1078	415205	3088691	37.0
40 - 44	0.003455	0.0171	82502	1413	408978	2673486	32.4
45 - 49	0.004936	0.0244	81089	1977	400503	2264508	27.9
50 - 54	0.007261	0.0357	79112	2821	388508	1864005	23.6
55 - 59	0.010802	0.0526	76291	4012	371424	1475497	19.3
60 - 64	0.017053	0.0818	72279	5911	346617	1104073	15.3
65 - 69	0.028022	0.1309	66368	8690	310115	757456	11.4
70 - 74	0.047333	0.2116	57678	12206	257874	447341	7.8
75 +	0.24	1	45472	45472	189467	189467	4.2

Sources : The researcher from applied study(Sudan Census Data .,2008)

Table (4-12) above shows life table for females-Census (2008) -Sudan. From this table we observe that, the highest life expectation in the second age group (e_4) and the lowest value of (nq_x) (nM_x) at (10 q_{14}) (10M_{14}).

4 -3Mortality Levels, Trends and Patterns in Sudan:

The following section will discussed levels, trends and patterns of mortality in Sudan .The indicators which we used to measure the levels , trends and patterns of mortality are: life expectation at birth(e_0) , life expectation at age 20 (e_{20}) ,infant mortality rate(q_1), under five mortality rate (q_5) and central death rate(nM_x).The first indicator (e_0) reflects the general health conditions of the population, the second one (e_{20}) reflects the adult health conditions, the third one (q_1) reflects the infant health conditions and socioeconomic situations of the population, (q_5) reflects the child hath conditions finally(nM_x) as indicator to reflects the level of mortality . We used (q_1), (q_5) because these indicters more influence in level of mortality and reflect the improvement and development of population, All these indicters we obtained from the previous tables from table (4 - 1) up to table (4 – 12).

4 - 3 - 1 Mortality Levels In Sudan:

Figure (4 - 1), (4 - 2), (4 -3) and (4-4) show nM_x for both sexes, males and females respectively, during the period (1973 -1983- 1993 - 2008) for the Sudan, which it reflects the level of mortality.

Note that these curves take the form of the a letter (u) which means higher mortality rates in the age groups small and large, and be moderate slightly in the middle of the age distribution and more moderate in the broad age group (10-39). These feature of deaths in developing countries.

Sudan census., 1973 explained that, with respect to infant and child mortality rate, there was about 132 per 1000 live births for both sexes, 141 per 1000 live births for males and122 per 1000 live births for females expected to die before celebrating their first birth day. Under five mortality rate was158 per 1000 live births for both sexes, 171 per 1000 live births for males and 146 per 1000 live births for females die before reaching age five. For the adult health conditions, the life expectancy age 20 about 41.9, 40.7 and 43.1 years for sexes, males and females respectively.

The life expectation at birth (e_0) is used as indicter for the general health of population as we mentioned above, the figures for this indicator show that, the average expected years for new born to be in life are 47.3, 45.5and 49.1 years for both sexes , males and females.

Sudan census, 1983 explained that, the infant mortality rate was about 110 per 1000 live births for both sexes, 116 per 1000 live births for males and141 per 1000 live births for females expected to die before celebrating their first birth day. Child mortality rate was137 per 1000 live birth for both sexes, 132 per 1000 live births for males and 144per 1000 live births for females die before reaching age five. For the adult health conditions, the life expectation age 20 about 45.3, 44.3 and 46.2 years for sexes, males and females respectively.

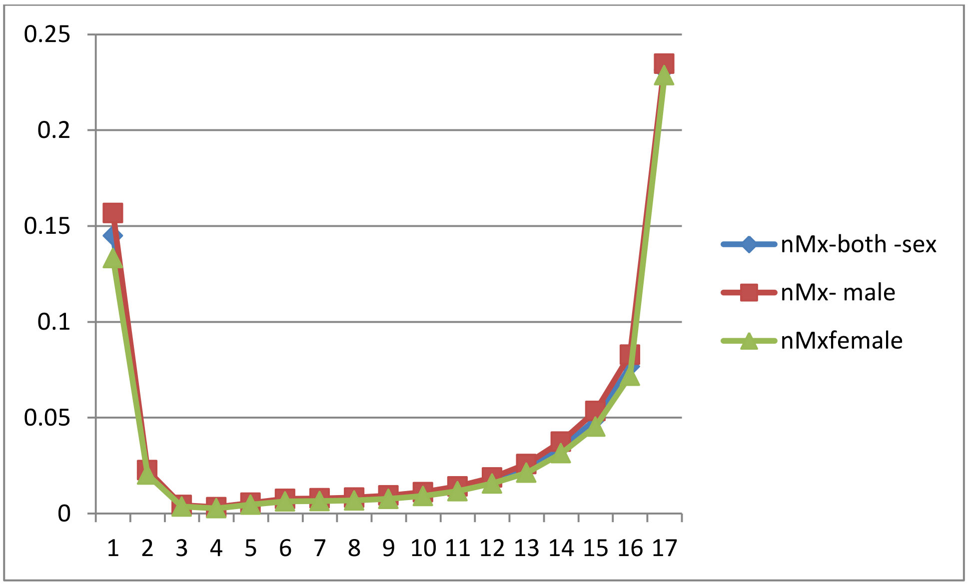
The life expectation at birth (e_0) was 53.5 years for both sexes, 52.7 years for males and 54.4 years for females.

Sudan census, 1993 show that, the infant mortality rate was about 108 per 1000 live births for both sexes, 116 per 1000 live births for males, there is no change in mortality among infants males during the period from 1983 to 1993, and females were expected to die before completing their first year was103 per 1000 live births. Child mortality rate was157 per 1000 live birth for both sexes, 163 per 1000 live births for males and 150 per 1000 live births for females die before reaching age five. For the adult health conditions, the life expectation age 20 about 47.8, 46.3 and 49.2 years for both sexes, males and females respectively. The life expectation at birth were 55.4 years for both sexes,و 55.34 years for males and 55.31 years for females.

The last census, 2008 show that, the infant mortality rate was about 79 per 1000 live births for both sexes, 86 per 1000 live births for males and72 per 1000 live births for females expected to die before celebrating their first birth day. Child mortality rate was111 per 1000 live birth for both sexes, 117 per 1000 live births for males and 104 per 1000 live births for females die before reaching age five. The life expectation age 20 about 49.0, 47.6 and 50.3 years for both sexes, males and females respectively .The life expectation at birth was 59.8 years for both sexes, 58.1 years for males and 61.4 years for females. See table (4 - 13).

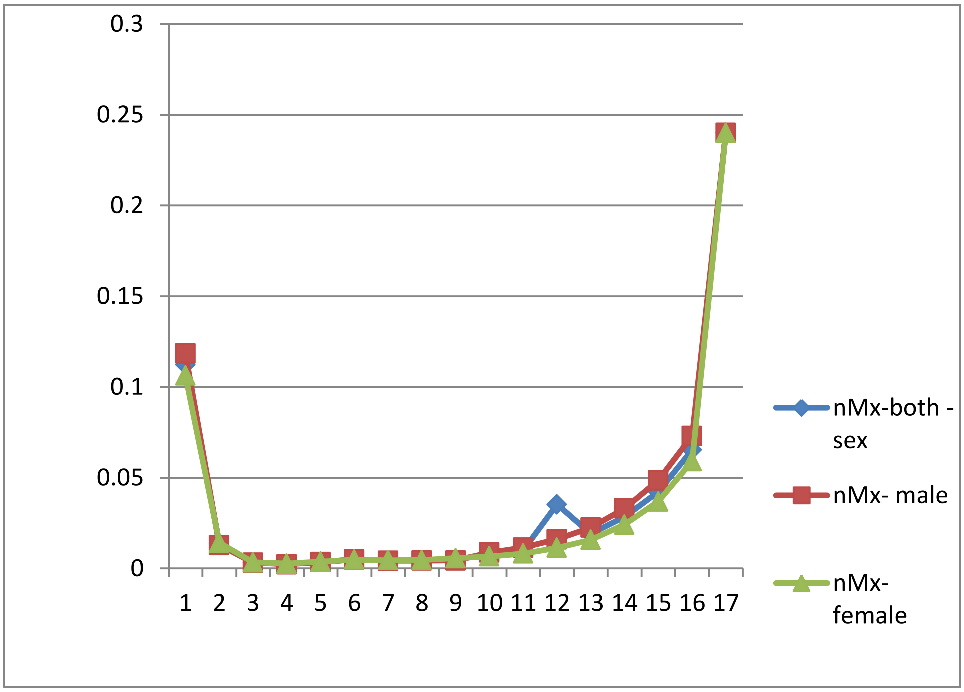
These levels of mortality in Sudan are considered to be high compared to the average levels in developed country.

Figure (4 - 1): nM_x by Sex Sudan - 1973



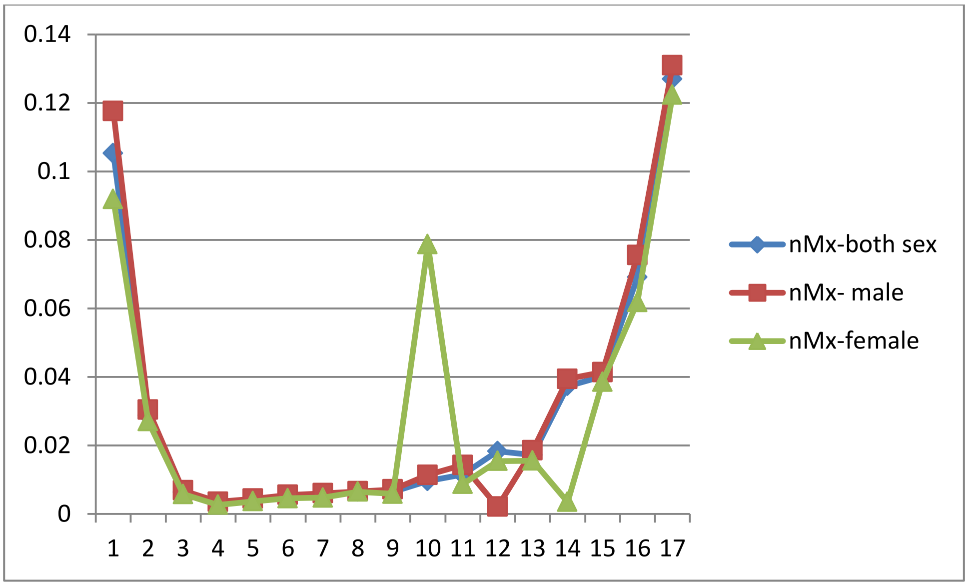
Sources : The researcher from applied study(Sudan Census Data .,2008)

Figure (4 - 2): nM_x by Sex Sudan – 1983



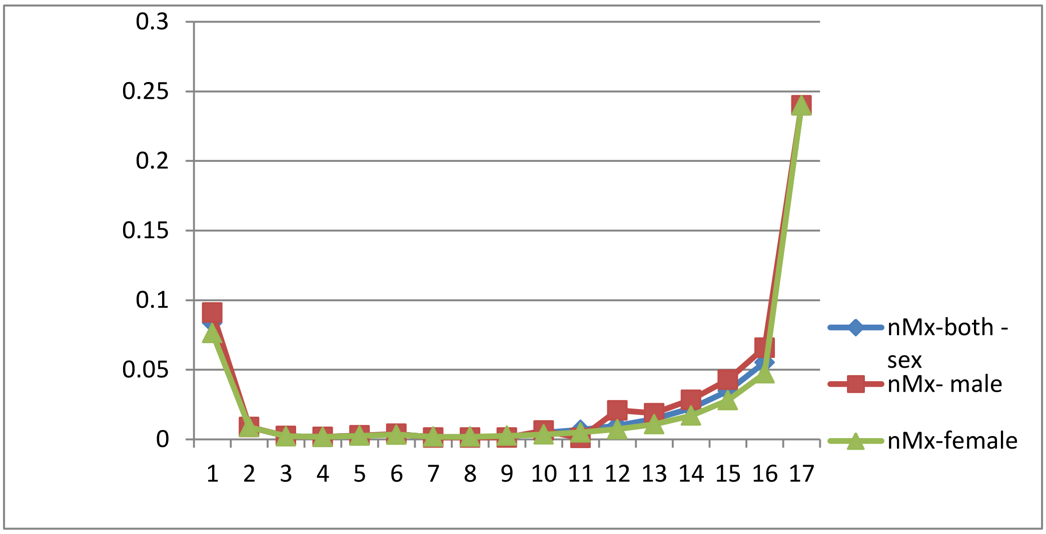
Sources : The researcher from applied study(Sudan Census Data .,2008)

Figure (4 - 3):nM_x by Sex Sudan - 1993



Sources : The researcher from applied study(Sudan Census Data .,2008)

(4 -4) nMx by Sex Sudan – 2008



Sources : The researcher from applied study(Sudan Census Data .,2008)

4 -3 – 2Mortality Trends in Sudan:

Table (4 – 13) show that, infant mortality rate (q_1) for both sexes in 2008 decreased by 40% from that in 1973, 39% for males and 41% for females and the under five mortality rate (q_5) decreased by 30% for both sexes, 32% for males and 29% for females during the same period, which indicter to steadily improvement in infant and child health conditions during this period.

Figure (4 - 7) explained that, no significant change is observed in the life expectation at age 20 (e_{20}) and increased by only around 7 years for both sexes, males and females during this period , that indicates to a tendency of leveling improvement in the adult health conditions in Sudan. This figure will be less if the data for southern Sudan is included.

As a result the life expectation at birth (e_0) increased by around 12.5 years for both sexes, 12.6 years for males and 12.3 years for females during 35 years in Sudan. See figure (4 - 8).

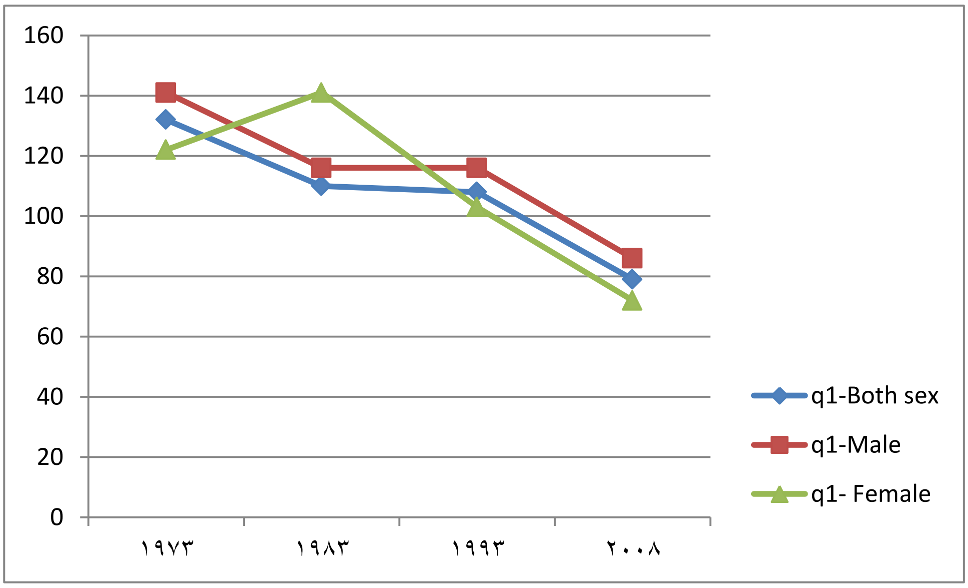
While life expectation at opened age group (e_{75+}) decreased by around 0.12 years for both sexes, 0.06 years for males and 0.17 years for females during same period in Sudan.

Figures (4 - 5) and (4 - 6) show that, (q_1) for females increased in the 1983 census compared to (q_1) of 1973 census. Also (q_5) for both sexes, females and males rose in the 1993 census compared to (q_5) of 1983 census. During the period (1993 - 2008) there was a decrease in (q_1) by 27%,26% and30% for both sexes, males and females, Also there was a decrease in (q_5) during that period the decline about 29%,28% and 31% for both sexes, males, females.

The fall in health levels in Sudan in 1983 was obviously due to the drought at that time and the deterioration after 1987 is probably due to decline in the general economic situation, decline in the value of the pound and inflation.

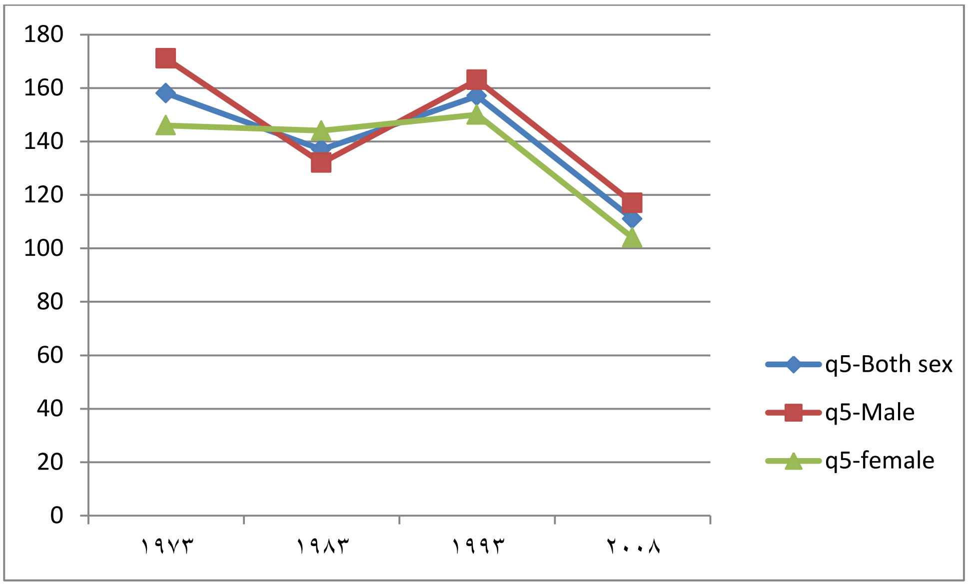
Compared to the regions, we find that life expectation at birth in Sudan is low compared to each regions of the world except Africa. Where life expectation at birth in Africa during the period (1970 - 2000) has increased by about 8 years only, while in Sudan, as previously reported increased by about 14 years during that period. See table (4 -14).

Figure (4 - 5): Estimates of Infant Mortality Rates (per 1000) as Obtained from the (1973, 1983, 1993, 2008) Censuses Data, Sudan



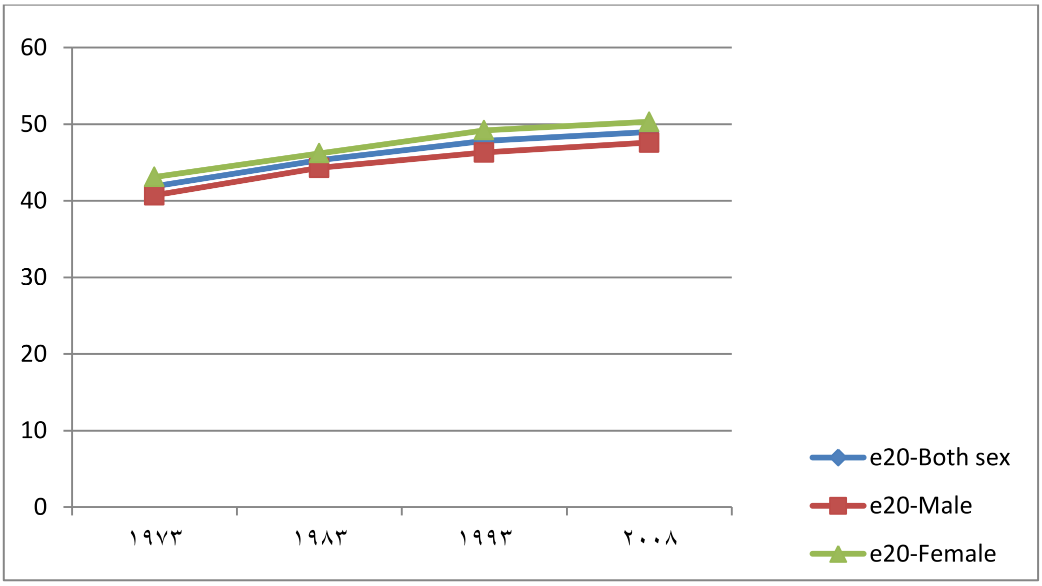
Sources : The researcher from applied study(Sudan Census Data .,2008)

Figure (4 - 6): Estimates of Child Mortality Rates(per 1000) as Obtained from the (1973, 1983, 1993, 2008) Censuses Data, Sudan



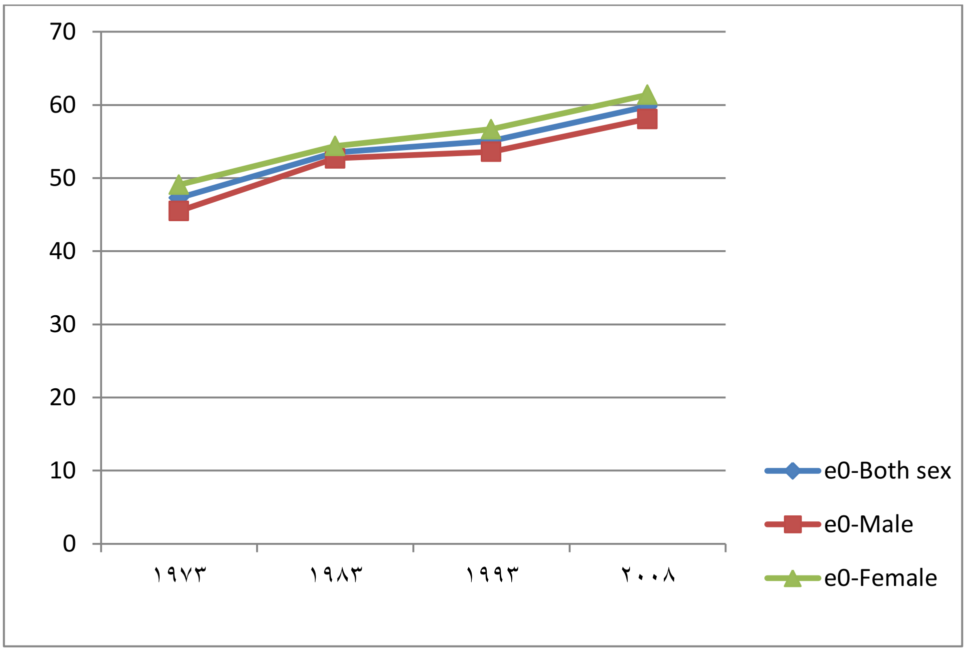
Sources : The researcher from applied study(Sudan Census Data .,2008)

Figure (4 - 7): Estimates of Life Expectation Age 20 as Obtained from the (1973, 1983, 1993, 2008) Censuses Data, Sudan



Sources : The researcher from applied study(Sudan Census Data .,2008)

Figure (4 - 8): Estimates of Life Expectation at Birth as Obtained from the (1973, 1983, 1993, 2008) Censuses Data, Sudan



Sources : The researcher from applied study(Sudan Census Data .,2008)

4 – 3- 3 Mortality Patterns in Sudan:

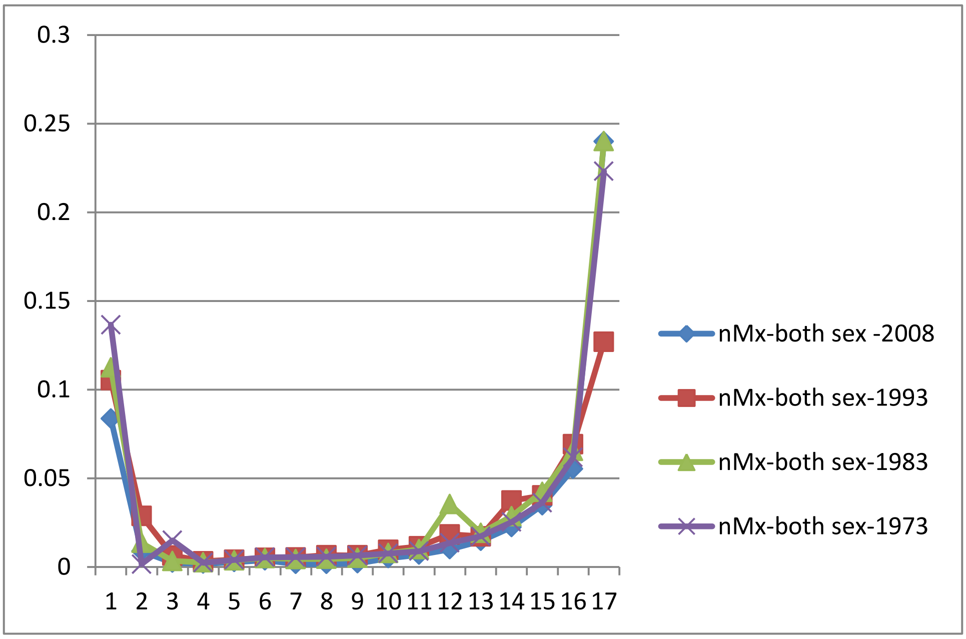
Given figures (4-1) (4-2) (4-3) (4-4) As mentioned previously, The central death rate (nM_x) by age groups takes the form of the letter (u), which means higher rates of death in earlier and later age groups, with steady shape for the middle age groups . It was observed that, the age group (50-54) for both sexes of the 1983 census showed significant rise in central death rate is likely to be occurred due to some mistakes made at any stage of data collection for that category. Also there is a sharp rise in the central death rate for females in the age group (40-44) in the 1993.

The curve fells particularly in age group (50 - 54) for males and (60 - 64) for females which it is nearly zero and then rose again in the last opened age group (75+), this pattern is familiar in the third World countries, but it was more severe in the 1993 census .With respect of the central death rate by sexes of death is clear that, the risk of death more density among men than females in most age groups. See figure (4-9) (4-10) (4-11).

Table (4-13) shows that, the infant mortality rate (q_1) for males is higher than females except for Census 1983 which increased the proportion of females than males by 21.6% .Also child mortality rate (q_5) for males is higher than females, the increase about 9.1%.

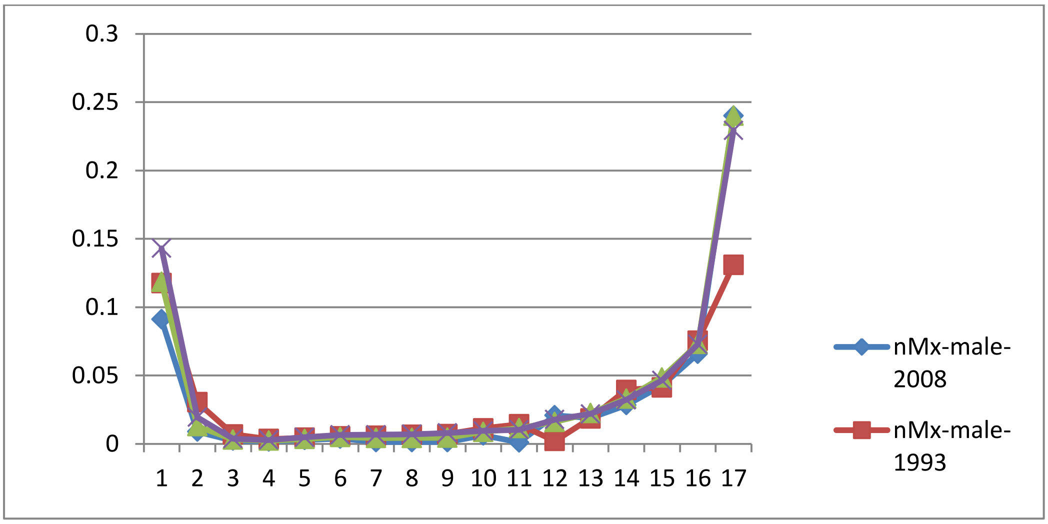
Also it's observed that, the level of life expectation at birth (e_0) and life expectation age 20 (e_{20}) is higher for females than males in every census(See figure (4 -5),(4- 6), (4 -7)and (4 -8).

Figure (4 - 9) Estimates of nM_x for Both Sex as Obtained from the (1973, 1983, 1993, 2008) Censuses Data, Sudan



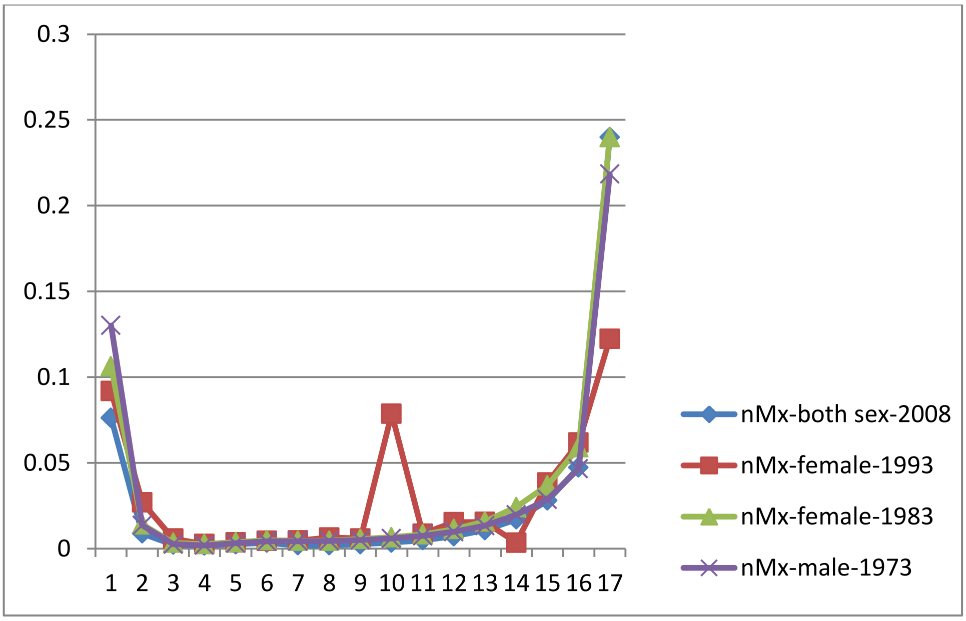
Sources : The researcher from applied study(Sudan Census Data .,2008)

Figure (4 - 10): Estimates of Life nM_x for Males as Obtained from the (1973, 1983, 1993, 2008) Censuses Data, Sudan



Sources : The researcher from applied study(Sudan Census Data .,2008)

Figure (4 - 11): Estimates of nM_x for Females as Obtained from the (1973, 1983, 1993, 2008) Censuses Data, Sudan



Sources : The researcher from applied study(Sudan Census Data .,2008)

Table (4 - 13): Mortality Levels, Trends and Patterns in Sudan from Population Censuses (1973 – 1983- 1993- 2008)

Both sexes					Males				Females			
year	**$q_{(1)}$**	**$q_{(5)}$**	**e_0**	**e_{20}**	**$q_{(1)}$**	**$q_{(5)}$**	**e_0**	**e_{20}**	**$q_{(1)}$**	**$q_{(5)}$**	**e_0**	**e_{20}**
1973	132	158	47.3	41.9	141	171	45.5	40.7	122	146	49.1	43.1
1983	110	137	53.5	45.3	116	132	52.7	44.3	141	144	54.4	46.2
1993	108	157	55.4	47.8	116	163	55.34	46.3	103	150	55.31	49.2
2008	79	111	59.8	49.0	86	117	58.1	47.6	72	104	61.4	50.3

Sources : The researcher from applied study (1973 – 1983- 1993- 2008) population censuses

Table (4 - 14): Life expectation at Birth by World Region, 1950–2000

Area	**Years**					
	1950 – 1955	**1960 – 1965**	**1970 – 1975**	**1980 – 1985**	**1990 – 1995**	**1995 – 2000**
World	46	52	58	61	64	66
Developed Countries	67	70	71	73	74	74
Less Developed Countries	41	48	55	59	62	64
Africa	38	42	46	49	53	54
Asia	41	48	56	60	65	66
Latin America (and Caribbean)	51	57	61	65	69	70
Europe	66	70	71	72	73	73
North America (U.S. and Canada)	69	70	72	75	76	77

Source **:** Yaukey, David, and Douglas L. Anderton. Demography: The Study of Human Population. Prospect Heights, IL: Waveland, 2001.

4- 4 Comparison of Population Censuses Using Life Tables:

Population census1956 was the most accurate census in Sudan compared with the others. This population census has become the standard basis for all subsequent censuses. All the subsequent censuses were adjusted according to that census, However there are some variations in the population growth within the regions of Sudan. That was mainly due to the open immigration from the West African countries.

On the other hand if we look at South Sudan and nomadic populations, there is illogical and inconsistent population growth. It is oscillating in the case of South Sudan and there is a big drop in the case of the nomads specifically in the 1993 census. All censuses Sudan were suffering from some states not covered by the census except the 2008 census, difficulties in reaching the households and incomplete coverage. Also there was lack of accuracy in the information given by the households. It is commonly known in Sudan that some tribes report lesser number of their children believing that the evil eyes of the enumerators will kill their children when they report many.

The third population census in Sudan, there was concerns about accuracy in the South due to weak logistics. It was noted that some enumerators who found it difficult to walk for long distances had used the chiefs of the tribes to fill the forms on behalf of the households. So 1983 census it was considered the worst in terms of accuracy, so in this study we faced many problems to estimate life expectation.

The 5th population census was one of the most important censuses in the history of Sudan. It was based on the comprehensive peace agreement. To achieve this goal a population census with a high accuracy and a full coverage is a necessity.

When we look to the life tables (1973-1983-1993-2008) we observe the following:

1- Life expectation at birth (e_0) and life expectation in the age group 20 (e_{20}) increased. Also, (q_1) (q_5) decreased during that period (1973-2008) except population census 1993 and (q_1) for females in population census 1983 which is likely to have occurred due to some mistakes made at any stage of data collection for that category. We also notice a rise (q_5) in the population census 1993 compared to population census 1983 despite improved in health situation. Also (q_1) is higher among males than females in all population censuses except population census 1983.See table (4-13).

2- The greatest improvement in life expectation at birth (e_0) and (e_{20}) occurred between the population censuses (1973 -1983) maybe due to the population census 1983 was an estimated in a number of areas. See table (4-13).

3-There was high difference in life expectation at birth between the first and second age groups in population census 1993 compared to other population censuses may be due to errors in reporting. The population census 1993 also recorded the lowest life expectation in the open age group (e_{75+}) compared with other population censuses.

4 – 5Forecasting of Life Expectation at Birth (e_0) by Using Demographic Models:

To forecast life expectation at birth during the period (2009 - 2020), we used several criteria based on life expectation at birth (e_0) during the period (1973-2008) to find them. See table (4 - 32), such as:

1- Annual Growth Rate:

$$AGR = e^{rt} \ldots\ldots\ldots\ldots\ldots (4 -1)$$

2- Sudan standardization.

3- The Spectrum program.

The results were as follows:

Table (4 - 15): Forecasting of Life Expectation at Birth (e_0) by Using Demographic Models for (Both Sexes- Males – Females) during period (2009 - 2020)

year	e_0 for Both Sexes	e_0 for Male	e_0 for Female
2009	60.0	58.0	61.8
2010	60.3	58.2	62.3
2011	60.6	58.4	62.7
2012	60.9	58.5	63.1
2013	61.2	58.7	63.6
2014	61.5	58.9	64.0
2015	61.8	59.1	64.5
2016	62.1	59.2	64.9
2017	62.4	59.4	65.4
2018	62.8	59.6	65.8
2019	63.1	59.8	66.3
2020	63.4	60.0	66.8

Source: The researcher from applied study, E-views Package. 2015

Table (4-15) shows that, forecasting of life expectation at birth (e_0) by using demographic models for (both sexes- males – females) during period (2009 - 2020).We note that, the life expectation at birth (e_0) during that period increased about 3.4, 2.0 and 5.0 years for both sexes, males and females respectively.

4 – 6 Statistical Model (Box – Jenkes time series model):

In this section Box – Jenkes time series model has been applied for life expectation at birth(e_0)for (both sexes, males and females) in the Sudan during the period (1973-2008) in order to forecasting (e_0) during the period (2009-2020).

4 -6- 1Time Series Box – Jenkes Model for Both Sexes:

4-6-1-1 Description of the Series for Both Sexes:

Table (4 - 16): Descriptive Statistics for Both sexes (e_0)

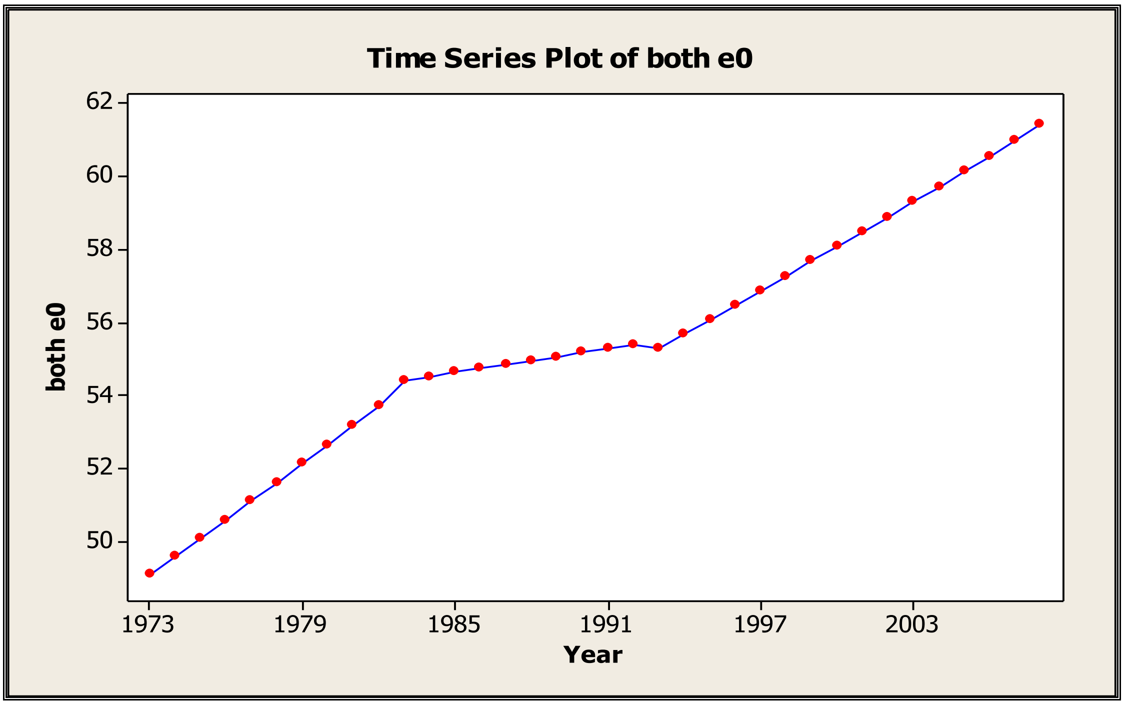
Minimum	Maximum	Mean	Std. Deviation
47.40	59.80	54.4694	3.47463

Source: The researcher from applied study , E-views Package., 2015

From the above table (4- 16), which it shows the life expectation at birth for both sexes (e_0), we see that, the maximum value is (59.80), the minimum value is (47.40), the mean of (e_0) is (54.4694) and the standard deviation of (e_0) is (3.47463).

4 – 6 - 1 – 2 Time Series Plot for Both Sexes (*e_0*):

Figure (4 - 12): Life Expectation at Birth (e_0) for Both Sexes (1973 - 2008)



Source: The researcher from applied study , E-views Package., 2015

The previous figure (4- 12) shows that, the original data series of (e_0) for both sexes, it is clear that it has steadily increasing trend with not fixed mean and variance , so that we can say the series is not stationary . To attain stationary series we use suitable transformation such as differences.

4 – 6 – 1- 3 Diagnosis of (*e_0*) for Both Sexes:

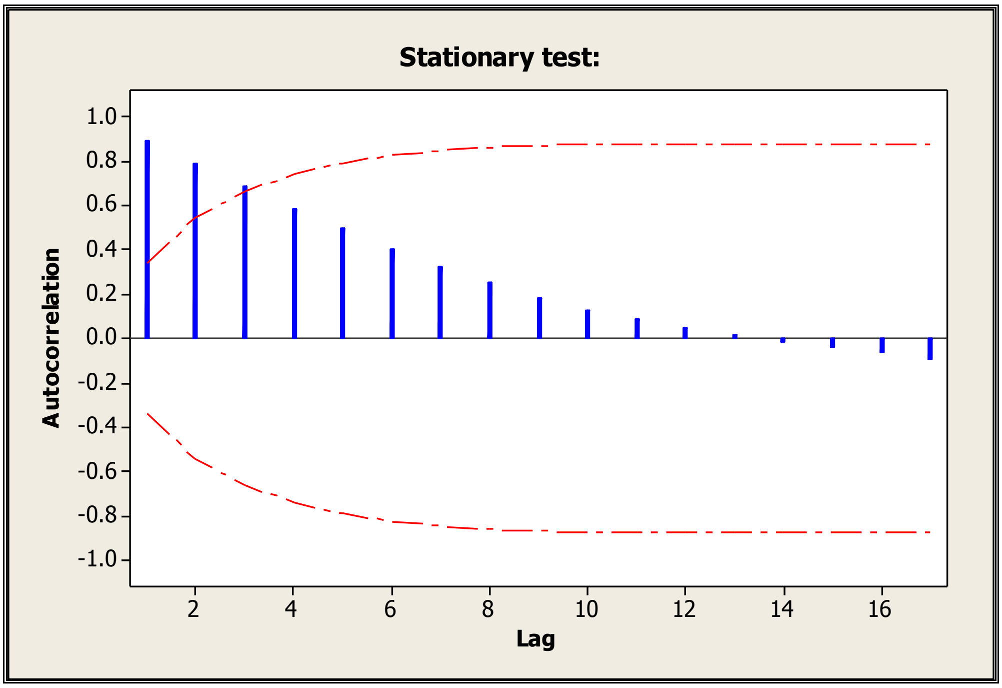
Table (4 - 17): Stationary Test of Original Data of (e_0) for Both Sexes

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-8.055114	0.0000
Test critical values:	1% level	-4.262735	
	5% level	-3.552973	
	10% level	-3.209642	

*MacKinnon (1996) one-sided p-values.

Source: The researcher from applied study , E-views Package., 2015

Figure (4 - 13): Autocorrelation Coefficient Function (ACF) of Original Data of (e_0) for Both Sexes



Source: The researcher from applied study , E-views Package., 2015

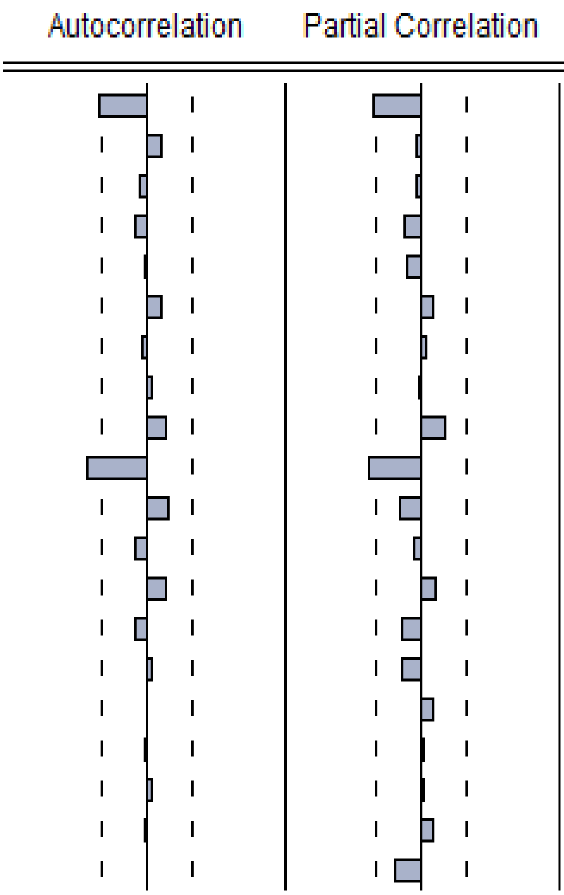
We use the Augmented Dickey-Fuller test statistic (ADF) to test the stationary of the series.

Table (4 - 17) shows that, the value of t – calculated of original data is greater than t – tabulated, so we reject the null hypothecs {H_0: $P_k = 0$ (the series is stationary)}, hence the series is not stationary, also figure (4-13) explains that, all lags of Autocorrelation Coefficient Function (ACF) of original data lying between upper and lower confidence limits except the first ,second and third lags are outside of limits ,so also indicated that , the original data series is not stationary.

4 – 6– 1- 4 Stationary of Series after Second Difference for Both Sexes (e_0):

Figure (4 - 14): (ACF), (PACF) After Second Difference of (e_0) for Both Sexes

Date: 05/28/15 Time: 15:40
Sample: 1973 2008
Included observations: 34



	AC	PAC	Q-Stat	Prob
1	-0.360	-0.360	4.8111	0.028
2	0.106	-0.027	5.2432	0.073
3	-0.054	-0.028	5.3574	0.147
4	-0.094	-0.138	5.7165	0.221
5	-0.014	-0.108	5.7245	0.334
6	0.106	0.085	6.2180	0.399
7	-0.041	0.025	6.2937	0.506
8	0.026	-0.010	6.3260	0.611
9	0.145	0.182	7.3617	0.600
10	-0.452	-0.387	17.802	0.058
11	0.161	-0.166	19.183	0.058
12	-0.092	-0.055	19.658	0.074
13	0.134	0.113	20.711	0.079
14	-0.092	-0.151	21.233	0.096
15	0.025	-0.150	21.274	0.128
16	-0.000	0.092	21.274	0.168
17	-0.014	0.009	21.287	0.214
18	0.026	0.016	21.341	0.263
19	-0.012	0.085	21.353	0.318
20	-0.001	-0.201	21.353	0.377

Source: The researcher from applied study , E-views Package., 2015

Table (4 - 18):Statistical – Tests after Second Difference of *(e_0)* for Both Sexes

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	55.63024	0.556026	100.0497	0.0000
MA(1)	0.953769	0.020422	46.70210	0.0000
R-squared	0.738143	Mean dependent var		55.43333
Adjusted R-squared	0.730442	S.D. dependent var		3.289637
S.E. of regression	1.707947	Akaike info criterion		3.962414
Sum squared resid	99.18081	Schwarz criterion		4.050387
Log likelihood	-69.32345	Hannan-Quinn criter.		3.993119
F-statistic	95.84205	Durbin-Watson stat		0.078473
Prob(F-statistic)	0.000000			

Source: The researcher from applied study , E-views Package., 2015

If we look to the figure (4-14) we see that, all lags of Autocorrelation Coefficient Function (ACF) and Partial Autocorrelation Coefficient Function (PACF) laying between upper and lower confidence limits after second difference. Also table (4 - 18) shows statistical tests after second difference from this table we can note that, t-test and F-test are significant for that we accept null hypothesis, so this refers that, the series of second difference is stationary.

So the best model is ARIMA (0, 2, 1), because it the lowest value to (Akaike) and the biggest value of Coefficient of Determination (R2).

4 – 6– 1- 5 Estimates of Parameters Model for Both Sexes (e_0):

Table(4- 19):MA(1)Result of Model for Both sexes

Type	Coefficient	Estimate	t-Statistic	P-value
ϕ_1	55.63024	0.1218	100.0497	0.0000
μ	0.953769	0.02189	46.70210	0.0000

Source: The researcher from applied study , E-views Package., 2015

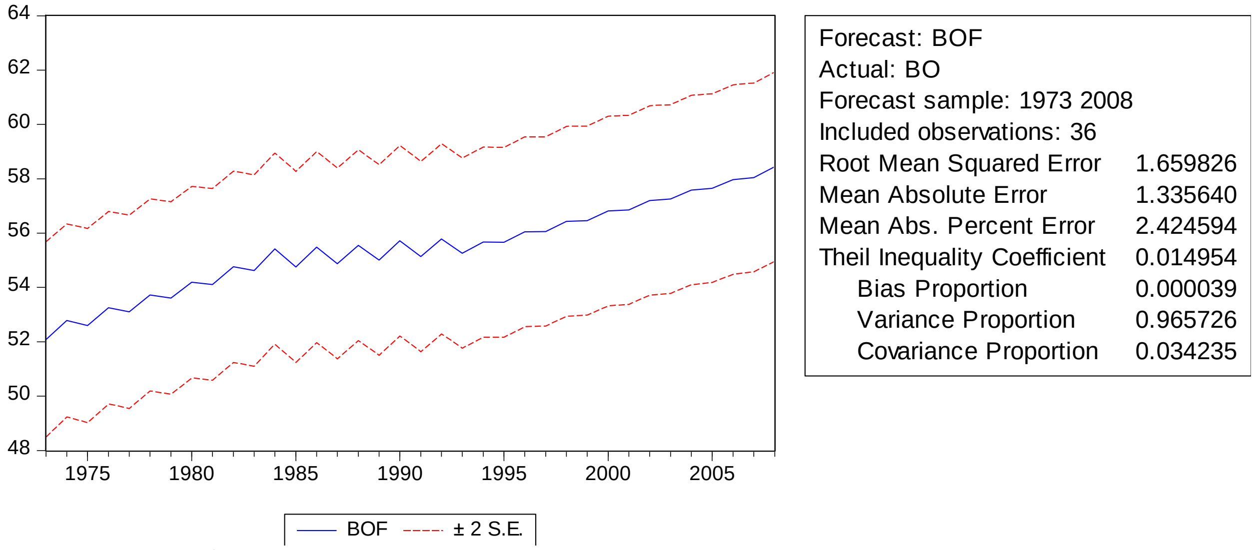
From the Previously we determined the appropriate model for a series ARIMA(0,2,1), and estimated as follows:

$$Z_t = \mu - \theta_1 \in_{t-1} - \theta_2 \in_{t-2} + \in_t$$

$$z_{t=} 0.1218 - 0.02189 \in_{t\text{-}1} - 0.9781 \in_{t\text{-}2} + \in_t \text{ ………….. } (4\text{-}2)$$

4 – 6– 1- 6 Forecasting of Series for Both Sexes (e_0):

Figure (4 - 15): Forecasting of (e_0) for Both Sexes (2009 - 2020)



Source: The researcher from applied study , E-views Package., 2015

Figure (4-15) shows that, the results of applied to (RMSE) confirm the validity of the estimated model to represent the time series and then use it to forecast.

Table (4 - 20): Forecasting of (e_0) for Both Sexes (2009 - 2020)

year	Forecasting of(e_0) for Both Sex
2009	61.68
2010	61.97
2011	62.26
2012	62.55
2013	62.83
2014	63.14
2015	63.39
2016	63.67
2017	63.94
2018	64.22
2019	64.48
2020	64.75

Source: The researcher from applied study , E-views Package., 2015

We estimated the future values of life expectation at birth (e_0) for both sexes to about12years (2009-2020) subsequent to the last Sudan population and housing census as sufficient to induce any change in the program of economic development-social -healthy to give its features .Table (4 - 20) and figure (4 - 15)illustrate it.

4 – 6 – 2Time Series Box – Jenkes Model for Males:

4 – 6– 2-1 Description of the Series for Males :

Table (4 - 21): Descriptive Statistics for Males (e_0)

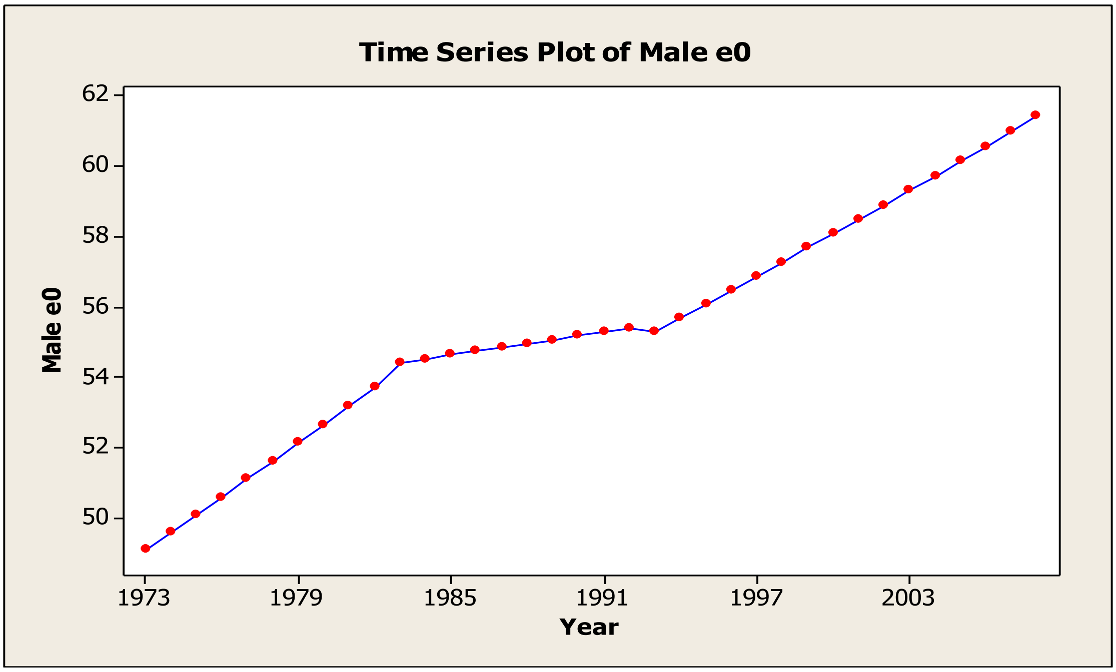
Minimum	Maximum	Mean	Std. Deviation
45.50	58.10	53.6500	3.55291

Source: The researcher from applied study , E-views Package., 2015

Table (4-21) shows that, the maximum value of life expectation at birth for males (e_0) is (58.10), the minimum value is (45.50), the mean of (e_0) is (53.65) and the standard deviation of (e_0) is (3.55291).

4 – 6 - 2 – 2 Time Series Plot for Males (e_0):

Figure (4 - 16): Life Expectation at Birth (e_0) for Males (1973 - 2008)



Source: The researcher from applied study , E-views Package., 2015

Figure (4- 16) shows that, the original data series of **(e_0)**for males it has steadily increasing trend with not fixed mean and variance , so that we can say the series is not stationary .

4 – 6- 2 – 3 Diagnosis of (e_0) for Males:

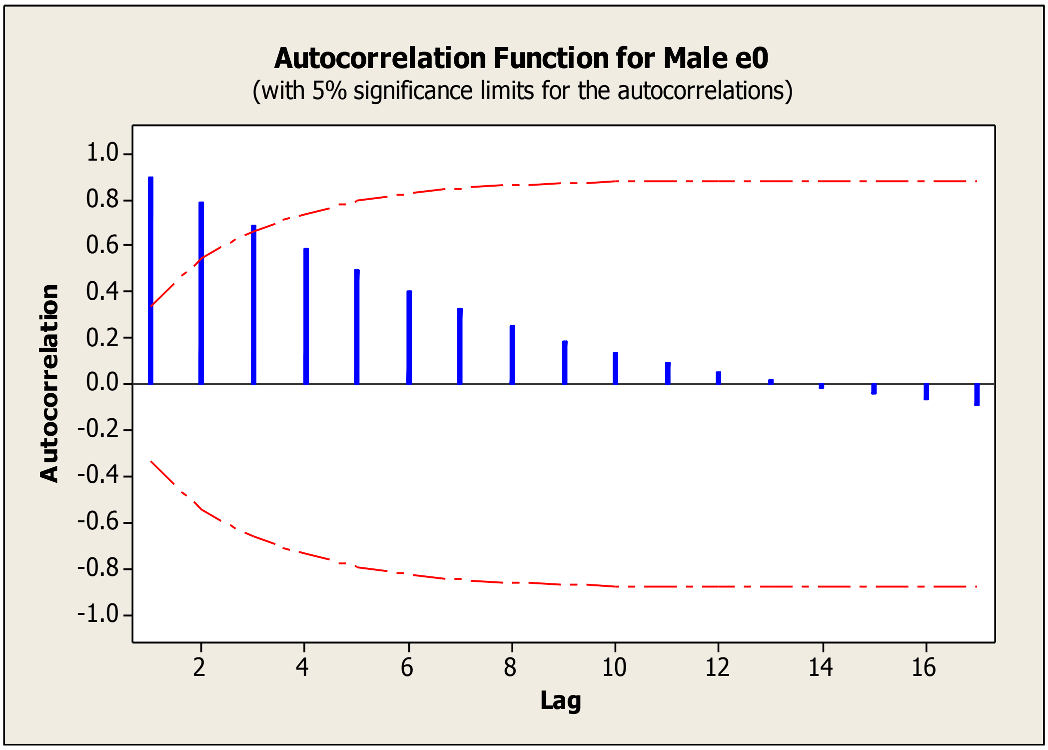
Table (4 - 22): Stationary Test of Original Data of (e_0) for Males

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.143383	0.0128
Test critical values:	1% level	-4.243644	
	5% level	-3.544284	
	10% level	-3.204699	

*MacKinnon (1996) one-sided p-values.

Source: The researcher from applied study , E-views Package., 2015

Figure (4 - 17): Autocorrelation Coefficient Function (ACF) of Original Data of (e_0) for Males



Source: The researcher from applied study , E-views Package., 2015

Table (4 - 22) shows that, the value of t – test of original data is not stationary so we reject the null hypothecs {H_0: P_k = 0 (the series is stationary)}, hence the series is not stationary, also figure (4 - 17) explains that , some lags of Autocorrelation Coefficient Function (ACF) of original data are laying outside of limits it indicated that , the original data series is not stationary.

4 – 6 - 2 – 4 Stationary of Series After First Difference for Males (*e_0*):

Figure (4 - 18): (ACF), (PACF) After First Difference of (*e_0*) for Males

Date: 05/28/15 Time: 15:35
Sample: 1973 2008
Included observations: 36

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.895	0.895	31.313	0.000
		2	0.794	-0.038	56.649	0.000
		3	0.694	-0.045	76.628	0.000
		4	0.598	-0.043	91.918	0.000
		5	0.505	-0.045	103.16	0.000
		6	0.416	-0.037	111.06	0.000
		7	0.335	-0.023	116.37	0.000
		8	0.261	-0.026	119.70	0.000
		9	0.196	-0.016	121.64	0.000
		10	0.138	-0.012	122.65	0.000
		11	0.090	-0.005	123.09	0.000
		12	0.043	-0.041	123.20	0.000
		13	-0.003	-0.046	123.20	0.000
		14	-0.047	-0.037	123.33	0.000
		15	-0.089	-0.040	123.85	0.000
		16	-0.130	-0.044	125.01	0.000

Source: The researcher from applied study , E-views Package., 2015

Table (4 - 23):Statistical – tests of (*e_0*) after First Difference for Males

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	59.60183	0.698691	85.30503	0.0000
AR(1)	0.940779	0.005911	159.1501	0.0000

R-squared	0.998699	Mean dependent var	53.88286
Adjusted R-squared	0.998659	S.D. dependent var	3.314361
S.E. of regression	0.121353	Akaike info criterion	-1.324785
Sum squared resid	0.485974	Schwarz criterion	-1.235908
Log likelihood	25.18374	Hannan-Quinn criter.	-1.294105
F-statistic	25328.76	Durbin-Watson stat	0.760526
Prob(F-statistic)	0.000000		

Inverted AR Roots	.94

Source: The researcher from applied study , E-views Package., 2015

To make stationary series we took the first difference, after taking the first difference series has become stationary, as is evident in the figure (4-18) and the table (4-23) where, most lags of Autocorrelation Coefficient Function (ACF) and Partial Autocorrelation Coefficient Function (PACF) laying between upper and lower confidence limits after first difference. Also t-test and F-test are significant. Akaike value is smallest (-1.324785) and R^2 is highest value (0.998699) .So the best model is ARIMA (1, 1, 0).

4 – 6– 2- 5 Estimates of Parameters Model for Males (e_0):

Table (4- 24): AR (1) Result of Model for Males

Type	Coefficient	Estimate	t	P-value
ϕ_1	0.940779	0.1101	85.30503	0.0000
μ	59.60183	0.5287	159.1501	0.0000

Source: The researcher from applied study , E-views Package., 2015

From the previously we determined the appropriate model for a series ARIMA (1, 1, 0) .It has been estimated as follows:

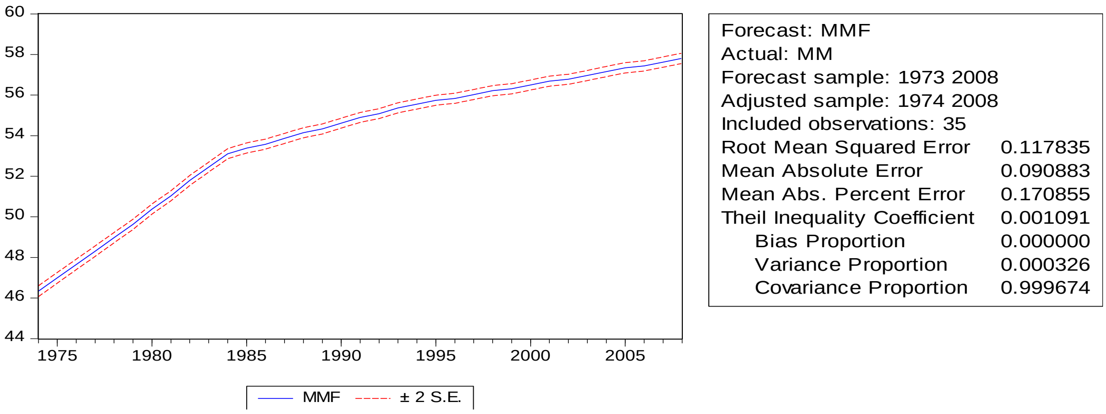
$$z_{t=}\ \mu + \emptyset_1 Z_{t-1} + \ \ \epsilon_t$$

$$z_{t=}\ 0.5287 + 0.1101 Z_{t-1} + \epsilon_t \ \ldots\ldots \qquad (4\text{-}3)$$

4 – 6 - 2 – 6 Forecasting of Series for Males (e_0):

Figure (4 - 19): Forecasting of Life Expectation for Males (e_0)

(2009 - 2020)



Source: The researcher from applied study , E-views Package., 2015

Figure (4-19) shows that, the results of applying to (RMSE) confirm the validity of the estimated model to represent the time series and then use it to forecast.

Table (4 - 25): Forecasting of (e_0) for Males (2009 - 2020)

year	Forecasting of(e0) for Males
2009	58.04
2010	58.13
2011	58.22
2012	58.30
2013	58.37
2014	58.45
2015	58.52
2016	58.58
2017	58.64
2018	58.70
2019	58.75
2020	58.80

Source: The researcher from applied study , E-views Package., 2015

Table (4- 25) illustrate that, the estimation of life expectation at birth for males to 12 years later to the last Sudan population and housing census (2009 - 2020).

4 – 6 – 3Time Series Box – Jenkes Model for Females:

4- 6- 3 -1 Description of the Series for Females:

Table (4 - 26): Descriptive Statistics for Females (e_0)

Minimum	Maximum	Mean	Std. Deviation
49.10	61.40	55.4333	3.28964

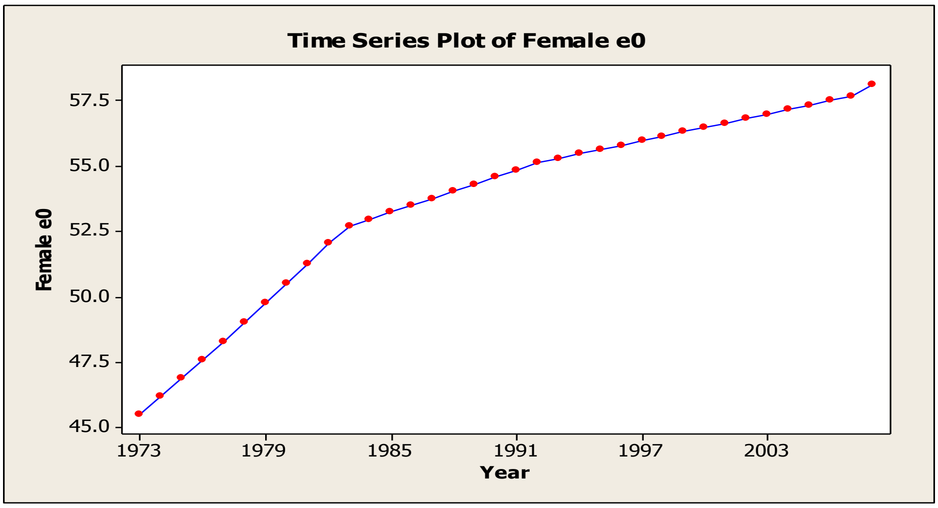
Source: The researcher from applied study , E-views Package., 2015

Table (4-26) explains that, the maximum value of life expectation at birth for females (e_0) is (61.40), the minimum value is (49.10), the mean of (e_0) is (55.4333) and the standard deviation of (e_0) is (3.28964).

4-6-3-2 Time Series Plot for Females (e_0):

Figure (4 - 20): Life Expectation at Birth (e_0) for Females

(1973 - 2008)

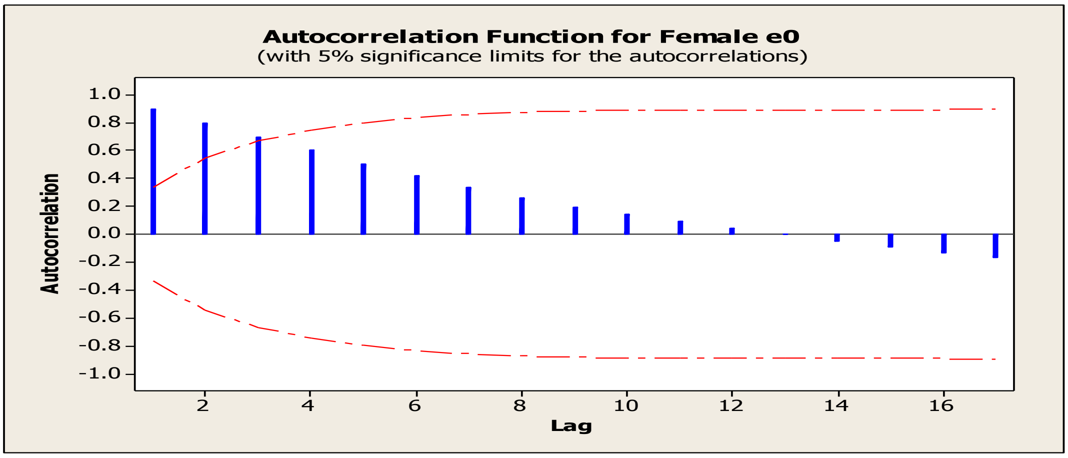


Source: The researcher from applied study , E-views Package., 2015

The figure (4-20) shows that, the life expectation at birth for females increasing with time, with fixed mean and variance, so that we can say the series is stationary.

4-6 - 3 – 3 Diagnosis of (e_0) for Females:

Figure (4 - 21): Autocorrelation Coefficient Function (ACF) of Original Data of (e_0) for Females



Source: The researcher from applied study , E-views Package., 2015

Table (4 - 27): Stationary Test of Original Data of (e_0) for Females:

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.2226716	0.0266
Test critical values:	1% level	-3.632	
	5% level	-2.948	
	10% level	-2.612	

*MacKinnon (1996) one-sided p-values.

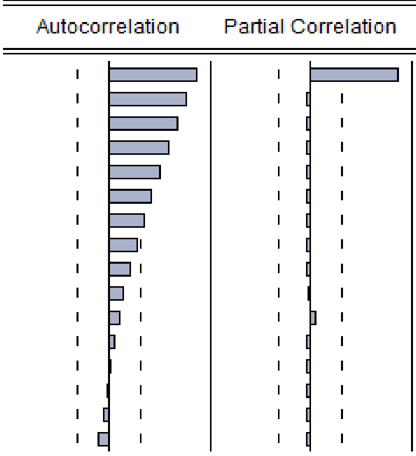
Source: The researcher from applied study , E-views Package., 2015

Table (4 - 27) shows that, the value of t – test of original data is stationary, also figure (4 - 21) explains that , lags of Autocorrelation Coefficient Function (ACF) of original data are laying between upper and lower confidence limits This means that , the original data series is stationary.

4-6-3 – 4 Stationary of Series for Females (e_0):

Figure (4 - 22): (ACF), (PACF) of (e_0) for Females

Date: 05/28/15 Time: 15:28
Sample: 1973 2008
Included observations: 36



	AC	PAC	Q-Stat	Prob
1	0.900	0.900	31.679	0.000
2	0.802	-0.043	57.587	0.000
3	0.706	-0.046	78.253	0.000
4	0.613	-0.037	94.337	0.000
5	0.524	-0.042	106.43	0.000
6	0.439	-0.033	115.20	0.000
7	0.358	-0.037	121.23	0.000
8	0.282	-0.032	125.11	0.000
9	0.211	-0.031	127.36	0.000
10	0.148	-0.018	128.52	0.000
11	0.103	0.043	129.10	0.000
12	0.061	-0.034	129.31	0.000
13	0.018	-0.044	129.33	0.000
14	-0.023	-0.041	129.36	0.000
15	-0.064	-0.045	129.63	0.000
16	-0.103	-0.037	130.36	0.000

Source: The researcher from applied study , E-views Package., 2015

Table (4 - 28):Statistical – Tests of (e_0) for Females

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	66.09405	3.799296	17.39639	0.0000
AR(1)	0.969917	0.009334	103.9102	0.0000
R-squared	0.996953	Mean dependent var		54.67143
Adjusted R-squared	0.996861	S.D. dependent var		3.303970
S.E. of regression	0.185121	Akaike info criterion		-0.480174
Sum squared resid	1.130897	Schwarz criterion		-0.391297
Log likelihood	10.40305	Hannan-Quinn criter.		-0.449494
F-statistic	10797.34	Durbin-Watson stat		1.365851
Prob(F-statistic)	0.000000			
Inverted AR Roots	.97			

Source: The researcher from applied study , E-views Package., 2015

The series of e_0 is stationary, as is evident in the figure(4 -22) and the table(4 - 28) where, most lags of Autocorrelation Coefficient Function (ACF) and Partial Autocorrelation Coefficient Function (PACF) laying between upper and lower confidence limits AR(1). t- test and F-test are significant, Akaike value is small and R^2 is high (0.996953) .So the best model is ARMA(1, 0).

4 – 6– 3- 5 Estimates of Parameters Model for Females (e_0):

Table (4- 29): AR (1) Result of Model for Females

Type	Coefficient	Estimate	t-Statistic	P-value
ϕ_1	0.969917	0.0978	17.39639	0.0000
μ	66.09405	0.5529	103.9102	0.0000

Source: The researcher from applied study , E-views Package., 2015

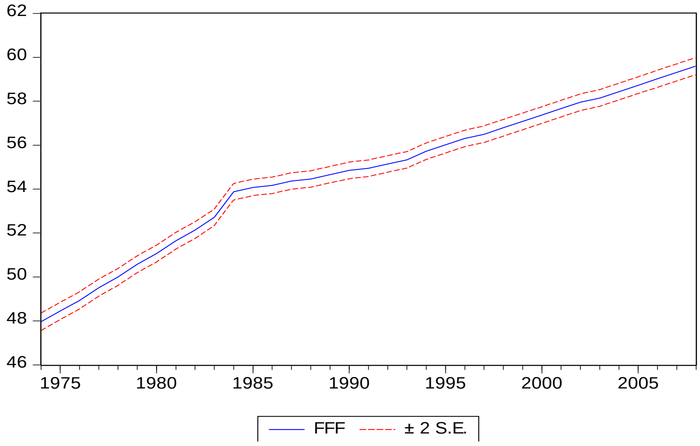
From the previously we determined the appropriate model for a series ARMA (1.0) .It has been estimated as follows:

$$z_{t=} \mu + \emptyset_1 Z_{t-1} + \in_t$$

$$z_{t=} 0.5529 + 0.0978\, Z_{t-1} + \in_t \text{ ……..} \quad (4\text{-}4)$$

4 –6 – 3 - 6 Forecasting of Series for Females (e_0):

Figure (4 - 23): Forecasting of (e_0) for Females (e_0) (2009 - 2020):



Forecast: FFF	
Actual: FF	
Forecast sample: 1973 2008	
Adjusted sample: 1974 2008	
Included observations: 35	
Root Mean Squared Error	0.179754
Mean Absolute Error	0.117134
Mean Abs. Percent Error	0.215842
Theil Inequality Coefficient	0.001641
Bias Proportion	0.000000
Variance Proportion	0.000763
Covariance Proportion	0.999237

Source: The researcher from applied study , E-views Package., 2015

Figure (4-23) shows that, the results of applied to (RMSE) confirm the validity of the estimated model to represent the time series and then use it to forecast.

Table (4 - 30): Forecasting of (e_0) for Females (2009 - 2020)

year	Forecasting of(e_0) for Females
2009	59.87
2010	60.06
2011	60.24
2012	60.41
2013	60.58
2014	60.75
2015	60.91
2016	61.07
2017	61.22
2018	61.37
2019	61.51
2020	61.65

Source: The researcher from applied study , E-views Package., 2015

Table (4-30) explains that, the forecasting of life expectation at birth for females to 12 years later to the last Sudan population and housing census (2009 - 2020).

4 –7Comparsion between Demographic and Statistical Models to Forecast Life Expectation at Birth (e_0):

One of the objectives of the research is to choose the best model (demographic model or statistical model) to forecast (e_0). So after we applied the statistical and demographic models to forecast (e_0) for (both sexes – males - females) data series, so we found the statistical model is better than demographic model when data are homogeneous, While in the case of heterogeneous data, the demographic model gives better results than the statistical model, it provides more accurate forecasting life expectation at birth (e_0) data series in Sudan during the period (2009 - 2020) for (both sexes- males - females) when data homogeneous. There are several measures that can be used for comparison between the two models (demographic model or statistical model) such as:

Mean Absolute Error, Rood Mean Square Error, Mean Absolute Percentage Error and Standard Error of Mean.

Tables (4 -32), (4 - 33), (4 - 34) below show the comparison between the two models when data are homogeneous according to some criteria mentioned we choose it.

Table (4-31): Forecasting of Life Expectation at Birth (e_0) by Using Demographic & Statistical Models for (Both Sexes- Males – Females) during period (2009 - 2020)

Year	Forecasting of(e_0) by Using Demographic Model for (Both Sexes- Males – Females)			Forecasting of (e_0) by Using Statistical Model for (Both Sexes- Males – Females)		
	(e_0) for Both Sex	(e_0) for Males	(e_0) for Females	(e_0) for Both Sex	(e_0) for Males	(e_0) for Females
2009	60.0	58.0	61.8	61.68	58.04	59.87
2010	60.3	58.2	62.3	61.97	58.13	60.06
2011	60.6	58.4	62.7	62.26	58.22	60.24
2012	60.9	58.5	63.1	62.55	58.30	60.41
2013	61.2	58.7	63.6	62.83	58.37	60.58
2014	61.5	58.9	64.0	63.14	58.45	60.75
2015	61.8	59.1	64.5	63.39	58.52	60.91
2016	62.1	59.2	64.9	63.67	58.58	61.07
2017	62.4	59.4	65.4	63.94	58.64	61.22
2018	62.8	59.6	65.8	64.22	58.70	61.37
2019	63.1	59.8	66.3	64.48	58.75	61.51
2020	63.4	60.0	66.8	64.75	58.80	61.65

Source: The researcher from applied study , E-views Package., 2015

Table (4 - 32): Comparison between Demographic and Statistical Models to Forecast (e_0): for Both Sexes

The Model	Mean Absolute Error	Rood Mean Square Error	Mean Absolute Percentage Error	Standard Error of Mean
Statistical Model(ARIM)	1.335640	1.659826	2.424594	0.29046
Demographic Model	1.350640	1.682725	2.447496	0.32219

Source: The researcher from applied study , E-views Package., 2015

Table (4 - 33): Comparison between Demographic and Statistical Models to Forecast (e_0): for Males

The Model	Mean Absolute Error	Rood Mean Square Error	Mean Absolute Percentage Error	Standard Error of Mean
Statistical Model(ARIMA)	0.090883	0.117835	0.170855	0.07200
Demographic Model	0.118830	0.143803	0.205840	0.18580

Source: The researcher from applied study , E-views Package., 2015

Table (4 – 34): Comparison between Demographic and Statistical Models to Forecast (e_0):for Females

The Model	Mean Absolute Error	Rood Mean Square Error	Mean Absolute Percentage Error	Standard Error of Mean
Statistical Model(ARIM)	0.117134	0.179754	0.215842	0.16837
Demographic Model	0.140134	0.203743	0.248851	0.46813

Source: The researcher from applied study , E-views Package., 2015

From the above tables(4 -32), (4 - 33), (4 – 34) we see that, the statistical model(ARIM) has minimum Mean Absolute Error, Rood Mean Square Error, Mean Absolute Percentage Error and Standard Error of Mean when the data are homogeneous. So the statistical model (Box-Jenkins) is the best to forecasting for life expectation at birth (e_0) for (both sexes – males - females) data series in Sudan during period (2009 – 2020) when data are homogeneous, While in the case of heterogeneous data, the demographic model gives better results than the statistical model.

Table (4 - 35): Life Expectation at Birth (e_0) by Years for (Both Sexes- Males- Females)(1973-2008)

Years	e_0 Both Sexes	e_0 Males	e_0 Female
1973	47.3	45.5	49.1
1974	47.9	46.2	49.6
1975	48.4	46.9	50.1
1976	49.0	47.6	50.6
1977	49.5	48.3	51.1
1978	50.1	49.0	51.6
1979	50.6	49.8	52.1
1980	51.2	50.5	52.6
1981	51.7	51.3	53.2
1982	52.3	52.0	53.7
1983	53.5	52.7	54.4
1984	53.7	53.0	54.5
1985	53.8	53.2	54.6
1986	54.0	53.5	54.7
1987	54.1	53.8	54.8

1988	54.3	54.0	54.9
1989	54.5	54.3	55.1
1990	54.6	54.6	55.2
1991	54.8	54.8	55.3
1992	55.0	55.1	55.4
1993	55.4	55.3	55.3
1994	55.7	55.5	55.7
1995	56.0	55.6	56.1
1996	56.2	55.8	56.5
1997	56.5	56.0	56.9
1998	56.8	56.1	57.3
1999	57.1	56.3	57.7
2000	57.4	56.5	58.1
2001	57.7	56.6	58.5
2002	57.9	56.8	58.9
2003	58.2	57.0	59.3
2004	58.5	57.2	59.7
2005	58.8	57.3	60.1
2006	59.1	57.5	60.5
2007	59.4	57.7	61.0
2008	59.8	58.1	61.4

Sources :The researcher from applied study (Sudan Life Table., 1973,1983,1993,2008)

Chapter Five

Conclusions and Recommendations

5 – 1 Preface:

In this chapter we will discuss the conclusion of findings and recommendations of the study.

5 – 2 Conclusion of Findings:

Through this study, we have reached many results which can be considered as basis for other researches in the same field and other related fields, and this can be identified in the following:

1- In general, the mortality data obtained from Sudan Population Censuses (1973 -1983 – 1993 – 2008) used to estimate the mortality indicators are found to be acceptable in quality, so indirect techniques are sound enough to reflect the mortality indicators.

2-The study found life expectation at birth (e_0) for Sudan Population Censuses (1973-1983-1993-2008), using demographic and statistical models, such as Brass technique methods and the (q_1) is estimated from Coale and Demeny model (the West region).

3-The study found levels, trends and patterns of mortality are based on the results obtained from life tables. The indicators which are used to measure the levels , trends and patterns of mortality in this study are: life expectation at birth(e_0) , life expectation at age 20 (e_{20}) ,infant mortality rate (q_1), under five mortality rate (q_5) and central death rate(nM_x) for both sexes, males and females during the period (1973 – 1983 – 1993 - 2008).

4- The study found that, the levels of mortality in Sudan for both sexes, for males and for females during the period (1973 – 1983 – 1993 - 2008) are considered to be high compared to the average levels in developed country.

5- The main findings regarding of mortality trends in Sudan during this period (1973 – 1983 – 1993 - 2008) are:

5-1 The life expectation at birth (e_0) increase by around 14 years for both sexes, 13 years for males and 12 years for females during 35 years (1973 - 2008) in Sudan.

5-2 There is no significant change observed in the life expectation at age 20 (e_{20}) only an increase by around 7 years for both sexes, for males and for females during this period (1973 - 2008) , which indicates a tendency of significant improvement in levels of the adult health conditions in Sudan.

5-3 The infant mortality rate (q_1) and the under five mortality rate (q_5) decreased during the period (1973 - 2008) which indicated that, there's steady improvement in infant and child health conditions.

6- The main findings of the study with respect to patterns of mortality in Sudan during the period (1973 – 1983 – 1993 -2008) are:

6-1 The central death rate (nM_x) by age groups takes the form of the letter (u). It was observed that, the age group (50-54) for both sexes of the 1983 census showed significant rise in central death rate which is likely to have occurred due to some mistakes made at any stage of data collection for that category, also there is a sharp rise in the central death rate for females in the age group (40-44) in the 1993.The curve fell particularly in age group (50 - 54) for males and (60 - 64) for females which it is nearly zero and then rose again in the last opened age group (75+), this pattern is familiar in the third world countries, but it was more severe in the 1993 population census. With respect to the central death rate by sex of death it is clear that, the risk of death is of more density among males than females in most age groups.

6-2 The infant mortality rate (q_1) and under five mortality rate (q_5) for males is higher than females except in the 1983 population census which increased the proportion of females than males.

6-3 Also it's observed that, the level of life expectation at birth (e_0) and life expectation at age 20 (e_{20}) is higher for females than for males in every census.

7- The study also compared the levels of mortality between the Population Censuses of Sudan (1973 - 1983 - 1993 - 2008),hence the study found that:

7-1 Life expectation at birth (e_0) and life expectation in the age group 20 (e_{20}) increased. Also, (q_1) (q_5) decreased during that period (1973-2008) except Population Census 1993 may be due to errors in any stage of data collection. We also notice a rise (q_5) in the population census 1993 compared to Population Census 1983 despite improved in health situation. Also (q_1) is higher among males than females in all population censuses except population census 1983.

7-2 The study also found the greatest improvement in life expectation at birth (e_0) and (e_{20}) occurred between the Population Censuses (1973 - 1983).

8- The study also calculated the forecast of life expectation at birth (e_0) (both sexes-males-females) for the years (2009-2020), using demographic and statistical models and the results were as follows:

For forecasting life expectation at birth (e_0) for the years (2009 -2020), the study applied several demographic criterias such as: {Annual Growth Rate(AGR $= e^{rt}$), Sudan standardization and the Spectrum program}.It has been found that, life expectation at birth (e_0) between (60.0 - 63.4) for both sexes,(58.0 - 60.0) for males and (61.8 - 66.8) for females during the period (2009 - 2020) .

9- The study has also applied Box – Jenkes time series model for life expectation at birth(e_0)for (both sexes, males and females) in the Sudan during the period (1973-2008) in order to forecast (e_0) during the period (2009-2020) and the results are as follows:

9-1 The life expectation at birth (e_0) in Sudan follows a non – stationary time series, for both sexes and for males but it follows stationary time series for females, these are decided by applying the values by Augmented Dickey-Fuller criteria, Autocorrelation Coefficient Function (ACF) and Partial Autocorrelation Coefficient Function (PACF).

9-2 The series of (e_0) for both sexes after second difference becomes stationary. The best model is ARIMA (0,2,1) to forecast (e_0) for both sexes and the value of (e_0) for both sexes, is ranging from (61.68 to 64.75) during the period (2009 - 2020).

9-3 The series of (e_0) for males after the first difference become stationary. The best model is ARIMA (1,1,0) to forecast (e_0) for males and the value of (e_0) for males is ranging from (58.04 to 58.80) during the period (2009 - 2020).

9-4 The series is stationary for females and the best model is ARMA (0,1) to forecast (e_0) for females and the value of (e_0) for females is ranging from (59.87 to 61.65) during the period (2009 - 2020).

10- While one objective of the study is to make a comparison between statistical and demographic models, and which is best to forecast the life expectation at birth (e_0) for (both sexes, males, females).The study used several criteria such as{ Mean Absolute Error, Rood Mean Square Error, Mean Absolute Percentage Error and Standard Error of Mean }, The study found that, statistical model (Box-Jenkins) is better than the demographic model in forecasting (e_0) when data are homogeneous.

5-3Recommendations of the Study:

From the above results, the study recommended:

1 - This study is considered to be one of the rare studies in Sudan which focuses on using statistical models , it contains rich and important demographic indicators , it must be given more attention and make use of its findings in all the relevant institutions such as the Ministry of Health, Ministry of Finance and Economics , insurance companies and other institutions.

2 – It is necessary to develop programs such as demographic packages (MORTPAK Package and CHADMOR Package) in order to apply the modern demographic models in a country like Sudan suffering from problems of lack and scarcity of data.

3-The necessity to increase the volume of data and eliminate all the discrepancies of time series data before the application to raise the efficiency of each of the models(Box -Jankins , demographic models) in analyzing life expectation at birth(e_0).

4 - More applications from other statistical models must be used in forecasting the series models of (e_0) in order to compare the results to

demographic models such as linear regression, Artificial Neural Networks (ANN) models .

5- It is necessary to conduct similar studies after a number of years to identify statistical or demographic models better than those used in the study.

6 - It is necessary to use both statistical and demographic models to forecast many of the other important demographic indicators.

7 - To conduct more broader and comprehensive studies with respect to the life expectation forecast and to formulate all state strategies and polices in accordance with these study .

8- The study left many questions about the demographic indicators, especially in the Population Census 1993, further studies are needed.

9- There should be a deeper study to find out the reasons for the rise (q_5) and decreased (e_{75+}) in the Population Census 1993 despite the improvement of health conditions.

10- (nM_x) in the (50 M_{54}) age group for both sexes in the 1983 census significantly increased. The study concluded that this finding is likely to have been an error at any stage of data collection also (q_1) is higher among males than females, except for the population 1983 census, a deeper study is needed to ascertain the cause.

11- Other studies should focus on comparing statistical and demographic models to forecast the life expectancy at birth when the data are heterogeneous using other models not used in the study.

12- A deeper study needs to be done to find out why there has been a significant improvement in (e_0)and(e_{20}) between the Population Censuses

(1973-1983).

13 - The study showed that, the difference between(e_0) and (e_4) of the Population Census 1993 requires a deeper study of the cause.

14 –The necessity of training for Central Bureau of Statistics (Sudan) staff on the importance of the data and how to make use of the Population and Housing Census data.(Awareness of the importance of complete vital records.)

15 – The use of research units manned by specialized statistical staff to carry out research in the field of time series demographic studies.

16- To conduct Future studies in the shape of mortality distribution by causes of death which can improve our understanding of the dynamics of mortality and its distribution.

17- Focuses on ensuring the provision of health care to the citizens of Sudan, especially poor and vulnerable population and also in reducing inequalities in access to health care, fostering community participation and developing health facility infrastructure.

REFERENCES:

1-Andersson, G., Philipov, D. (2002), " study, Life-table representations of family dynamics in Sweden, Hungary, and 14 other FFS countries", www.demographic-research.org.

2- Ediev D. (2003), "Study on Monotonic Convergence to Stability", www.demographic-research.org/Volumes/Vol8/2/.

3- Merrill, R. M. (2004), "Life Expectancy among LDSand Non-LDS in Utah",www.demographic-research.org/Volumes/Vol10/3/ 2004.

4- Vladimir M. Shkolnikov, Evgueni E. Andreev, Alexander Z. Begun. (2003), "Gini coefficient as a life table function: computation from discrete data, decomposition of differences and empirical examples", www. demographic- research.org/Volumes/Vol8/11/, 11-8-2003.

5- Maxim Finkelstein. (2005), "On stochastic comparisons of population densities and life expectancies", http://www.demographic-research.org / Volumes/Vo113/6/.,2005.

6- Joshua R. Goldstein.(2006),"A cohort perspective on tempo-adjusted life expectancies", http://www.demographic-research.org/ Volumes Vo114/5/.,2006.

7- Congdon P. *(*2006), "A model for geographical variation in health and total life expectancy",http://www.demographic- research.org/ Volumes/ Vol14/9/,Research Article.

8-Carlson E.(2006), "Ages of origin and destination for a difference in life expectancy", http://www.demographic- research.org/Volumes/ Vol14/ 11/., 2006

9- Beltrán- Sánchez H , Samuel H. Preston.(2008), “An integrated approach to cause-of-death analysis: cause-deleted life tables and decompositions of life expectancy”, http://www.demographic-research. org/ Volumes/ Vol19/ 35, . 2008.

10-Tymicki K.(2009), “Correlates of infant and childhood mortality: A theoretical overview and new evidence from the analysis of longitudinal data of the Bejsce (Poland) parish register reconstitution study of the 18th-20th centuries”, VOLUME 20, ARTICLE 23, PAGES 559-594 .

11- Cai L , Mark D.Hayward, Saito A , Lubitz J, Hagedorn A , Crimmins E.(2010), “Estimation of multi-state life table functions and their variability from complex survey data using the SPACE Program”, http:// www.demographic-research.org/Volumes/Vol22/6/ .,2010.

12- Canudas.(2008), “The modal age at death and the shifting mortality hypothesis, Vladimir”.

13- Lee R, Lawrence R. Carter. (1992), "Modeling and Forecasting U.S. Mortality.

14- Vladimir Canudas-Romo.(2008) , “What can we learn from indirect estimations on mortality in Mongolia, 1969-1989?”, http://www. demographic- research.org/Volumes/Vol18/10/.,2008

15- Waddad A.(2010), “Combined Maternal Mortality Ratio and Orphan hood Technique to estimate Maternal Mortality logit System From Sudan 2008 Census data”, Ministry of the Cabinet Central Bureau of Statistics.

16- Abu El-Yamen S, Musafa I, Ibrahim S.(2010), “Analysis of Mortality Levels Trends and differentials Census Data” ,Ministry of the Cabinet Central Bureau of Statistics.

17- TuljapurkarS , Ryan D. Edwards. (2007) ,”Variance in Death and Mortality Decline”.

18- Carter .Rlawrence. (2008),"Forecasting U.S. Mortality: AComparison of Box-Jenkins ARIMA and Structural Time Series Models".

19- Nan Li and Patrick Gerland. (2012), "The Lee Carter method for estimating and forecasting of mortality: an application for Argentina" , National University of Rosario, Argentina , Faculty of Economics and Statistics.

20- Smart A. Sarpong.(2013), "Modeling and Forecasting Maternal Mortality; an Application of ARIMA Models" ,International Journal of Applied Science and Technology Vol. 3 No. 1., January 2013.

21- Adhistya Erna Permanasari at el .(2009), "Prediction of Zoonosis Incidence in Human using Seasonal Auto Regressive Integrated Moving Average (SARIMA), (IJCSIS)", International Journal of Computer Science and Information Security, Vol. 5, No. 1, 2009 .

22- Smart A. Sarpong.(2013), "Modeling and Forecasting Maternal Mortality; an Application of ARIMA Models", School of Graduate Studies, Research and Innovations Kumasi Polytechnic –Ghana.

23- Ekezie Dan Dan, Opara Jude, Okenwe Idochi.(2014), "Modelling and Forecasting Malaria Mortality Rate using SARIMA Models", A Case Study of Aboh Mbaise General Hospital, Imo State Nigeria, Science

Journal of Applied Mathematics and Statistics. Vol. 2, No. 1, 2014, pp. 31-41. Doi.

24- Musa Mohammed.(2013), "Malaria Disease Distribution in Sudan Using Time Series ARIMA Model" .

25- Ahmed I. Abd elghani at el. (2013)"Predicting Cerebrospinal Meningitis in Sudan Using Time Series Molding", Sudanese Journal of Public Health VOL. 8 No. 2., April 2013.

26- C.J.L. Murray et al.(2011), "GPE Discussion Paper Series: No. 8EIP/GPE/EBD, World Health Organization".

27- Brass W.(1971), "On the Scale of Mortality, in W. Brass (ed.)", Demography London :Taylor & Fracis.

28- Scand J.(2007), “Public Health Suppl ,Mortality trends in a new South Africa: Hard to make a fresh start”.

29- Orlaith Burke.(2011), "Statistical Methods Autocorrelation Box-Jenkins Modeling Introduction", University of Oxford, Department of Statistics.

30- Jack Johnston& John Dinardo," Econometric Methods, Four Edition, Univariate Time Series Moddling",http :// www.mh college.com .

31- Coale A and Demeny P (1966). "Regional Life Tables Model and Stable Populations. Princeton": Princeton University Press.

32- Brass, W.(1971),"On the scale of mortality. In W. Brass (ed.)". Biological Aspects of Demography. London: Taylor and Francis.

33- United Nations. (1981), "Life Tables Model for Developing Countries", New York: United Nations. (Sales No. E.81.XIII.7).

34- Lopez AD, Ahmad OB, Guillot M, Inoue M, Ferguson BD.(2001), "Life Tables for 191 .Countries for 2000: Data, Methods and Results", Geneva, World Health Organization (GPE Discussion Paper No. 40).

35- Bartlett, M. S. (l956),"An Introduction to Stochastic Processes", Cambridge University Press.

36- Aa1en, O. (1976),"Nonparametric inference in connection with multiple decrement modes",Scandel.J.Stat.,3,15-27.

37- Ad1akha, A. (1972)," Hodel life tables: an empirical test of their applicability to less developed countries", Demography ,9,589 – 601.

38- Coale. Ansley J: Coale.(July 1991), "A patterns of marriage, Population Studies", vol. XXV, No. 2. pp. 193-114.

39- Maml 111.(1998),"Methods fw Population Ro'rrriom by Sex and Age" ,United Nations publication. Sales ~0.56.~111.3.

40- Colin Newell(1988 – 1989), "Methods and Models in Demography", Belhaven Press a division o Printer Publishers, London.

41- National Statistics (ONS).(January 2012) ,"Quality and Methodology Information", lifetables@ons.gov.uk.

42- Bongaarts, J. (2004)," Six Period measures of longevity", Unpublished manuscript.

43- Bongaarts, J., & Feeney, G. (2002)," How long do we live?" Population and Development Review, 28(1), 13-29.

44- Bongaarts, J., & Feeney, G. (2003),"Estimating mean lifetime", Proceedings of the National Academy of Sciences, 100(23), 13127-13133.

45- Brouard, N. (1986),"Structure et dynamique des populations. La pyramide des ann´ees `a vivre, aspects nationaux et exemples r´egionaux", Espaces, Populations, Soci´et´es, 2(14-5), 157-68.

46- Goldstein, J. R., & Wachter, K. W. (2004),"Gaps and lags: Relationships between cohort and period life expectancy", Unpublished manuscript presented at the annual meeting of the Population Association of America, Boston.

47- Eric V. Slud. (2006),"Actuarial Mathematics and Life-Table Statistics", Mathematics Department University of Maryland, College Park.

48- Arias, E. (2007),"United States Life Tables, 2004", National Vital Statistics Reports 56(9). Hyattsville, MD: National Center for Health Statistics. http://www.cdc.gov/ nchs/data/nvsr/nvsr56/nvsr56_09.pdf.

49- Keyfitz, N. (1968),"Introduction to the Mathematics of Population", Reading, MA: Addison-Wesley.

50- Vladimir M. Shkolnikov, Evgueni E. Andreev, Alexander Z. Begun. (17 JUNE 2003), "Gini coefficient as a life table function: computation from discrete data, decomposition of differences and empirical examples", VOLUME 8, ARTICLE 11, PAGES 305-358 Published, www.demographic-research.org/Volumes/Vol8/11/DOI: 10.4054/DemRes.2003.8.11.

51- Joel E. Cohen. (22 JANUARY 2010)," Life expectancy is the death-weighted average of the reciprocal of the survival-specific force of mortality", VOLUME 22, ARTICLE 5, PAGES 115-128 Published, http://www.demographic-research.org/Volumes/Vol22/5/DOI: 10.4054/DemRes.2010.22.5 Formal Relationships 7.

52- Chien Long Chlang. (2003), "Life Table and Mortality Analysis", World Health Organization .

53- Joshua R. Goldstein. (2006),"Found in Translation? A cohort perspective on Tempo-adjusted life expectancy", Volume 14, Article 5, Pages 71-84 Published 10 February 2006, http:// www. demographic-research.org/Volumes/Vol14/5/DOI: 10.4054/ Dem Res.2006.14.5.

54- Maxim Finkelstein.(2005)," On stochastic comparisons of population densities and life expectancies", Volume 13, Article 6, Pages 143-162 Published 19 October 2005, http://www.demographic-research. org/ Volumes/ Vol13/6/DOI: 10.4054/Dem Res.2005.13.6.

55- Anastasia Kostaki, Vangelis Panousis. (2001), "Expanding an abridged life table", Volume 5, Article 1, Pages 1-22, Published 19 September 2001,www.demographic-research.org/Volumes/Vol5/1/DOI: 10.4054.

56- Alfred J et al. (1972),"having been subject to unchanging fertility and mortality for many years", establishes its value. Alfred J. Lotka and F. R. Sha r,A problem in age distribution.

57-Chiang, C. L. (1960), "A stochastic study of the life table and its applications:1. Probability distributions of the biometric functions. Biometrics", 16, 618-635.

58- Ad1akha, A. (1972)," Hodel life tables: an empirical test of their applicability to less developed countries", Demography9, 589-6~H.

59- UN Population Division. (2010),"World Population Prospects 2012: Extended Model Life Tables", New York: United Nations, Department of Economic and Social Affairs. http:// esa.un . org/wpp/Model-Life-Tables/data/MLT_UN2010-130_1y.xls

60-C.J.L. Murrary et al.(2010),"GPE Discussion Paper Series",:No8,EIP / GPE / EBD World Health Organization

61- Toson, B., Baker, A., and the Office of National Statistics. (20013), "Life expectancy at birth: Methodological options for small populations", *National Statistics Methodological Series* 20013(33): 1-27. http:// www.statistics.gov.uk/downloads /theme_ other/ GSS Methodology _No_ 33.pdf.

62- Williams, E., Dinsdale, H., Eayres, D., and Tahzib, F. (2005)," Technical Report: Calculating Life Expectancy in small areas",South East Public Health Observatory.

63- Bongaarts, J. (2005),"Long-range trends in adult mortality: Models and projection methods",*Demography* 42(1): 23-49. doi:10.1353/ dem. 2005.0003.

64- Bongaarts, J. (2009),"Trends in senescent life expectancy", *Population Studies* 63(3): 203-213. doi:10.1080/00324720903165456.

65- Brass, W. (1971),"On the scale of mortality", In: Brass, W. (ed.). Biological Aspects of Demography. London: Taylor and Francis: 69-110.

66- Merrill RM, Hilton SC, Daniels M. (2003),“Impact of the LDS Church’s health doctrine on deaths from diseases and conditions attributed to cigarette smoking.” *Ann Epidemiol*, 13 .

67- CHIN Long Chiang .(2007**),"**Life Table and Mortality Analysis", World Health Organization.

68- Aa1en, O.(2003),"Nonparametric inference in connection with multiple decrement models", Scand. J. Stat., 3, 15 -27.

69- Anderson, T.W. (1971), *"The Statistical Analysis of Time Series*", New York: John Wiley & Sons, Inc.

70- Box, G.E.P., Jenkins, G.M., and Reinsel, G.C. (1994), "*Time Series Analysis: Forecasting and Control"*, Third Edition, Englewood Cliffs, NJ: Prentice Hall, 197-199.

71- Orlaith Burke.(2011)," Statistical Methods Autocorrelation- Box-Jenkins Modeling Introduction", Department of Statistics, 1 South Parks Road, Oxford OX1 3TG.

72- Kirchgassner G. and Wolters J. (2007) "Introduction to Modern Time Series Analysis", SPRINGER-Verlag, Berlin Heidelberg.

73- YAFFEE R. and McGee M. (1999) "Introduction to Time Series Analysis and Forecasting". Academic Press S, INC, New York.

74- Box Jankins.(1976)," Time Series Analysis Forecasting and Control Holden Day. This covers Both Theory and practices if not essential, for practicing forecasting ".

75- John Frain.(January 1999), " Lecture Notes on Univariate Time Series Analysis and Box Jankins Forecasting ", Economic Analysis, Research and Publications.

76- Coale A and Demeny P (1966). Regional Model Life Tables and Stable Populations. Princeton: Princeton University Press.

77- United Nations (1981). Model Life Tables for Developing Countries. New York: United Nations. (Sales No. E.81.XIII.7).

78- Colin Newell. (July 1987),"Methods and Models in Demography".

79- Al-Saleh, M.F. (2004), "Steady-state ranked set sampling and parametric inference", *Journal of Statistical Planning and Inference*, 123, 83-95. C. Chatfield, “Model uncertainty and forecast accuracy”, J. Forecasting 15 (1996), pages: 495–508.

80- Bowerman, Bruce L., Richard T. O’Connell, and Anne B. Koehler. (2005), "Forecasting, Time Series, and Regression", 4th ed. Belmont, CA: Thomson Brooks/Cole.

81- Smart A. Sarpong.(2013), "Modeling and Forecasting Maternal Mortality; an Application of ARIMA Models" , *International Journal of Applied Science and Technology Vol. 3 No. 1; January 2013*

82- Coale, A., and P. Demeny. (1983), *"Regional Model Life Tables and Stable Populations"*, New York: Academic Press.

83- United Nations. (2005), *"World Population Prospects. The 2004 Revision"*, New York: United Nations.

84- World Bank.(2005), *"World Development Indicators.* Washington", DC: World Bank.

85- Nan Li and Patrick Gerland, (2011)"Modifying the Lee-Carter method to project mortality changes up to 2100", Session 125: Formal Demography I: Mathematical Models and Methods (Friday 1 April 2011, 12:30 PM - 2:20 PM) Latest revision 6 April 2011 , United Nations Population Division. Department for Economic and Social Affairs, United Nations, New York, NY 10017, USA.

86- International Journal of Science and Research (IJSR) ISSN.(2013) (Online) (2013): 2319-7064.Index Copernicus Value (2013): 6.14 Impact Factor (2013): 4.438 Volume 4 Issue 2, February 2015 www.ijsr.

Sources of Indicators Data:

1- Department of Statistics, the Second Population Census of Sudan., 1973.

2- Department of Statistics, the Third Population Census of Sudan., 1983.

3- Department of Statistics, the Fourth Population Census of Sudan., 1993.

4 - Central Bureau of Statistics, Sudan & UNPFA, Base Line on Surveys.

5-Department of Statistics, the Fifth Population Census of Sudan.,2008.

5 -Demographic & Health surveys.