ANALYSIS OF MATERNAL MORTALITY WITH TIME; A CASE STUDY OF THE OKOMFO ANOKYE TEACHING HOSPITAL – KUMASI (2000-2010)

By

Sarpong Smart Asomaning (BSc. Mathematics and Statistics)

A Thesis submitted to the Department of Mathematics, Kwame Nkrumah University of Science and Technology in partial fulfilment of the requirements for the degree of

MASTER OF PHILOSOPHY

Supervisor

Nana Kena Frempong (Lecturer)

College of Science

May, 2012

DECLARATION

I hereby declare that this submission is my own work towards the MPhil and that, to the best of my knowledge, it contains no material previously published by another person nor material which has been accepted for the award of any other degree of the University, except where due acknowledgement has been made in the text.

SARPONG, SMART ASOMANING _____ _____

(PG5137110) SIGNATURE DATE

Certified by:

NANA KENA FREMPONG _____ _____

(SUPERVISOR) SIGNATURE DATE

Certified by:

F. K. DARKWA _____ _____

(HEAD OF DEPARTMENT) SIGNATURE DATE

DEDICATION

This thesis is dedicated to my Mom Sophia Afua Bassanyin, Dad Emmanuel Agyei Sarpong, Brothers; Frank, Fred, Prince, Lord, Sammy and Maame Abena Kyeraah Brobbey my best friend. These and many other persons have been and continue to be a blessing to my life. God Bless you all richly... AMEN

ABSTRACT

This study examined the occurrence and incidence of Maternal Deaths as well as maternal mortality ratios at the Okomfo Anokye Teaching Hospital in Kumasi from 2000 to 2010. The study explored the feasibility for application of Poisson models and time series autoregressive integrated moving average (ARIMA) in the study of occurrence and incidence of Maternal Deaths and to predict Maternal Mortality ratios respectively. Analyses were based on data available at the Bio-Statistics Department of the Obstetrics & Gynaecology directorate of the Okomfo Anokye Teaching Hospital in Kumasi for the period 2000-2010. The Statistical Analysis software (SAS) as well as the R-consol statistical analysis software was used in analysing the data. We found that, the mean number of occurrence of maternal death cases were high for all the years considered and established that the mean number of occurrence of maternal death cases has not significantly reduced over the period 2000 to 2010. The result also shows that there was a statistically significant in the incidence of maternal deaths difference between year 2010 (the referenced year) and years 2004, 2005 and 2008. Their chi-square values were 3.95, 5.12 and 5.83 with p-values of 0.0469, 0.0236 and 0.0158 respectively. Finally, the hospitals Maternal Mortality Ratio (MMR) is relatively stable but has a very high average MMR of 967.7 per 100,000 live births which is about twice the National ratio of 451 per 100,000 live births. An ARIMA model fitted was used to predict maternal mortality ratios (MMRs) for the next eight quarters. We conclude that statistically the mean rate of maternal death cases is not significant over the period of time under study, mean number of occurrence of maternal death cases has not significantly reduced over the period 2000 to 2010 and that the ARIMA model is adequate for forecasting quarterly maternal mortality ratios at the hospital.

Key words: ARIMA, Poisson Regression Model, Bio-statistics, MMR

TABLE OF CONTENTS

CHAPTER 3: METHODOLOGY

CHAPTER 4: RESULTS AND DISCUSSION

CHAPTER 5: CONCLUSION AND RECOMMENDATION

APPENDIX

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENT

"For I am the Lord, your God, who takes hold of your right hand and says to you,

Do not fear; I will help you (Isaiah 41:13)"

Truly God Almighty holds my hand throughout all these years and helps me. I would therefore like to express my sincere gratitude to Him for seeing me through this MPhil programme. My special thanks also go to my dependable supervisor Nana Kena Frempong for offering directions, suggestions, and encouragement during the entire duration of this thesis. I am also grateful to the Gynaecology and Obstetrics Directorate of the Okomfo Anokye Teaching Hospital-Kumasi especially Dr. Morhea and Miss Irene O. Attrams for their enthusiastic and technical support. To the Heads of Department of Mathematics (KNUST) and (Kumasi Polytechnic), Mr. F.K. Adarkwa and Dr. Z. K. M. Batse respectively, I say thank you for your guidance and comments which have help shaped many aspects of this work. I am grateful to all lecturers especially Mr. Emmanuel Harris and the entire staff of the Department Mathematics KNUST for the knowledge imparted unto me and for making the process of Higher learning a pleasant one. I reserve this heartfelt gratitude to Mr. Eric Aidoo, of CSIR - Kumasi for his timely contributions to this thesis. I owe a debt of gratitude to all my class mates, most especially Wahab, Omane, Ababio and his wife Rita, Alex, Olivia, Agyemang, James, Garkus and all friends and family who in one way or the other contributed towards the successful completion of this MPhil.

CHAPTER 1

1.0 INTRODUCTION

This study seeks to examine the occurrence and incidence of Maternal Deaths as well as maternal mortality ratios at the Okomfo Anokye Teaching Hospital in Kumasi from 2000 to 2010. The study explore the feasibility for application of Poisson models and time series ARIMA in the study of occurrence and incidence of Maternal Deaths and to predict Maternal Mortality ratios respectively so as to provide the theoretical basis for continuing programs to reduce this problem. This chapter takes a look at the background of the study, the general socio-economic profile of the study area. The problem statement, research questions and objectives, research methodology, justification of the study as well as scope and limitations of the study are also discussed.

1.1 BACKGROUND OF THE STUDY

Global attention began to focus more seriously on maternal mortality when in 1985; Rosenfield and Maine (1985) published a thought-provoking article in the Lancet. In this classic article titled 'Maternal Mortality - a neglected tragedy - where is the M in MCH?', Rosenfield and Maine alerted the world to the fact that many developing countries were neglecting this important problem and that existing programs were unlikely to reduce the high maternal mortality rates in the developing world. Another significant contribution to the crusade against maternal mortality was the WHO (1986) publication, 'Maternal Mortality: helping women off the road to death.'

All these led to the Safe Motherhood Conference in Nairobi, Kenya in 1987. Speakers at this conference presented global statistics on death and complications resulting from

pregnancy. They also showed that in sub-Saharan African, the lifetime risk that a woman would die in childbirth is 1 in 21 and that this is 400 times higher than the lifetime risk for her counterpart in Western Europe or North America. The conference concluded with strong recommendations about maternal health and that was when the Safe Motherhood Initiative was born.

The reduction of maternal mortality by half the 1990 levels by the year 2000 was a goal common to several of such conferences including in particular, the 1990 World Summit for Children, the 1994 Cairo International Conference on Population and Development and the 1995 Fourth World Conference on Women. At the Millennium Summit in September, 2000, the largest gathering of world leaders in history adopted the United Nations (UN) Millennium Declaration, committing their nations to a new global partnership aimed at eradicating extreme poverty and hunger, achieving universal primary education, promoting gender equality and empowerment of women, reducing child mortality, improving maternal health, combating HIV/AIDS, malaria and other diseases and ensuring environmental sustainability. These targets were all given a deadline of 2015 and have since been known and referred to as the Millennium Development Goals (MDGs).

Globally, 529,000 women die each year from pregnancy-related complications, of which about 90% occur in developing countries, the worst affected being West Africa, including Ghana (UN Millennium Project, 2006). In Africa 1 out of 16 women stand the risk of dying through pregnancy and childbirth. The risk of maternal deaths is highest in Africa, where countries struggle to provide health services for large number of its populations.

The World Health Organization (WHO) states that worldwide 1500 women die each day, or one a minute, in pregnancy or due to childbirth related complications. It is estimated that over half of these deaths are in sub-Saharan Africa, with maternal mortality ratio of 910 deaths per 100,000 live births (WHO, 2006). Maternal haemorrhage, obstructed labour,

postpartum sepsis, eclampsia, unsafe abortion and anaemia are among the leading causes of death among pregnant women in developing countries. These complications of pregnancy contribute significantly to the high levels of maternal and neonatal mortality in Sub-Saharan Africa.

In addition to maternal deaths, millions more women suffer from near death complications and long-term disabilities as a result of pregnancy-related complications which also affect the lives of numerous babies. Contributory factors include lack of access to good quality maternal and neonatal health services and strong adherence to negative cultural beliefs and practices (AbouZahr & Wardlaw, 2001 and WHO, 2005). Globally, there is increasing evidence that reduction of maternal deaths is achievable with the timely provision of quality emergency obstetric care (EmOC). The challenge therefore is to focus and concentrate on improving efficient and timely delivery of emergency obstetric care. Studies have shown that most life-threatening obstetric complications cannot be predicted or prevented but can be successfully treated if prompt access to quality Emergency Obstetric services and skilled attendance are available (UN Millennium Project, 2005).

In recent years, Ghana's maternal mortality ratio (MMR) declined from 560 deaths out of 100,000 live births in 2005, to 451 deaths in 2007 (WHO, 2005; Ghana DHS, 2007). However, statistics from Kumasi show that maternal deaths then began rising. In 2007, Kumasi's MMR was 359 out of 100,000 live births while in 2008 it was 397 out of 100,000 live births (KMHD, 2009). The majority of maternal deaths in Kumasi (about 93%) occurred at KATH, most likely because this hospital is the referral hospital for complicated medical emergencies. The table below shows Maternal Mortality Ratios in Ashanti and Kumasi.

Table 1.1 Maternal Mortality Ratios in Ashanti and Kumasi

YEAR	2005	2006	2007	2008	2009	2010 (Projected)

KUMASI	395	383	359	397	332	273
ASHANTI REGION	**200**	**208**	**246**	**253**	**210**	

Source: World Health Organization (2007), Ghana Maternal Health Survey (2007), GHS Ashanti Region Annual Report (2006), KMHD (2009), KMHD (2010a), KMHD (2010b).

1.1.1 PROFILE OF STUDY AREA

- **Kumasi Metropolitan Assembly (KMA)**

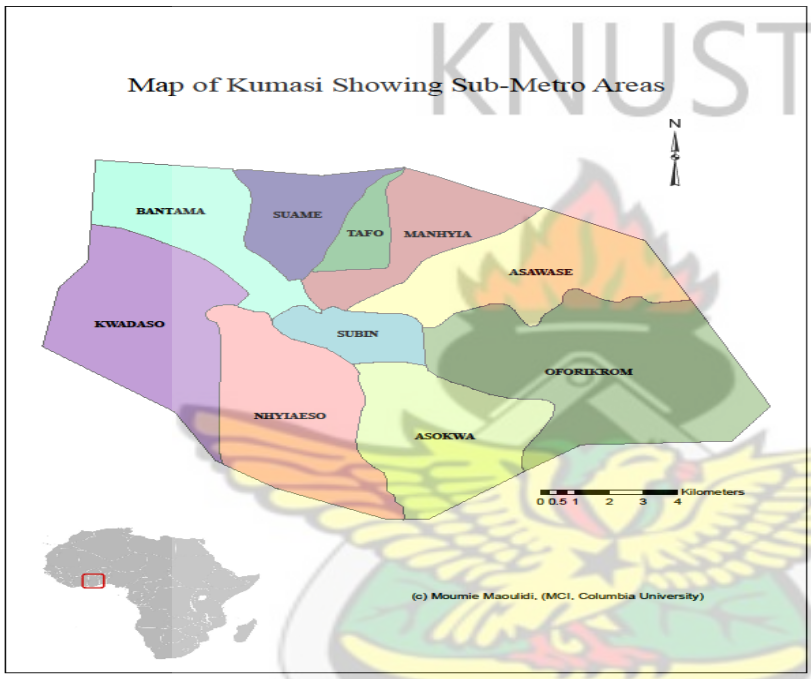


Figure 1.1: Map of Kumasi showing sub-metro areas *(Source: Moumie Maoulidi, Millennium City Initiative, Columbia University)*

The Kumasi metropolis is the most populous district in the Ashanti Region. The city accounts for nearly one-third of the region's population and is divided, in turn, into 10 sub-metropolitan areas, as shown in the Figure above. It has a population of 1,170,270, a total surface area of 254 sq km, and a population density of 5,419 persons per sq. km (2000 population census). It has been projected to have a population of 1,625,180 in 2006 based on a growth rate of 5.4% per annum. The Asantehene is the embodiment of the culture of the people. Ashanti has thirty-three (33) Traditional Council areas. A Paramount Chief heads each council. All the Paramount chiefs also owe allegiance to Otumfuo, the Asantehene who is the

head of the Asanteman Council. Otumfuo has His Royal Palace at Manhyia in the heart of Kumasi.

- **Major Economic Activities**

Although many people in Kumasi are engaged in a form of employment, either with the private or public sector, about 60 per cent of residents still have a lower standard of living resulting from low incomes. The famous Suame Magazine where small engineering based industries are sited contributes immensely to the economy of the metropolis. The woodworking business at Sokoban produce to meet the needs of residents as well as clients from Accra and other parts of the country and neighbouring countries of Burkina Faso, Ivory Coast and Mali. Agriculture is practised on a limited scale. Crop farming is along valleys of rivers and streams that traverse the metropolis. It is also carried out in open backyards and in the peri-urban areas as well as animal production in sheep/goats, cattle, poultry and fish farming.

- **Health Care Delivery**

The Metropolitan Health Services are organized around five (5) Sub Metro Health Teams; namely, Bantama, Asokwa, Manhyia North, Manhyia South and Subin. Komfo Anokye Teaching Hospital (KATH) is an autonomous facility. Kumasi Metro Health Directorate (KMHD) oversees all sub-metro district hospitals. In addition to these sub-metro district hospitals, there are also quasi-governmental, private and mission hospitals, which offer similar levels of care to the government hospitals. There are also several clinics, maternity homes, laboratories and other health care providers. There are also 122 outreach stations in Kumasi, located throughout the five sub-metro areas (KMHD, 2008). Hence, most Kumasi residents are geographically situated well within the vicinity of a health care facility.

- **Komfo Anokye Teaching Hospital**

The Komfo Anokye Teaching Hospital (KATH) in Kumasi is the second-largest hospital in the country and the only tertiary health institution in the Ashanti Region. It is the main referral hospital for the Ashanti, Brong Ahafo, Northern, Upper East and Upper West Regions. The hospital was built in 1954 as the Kumasi Central Hospital. It was later named Komfo Anokye Hospital after Okomfo Anokye, a legendary fetish priest of the Ashanti. It was converted into a teaching hospital in 1975 affiliated to the medical school of the Kwame Nkrumah University of Science and Technology. The hospital is also accredited for postgraduate training by the West African College of Surgeons in surgery, obstetrics and gynaecology, otorhinolaryingology, ophthalmology and radiology. The hospital currently has about 1000 beds, up from the initial 500 when first built. The facility has eleven (11) Clinical Directorates including; Anaesthesia and Intensive Care Unit (ICU), Child Health, Dental-Eye-Ear-Nose and Throat (DEENT), Diagnostics, Medicine, Obstetrics & Gynaecology, Oncology, Polyclinic, Surgery, Accident and Emergency department and Pharmacy. The Non-Clinical Directorates includes the Domestic Services, Security, Supply Chain Management and Technical Services.

The Obstetrics and Gynaecology directorate has six wards of which three are designated as labour wards. The three labour wards are general, official and high risk wards. The wards in the directorate have a bed capacity of 160. The average turnover per bed is 118. The average length of stay in the ward is 4 days with a bed occupancy rate of 132%. Specialist clinics in asthma in pregnancy, sickle cell clinic and diabetes mellitus commenced in 2008. Stringent quality assurance measures have been adopted to ensure the discharge of uncomplicated major surgical cases within five days.

The Komfo Anokye Teaching Hospital (KATH) as a policy requires audits to be conducted on all maternal deaths occurring within the hospital. It includes the examination of the standards of care given, parity and age of deceased, gestational age and the cause of death. A

total of 27 maternal deaths were examined in 2008. The commonest causes of death include Eclampsia, Malignant of cervix and Post partum haemorrhage. Majority of these deaths occurred less than 24 hours upon arrival in KATH. Family Planning services are comprehensively organised by the Family Planning unit of the directorate. (KATH official website)

1.2 PROBLEM STATEMENT

Maternal mortality is one of the most sensitive indicators of the health disparity between richer and poorer nations. The lifetime risk of dying due to maternal causes is about one in six in the poorest countries, compared with about one in 30,000 in Northern Europe (Ronsmans and Graham, 2006). Every minute, a woman dies from complications of pregnancy and childbirth. Available record indicates that about 530,000 deaths per year worldwide are caused by complications associated with pregnancy and childbirth. In Africa 1 out 16 women stand the risk of dying through pregnancy and childbirth. Ghana's maternal mortality ratio (MMR) declined from 560 deaths out of 100,000 live births in 2005, to 451 deaths in 2007 but this rate is still too high (WHO, 2005; Ghana DHS, 2007). Statistics from Kumasi show that maternal deaths are rising. In 2007, Kumasi's MMR was 359 out of 100,000 live births, while in 2008 it was 397 out of 100,000 (KMHD, 2009).

The national target is to reduce maternal mortality rate to 54 per 100,000 live births by 2015. It is observed by most health professionals and institutions, that Ghana, like many countries, is off-track with respect to this goal. The target, according to them, is not likely to be achieved within the expected period. However, there has not been any statistical evidence to justify this accession or otherwise especially in the Okomfo Anokye Teaching Hospital. Again, very little studies have gone into mortality patterns in Okomfo Anokye Teaching Hospital and

the Kumasi Metropolitan Assembly especially that relating to the death of pregnant women? Also, questions as to what statistical model would be reliable for a comprehensive study of Maternal Mortality incidence in the facility needs to be made evident. It is against this background that this study is being undertaken to assess the pattern and of Maternal Mortality and to predict future incidence as well future maternal mortality rates.

1.3 RESEARCH QUESTIONS

- What is the pattern of maternal mortality ratios (MMR) in the Okomfo Anokye Teaching Hospital Kumasi?
- To what extent can Okomfo Anokye Teaching Hospital as a facility achieve a reduction in the Maternal Mortality ratio?
- What is the significance of the incidence and occurrence of maternal mortality of the facility?

1.4 RESEARCH OBJECTIVES

Generally, the study aims examining the occurrence and incidence of Maternal Deaths as well as the patterns of maternal mortality ratios at the Okomfo Anokye Teaching Hospital, Kumasi in line with the Millennium Development Goal five (5) using Poisson Regression model and ARIMA model respectively.

1.4.1 SPECIFIC OBJECTIVE

The specific objectives which the study shall focus on are as follows:

- To examine the significance of the incidence and occurrence of maternal mortality over the period of time under study using Poisson Regression models,

- To examine the pattern of maternal mortality ratios (MMR's) using time series plot
- To assume a Time Series ARIMA model to forecast MMR's over eight (8) quarters.

1.5 METHODOLOGY

Time series ARIMA models would be adopted for the estimation of patterns and predictions of future maternal mortality ratios. ARIMA models are, in theory, the most general class of models for forecasting a time series which can be stationarized by transformations such as differencing and logging. Poisson Regression shall also be used to examine the significance of the incidence and occurrence of maternal mortality over the period of time under study. Poisson regression is a form of a generalized linear model where the response variable is modelled as having a Poisson distribution. Poisson regression models are appropriate when the response variable is count data. Counts are all positive integers and for rare events the Poisson distribution (rather than the Normal) is more appropriate since the Poisson mean > 0.

1.5.1 DATA

The analyses would be based on data available at the Bio-Statistics Department of the Obstetrics & Gynaecology directorate of the Okomfo Anokye Teaching Hospital in Kumasi for the period 2000-2010. The required data are;

- Quarterly recorded maternal mortally ratios (MMR)
- Monthly recorded maternal deaths
- Monthly recorded deliveries

1.5.2 SOFTWARE TO BE USED

The Statistical Analysis software (SAS) would be used in examining the occurrence and incidence of maternal mortality and fitting the Poisson model while the R-Consol statistical software would be used in analysing and fitting ARIMA models. Other statistical and numerical simulation methods of parameter estimation would be utilized as and when deemed fit.

1.5.3 SOURCE OF KNOWLEDGE

The main source of knowledge for the successful completion of this study would be the Okomfo Anokye Teaching Hospital Library, the Kwame Nkrumah University of Science and Technology Library and the Ministry of Health Research Centre – Accra. However, the internet and other obstetrics and gynaecology professionals have and would continue to help enrich the progress and outcome of the study

1.6 JUSTIFICATION

According to WHO, Ghana has one of highest maternal mortality rates in the West African region, 540 deaths per 100,000 live births (WHO, 2007). Kumasi's challenges in the public health arena remain formidable. In addition, data from 2007 - 2008 for Kumasi show that maternal and infant deaths were on the rise and diseases such as malaria and tuberculosis continued to be major causes of morbidity. However, on a positive note, Maternal Mortality Ratio fell significantly in 2009, and projections for 2010 show that it will most likely continue to decline (KMHD, 2010). Selection of the maternal mortality ratio (MMR) as the primary indicator for Millennium Development Goal number 5 (MDG-5) on improving maternal health

has increased interest in programs to improve maternal health and in having reliable sources of data on maternal mortality rates.

The study would therefore provide the needed statistical evidence to justify the success or failure of the set target with respect to the Okomfo Anokye Teaching Hospital. The forecast models to be generated would also be useful for predicting future maternal deaths at Okomfo Anokye Teaching Hospital and the Kumasi Metropolitan Area.

On the basis of available empirical evidence, this study would seek to furnish decision makers and other stakeholders with vital information regarding the trend of maternal mortality in the facility for possible policy interventions. Additionally, this study would also contribute to knowledge on the use of Poisson Regression and ARIMA Models in the area of maternal health and other related areas with a view to, among other things, stimulating further research.

1.7 SCOPE AND LIMITATION

The scope of this study is restricted to the Okomfo Anokye Teaching Hospital in Kumasi. It is the second-largest hospital in the country and the only tertiary health institution in the Ashanti Region. It is the main referral hospital for the Ashanti, Brong Ahafo, Northern, Upper East and Upper West Regions. Particularly in the Kumasi Metropolitan Area, this facility record a high number of all pregnancy related cases and has an Obstetrics & Gynaecology department that takes care of maternal health issues in and around Kumasi. This may be a limitation to the study. The unavailability of 2011 quarterly data which could have been used for verification of ARIMA model is also a limitation to the study. It is also envisage that, the study will suffer constraints of time, resource inadequacy, unavailability of relevant literature and others of the like. The study would be structured within the confines of the thesis study matter.

1.8 THESIS ORGANIZATION

This report is organised in five chapters. Chapter 1 is the introductory chapter to the entire study. It takes a critical look at the general background of maternal mortality as well as maternal health and also looks at the general socio-economic profile of the study area. The problem statement, research questions and objectives, research methodology, justification of the study as well as scope and limitations of the study are discussed in this chapter. Chapter 2 reviews related literature based on the thesis objectives and preferred models to be used in achieving these objectives. Expected outcome of the study and other comparative results of similar studies are also discussed in this chapter. Chapter 3 describe the theory of history model to be used, formulations and methods of solution. Chapter 4 is dedicated to data collection, analysis and results. Chapter 5 concludes the entire study by stating specific recommendations to stakeholders based on the major findings made in the study.

CHAPTER 2

LITERATURE REVIEW

2.0 INTRODUCTION

This chapter discusses the literature available on Maternal Mortality in general. It also looks at summary of abstracts on various literatures with regard to the model being used and the general working title.

2.1 THE MATERNAL MORTALITY SITUATION

Maternal mortality is defined as the death of a woman while pregnant or within 42days of termination of pregnancy, irrespective of the duration and site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management(WHO–ICD 10).

Globally, the lifetime risk for women of maternal death is 1 in 74. In industrialised countries this risk is 1 in 2,800. In the least developed countries, they face a 1 in 16 chance of dying in childbirth in their lifetime (DFID, 2004). The fifth Millennium Development Goal is to improve maternal health, with a target to reduce the maternal mortality ratio by three quarters, between 1990 and 2015. Yet maternal mortality in developing countries has barely decreased over the past decade, and in parts of Africa it has increased. The national target was to reduce the 1990 maternal mortality rate of 740 per 100,000 live births by 3/4 to 185 per 100,000 live births by 2015. In this thesis, we also looked at the prevalence of maternal mortality using the Poisson Regression and ARIMA models.

The international definition of the maternal mortality ratio (MMR) is the number of Direct and Indirect deaths per 100,000 live births (ICD-10).

$$\text{Maternal Mortality Ratio} = \frac{\text{Total Maternal Deaths}}{\text{Total Live Birth}} \times 100{,}000 \text{ live births} \qquad (2.0)$$

In many countries of the world this is difficult to measure due to the lack of death certificate data as well as a lack of basic denominator data, since baseline vital statistics are also not available or unreliable.

Qin Yi Lee (2010) of the University Of Leicester conducted a Hospital-Based Review of Maternal Mortality in Ghana. This retrospective review was undertaken at the Obstetrics and Gynaecology Department of Komfo Anokye Teaching Hospital (KATH) in Kumasi, Ghana. Data from biostatistics unit as well as all maternal deaths following admission from the period 1st January 2008 to 31st May 2010 were analyzed. The result revealed an estimated maternal mortality ratio of 1021.9 per 100 000 live births (95% CI: 906.6 - 1130.8). Many of such studies have given specific MMR's however we seek to access the patterns of the ratios over the period under study.

Aikins L. (2010) in her thesis titled "Female Health and Development; A case study regarding a maternal Health Scheme in Ghana". The study conducted at Komfo Anokye Teaching Hospital (KATH) in Kumasi, revealed an unstable situation regarding maternal deaths even though pregnant women have free access to antenatal care. The study aimed at finding out whether the free antenatal and delivery care provided by the Ghanaian government is encouraging pregnant women to access the facility in order to improve maternal health and also, whether it is aiding in the reduction in maternal deaths. A total of three health personnel's and fifty-five pregnant women participated in the study. Semi structured interviews and observation were the main tools employed for data collection.

The study indicated that the presence of skilled birth attendants at delivery as per pregnant woman ratio is quite poor. The ratio is one doctor is to 17,733 and the Nurse-Population ratio is 1: 1,510 with disparities between urban and rural settings and dwellers. In Kumasi for instance, the Ratio of Midwives to Women of Reproductive Age is 1:427.

Yoko L, et al (2011) did a study titled "Make it happen 2015: validation of the maternal mortality ratio in Trinidad and Tobago for 2000–06". The aim of their work was to examine the quality of the data used for the estimates of MMR provided by the Trinidad and Tobago Central Statistical Office (CSO). A retrospective reproductive age mortality survey (RAMOS) was applied for 2000–06 to evaluate national estimates. They found that, data from CSO and external data sources yield conflicting results. The CSO estimate of MMR in 2005 was 34·8, while those provided by UNICEF and the World Bank were 45·0 and 55·0, respectively. They recommended that specific maternal death review committee be established as the ideal maternal death review mechanism across all health jurisdictions in Trinidad and Tobago.

Hogan and Foreman (2010) assessed levels and trends in maternal mortality for 181 countries in their report titled "Maternal mortality for 181 countries, 1980—2008: a systematic analysis of progress towards Millennium Development Goal 5". They constructed a database of 2651 observations of maternal mortality for 181 countries for 1980—2008, from vital registration data, censuses, surveys, and verbal autopsy studies. They used robust analytical methods to generate estimates of maternal deaths and the MMR for each year between 1980 and 2008. They explored the sensitivity of the data to model specification and show the out-of-sample predictive validity of the methods. As a result, they estimated that there were 342 900 maternal deaths worldwide in 2008, down from 526 300 in 1980. The global MMR decreased from 422 in 1980 to 320 in 1990, and was 251 per 100 000 live births in 2008. The yearly rate of decline of the global MMR since 1990 was 1·3%. They found out that in the absence of HIV, there would have been 281 500 maternal deaths worldwide in 2008.The report concluded that, although substantial progress has been made towards MDG 5, only 23 countries are on track to achieving a 75% decrease in MMR by 2015 and that countries such as Egypt, China, Ecuador, and Bolivia have been achieving accelerated progress.

Worawan et al. (2010) undertook a study aimed at using multiple sources of data to calculate the maternal mortality ratio (MMR) in 2004–09, and to illustrate the difference between the official causes of death with the research findings. In their research titled "Thailand's approach to measuring maternal mortality ratio" individual data from civil registration and inpatient records from all public hospitals were used. The civil registration contains data about individual's personal identification (PID) etc. Their result shown that, the number of maternal deaths declined from 362 in 2004 to 269 in 2009. The country's MMR declined from 44·5 to 35·2, a 21% reduction. Their conclusion was that, using matching technique together with individual data, policy makers can get reliable information about the causes of maternal death.

2.2 USE OF POISSON REGRESSION AND ARIMA MODELS IN MATERNAL MORTALITY

In the 2009 December Bulletin of World Health Organization, Yanqiu, et al. (2009) studied Time trends and regional differences in maternal mortality in China from 2000 to 2005. They used Poisson regression analysis of data from the Chinese National Maternal and Child Health Routine Reporting System between 2000 and 2005 to identify time trends in the maternal mortality ratio (MMR) by province and region. They found that MMR declined by an average of 5% per year. There was no interaction between region and year (P = 0.2311). Mortality declined by 5% per year in the eastern region, by 5% per year in the central region, and by 4% per year in the western region. The absolute difference in MMR between the western and eastern regions declined from 65.4 deaths per 100 000 live births in 2000 to 49.4 per 100 000 live births in 2005. Their conclusion was that, China is making good progress towards achieving the fifth Millennium Development Goal, and there is no evidence of a widening gap between better-off and economically more deprived provinces.

In their study titled "Maternal mortality in the former East Germany before and after reunification: changes in risk by marital status", Razum, et al (1998) examined the impact of marital status on maternal mortality in the period before and the period after German reunification in the area covered by the former East Germany. They calculated the maternal mortality ratio by relating the number of maternal deaths among women resident in eastern Germany in 1980-96 to the respective number of live births, using national register data. They then investigated the effect of marital status, controlling for maternal age and year of death, in a Poisson regression model. Altogether, 413 maternal deaths and 2.99 million live births were reported. The overall maternal mortality ratio was stable before, and declined after, reunification. Before reunification, unmarried women had a risk of maternal death equal to that of married women; after reunification, they had 2.6 times the age adjusted risk of married women. Unmarried status thus became a significant risk factor for maternal mortality in eastern Germany after reunification.

The European Journal of Obstetrics & Gynecology and Reproductive Biology 2003, published a study titled "Maternal mortality in Northern Nigeria: a population-based study" by Adamu, et al (2002). They sort to determine the incidence and causes of maternal mortality as well as its temporal distribution over the last decade (1990–1999). All maternal deaths recorded within the study period in the State of Kano, Northern Nigeria, were analyzed. Maternal mortality ratios (MMR) were computed using the Poisson assumption to derive confidence intervals around the estimates. A non-linear regression model was fitted to obtain the best temporal trajectory for MMR across the decade of study. Their study reveals that a total of 4154 maternal deaths occurred among 171 621 deliveries, yielding an MMR of 2420 deaths per 100 000. Eclampsia, ruptured uterus and anemia were responsible for about 50% of maternal deaths. They concluded that the area had one of the highest maternal mortality ratios in the

world and suggested that maternal mortality could be reduced by half at study site with effective interventions targeted to prevent deaths from eclampsia, ruptured uterus and anemia.

In the journal of Epidemiology Community Health 2009, Fernández, et al, (2008) published a research report titled "Increase in maternal mortality associated with change in the reproductive pattern in Spain: 1996–2005". Their study aimed at analyzing the age-related trend in the maternal mortality ratio among mothers in Spain for the decade 1996–2005, and to describe the causes of death and associated socio-demographic factors for the years with highest mortality. An ecological study on trends, for the age-related trend in the maternal mortality ratio; an indirect standardization and Poisson regression model was used. They found that, Prevalence of live births among mothers aged 35 years and over was 15% higher in Spain than in Europe. The maternal mortality rate increased by 20% (standardized mortality ratio of 1.2, 95% CI 0.9 to 1.4) in 2005 with respect to 1996. The age-related risk of maternal mortality was three times higher (relative risk of 2.90, 95% CI 2.01 to 4.06) among mothers aged 35–44 years versus those aged under 35 years. The highest mortality was detected during 2003–2004. The study therefore concluded that there was a change in the maternal mortality trend characterized by an increase in deaths, associated with advanced maternal age, as well as an increase in the prevalence of live births among mothers aged 35 years and over.

Sullivan et al, (2003) published the report of their study titled "Maternal-fetal medicine specialist density is inversely associated with maternal mortality ratios." in the American Journal of obstetric and Gynecol. 2005. Their objective was to determine the relationship between state-specific maternal mortality ratios and the density of maternal-fetal medicine specialists. State maternal mortality ratios from 1994 to 2001 were calculated from the Centers for Disease Control and Prevention WONDER database. Practitioner distribution data were obtained from professional associations. Demographic information regarding states was

gathered from the 2000 US census data. Bivariable and multivariable analyses were conducted with the use of Spearman correlations and Poisson regression, respectively. The study showed that an increase of 5 maternal-fetal specialists per 10,000 live births results in a 27% reduction in the risk of maternal death (relative risk [RR] = 0.73, 95% CI = 0.58-0.93, P = 0.012). This risk reduction was based on a multivariable Poisson regression model that included the following variables and their significant interactions: state-specific percentages of mothers in poverty, mothers without a high school diploma, minority mothers, and teenage mothers. The density of maternal-fetal medicine specialists is significantly and inversely associated with maternal mortality ratios, even after controlling for state-level measures of maternal poverty, education, race, age, and their significant interactions.

A study titled "Use of Poisson Regression and Time Series Analysis for Detecting Changes over Time in Rates of Child Injury following a Prevention Program" was undertaken by Louise Kuhn, Leslie L. Davidson and Maureen S. Durkin in 1994 (Louise, et al, 1994). The two analytical methods were used to quantify changes in the rate of injury following the program, while taking into account the underlying annual and seasonal trends. Rates of severe injury during the period from 1983 to 1991 among children under the age of 17 years living in Central Harlem and in the neighboring community of Washington Heights were analyzed. The two methods provide similar point estimates of the effect of the intervention and had a good fit to the data. They concluded that, although time series analysis has been promoted as the method of choice in analysis of sequential observations over long periods of time, their illustration suggests that Poisson regression is an attractive and viable alternative. According to their study, Poisson regression provides a versatile analytical method for quantifying the time trends of relatively rare discrete outcomes, such as severe injuries, and provides a useful tool for epidemiologists involved with program evaluation.

The Journal of China Medical University in March 2011 conducted a study. The study was to explore the feasibility for application of time series ARIMA model to predict the maternal mortality ratio (MMR) in china so as to provide the theoretical basis for continuing to reduce the MMR. ARIMA model was established based on the MMR of China from 1991 to 2009. Using difference method to smooth the sequence, they determined the order and established the 2010 national maternal mortality ratio forecast model to evaluate the predicting results. It was found that ARIMA model fitted very well, the residual autocorrelation function graph showed the residuals were white noise sequences, the prediction results showed that maternal mortality ratio in national urban and rural areas would be 30.39 ‰, 24.73 ‰ and 28.80 ‰ in 2010, which showed MMR, would decline and reach a lower level. The researchers concluded that the fitting result in ARIMA model of the incidence of the MMR is satisfactory, the forecasting achieve good effects, which also provides scientific basis for the prevention and control of maternal mortality ratio.

Koch (2009) on behalf of The Chilean Maternal Mortality Group, Faculty of Medicine, University of Chile, wanted to find out whether there exist an association between maternal mortality reduction and abortion legalization? Time series of maternal mortality ratio (MMR) and abortion mortality ratio (AMR) from 1960 to 2007 were analyzed using multiple autoregressive moving average (ARIMA) models. Therapeutic abortion was legal until 1989 and was considered as a dummy variable in statistical analyses along time series of social and demographic factors and maternal health facilities. During the study period, MMR was found to have decreased from 293.7 to 18.2 per 100,000 live births (-93.8%); AMR decreased from 92.5 to 1.7 per 100,000 live births (-98.1%). No significant effect of legal and illegal abortion periods on these decreasing trends was observed in ARIMA models. After abortion was fully prohibited, MMR and AMR decreased from 41.3 to 18.2 (-44.1%) and 16.5 to 1.7 (-10.3%) per 100,000 live births respectively. The average of education years, illiteracy rate, GDP per-

capita, and the percentage of delivery by skilled attendants were all significant predictors of MMR. The same factors along decreasing fertility rate were significant predictors of AMR trends. The study concluded that reductions in MMR and AMR are not related with legal/illegal therapeutic abortion periods in Chile. The increasing education level appears as the most important factor predicting maternal mortality reduction in this developing country, likely influencing other factors such as fertility and maternal health facilities.

The Ethiopian Government through their Ministry of Health (MOH Ethiopia, 2000) undertook a study to analyze trends and develop model for prediction of Health and Health related indicators. Key indicators of Mortality and Morbidity, Health service coverage, Health systems resources, Demographic and socio-economic, and Risk factor indicators were extracted and analyzed. The trends in these indicators were established using trend analysis techniques. The determinants of the established trends were identified using ARIMA models in STATA. The trend-line equations were then used to predict future values of the indicators. Among the mortality indicators considered in this study, it was only Maternal Mortality Ratio that showed statistically significant decrement within the study period. The trends of Total Fertility Rate, physician per 100,000 population, skilled birth attendance and postnatal care coverage were found to have significant association with Maternal Mortality Ratio trend. Based on the prediction from the current trend, they concluded that the Millennium Development Goal target for under-five mortality rate and proportion of people having access to basic sanitation can be achieved.

2.3 OTHER USE OF POISSON REGRESSION AND ARIMA MODELS

Yun-li, et al (2004), tried to understand the prevalence and changing trend of Infant Mortality Rates (IMR) in China from 1991 to 2004, and to forecast the IMR in China from 2005 to 2007. The average velocity of increase was calculated and trend analysis was done

simultaneously; ARIMA model was used to forecast IMR in China from 2005 to 2007. They found that, from 1991 to 2004, the IMR of China has decreased, the average rate of decay of the urban IMR and the rural IMR is 3.77% (u = 3.996 4, P< 0.001) and 5.97% (u = 4.762 8, P< 0.001) respectively. And the rural IMR is obviously higher than the urban IMR. From 2005 to 2007, the rural IMR would descend obviously; while the urban IMR would keep steady. They concluded that though the IMR in urban areas and rural areas have both decreased, the urban-rural difference was obvious. Compared with the urban IMR, the rural IMR has greater space to descend because of the previous higher IMR.

Kuhn, et al (1993) used Poisson regression and time series analysis to analyze changes in child injury incidence after implementation of community-based injury prevention program in Central Harlem, New York City. Rates of severe injury during the period from 1983 to 1991 among children under the age of 17 years living in Central Harlem and in the neighbouring community of Washington Heights are analyzed. The two methods provide similar point estimates of the effect of the intervention and have a good fit to the data. Although time series analysis has been promoted as the method of choice in analysis of sequential observations over long periods of time, this illustration suggests that Poisson regression is an attractive and viable alternative. Poisson regression provides a versatile analytical method for quantifying the time trends of relatively rare discrete outcomes, such as severe injuries, and provides a useful tool for epidemiologists involved with program evaluation.

Fernández, et al (2009), aimed at analysing the age-related trend in the maternal mortality ratio among mothers in Spain for the decade 1996–2005. An ecological study on trends, for the age-related trend in the maternal mortality ratio; an indirect standardisation and Poisson regression model was used. Prevalence of live births among mothers aged 35 years and over was 15% higher in Spain than in Europe. The maternal mortality rate increased by 20%

(standardised mortality ratio of 1.2, 95% CI 0.9 to 1.4) in 2005 with respect to 1996. The age-related risk of maternal mortality was three times higher (relative risk of 2.90, 95% CI 2.01 to 4.06) among mothers aged 35–44 years versus those aged under 35 years. The highest mortality was detected during 2003–2004. The risk of maternal mortality was higher in foreign mothers. This study confirms that there was a change in the maternal mortality trend characterised by an increase in deaths, associated with advanced maternal age, as well as an increase in the prevalence of live births among mothers aged 35 years and over.

Mulu and Tilahun (2009) did a study that sort to analyze trends of and develop model for prediction of Health and Health related indicators of Ethiopia from the year 1987 to 2000. The determinants of the established trends were identified using ARIMA models in STATA. Among the mortality indicators considered in this study, it was only Maternal Mortality Ratio that showed statistically significant decrement within the study period. The trends of Total Fertility Rate, physician per 100,000 population, skilled birth attendance and postnatal care coverage were found to have significant association with Maternal Mortality Ratio trend. They concluded that current trend indicates the need to accelerate the progress of the indicators to achieve MDGs at or before 2015, particularly for Maternal Health and access to safe water supply.

CHAPTER 3

3.0 INTRODUCTION

This chapter describes the theory of models to be used, formulations and methods of analyzing the available data to satisfy the objectives of the study. It focuses on the detail and comprehensive understanding of The Box-Jenkins methodology for ARIMA models as well as the definition of Poisson distribution and the Poisson Regression model. Among the aspects that will come under discussion include the methodologies used in modelling, the software specifications, and the features that are incorporated in the model.

3.1 DATA SOURCE AND TYPE

This study essentially seeks to model the maternal mortality patterns to be used in predicting future maternal mortality ratios using the Okomfo Anokye Teaching Hospital as the case study. The analysis is based on secondary data available at the Bio-Statistics Department of the Obstetrics & Gynaecology directorate of the Okomfo Anokye Teaching Hospital in Kumasi for the period 2000-2010. The required data include: Quarterly recorded live births, Quarterly recorded maternal deaths and Quarterly recorded deliveries. Maternal mortality ratios were also computed using the tenth revisions of the International Classification of Diseases (ICD-10). The data has only one variable under study which is time in Quarters. In all, there are 44 quarterly observations under the eleven years duration.

3.2 THE CONCEPT OF TIME SERIES

Time series is a time dependent sequence Y_t, where t belongs to the set of integers and denotes the time steps. If a time series can be expressed as a known function, $Y_t = f(t)$, then it is said to be a deterministic time series. If it is however expressed as $Y_t = X(t)$, where X is a random variable then $\{Y_t\}$ is a stochastic time series. The basis for analyzing a time series may be classified as prediction, description and control. Time series analysis mainly decomposes the variations in a series into the various components of trend, periodic and stochastic.

3.2.1 Stationary and Non-stationary Series

A time series is said to be strictly stationary if the joint distribution of X_{t1}, X_{t2}, … X_{tn} is the same as the joint distribution of X_{t1+T}, X_{t2+T}, … X_{tn+T}, for all t_{1+T} …t_{n+T}. Thus, shifting the time position by T periods has no effects on the joint distributions, which therefore depends on the interval between t_1… t_n. If a time series is not stationary then it is said to be non-stationary. A simple non-stationary time series model is given by

$$Y_t = \mu_t + e_t \qquad (3.0)$$

where the mean μ_t is a function of time and e_t is a weakly stationary series. Unlike the stationary time series, the mean and variance of the non-stationary process changes with time. If a non-stationary series is differenced one or more times it becomes stationary and that series is then said to be homogeneous.

The stationarity condition ensures that the autoregressive parameters in the estimated model are stable within a certain range as well as the moving average parameters in the model are invertible. If this condition is assured then, the estimated model can be forecasted (see Hamilton, 1994). To check for stationarity, we usually test for the existence or nonexistence of

what we called unit root. Unit root test is performed to determine whether a stochastic or a deterministic trend is present in the series. There are several statistical tests in testing for presence of unit root in a series. For series with seasonal and non-seasonal behaviour, the test must be conducted under the seasonal part as well as the non-seasonal part. Some example of the unit root test for the non-seasonal time series are the Dickey-Fuller and the Augmented Dickey- Fuller (DF, ADF) test, Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test and Zivot-Andrews (ZA) test (see Dickey & Fuller, 1979; Kwiatkowski et al, 1992; Zivot & Andrews, 1992).

- **Unit Root Test**

Unit Root Test was derived in 1979 by Dickey and Fuller to test the presence of a unit root vs. a stationary process. The unit root process and a stationary process are given by equations 3.1 and 3.2 below;

$$\rho_t = \emptyset_1 \rho_{t-1} + e_t \quad (3.1)$$

$$\rho_t = \emptyset_0 + \emptyset_1 \rho_{t-1} + e_t \quad (3.2)$$

If $\emptyset_1 = 1$ then the series is said to have unit root and is not stationary.

The Unit Root Test as proposed by Kwiatkowski-Phillips-Schmidt-Shin (KPSS), test the hypothesis below

$$H_0: \emptyset_1 = Series\ is\ level\ or\ trend\ stationary$$

$$H_A: \emptyset_1 = series\ is\ level\ or\ trend\ non-stationary$$

If test statistic value of the KPSS test is less than critical value we accept the null hypothesis that data is level or trend stationary.

Similarly, the Unit Root Test as proposed by Dickey and Fuller (ADF), test the hypothesis below

$$H_0: \emptyset_1 = Series\ has\ unit\ root$$

$$H_A: \emptyset_1 = series\ has\ no\ unit\ root$$

If test statistic of the ADF test is less than critical value we reject the null hypothesis that data has a unit root.

3.2.2 Lag

Lag is a difference in time between an observation and a previous observation. Thus Y_{t-k} lags Y_t by k periods

3.2.3 White Noise

A collection of uncorrelated random variables, ω_t, with mean 0 and finite variance ${\sigma_\omega}^2$ was first used as a model for noise in engineering applications, where it is called white noise. We denote this process as $\omega_t \sim wn(0, {\sigma_\omega}^2)$. The designation white originates from the analogy with white light and indicates that all possible periodic oscillations are present with equal strength. We will, at times, also require the noise to be iid random variables with mean 0 and variance ${\sigma_\omega}^2$. We shall distinguish this case by saying white independent noise, or by writing $\omega_t \sim iid(0, {\sigma_\omega}^2)$. A particularly useful white noise series is Gaussian white noise, wherein the $\omega_t(s)$ are independent normal random variables, with mean 0 and variance ${\sigma_\omega}^2$; or more succinctly, $\omega_t \sim iid$ N $(0, {\sigma_\omega}^2)$.

3.2.4 Autocorrelation Function (ACF)

The Autocorrelation function is extremely useful for describing the general process used to develop a forecasting model. It measures the degree of correlation between neighbouring

observations in a time series. The autocorrelation at any lag k is defined as $COR(Y_t, Y_{t-k})$ and is measured by

$$\rho_k = \frac{COV(Y_t, Y_{t-k})}{\delta_{Y_t}\delta_{Y_{t-k}}} = \frac{E[(Y_t-\mu_\gamma)(Y_{t-k}-\mu_\gamma)]}{[E(Y_t-\mu_\gamma)^2(Y_{t-k}-\mu_\gamma)^2]} \quad (3.3)$$

Where Y_t, the observation at time t, Y_{t-k} is observation at time t-k and μ_γ is the observed mean.

The theoretical autocorrelation function is generally unknown, but may be estimated from the sample autocorrelation function as follows:

$$\widehat{\rho_k} = \frac{\sum_{t=1}^{n}(Y_t-\bar{Y})(Y_{t-k}-\bar{Y})}{\sum_{t=1}^{n}(Y_t-\bar{Y})^2} \quad (3.4)$$

Where t is the length of the time series under study $\bar{Y}$ is the mean of the Y_t observation and k = 1, 2... k.

Normally we compute the first K to be less than N/4 sample autocorrelations.

If we define the covariance between Y_t and Y_{t-k} as γ_k, then

$$\rho_k = \frac{\gamma_k}{\gamma_0}, \quad (3.5)$$

Thus for any stochastic process $\rho_0 = 1$. It is also true that $\rho_k = \rho_{-k}$

For a series to be white noise all $\rho_k = 0$, for k > 0. This may be tested using the Box-Pierce test statistic:

$$Q = n\sum_{k=1}^{k}\widehat{\rho_k}^2 \quad (3.6)$$

which is approximately distributed as chi-square with k degrees of freedom? Thus a value below the critical value would lead to acceptance of white noise. If a series is stationary, the sample autocorrelations will tail off quickly as k increases. The

expression tails off means that the function decays in an exponential, sinusoidal, or geometric fashion, with a relatively large number of nonzero values. A function is said to cut off if the function truncates abruptly with only a few nonzero values.

3.2.5 Partial Autocorrelation Function

A partial autocorrelation coefficient measures the degree of association between an observation Y_t and Y_{t-k} when the effects of the other time lags are held constant. We consider partial autocorrelation when we are unaware of the appropriate order of the autoregressive process to fit the time series. PACF is denoted by Q_{kk} and is defined as

$$Q_{kk} = \frac{|P_k{}^*|}{|P_k|} \tag{3.7}$$

Where P_k is a k x k auto correlation matrix and $P_k{}^*$ is P_k with the least column replaced by $[\rho_1\ \rho_2 \dots\ \rho_k]^T$

The partial autocorrelation coefficient of order k is denoted by $\propto_k$ and that can be calculated by regressing Y_t against $Y_{t-1} \dots Y_{t-k}$

$$Y_t = b_0 + b_1 Y_{t-1} + b_2 Y_{t-2} + \cdots + b_k Y_{t-k} \tag{3.8}$$

Here, explanatory variables on the right hand side are previous values of the forecast variable Y_t whereas the partial autocorrelation $\propto_k$ is the estimated coefficient b_k from the regression equation above.

3.2.6 Auto regressive model AR (p)

Autoregressive models are based on the idea that the current value of the series X_t, can be explained as a function of p past values, X_{t-1},X_{t-2}, $\dots X_{t-p}$, where p determines the number of

steps into the past needed to forecast the current value. An autoregressive model of order p, abbreviated AR (p), is of the form

$$X_t = \alpha + \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \cdots + \emptyset_p X_{t-p} + \omega_t \quad (3.9)$$

Where X_t is stationary, $\emptyset_1, \emptyset_2, \ldots \emptyset_p$ are constants. Unless otherwise stated, we assume that ω_t is a Gaussian white noise series with mean μ zero and variance $\sigma_\omega{}^2$. However, if the mean μ of X_t is not zero, replace X_t with $X_t - \mu$ in the AR (p) model.

$$X_t - \mu = \emptyset_1(X_{t-1} - \mu) + \emptyset_2(X_{t-2} - \mu) + \cdots + \emptyset_p(X_{t-p} - \mu) + \omega_t \quad (3.10)$$

or we write as

$$X_t = \alpha + \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \cdots + \emptyset_p X_{t-p} + \omega_t \quad (3.11)$$

Where $\alpha = \mu(1 - \emptyset_1 - \cdots - \emptyset_p)$

A more concise form of the AR (p) model is $\emptyset(B)X_t = \omega_t$

Where $\emptyset(B)$ is an autoregressive backward shift operator defined as

$$\emptyset(B) = 1 - \emptyset_1 B - \emptyset_2 B^2 - \cdots - \emptyset_p B^p \quad (3.12)$$

3.2.7 Moving Average Model MA (q)

The moving average model of order q, or MA (q) model, is defined to be

$$X_t = \omega_t + \theta_1 \omega_{t-1} + \theta_2 \omega_{t-2} + \cdots + \theta_q \omega_{t-q} \quad (3.13)$$

Where there are q lags in the moving average and θ_1, θ_2,…, θ_q ($\theta_q \neq 0$) are parameters. The noise ω_t is assumed to be Gaussian white noise. We may also write the MA (q) process in the equivalent form

$$X_t = \theta(B)\omega_t \quad (3.14)$$

Where $\theta(B)$ is a moving average backward shift operator defined as

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_p B^p \quad (3.15)$$

3.2.8 ARMA model

A time series $\{ X_t; t = 0, \pm1, \pm2, ...\}$ is ARMA (p, q) if it is stationary and

$$X_t = \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \cdots + \emptyset_p X_{t-p} + \omega_t + \theta_1 \omega_{t-1} + \theta_2 \omega_{t-2} + \cdots + \theta_q \omega_{t-q} \quad (3.16)$$

$\emptyset_p \neq 0, \theta_q \neq 0$, and $\sigma_\omega{}^2 > 0$. The parameters p and q are called the autoregressive and the moving average orders, respectively.

If X_t has a non-zero mean μ, we set $\alpha = \mu(1 - \emptyset_1 - \cdots - \emptyset_p)$ and write the model as

$$X_t = \alpha + \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \cdots + \emptyset_p X_{t-p} + \omega_t + \theta_1 \omega_{t-1} + \theta_2 \omega_{t-2} + \cdots + \theta_q \omega_{t-q} \quad (3.17)$$

Unless otherwise stated, $\{ \omega_t; t = 0, \pm1, \pm2, ...\}$ is a Gaussian white noise sequence.

As previously noted, when $q = 0$, the model is called an autoregressive model of order *p*, AR (*p*), and when $p = 0$, the model is called a moving average model of order *q*, MA (*q*). The ARMA (*p*, *q*) model in (3.15) can then be written in concise form using the AR operator, and the MA operator, as

$$\emptyset(B) X_t = \Theta(B) \omega_t \quad (3.18)$$

3.2.9 ARIMA Models

The acronym ARIMA stands for "Auto-Regressive Integrated Moving Average." Lags of the differenced series appearing in the forecasting equation are called "auto-regressive" terms, lags of the forecast errors are called "moving average" terms, and a time series which needs to be differenced to be made stationary is said to be an "integrated" version of a stationary series. A non-seasonal ARIMA model is classified as an "ARIMA (p, d, q)" model, where

p is the number of autoregressive terms, d is the number of non-seasonal differences, and q is the number of lagged forecast errors (moving average) in the prediction equation.

A process, X_t is said to be ARIMA (p, d, q) if

$$\nabla^d X_t = (1 - B)^d X_t \quad (3.19)$$

is ARMA (p, q). In other words the process should be stationary after differencing a non-seasonal process d times. In general, we will write the model as

$$\emptyset(B)(1 - B)^d X_t = \theta(B)\omega_t \quad (3.20)$$

If $E(\nabla^d X_t) = \mu$ we write the model as

$$\emptyset(B)(1 - B)^d X_t = \propto + \theta(B)\omega_t \quad (3.21)$$

Where $\alpha = \mu(1 - \emptyset_1 - \cdots - \emptyset_p)$

3.3 THE BOX-JENKINS ARIMA MODEL

The Box-Jenkins methodology refers to the set of procedures for identifying, fitting, and checking ARIMA models with time series data. Forecasts follow directly from the form of the fitted model. By Box-Jenkins, a p^{th} order autoregressive model: AR (p), has the general form

$$X_t = \alpha + \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \cdots + \emptyset_p X_{t-p} + \omega_t \quad (3.22)$$

Where X_t= Response (dependent) variable at time t, $X_{t-1}, X_{t-2}, \ldots X_{t-p}$= Response variable at time lags t − 1, t − 2, . . . , t − p, respectively.

$\emptyset_1, \emptyset_2, \ldots \emptyset_p$= Coefficients to be estimated, and ω_t = Error term at time t.

Also, a q^{th}- order moving average model: MA (q), has the general form

$$X_t = \mu + \omega_t + \theta_1 \omega_{t-1} + \theta_2 \omega_{t-2} + \cdots + \theta_q \omega_{t-q} \quad (3.23)$$

Where X_t= Response (dependent) variable at time t, μ = Constant mean of the process, $\emptyset_1, \emptyset_2, \ldots \emptyset_p$= Coefficients to be estimated, ω_t = Error term at time t, and $\omega_{t-1}, \omega_{t-2}, \ldots, \omega_{t-p}$ = Errors in previous time periods that are incorporated in the response X_t.

Autoregressive Moving Average Model: ARMA (p, q), which has the general form

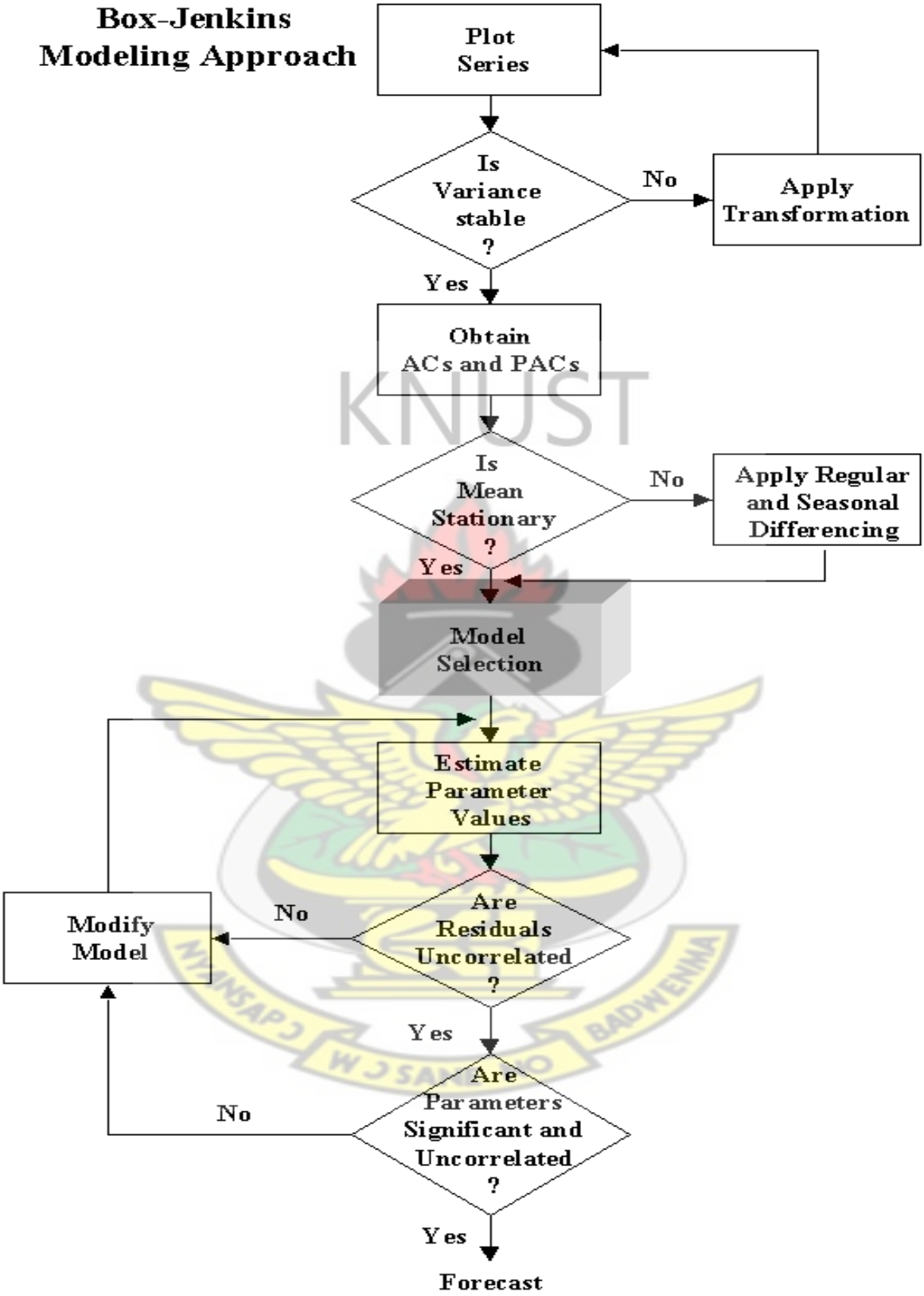
$$X_t = \alpha + \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \cdots + \emptyset_p X_{t-p} + \omega_t + \theta_1 \omega_{t-1} + \theta_2 \omega_{t-2} + \cdots + \theta_q \omega_{t-q} \quad (3.24)$$

We can use the graph of the sample autocorrelation function (ACF) and the sample partial autocorrelation function (PACF) to determine the model which processes can be summarized as follows:

Table 3.1: How to determine the model by using ACF and PACF patterns

MODEL	**ACF**	**PACF**
AR (p)	Dies down	Cut off after lag q
MA (q)	Cut off after lag p	Dies down
ARMA (p, q)	Dies down	Dies down

Box-Jenkins forecasting models consist of a four-step iterative procedure as follows; Model Identification, Model Estimation, Model Checking (Goodness of fit) and Model Forecasting. The four iterative steps are not straight forward but are embodied in a continuous flow chart depending on the set of data under study. The chart below describes the Box-Jenkins modelling approach for an ARIMA model.



Source: BOX, G.E.P. and G.M. JENKINS (1970) Time series analysis: Forecasting and control, San Francisco: Holden-Day.

- **STEP 1: Model Identification (Selecting an initial model)**

We first Determine whether the series is stationary or not by considering the graph of ACF. If a graph of ACF of the time series values either cuts off fairly quickly or dies down fairly quickly, then the time series values should be considered stationary. If a graph of ACF dies down extremely slowly, then the time series values should be considered non-stationary. If the series is not stationary; it would then be converted to a stationary series by differencing. That is, the original series is replaced by a series of differences. An ARMA model is then specified for the differenced series. Differencing is done until a plot of the data indicates the series varies about a fixed level, and the graph of ACF either cuts off fairly quickly or dies down fairly quickly. Once a stationary series has been obtained, then identify the form of the model to be used by using the theory in Table 3.1

- **STEP 2: Model Estimation and Evaluation**

Once a model is identified, the next stage for Box-Jenkins approach is to Estimate the parameters. In this study, the estimation of parameters was done using the R-Consol statistical package.

- **Method of moment's estimators**

We begin with method of moments estimators. The idea behind these estimators is that of equating population moments to sample moments and then solving for the parameters in terms of the sample moments. We immediately see that, if $E(\mathrm{X_t}) = \mu$, then the method of moment estimator of μ is the sample average $\bar{x}$. Thus, while discussing method of moments, we will assume $\mu = 0$. Although the method of moments can produce good estimators, they can sometimes lead to suboptimal estimators.

- **Maximum likelihood estimators**

Let $X_t = \mu + \emptyset_1(X_{t-1} - \mu) + \omega_t$ (3.25)

be a time series process, where $|\emptyset| < 1$ and $\omega_t \sim iid\ N(0, \sigma_\omega^2)$. Given the data $X_1, X_2, \ldots, X_n$, we seek the likelihood

$$\boldsymbol{L}(\mu, \emptyset, \sigma_\omega^2) = f_{\mu,\emptyset,\sigma_\omega^2}(X_1, X_2, \ldots, X_n). \quad (3.26)$$

In the case of an AR (1), we may write the likelihood as

$$\boldsymbol{L}(\mu, \emptyset, \sigma_\omega^2) = f(X_1) f(X_2 \,/\, X_1) \ldots f(X_n \,/\, X_{n-1}), \quad (3.27)$$

Where we have dropped the parameters in the densities, $f(.)$, to ease the notation.

We may then write the likelihood as

$$\boldsymbol{L}(\mu, \emptyset, \sigma_\omega^2) = f(X_1) \prod_{t=2}^{n} f_\omega[(X_t - \mu) - \emptyset(X_{t-1} - \mu)] \quad (3.28)$$

To find $f(X_1)$, we can use the causal representation

$$X_t = \mu + \sum_{j=0}^{\infty} \emptyset^j \omega_{1-j} \quad (3.29)$$

To see that X_1 is normal, with mean μ and variance $\frac{\sigma_\omega^2}{(1-\emptyset^2)}$

- **AIC, AICc, BIC**

The final model can be selected using a penalty function statistics such as **A**kaike **I**nformation **C**riterion (AIC or AICc) or **B**ayesian **I**nformation **C**riterion (BIC). See Sakamoto et. al. (1986); Akaike (1974) and Schwarz (1978). The AIC, AICc and BIC are a measure of the goodness of fit of an estimated statistical model. Given a data set, several competing models may be ranked according to their AIC, AICc or BIC with the one having the lowest information criterion value being the best. These information criterion judges a model by how close its fitted values tend to be to the true values, in terms of a certain expected value. The criterion value assigned to a model is only meant to *rank* competing models1 and tell you which the best among the given alternatives is. The criterion attempts to find the model that best explains the data with a

minimum of free parameters but also includes a penalty that is an increasing function of the number of estimated parameters. In the general case, the AIC, AICc and BIC is calculated as;

$$AIC = 2k - 2\log(L) \quad OR \quad 2k + nlog(\frac{RSS}{n}) \tag{3.30}$$

$$AICc = AIC + \frac{2k(k+1)}{n-k-1} \tag{3.31}$$

$$BIC = -2Log(L) + klog(n) \quad OR \ \log(\sigma_e^2) + \frac{k}{n}\log(n) \tag{3.32}$$

Where

k: is the number of parameters in the statistical model

L: is the maximized value of the likelihood function for the estimated model.

RSS: is the residual sum of squares of the estimated model.

n: is the number of observation, or equivalently, the sample size

σ_e^2: is the error variance

The AICc is a modification of the AIC by Hurvich and Tsai (1989) and it is AIC with a second order correction for small sample sizes. Burnham & Anderson (1998) insist that since AICc converges to AIC as *n* gets large, AICc should be employed regardless of the sample size.

- **STEP 3: Model Checking (Goodness of fit)**

In this step, model must be checked for adequacy by considering the properties of the residuals whether the residuals from an ARIMA model must have the normal distribution and should be random. An overall check of model adequacy is provided by the Ljung-Box Q statistic. The test statistic Q is

$$Q_m = n(n+2)\sum_{k=1}^{m}\frac{r_k^2(e)}{n-k} \sim \chi^2_{m-r} \tag{3.33}$$

Where $r_k(e)$= the residual autocorrelation at lag k, n= the number of residuals and m = the number of times lags includes in the test. If the p-value associated with the Q statistic is small

(P-value $<\propto$), the model is considered inadequate. We then consider a new or modified model and continue the analysis until a satisfactory model has been determined. Moreover we can check the properties of the residual with the following graph:

(1) We can check about the normality by considering the normal probability plot or the p-value from the One-Sample Kolmogorov – Smirnov Test.

(2) We can check about the randomness of the residuals by considering the graph of ACF and PACF of the residual. The individual residual autocorrelation should be small and generally within $\pm 2/\sqrt{n}$ of zero.

Residuals should at all cost be white noise. A test of whether the residuals form a white process is given by a modified version of the Box-Pierce Q statistic in the form:

$$Q = (T - d) \sum_{k=1}^{k} \hat{\rho}_k^2 \qquad (3.34)$$

Where rho hat is the autocorrelations of the residuals, d is the order of differencing to obtain a stationary series, T is the length of the series, and K is the number of autocorrelations being checked. Here if Q is larger than the critical value for the chi squared distribution with K-p-q degrees of freedom, the model should be considered inadequate.

- **STEP 4: forecasting**

Once the model has been selected, estimated and checked, and once residuals are carefully examine and seen to be uncorrelated and parameters assessed to be significant and uncorrelated, then we can go ahead and forecast. Forecasting with this system is straight forward; the forecast is the expected value, evaluated at a particular point in time. Confidence intervals may also be easily derived from the standard errors of the residuals.

3.4 THE POISSON DISTRIBUTION

The Poisson distribution is a discrete probability distribution that expresses the probability of a number of events occurring in a fixed period of time if these events occur with a known average rate and independently of the time since the last event. The distribution was first introduced by Simeon-Denis Poisson (1781–1840) and published, together with his probability theory, in 1838 in his work *Recherchessur la probabilite des jugements en matierecriminelle et enmatierecivile* ("Research on the Probability of Judgments in Criminal and Civil Matters"). The work focused on certain random variables *N* that count, among other things, the number of discrete occurrences (sometimes called "arrivals") that take place during a time-interval of given length.

If the expected number of occurrences in this interval is λ, then the probability that there are exactly *k* occurrences

(*k* being a non-negative integer, *k* = 0, 1, 2, ...) is equal to

$$f(k,\lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \tag{3.35}$$

Where *e* is the base of the natural logarithm (*e* = 2.71828...), *k* is the number of occurrences of an event - the probability of which is given by the function, *k*! is the factorial of *k*, λ is a positive real number, equal to the expected number of occurrences that occur during the given interval.

The parameter λ is not only the *mean* number of occurrences (k), but also its variance $\sigma_k^2 = (k^2) - (k)^2$

Thus, the number of observed occurrences fluctuates about its mean λ with a standard deviation $\sigma^2 = \sqrt{\lambda}$

As a function of *k*, this is the probability mass function. The Poisson distribution can be derived as a limiting case of the binomial distribution.

Assumptions of Poisson distribution are:

- Observations are independent.
- Probability of occurrence in a short interval is proportional to the length of the interval.
- Probability of another occurrence in such a short interval is zero.

3.5.1 Generalized Linear Models (GLM)

Generalized linear models (GLM) was first introduced by Nelder and Wedderburn (1972, JRSSA). They provided a unified framework to study various regression models, rather than a separate study for each individual regression. Generalized linear models (GLM) are extensions of classical linear models. It includes linear regression models, analysis of variance models, logistic regression models, Poisson regression models, log-linear models, as well as many other models. The above models share a number of unique properties, such as linearity and a common method for parameter estimation. A generalized linear model consists of three components:

1. A *random component*, specifying the conditional distribution of the response variable, Y_i given the explanatory variables.
2. A linear function of the regressors, called the *linear predictor*,

   $$\eta_i = \alpha + \beta_1 X_{i1} + \cdots + \beta_k X_{ik} = x_i^{'}\beta \qquad (3.36)$$

   on which the expected value μ_i of Y_i depends.
3. An invertible $link\ function\ g(\mu_i) = \eta_i$, which transforms the expectation of the response to the linear predictor. The inverse of the link function is sometimes called the *mean function*: $g^{-1}(\eta_i) = \mu_i$.

For traditional linear models in which the random component consists of the assumption that the response variable follows the Normal distribution, the canonical link function is the identity link. The identity link specifies that the expected mean of the response variable is identical to the linear predictor, rather than to a non-linear function of the linear predictor.

The Generalized Linear Model is an extension of the General Linear Model to include response variables that follow any probability distribution in the exponential family of distributions. The exponential family includes such useful distributions as the Normal, Binomial, Poisson, Multinomial, Gamma, Negative Binomial, and others.

3.4.2 The Link Function

In theory, link functions $\eta_i = g(\mu_i)$ can be any monotonic, differentiable function. In practice, only a small set of link functions are actually utilized. In particular, links are chosen such that the *inverse link* $\mu_i = g^{-1}(\eta_i)$ is easily computed, and so that g^{-1}maps from $X_i\beta = \eta_i \in \square$ into the set of admissible values for μ_i. A log link is usually used for the Poisson model, since while $\eta_i = g(\mu_i) \in \square$, because Y_i is a count, we have $\eta_i \in 0,1 \ldots$. For binomial data, the link function maps from $0 < \mu_i < 1$ to $\eta_i \in \square$.

Examples of link functions that are used are the identity, log, inverse, logit, probit, log -log, complementary log – log, etc. Note that binary data are handled in the GLM framework as special cases of binomial data.

3.4.3 Poisson Regression model

The Poisson regression model is a technique used to describe count data as a function of a set of predictor variables. It is often applied to study the occurrence of small number of counts or events as a function of a set of predictor variables, in experimental and observational study in

many disciplines, including Economy, Demography, Psychology, Biology and Medicine (Gardener et al, 1995). The Poisson regression model may be used as an alternative to the Cox model for survival analysis, when hazard rates are approximately constant during the observation period and the risk of the event under study is small (*e.g.*, incidence of rare diseases). In spite of its recent wide application, Poisson regression model remains partly poorly known, especially if compared with other regression techniques, like linear, logistic and Cox regression models.

The Poisson regression model assumes that the sample of n observations y_i are observations on independent Poisson variables Y_i with mean μ_i. Note that, if this model is correct, the equal variance assumption of classic linear regression is violated, since the Y_i have means equal to their variances. So we fit the generalized linear model,

$$\log(\mu_i) = x_i'\beta \tag{3.37}$$

We say that the Poisson regression model is a generalized linear model with Poisson error and a log link so that

$$\mu_i = exp(x_i'\beta) \tag{3.38}$$

This implies that one unit increases in an X_j are associated with a multiplication of μ_j by $exp(\beta_j)$.

3.4.4 Specification of the model

The primary equation of the model is

$$\Pr(Y_i = y_i) = \frac{e^{\mu}\mu_i{}^{y_i}}{y_i!}, y_i = 0,1,2 \tag{3.39}$$

The most common formulation of this model is the log-linear specification:

$$\log(\mu_i) = x_i'\beta \tag{3.40}$$

The expected number of events per period is given by

$$E(y_i/x_i) = \mu_i = e^{x_i'\beta} \tag{3.41}$$

Thus:

$$\frac{dE(y_i/x_i)}{dx_i} = \beta e^{x_i'\beta} = \beta_i \mu_i \tag{3.42}$$

The major assumption of Poisson model is

$$E(y_i/x_i) = \mu_i = e^{x_i'\beta} = Var(y_i/x_i) \tag{3.43}$$

3.4.5 Count Regression for Rate Data

When events occur over time, space, or some other index of size, models can focus on the rate at which the events occur. When a response count Y has index (such as population size) equal to t , the sample rate is $/t$. The expected value of the rate is μ/t , where $\mu = E(Y)$. A log linear model for the expected rate has form

$$\log\left(\frac{\mu}{t}\right) = \alpha + \beta_i X_i \quad \text{Where } i = 1,2,3,\ldots..i \tag{3.44}$$

This model has equivalent representation

$$\log\mu - logt = \alpha + \beta_i X_{ij} \ . \qquad \text{Where } i = 1,2,3,\ldots..i \ \ and \ j = 1,2,3,4. \tag{3.45}$$

The adjustment term, $-logt$ to the log of the mean is called an ***offset***. Standard GLM software can fit a model having an offset term. For such log linear model, the expected number of outcomes satisfies

$$\mu = t\exp(\alpha + \beta_i X_i) \tag{3.46}$$

The mean μ is proportional to the index t, with proportionality constant depending on the value of the explanatory variable. For a fixed value of X, for example, doubling the population size t also doubles the expected number of murders μ.

3.4.6 Estimation

Estimation involves estimating the regression parameters specifically using the maximum likelihood estimation

- **Maximum Likelihood Estimation**

The likelihood function for n independent Poisson observations is a product of probabilities given by

$$prob(y_i) = \frac{e^{-\lambda_i}\lambda_i{}^{y_i}}{y_i!} \quad (3.47)$$

Taking logs and ignoring a constant involving$\log(y_i!)$, we find that the log-likelihood function is

$$logL(\beta) = \sum_{i-1}^{n}[-\lambda_i + y_i x_i'\beta - log\ y_i!] = \sum_{i-1}^{n}\left[-e^{x_i'\beta} + y_i x_i'\beta - log\ y_i!\right] \quad (3.48)$$

Where $\lambda_i = \mu_i = e^{x_i'\beta}$

The parameters of this equation can be estimated using maximum likelihood method

$$\frac{\partial L}{\partial \beta} = \sum_{i-1}^{n}(y_i - e^{x_i'\beta})\, x_i = 0 \quad (3.49)$$

And

$$\frac{\partial^2 L}{\partial \beta \partial\, \beta'} = -\sum_{i-1}^{n}[e^{x_i'\beta} x_i' x_i], \quad (3.50)$$

which is the Hessian of the function and with typical element

$$\frac{\partial^2 L}{\partial \beta_j \partial \beta_l} = -\sum_{i-1}^{n}\left[e^{x_i'\beta} x_{ij}' x_{il}\right]; j, l = 1,2, \dots p \quad (3.51)$$

As $\frac{\partial^2 L}{\partial \beta_j \partial \beta_l} = -\sum_{i-1}^{n}\left[e^{x_i'\beta} x_{ij}' x_{il}\right]$ does not involve the y data

$$k_{jl} = (\frac{\partial^2 L}{\partial \beta_j \partial \beta_l}) = -\sum_{i-1}^{n}\left[e^{x_i'\beta} x_{ij}' x_{il}\right]; j, l = 1,2, \dots p \quad (3.52)$$

And the information matrix is

$$K = \sum_{i-1}^{n}[e^{x_i'\beta} x_i' x_i] \quad (3.53)$$

There is no closed form solution to $\frac{\partial L}{\partial \beta} = \sum_{i-1}^{n}(y_i - e^{x_i'\beta})\, x_i = 0$, so the MLE for β must be obtained numerically.

However, as the Hessian is negative definite for all x and β, the MLE ($\hat{\beta}$) is unique, if it exists. From $\frac{\partial^2 L}{\partial \beta_j \partial \beta_l} = -\sum_{i-1}^{n}\left[e^{x_i'\beta} x_{ij}' x_{il}\right]$ and $k_{jl} = E(\frac{\partial^2 L}{\partial \beta_j \partial \beta_l}) = -\sum_{i-1}^{n}\left[e^{x_i'\beta} x_{ij}' x_{il}\right]$;

$$k_{jlr} = (\frac{\partial^2 L}{\partial \beta_j \partial \beta_l \partial \beta_r}) = -\sum_{i-1}^{n}\left[e^{x_i'\beta} x_{ij}' x_{il} x_{ir}\right]; j, l, r = 1,2, \dots p \quad (3.54)$$

And

$$k_{jl}^{(r)} = [\frac{\partial k_{jl}}{\partial \beta_r}] = -\sum_{i-1}^{n}\left[e^{x_i'\beta} x_{ij}' x_{il} x_{ir}\right]; j, l, r = 1,2, \dots p \quad (3.55)$$

To make matters more transparent, consider the case of a single covariate and an intercept. Then x_i is a scalar observation and

$$l = \sum_{i-1}^{n}[-\lambda_i + y_i(\beta_1 + \beta_2 x_i) - \log(y_i)] \quad (3.56)$$

Where $\lambda_i = exp(\beta_1 + \beta_2 x_i), for\ i = 1,2, \dots, n$

The first order conditions, $\frac{\partial L}{\partial \beta} = 0$ yield a system of K equations (one for each β) of the form

$$\sum_{i-1}^{n}(y_i - e^{x_i'\beta})\, x_i = 0$$

Where $\hat{y}_i = e^{x_i'\hat{\beta}}$ is the fitted value of y_i. The predicted/fitted value has as usual been taken as the estimated value of $E(y_i/x_i)$. This first order condition tells us that the vector of residual is orthogonal to the vectors of explicative variables.

- **Fisher Scoring in Log - Linear Models**

Fisher scoring algorithm is a form of Newton-Raphson method used in statistics to solve maximum likelihood equations numerically. Nelder and Wedderburn (1972) applied Fisher scoring algorithm to estimate $\hat{\beta}$ in generalized linear models. The Fisher scoring algorithm for Poisson regression models with canonical link would be considered, where it would be modelled as:

$$g(\eta_i) = \log(\mu_i) \tag{3.57}$$

The derivative of the link is easily seen to be

$$g'(\eta_i) = \frac{1}{\mu_i} \tag{3.58}$$

Specifically, given an initial estimate β, the algorithm updates it to β^{new} by

$$\beta^{new} = \beta + \{E(-\frac{\partial L}{\partial \beta \partial \beta^T})\}^{-1} \frac{\partial L}{\partial \beta} \tag{3.59}$$

where both derivatives are evaluated at β, and the expectation is evaluated as if β were the true parameter values.

β is then replaced by β^{new} and the updating is repeated until convergence.

It can be shown that for a GLM, the updating equation (3.54) can be rewritten as

$$\beta^{new} = \beta + (X^T W X)^{-1} X^T W z \tag{3.60}$$

where z is the n-vector with i th component

$z_i = (Y_i - \mu_i)g'(\mu_i)$

and W is the $n \times n$ diagonal matrix with

$W_i = \{g'(\mu_i)^2 b''(\theta_i)\}^{-1}$ $W_i =$

$(\mu_i . \frac{1}{\mu_i^2})^{-1}$

And this simplifies to

$$W_i = \mu_i \quad (3.61)$$

It is noted that the weight is inversely proportional to the variance of the working dependent variable.

- **Likelihood Ratio Test**

A simple test on the overall fit of the model, as an analogue to the F-test in the classical regression model is a Likelihood Ratio test on the "slopes". The model with only the intercept is nothing but the mean of the counts, or

$$\lambda_i = \bar{y}\forall \quad (3.62)$$

Where $\bar{y} = \sum_{i-1}^{n} \frac{y_i}{n}$

The corresponding log-likelihood is:

$$L_R = n\bar{y} + \log(\bar{y})\,[\sum_{i-1}^{n} y_i] - \sum_{i-1}^{n} log\, y_i! \quad (3.63)$$

where the R stands for the "restricted" model, as opposed to the "unrestricted" model with $K-1$ slope parameters. The last term in $\sum_{i-1}^{n} log\, y_i!$ can be dropped, as long as it is also dropped in the calculation of the maximized likelihood $L = \sum_{i-1}^{n} \left[-e^{x_i'\beta} + y_i x_i'\beta - log\, y_i!\right]$ for the unrestricted modelL_R, using $L = e^{x_i'\hat{\beta}_i}$. The Likelihood Ratio test is then:

$$LR = 2(L_U - L_R) \quad (3.64)$$

and follows a χ^2 distribution with K-1 degrees of freedom.

3.4.7 Goodness of Fit Test

In order to assess the adequacy of the Poisson regression model you should first look at the basic descriptive statistics for the event count data. If the count mean and variance are very different (equivalent in a Poisson distribution) then the model is likely to be over-dispersed. The model analysis option gives a scale parameter (sp) as a measure of over-dispersion; this is equal to the Pearson chi-square statistic divided by the number of observations minus the number of parameters (covariates and intercept).

The variances of the coefficients can be adjusted by multiplying by sp. The goodness of fit test statistics and residuals can be adjusted by dividing by sp. Using a quasi-likelihood approach sp could be integrated with the regression, but this would assume a known fixed value for sp, which is seldom the case. A better approach to over-dispersed Poisson models is to use a parametric alternative model, the negative binomial.

The deviance (likelihood ratio) test statistic, G^2, is the most useful summary of the adequacy of the fitted model. It represents the change in deviance between the fitted model and the model with a constant term and no covariates; therefore G^2 is not calculated if no constant is specified. If this test is significant then the covariates contribute significantly to the model.

The deviance goodness of fit test reflects the fit of the data to a Poisson distribution in the regression. If this test is significant then a red asterisk is shown by the P value, and you should consider other covariates and/or other error distributions such as negative binomial.

Technical validation

The deviance function is:

$$Deviance = 2\sum_{i-1}^{n} y_i \, ln\left[\frac{y_i}{\hat{\mu}_i}\right] - (y_i - \hat{\mu}_i) \tag{3.65}$$

where y is the number of events, n is the number of observations and $\hat{\mu}_i$ is the fitted Poisson mean. The first term is identical to the binomial deviance, representing `twice' a sum of observed times log of observed over fitted'. The second term, a sum of differences between observed and fitted values, is usually zero, because MLE's in Poisson models have the property of reproducing marginal totals, as noted above.

The log-likelihood function is:

$$L = \sum_{i-1}^{n} y_i \ln(\hat{\mu}_i) - \hat{\mu}_i - \ln(y_i!) \tag{3.66}$$

The maximum likelihood regression proceeds by iteratively re-weighted least squares, using singular value decomposition to solve the linear system at each iteration, until the change in deviance is within the specified accuracy.

The Pearson chi-square residual is:

$$r_p = \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i} \tag{3.67}$$

For large samples the distribution of the deviance is approximately a chi- squared with $n - p$ degrees of freedom, where n is the number of observations and p the number of parameters. Thus, the deviance can be used directly to test the goodness of fit of the model. An alternative measure of goodness of fit is Pearson's chi-squared statistic, which is defined as

The Pearson goodness of fit test statistic is:

$$\chi^2 = \sum_{i-1}^{n} \frac{y_i - \hat{\mu}_i}{\sqrt{\hat{\mu}_i}} \quad (3.68)$$

The deviance residual is (Cook and Weisberg, 1982):

$$r_d = sign(y_i - \hat{\mu}_i)\sqrt{deviance(y_i, \hat{\mu}_i)} \quad (3.69)$$

The standardized residual is:

$$r_s = \frac{y_i - \mu_i}{\sqrt{1-h_i}} \quad (3.70)$$

where h is the leverage (diagonal of the Hat matrix).

- **Akaike Information Criterion (AIC)**

The Akaike Information Criterion (AIC) is a way of selecting a model from a set of models. The chosen model is the one that minimizes the Kullback-Leibler distance between the model and the truth. It's based on information theory, but a heuristic way to think about it is as a criterion that seeks a model that has a good fit to the truth but few parameters. It is defined as:

AIC = -2 (ln (likelihood)) + 2 K (3.71)

where likelihood is the probability of the data given a model and K is the number of free parameters in the model. AIC scores are often shown as ΔAIC scores, or difference between the best model (smallest AIC) and each model (so the best model has a ΔAIC of zero).

The second order information criterion, often called AICc, takes into account sample size by, essentially, increasing the relative penalty for model complexity with small data sets. It is defined as:

AICc = -2 (ln (likelihood)) + 2 K * (n / (n - K - 1)) (3.72)

where n is the sample size. As n gets larger, AICc converges to AIC (n - K -1 -> n as n gets much bigger than K, and so (n / (n - K - 1)) approaches 1), and so there's really no harm in always using AICc regardless of sample size. In model selection in comparative methods, sample size often refers to the number of taxa (Butler and King, 2004; O'Meara et al., 2006).

3.5 STATISTICAL SOFTWARE TO BE USED

The Statistical Analysis software (SAS) would be used in examining the occurrence and incidence of maternal mortality and fitting the Poisson model while the R-Consol statistical software would be used in analysing and fitting ARIMA models. Other statistical and numerical simulation methods of parameter estimation would be utilized as and when deemed fit.

.

CHAPTER 4

RESULTS AND DISCUSSION

4.0 PRELIMINARY ANALYSIS

The data used in this thesis is the monthly observations of Maternal Mortality cases and total deliveries as well as the quarterly maternal mortality ratios at the Okomfo Anokye Teaching Hospital from 2000 to 2010. The data is obtained from the Bio-Statistics Department of the Obstetrics & Gynaecology directorate. Firstly, the raw data is plotted and the patterns of maternal deaths and deliveries for the facility over the eleven (11) years period under study are observed. Figure 4.0 presents the observed bar plot of the deliveries data while Figure 4.1 presents the observed bar plot of the maternal mortality data.

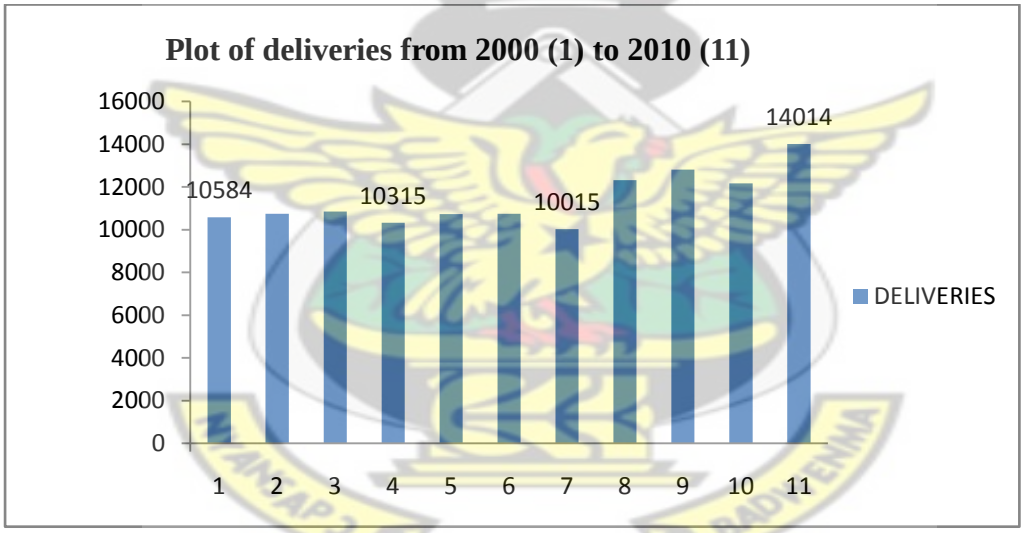


Figure 4.0: bar Plot of deliveries from 2000 to 2010

From Fig. 4.0, deliveries at the hospital appear to be in the range of 10584 from year one (2000) and 10315 mothers in a year until year seven (2006) when it dropped to 10015, the lowest within the eleven years period. Total deliveries went up in year eight (2007) through to the eleventh year (2010) which recorded the highest deliveries of 14014. From Fig. 4.1 also, the

hospital recorded 98 maternal deaths in the first year (2000), increased marginally to 106 and 107 in the following two years (2001 and 2002).

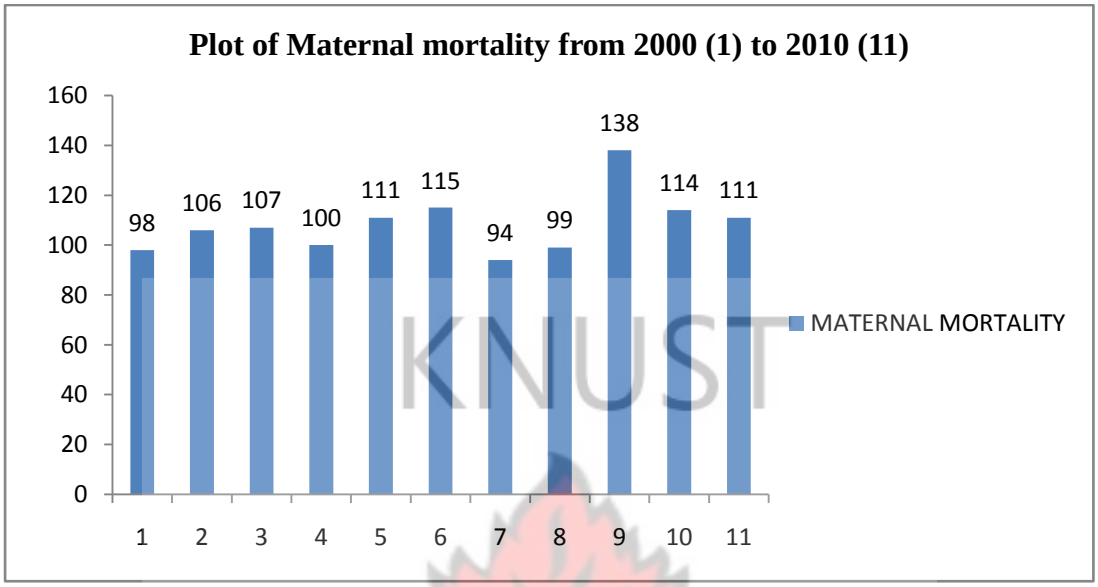


Figure 4.1, bar Plot of maternal mortality from 2000 to 2010

The drop in 2003 was just for a while as the deaths rise again in 2004, 2005 until 2006 and 2007 where the least deaths of 94 and 99 respectively were recorded. The highest maternal mortality figure of 138 was recorded in the ninth year (2008). The deaths then begin to decrease from 114 in 2009 to 111 in the eleventh year (2010).

4.1 MODELLING THE OCCURRENCE OF MATERNAL MORTALITY CASES

SAS statistical package version 9.1 was used in modelling the occurrence of maternal mortality cases. The distribution specified was Poisson with a log-link function. In all, 132 observations were used in 11 years period with maternal death cases as the response variable and time as the only predictor variable. The following model was obtained;

$$\log(mean\_Death) = \alpha + \beta * year_j \qquad \text{Where j = 2000, ... 2010} \qquad (4.0)$$

This implies that

$$mean\_Death = exp(\alpha + \beta * year_j) = e^{\alpha}.e^{\beta *^{year}} \quad (4.1)$$

The fitted model had deviance of 119.0390 which follows a Chi-square distribution with 121 degrees of freedom. The ratio of the deviance and the degrees of freedom is found to be 0.9838 indicating a clear absence of over dispersion in the data. The scaled deviance and the scaled Pearson chi-square had 119.03 and 116.97 values respectively and their values had the same degrees of freedom of 121. Detailed analysis of the Criteria for Assessing Goodness of fit is reported as Table 4.1 in appendix A.

Table 4.2 Analysis of Parameter Estimates for occurrence of maternal mortality

Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept		0	0.0000	0.0000	0.0000	0.0000	.	.
YEAR	2000	1	2.1001	0.1010	1.9021	2.2980	432.21	<.0001
YEAR	2001	1	2.1785	0.0971	1.9882	2.3689	503.08	<.0001
YEAR	2002	1	2.2156	0.0953	2.0287	2.4024	539.96	<.0001
YEAR	2003	1	2.1203	0.1000	1.9243	2.3163	449.55	<.0001
YEAR	2004	1	2.2246	0.0949	2.0386	2.4107	549.33	<.0001
YEAR	2005	1	2.2600	0.0933	2.0773	2.4428	587.39	<.0001
YEAR	2006	1	2.0584	0.1031	1.8562	2.2605	398.27	<.0001
YEAR	2007	1	2.1102	0.1005	1.9132	2.3072	440.85	<.0001
YEAR	2008	1	2.4423	0.0851	2.2755	2.6092	823.18	<.0001
YEAR	2009	1	2.2513	0.0937	2.0677	2.4349	577.79	<.0001
YEAR	2010	1	2.2246	0.0949	2.0386	2.4107	549.33	<.0001
Scale		0	1.0000	0.0000	1.0000	1.0000		

Table 4.2 represents the detailed analysis of the parameter estimates of the model. Time in years was treated as discrete variable with no intercept ($\alpha = 0$). We see clearly that the occurrence of maternal death cases were very significant throughout all the years. The parameters for the years were all positive. This indicates that the mean number of occurrence of

maternal death cases were high for all the years considered. However, the estimates are found to be approximately the same over the eleven years period. The year mean number of deaths for all the months in 2000 ($e^0 e^{2.1}$) = 8.166987, that of 2005 is 9.58309 while the highest mean number of maternal deaths of ($e^0 e^{2.4423}$) = 11.49946 was recorded in 2008. Also, the least mean number of maternal deaths of ($e^0 e^{2.0584}$) = 7.833426 was recorded in 2006.

From table 4.3 (see appendix), we observe that, there exists no significant occurrence of maternal death cases from 2000 to 2005 as well as from 2005 to 2010. The analysis of parameter estimates for the continuous time is also found to be insignificant (see tables 4.4 and 4.5 in appendix A). We can therefore establish that the mean number of occurrence of maternal death cases has not significantly reduced over the period 2000 to 2010

4.2 MODELLING THE INCIDENCE OF MATERNAL MORTALITY CASES

Again, SAS statistical package version 9.1 was used in modelling the incidence of maternal mortality cases. The distribution specified was Poisson with a log-link function. The offset variable used was the log of the total deliveries (population at risk). In all, 132 observations were used in the 11 years period with maternal death cases as the response variable and time as the only predictor variable. The model obtained is as below;

$$\log\left(\frac{mean\ \_Deaths}{Pop}\right) = \alpha + \beta * Year \qquad (4.2)$$

This model has equivalent representation

$\log(mean\ \_Death) - \log(Pop) = \alpha + \beta * Year_j$. Where j= 2000, ..., 2010.

Hence the final model is given as

$$mean\_Maternal\ Death = Pop \exp(\alpha + \beta * year_j); \quad (4.3)$$

For appropriateness of a Poisson model, the model mean and variance should be the same. An objective test for over-dispersion, which follows the Pearson's chi-square, was performed. The underlining condition for an appropriate Poisson is that the dispersion parameter should not be significantly greater than one. From the output, dispersion parameter (DP) was 1.0328 indicating a clear absence of over dispersion in the data. Hence, Poisson model is an appropriate model. Detailed analysis of the Criteria for Assessing Goodness of fit is reported as Table 4.6 in appendix A.

Table 4.7 Analysis of Parameter Estimates for incidence of maternal mortality

Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept		1	-4.8383	0.0949	-5.0243	-4.6523	2598.40	<.0001
YEAR	2000	1	0.1562	0.1386	-0.1155	0.4278	1.27	0.2599
YEAR	2001	1	0.2199	0.1358	-0.0463	0.4861	2.62	0.1054
YEAR	2002	1	0.2477	0.1345	-0.0160	0.5114	3.39	0.0656
YEAR	2003	1	0.2021	0.1379	-0.0681	0.4723	2.15	0.1427
YEAR	**2004**	**1**	**0.2667**	**0.1342**	**0.0036**	**0.5298**	**3.95**	**0.0469**
YEAR	**2005**	**1**	**0.3012**	**0.1331**	**0.0404**	**0.5620**	**5.12**	**0.0236**
YEAR	2006	1	0.1697	0.1402	-0.1050	0.4445	1.47	0.2259
YEAR	2007	1	0.0146	0.1382	-0.2564	0.2855	0.01	0.9160
YEAR	**2008**	**1**	**0.3078**	**0.1275**	**0.0579**	**0.5577**	**5.83**	**0.0158**
YEAR	2009	1	0.1681	0.1333	-0.0933	0.4294	1.59	0.2075
YEAR	2010	0	0.0000	0.0000	0.0000	0.0000	.	.
Scale		0	1.0000	0.0000	1.0000	1.0000		

Table 4.7 represents the detailed analysis of the parameter estimates of the model. The value of the intercept estimate was -4.8383. Positive parameters for 2000 to 2009 indicates that the

mean number of incidence of maternal deaths were all high compared to the mean of the year 2010 which happens to be the referenced year. The table also presented values of the standard errors, confident limits, the Wald chi-square and p-values of individual parameter estimates. The results show that there was a statistically significant difference between year 2010 (the referenced year) and years 2004, 2005 and 2008. Their chi-square values were 3.95, 5.12 and 5.83 with p-values of 0.0469, 0.0236 and 0.0158 respectively. Hence compared to year 2010, the rate of maternal death was significantly high in 2004, 2005 as well as 2008. For instance in 2004, 2005 and 2008, maternal death rate increased by $(e^{-4.8383}e^{0.2667})$ = 0.0103414, $(e^{-4.8383}e^{0.3012})$ = 0.01070440 and $(e^{-4.8383}e^{0.3078})$ = 0.01077529 respectively compared to 2010. In 2008, government showed much commitment to addressing the challenges of maternal health, specifically the problem of low coverage of supervised deliveries and high institutional maternal mortality rate among others. Maternal mortality was therefore declared a national emergency in 2008 (2008 Ghana Millennium Development Goals report. April, 2010) and the programme of free health care for pregnant women, including deliveries through the National Health insurance Scheme, was implemented since July 2008. Okomfo Anokye Teaching Hospital also experienced the nationwide increase in supervised deliveries rates hence the rise in maternal death rates. It is however obvious that mean incidence of maternal deaths at the facility has neither reduced nor increased over the period of time under study (see Table 4.8 in Appendix A). With chi-square value of 1.67 and p-value of 0.1963, we can conclude that statistically the mean rate of maternal death cases is not significant over the period of time under study.

4.3 PATTERNS OF MATERNAL MORTALITY RATIOS (MMR)

The data used in this section of the thesis is quarterly observations of Maternal Mortality Ratios at the Okomfo Anokye Teaching Hospital from 2000 to 2010. The data is obtained from the

Bio-Statistics Department of the Obstetrics & Gynaecology directorate. Firstly, the raw quarterly data are plotted and the patterns of MMR for the facility over the period under study are observed. Figure 4.2 presents the observed time series plot of the data

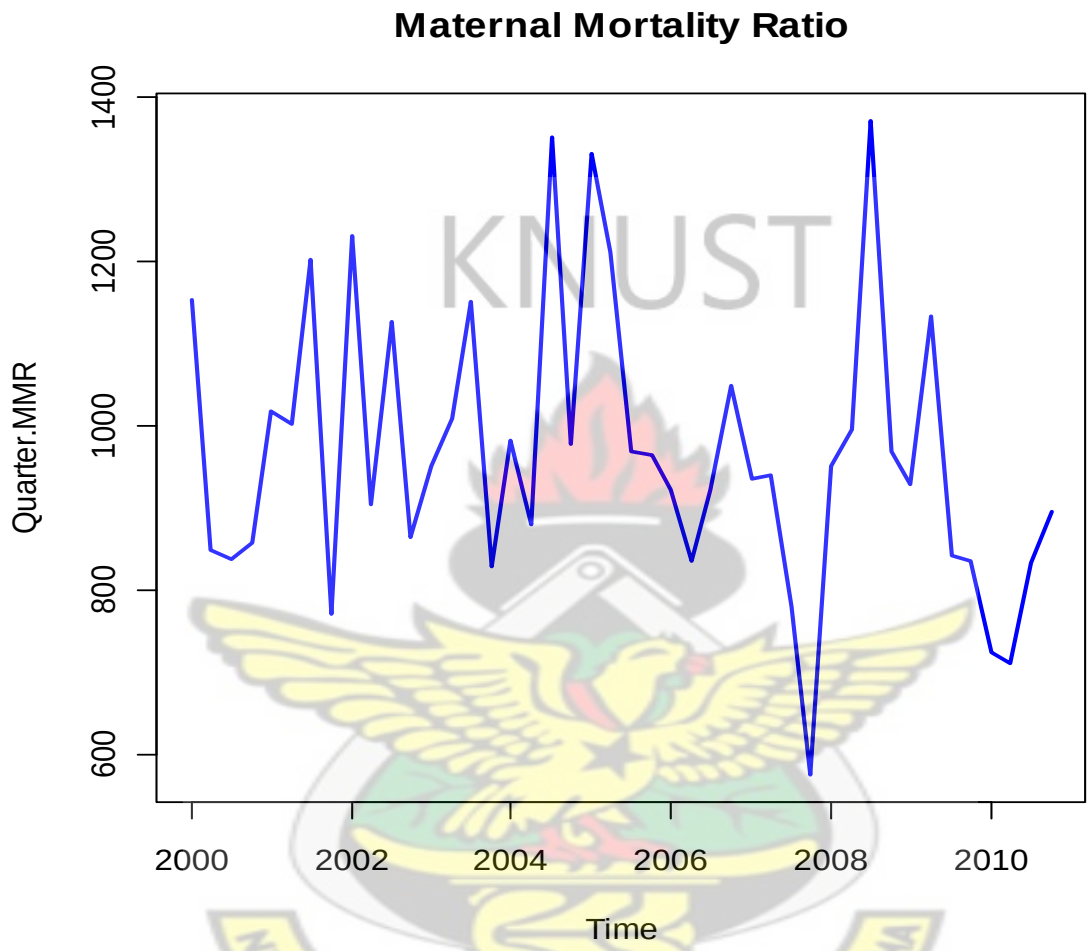


Fig. 4.2 plot of MMR patterns from 2000 to 2010

We realized from Figure 4.1 above that, the quarterly maternal mortality ratios recorded from 2000 to 2010 shows no particular trend, and behaves irregularly. The hospital recorded an appreciably high MMR of about 1152.9 per 100,000 live births in the first quarter of year 2000 but ended that year with mortality ratio of about 858.2 per 100,000 live births. The first three quarters of the following year also recorded high MMR figures of 1017.4, 1002.0 and 1203.0 per 100,000 live births respectively till the last quarter when it dropped sharply to 770.4 per

100,000 live births. Generally, all other years with the exception of 2000, 2001, 2002 and 2005 recorded relatively low MMRs in their first quarters ranging from 923.2 in 2006 to 724.0 in 2010. Maternal mortality ratios declined steadily from the third quarter of 2005 and increased marginally at the last quarter of 2006. The most significant decrease was recorded in 2007 and 2008. Unfortunately, in 2009 instead of continuous decline, the MMR was on the rise. However, it appears to decline in 2010.

The 2008 Ghana Millennium Development Goals report (April, 2010) reveals that by 2005, a number of initiatives were being implemented to positively affect maternal health outcomes including increased production of midwives through direct midwifery training (24.6% increase in enrolment), two new midwifery training schools opened in Tamale and Tarkwa and the implementation of free maternal health services. Antenatal care from health professional (i.e. nurse, doctor, and midwife or community health officer) increased from 82% in 1988 to 95% in 2008. This in no doubt accounts for the declines recorded from the third quarter of 2005. The highest MMR recorded by the hospital was 1373 per 100,000 live births and this was recorded in third quarter of 2008 while the lowest MMR was 574.5 per 100,000 live births, recorded in last quarter of 2007. The average quarterly MMR recorded within the period was 967.7 per 100,000 live births which is far higher than result from the Ghana Maternal Mortality Survey of 2008. The survey showed a slow decline of maternal deaths from 503 per 100,000 live births in 2005 to 451 per 100,000 live births in 2008, which is an average estimate for the seven-year period preceding the 2008 survey.

4.4 MODEL IDENTIFICATION

The model development process was begun by studying the original plot, ACF, PACF and objective test of the raw data to be sure that it is stationary. There are two relevant features from Fig. 4.1. First is that the mean appears to be stationary over the time period. Secondly,

with the exception of the extreme case at the fourth quarter of 2008, the rise and fall of the dispersion over the time period is quite stable. If the mean was changing, the trend is removed by differencing once or twice and if the variability was changing, the process may be made stationary by logarithmic transformation. However, as it stands now the data is said to be stationary in mean and in variance. Also, Kwiatkowski-Phillips-Schmidt-Shin (KPSS) and Augmented Dickey- Fuller (ADF) test were performed. KPSS test is used for verifying whether or not the series is stationary, while Augmented Dickey-Fuller test is used for verifying whether or not there is unit root. The p value of the KPSS test is greater than printed p-value (0.01), so it accepts the null hypothesis that data is level or trend stationary. This indicates that we may regard the time series to be stationary. While the p value of ADF test is smaller than printed p-value, so it rejects the null hypothesis that data has a unit root. From above results, we find the series to be stationary and there is no unit root. We also plot the graphs for sample autocorrelations function and sample partial autocorrelations function.

Table 4.9: Objective test (unit root test) for drift and trend stationarity of quarterly MMR's

	CONSTANT		CONSTANT + TREND	
TEST TYPE	Test statistic	Critical value	Test statistics	Critical value
ADF	-3.1419	-2.93	-3.3335	-3.5
KPSS	0.2503	0.463	0.0827	0.146

Figure 4.3 below, consists of plots of the ACF and the PACF for the quarterly Maternal Mortality Ratio from 2000 to 2010. 95% confidence brands are plotted in colour blue on the both panels. These two plots are useful in determining the p autoregressive terms and the q lagged error terms. Looking at the sample ACF and PACF plot of the series in Figure 4.3, we apply the Box-Jenkins approach to choose the value p and q by ACF and PACF plot. From PACF plot, it appears no significant spikes exist at 95% confidence. However, there exists a model at higher percentage intervals such as 99% confidence interval. Generally, we build an

AR (p) and compare the AIC, AICc and BIC of all the possible models and find out a model to fit the data better than others, which is the one has the lowest AIC, AICc and BIC values.

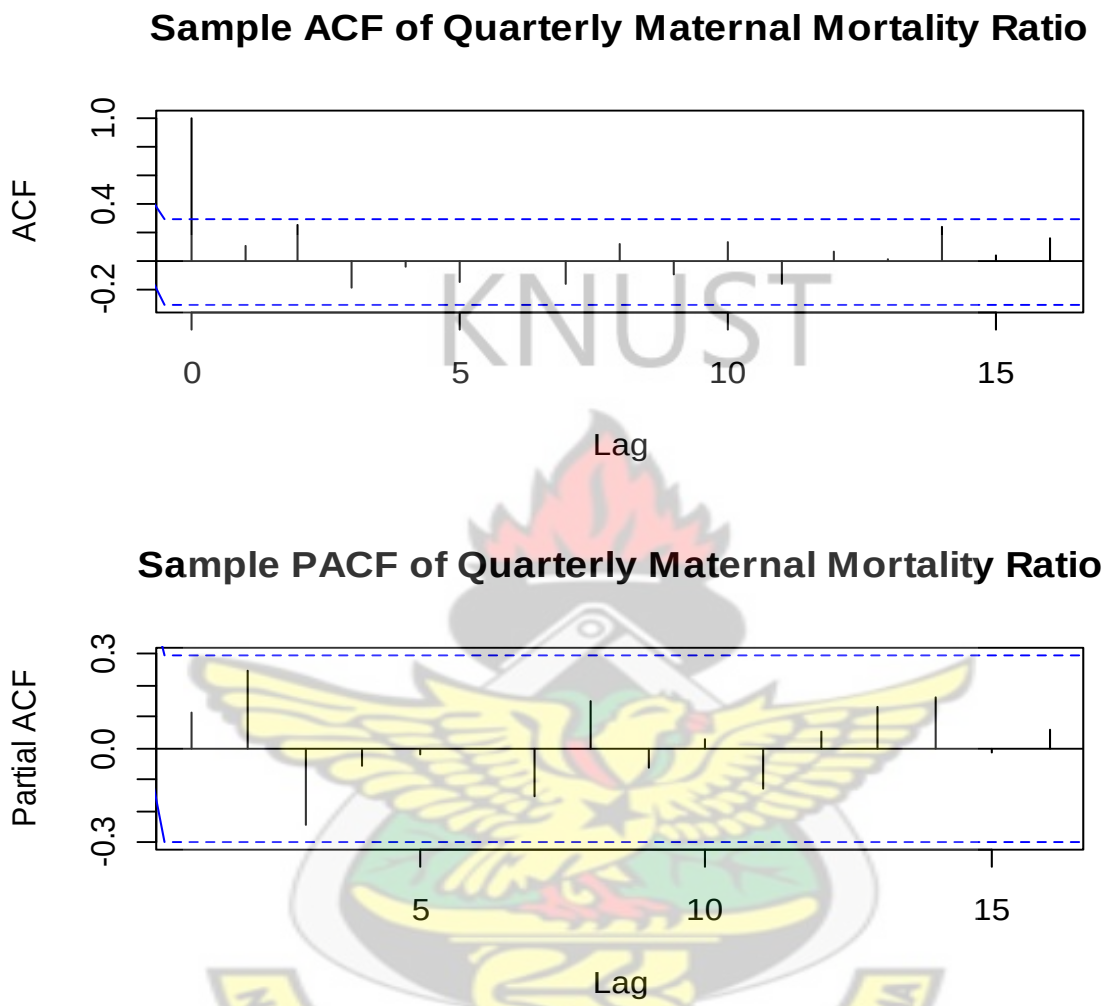


Fig. 4.3 ACF and PACF of quarterly MMRs from 2000 to 2010

4.5 MODEL ESTIMATION AND EVALUATION

The procedure for choosing these models relies on choosing the model with the minimum AIC, AICc and BIC. The models are presented in Table 4.5 below with their corresponding values of AIC, AICc and BIC. Among those possible models, comparing their AIC, AICc and BIC as

shown in Table 4.10, ARIMA (1, 0, 2) and ARIMA (2, 0, 1) were chosen as the appropriate model that fit the data well.

Table 4.10: AIC, AICc and BIC for the Suggested ARIMA Models

MODEL	AIC	AICc	BIC
ARIMA (1,0,0)	582.64	583.24	588.0
ARIMA (0,0,1)	582.84	583.44	588.19
ARIMA (1,0,1)	584.12	585.14	591.25
ARIMA (1,0,2)	**581.41**	**582.99**	**590.34**
ARIMA (2,0,1)	**581.41**	**582.99**	**590.34**
ARIMA (2,0,2)	583.35	585.62	594.05

From the two models, using the method maximum likelihood the estimated parameters of the models with their corresponding standard error is shown in the Table 4.11 and 4.12 below.

Table 4.11: Estimate of Parameters for ARIMA (2, 0, 1)

Variable	Estimate	Standard Error
AR(1)	**-0.6867**	**0.1754**
AR(2)	**0.2788**	**0.1498**
MA(1)	**0.8794**	**0.1262**
	σ^2 = 25218	

At 95% confidence level it can be seen that the MA (1) parameter for ARIMA (2, 0, 1) model is not significant different from one. Its estimate therefore fails to satisfy the invertibility condition. Hence after removing the MA (1) term from the model, we then settle on ARIMA (1, 0, 2) model

Table 4.12: Estimate of Parameters for ARIMA (1, 0, 2)

Variable	Estimate	Standard Error
AR(1)	**-0.308**	**0.356**
MA(1)	**0.4525**	**0.3246**
MA(2)	**0.3931**	**0.1452**
	σ^2 = 25369	

Therefore at 95% confidence level, we conclude that all the coefficients of the ARIMA(1,0,2) model are significantly different from zero and the estimated values satisfy the stability condition.

4.6 Goodness of fit

In time series modelling, the selection of a best model fit to the data is directly related to whether residual analysis is performed well. One of the assumptions of ARIMA model is that, for a good model, the residuals must follow a white noise process. That is, the residuals have zero mean, constant variance and also uncorrelated.

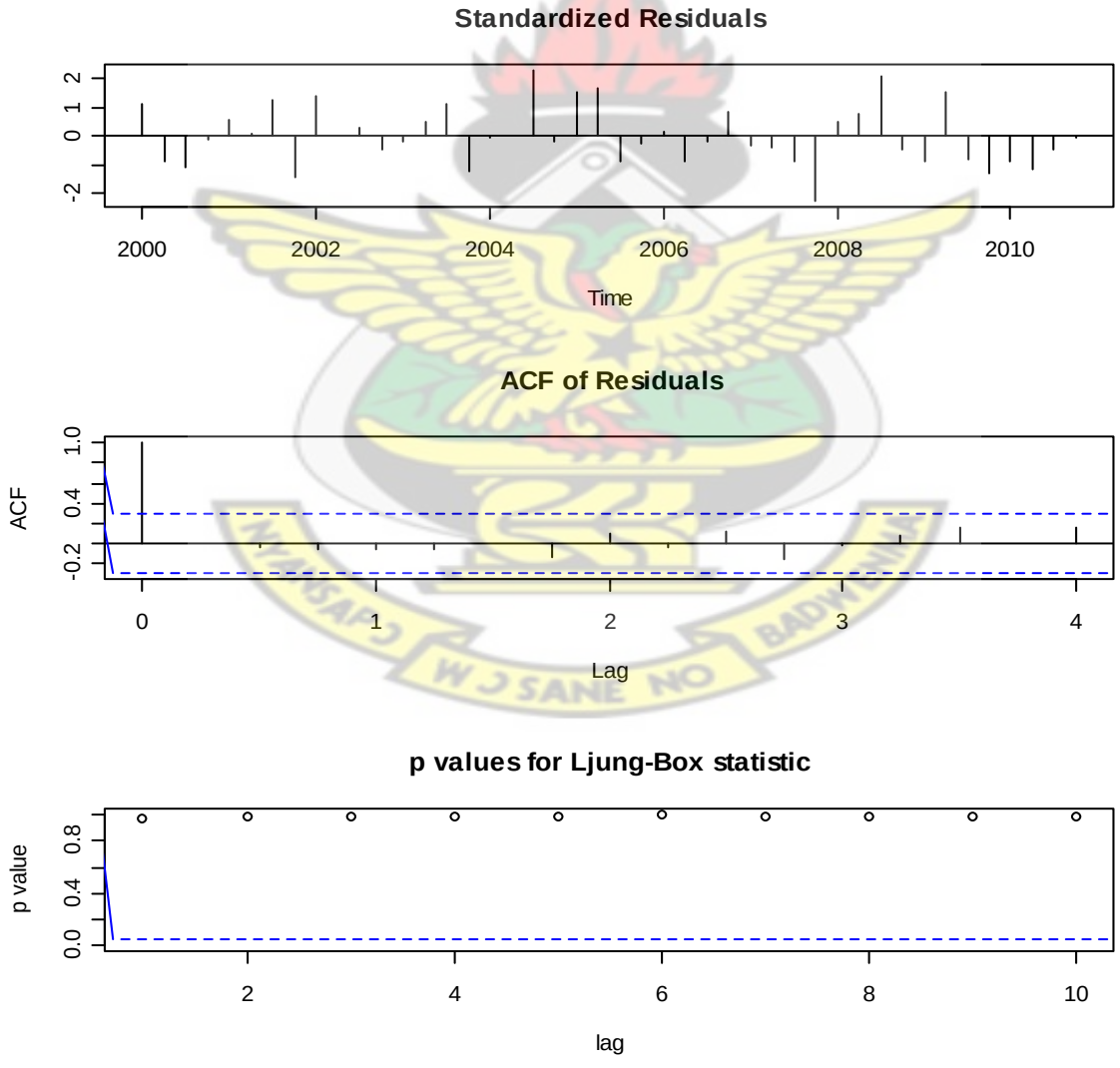


Fig. 4.4 plot of model residuals for quarterly MMRs from 2000 to 2010

From Figure 4.4 above, the standardized residual reveals that the residuals of the model have zero mean and constant variance. Also the ACF of the residuals depicts that the autocorrelation of the residuals are all zero, that is to say they are uncorrelated. Finally, the p-values for the Ljung-Box statistic in the third panel all clearly exceed 5% for all lag orders, indicating that there is no significant departure from white noise for the residuals. Thus, the selected model satisfies all the model assumptions. Since our model ARIMA (1, 0, 2) satisfies all the necessary assumptions, we can say that the model provide an adequate representation of the data. We therefore write our ARIMA (1, 0, 2) as:

$$X_t = 969.22 - 0.308X_{t-1} + 0.4525\omega_{t-1} + 0.3931\omega_{t-2} + \omega_t \quad (4.4)$$

4.7 FORECASTING

Forecasting is the process of estimation in unknown situations, which is commonly used in discussion of time-series data. It is a planning tool which helps decision makers to foresee the future uncertainty based on the behaviour of past and current observations. For intuitive notion, short-term forecasting should be more reliable than long-term forecasting. Using the model obtained above, we forecast 2011 to 2015 MMR's and compare to the observed values for 2011 from the hospital, with the statistical software R.

Comparing the predicted MMR for first quarter 2011 with the observed ratios, we can see that the predicted value (959.0) is close to the true value (996.7) recorded and published by the hospital. Also, this observed values fall inside the confidence interval. Hence, we can say that, ARIMA (1, 0, 2) model is adequate to be used to forecast quarterly Maternal Mortality ratios at the Okomfo Anokye Teaching Hospital Kumasi. The Table 4.13 below summarizes the forecasting results of the MMR's over the period 2011 to 2015 with 95% confidence interval.

Table 4.13: ARIMA (1,0,2) Forecasting Results for Quarterly Maternal Mortality Ratios

YEAR	QUARTER	predicted MMR (per 100,000 live births)	Actual MMR (per 100,000 live births)	lower limit	upper limit
	1	**959.0**	996.7	**647**	**1271**
	2	**969.6**		**654**	**1285**
	3	**969.1**		**635**	**1303**
2011	4	**969.3**		**634**	**1305**
	1	**969.2**		**634**	**1305**
	2	**969.2**		**634**	**1305**
	3	**969.2**		**634**	**1305**
2012	4	**969.2**		**634**	**1305**

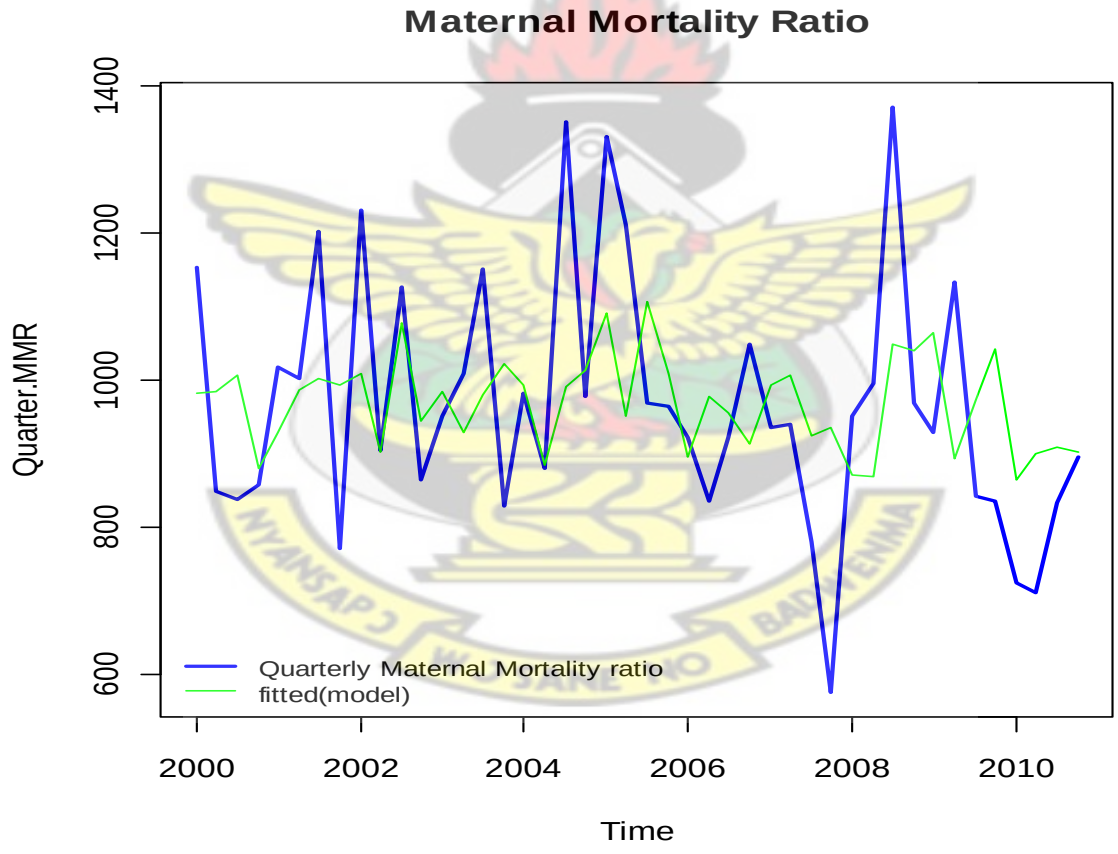


Fig. 4.5 plot of MMR patterns (blue) and fitted values (green) for 2000 to 2010

CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

5.1 CONCLUSION

This study examined the occurrence and incidence of Maternal Deaths as well as maternal mortality ratios at the Okomfo Anokye Teaching Hospital in Kumasi from 2000 to 2010. The study explored the feasibility for application of Poisson models and time series ARIMA in the study of occurrence and incidence of Maternal Deaths and to predict Maternal Mortality ratios respectively so as to provide the theoretical basis for continuing programs to reduce this problem. After careful analysis of data and available literature, we found the following;

In the Poisson regression model for occurrence of maternal mortality, the parameters for the years were all positive. This indicates that the mean number of occurrence of maternal death cases were high for all the years considered. However, the estimates are found to be approximately the same over the eleven years period. The year mean number of deaths for all the months in 2000 ($e^0 e^{2.1}$) is 8.166987, that of 2005 is 9.58309 while the highest mean number of maternal deaths of ($e^0 e^{2.4423}$) 11.49946 was recorded in 2008. Also, the least mean number of maternal deaths of ($e^0 e^{2.0584}$) 7.833426 was recorded in 2006. The analysis of parameter estimates for the continuous time is also found to be insignificant (see tables 4.4 and 4.5 in appendix A). We can therefore establish that the mean number of occurrence of maternal death cases has not significantly reduced over the period 2000 to 2010.

The Poisson regression model for incidence of maternal deaths considered the significant of the rate of maternal mortality in the hospital. It was found that the value of the intercept estimate was -4.8383. Positive parameters for 2000 to 2009 indicates that the mean number of incidence of maternal deaths were all greater than that of the mean of the year 2010

which happens to be the referenced year. The results show that there was a statistically significant difference between year 2010 (the referenced year) and years 2004, 2005 and 2008. Their chi-square values were 3.95, 5.12 and 5.83 with p-values of 0.0469, 0.0236 and 0.0158 respectively. Hence compared to year 2010, the rate of maternal death was high in 2004, 2005 as well as 2008. It is however obvious that mean incidence of maternal deaths at the facility has neither reduced nor increased over the period of time under study (see Table 4.8 in Appendix A). With chi-square value of 1.67 and p-value of 0.1963, we can conclude that statistically the mean rate of maternal death cases is not significant over the period of time under study.

The quarterly maternal mortality ratios recorded from 2000 to 2010 shows no trend in particular, and hence MMR's is relatively stable. The highest MMR recorded by the hospital was 1373 per 100,000 live births and this was recorded in third quarter of 2008 while the lowest MMR was 574.5 per 100,000 live births, recorded in last quarter of 2007. The average quarterly MMR recorded within the period was 967.7 per 100,000 live births which are far higher than result from the Ghana Maternal Mortality Survey of 2008. The survey showed a slow decline of maternal deaths from 503 per 100,000 live births in 2005 to 451 per 100,000 live births in 2008, which is an average estimate for the seven-year period preceding the 2008 survey.

ARIMA (1, 0, 2) model was selected as the appropriate model for predicting future Maternal mortality ratios for the hospital. The model satisfied all conditions of a good ARIMA model and was used to predict MMRs for the next 20 quarters. Comparing the predicted MMR for first quarter 2011 with the observed ratios, we can see that the predicted value (959.0) is close to the true value (996.7) recorded and published by the hospital. Also, this observed values fell within the confidence interval. Hence, we could say that, ARIMA (1, 0, 2) model is adequate to be used to forecast quarterly Maternal Mortality ratios at the Okomfo Anokye Teaching Hospital Kumasi.

5.2 RECOMMENDATIONS

This study found that both the mean number of occurrence and incidence of maternal death cases at the hospital were high for all the years considered. Much more interesting finding of this study is the establishment that the mean number of occurrence of maternal death cases has not significantly reduced over the period 2000 to 2010. We therefore recommend that management of the hospital to put in more efforts at implementing existing intervention programs aimed at reducing this problem.

Also with chi-square value of 1.67 and p-value of 0.1963, the study concluded that statistically, the mean rate of maternal death cases is not significant over the period of time under study. Such revelations are critical for monitoring and evaluating the existing intervention. We therefore recommend to management and government to as a matter of urgency review and evaluate all the existing intervention programs of maternal health since they seem not to have yielded the expected results over the past eleven years (2000 - 2010)

Although the ARIMA model adequately fits the data and is useful for predicting future mortality ratios, it is not recommended for medium and long term predictions. We therefore recommend that future studies look at other models which have the ability to do medium and long term predictions.

Finally, this study limited itself to the analysing maternal mortality cases in the hospital regardless of the cause. We recommend further studies into the actual causes of maternal mortality in the hospital.

REFERENCE

1. Aboagye, P.K., H. Gebreselassie, G.Q. Asare, E.M.H. Mitchell, and J. Addy, (2007). *An assessment of the readiness to offer contraceptives and comprehensive abortion care in the Greater Accra, Eastern and Ashanti regions of Ghana.* Chapel Hill, USA: Ipas.
2. AbouZahr, C. (1998). Maternal mortality overview. Chapter 3 in *Health dimensions of sex and reproduction*, ed. C.J.M. Murray and A. Lopez, Geneva: World Health Organization.
3. Adanu R. M. K., Ntumy, M. N. E., and Tweneboah E. (2005). Profile of women with abortion complications in *Ghana Trop Doct*; 35:139-142
4. AbouZahr C. "Global burden of maternal death and disability". *BMJ* 2003; 67:1-11.
5. Ahiadeke, C. (2001). Incidence of induced abortion in Southern Ghana; International *Family Planning Perspectives, 27* (2), 96-101,108.
6. Aikins L. (2010),"Female Health and Development; A case study regarding a maternal Health Scheme in Ghana". University of Agder, Faculty of Economics and Social Science
7. Butler, M. A., and A. A. King (2004), Phylogenetic comparative analysis: A modelling approach for adaptive evolution. American Naturalist 164:683-695.
8. Burnham, K.P., & Anderson, D.R. (1998). Model Selection and Inference. Springer-Verlag, New York.
9. Box, G.E.P. & Jenkins, G.M. (1976) Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco.
10. Campbell, O.M., and W.J. Graham (2006). Strategies for reducing maternal mortality: Getting on with what works. *The Lancet* [Maternal Survival Series] 368(9543): 1284-99.

11. Cook, R. D., Weisberg, S. (1982). Residuals and Influence in Regression. London: Chapman and Hall.

12. D'Ambruoso L, Abbey M., Hussein J. (2005). Please understand when I cry out in pain: women's accounts of maternity services during labour and delivery in Ghana.

13. Dickey, D.A. & Fuller, W.A. (1979). Distribution of the Estimators for Autoregressive Time Series with Unit Root. *Journal of the American Statistical Association*, 74: 427-431.

14. DFID (2004) "Reducing maternal deaths: Evidence and action"A strategy for DFID Published by the Department for International Development, September 2004

15. Dickey, D.A. & Fuller, W.A. (1981). Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. *Econometrica*, 49(4), 1057-1072.

16. Edwin Vicente C B, Yoko L, et al (March 2011), "Make it happen 2015: validation of the maternal mortality ratio in Trinidad and Tobago for 2000–06".

17. Elard Koch (2009), "maternal mortality reduction and abortion legalization", The Chilean Maternal Mortality Group, Faculty of Medicine, University of Chile

18. Fraser M.D. & Cooper A.M (2003) Myles textbook for Midwives, 14th Edition, Churchill, Livingstone

19. Gao Yanqiu, et al. (2009), "Time trends and regional differences in maternal mortality in China from 2000 to 2005", Bulletin of World Health Organization, 2009 December.

20. 2008 Ghana Millennium Development Goals report (April, 2010)

21. Ghana Health Service (GHS), (2007) a. *Reproductive Health Strategic Plan 2007-2011*. Accra, Ghana: Ghana Health Service.

22. Ghana Health Service (GHS), (2007) b. *Ghana Health Service Annual Report 2007*. Accra, Ghana: Ghana Health Service.

23. Ghana Millennium Development Goals Report, April 2011
24. Gardner, Howard (1995). Reflections on multiple intelligences. *Phi Delta Kappan*, 77 200-
25. Hamilton, J. D. (1994). Time Series Analysis. Princeton Univ. Press, Princeton New Jersey.
26. Hurvich, C.M., & Tsai, C.L (1989). Regression and Time Series Model Selection in Small Sample. *Biometrika*, 76: 297-307.
27. International Classification of Diseases (ICD), 10th Edition
28. Komfo Anokye Teaching Hospital, (KATH official website)
29. Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. & Shin, Y. (1992): Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root. *Journal of Econometrics* 54, 159–178.
30. Kumasi Metro Health Directorate, (KMHD. 2009) "Annual report 2009"
31. Louise Kuhn, et al (1993), "Use of Poisson Regression and Time Series Analysis for Detecting Changes over Time in Rates of Child Injury following a Prevention Program ", *America Journal of Epidemiology* 1994;140:943–55
32. Mayhew, S. (2004). Sexual and reproductive health in Ghana and the role of donor assistance. In *Progress and promises: Trends in international assistance for reproductive health and population*, ed. S. Ethelston, A. Bechtel, and C.G. Vogel. Washington, D.C.: Population Action International.
33. Margaret C Hogan and Kyle J Foreman (2010), "Maternal mortality for 181 countries, 1980—2008: a systematic analysis of progress towards Millennium Development Goal 5"

34. M Á Luque Fernández, et al, (2008), "Increase in maternal mortality associated with change in the reproductive pattern in Spain: 1996–2005". Journal of Epidemiology Community Health, 2009.
35. MOH Ethiopia, (2000), "to analyzing trends and developing a model for prediction of Health and Health related indicators." The Ethiopian Government through their Ministry of Health,2000
36. Mulu W. A. and Tilahun H. N. (2009), "Modelling trends of health and health related indicators in Ethiopia (1995-2008): a time-series study", Health Research Policy and Systems 2009, 7:29
37. Nelder, J .A. and Wedderburn R. W. M. (1972), Generalized linear models, Journal of the statistical society. Pp 370–384
38. Oliver Razum, et al (1998), "Maternal mortality in the former East Germany before and after reunification: changes in risk by marital status",
39. Qin Yi Lee (2010), "Maternal Mortality in Ghana; Hospital-Based Review" University Of Leicester
40. Ronsmans, C., and W. Graham. (2006). Maternal mortality: Who, when, where, and why. *The Lancet* 368: 1189-1200.
41. Samuel Mills, et al (2007) "Obstetric Care In Poor Settings In Ghana, India, And Kenya" , November 2007
42. Sven Cnattingius MD, James L, et al. (1997), "the paradoxical effect of smoking in preeclampsia pregnancies", The American journal of obstetrics and gynaecology, 1997
43. Sullivan SA, Hill EG, Newman RB, Menard MK (2003), "Maternal-fetal medicine specialist density is inversely associated with maternal mortality ratios." ,American Journal of obstetric and Gynecol. 2005.

44. Thaddeus, S., and D. Maine (1994). Too far to walk: Maternal mortality in context. *Social Science & Medicine* 38(8): 1091.

45. Tietjen L, D., Bossemeyer N. Mclntosh, (2003). Infection prevention guidelines for heath care facilities with limited resources. JHPIEGO Corporation: Baltimore.

46. The Journal of China Medical University, "Exploring the feasibility for application of time series ARIMA model to predict the maternal mortality ratio (MMR) in china", March 2011

47. United Nations Development Programme (UNDP), (2003). *Indicators for monitoring the Millennium Development Goals: Definitions, rationale, concepts, and sources.* New York: United Nations.

48. Unicef (2001) Situation Analysis of Children and Women in Ghana 2000. Accra: Unicef Ghana office, p.64

49. United Nations (2007) The Millennium Development Goals Report 2007, New York:

50. United Nations UNDP (2007) Ghana Human Development Report 2007, Accra: UNDP, Ghana office.

51. World Health Organization (WHO), UNICEF, UNFPA and The World Bank (2007). *Maternal mortality in 2005: Estimates developed by WHO, UNICEF, UNFPA and The World Bank.* Geneva: World Health Organization

52. Worawan C, et al. (2010) "Thailand's approach to measuring maternal mortality ratio"

53. Yusuf M. Adamu, et al (2002), **"**Maternal mortality in Northern Nigeria: a population-based study" European Journal of Obstetrics & Gynecology and Reproductive Biology 2003

54. Ye yun-li, et al (2004), "the trend and forecast of infant mortality rate in China from 1991 to 2004", Department of Epidemiology, West China School of Public Health, Sichuan University, Chengdu 610041,China

55. Zivot, E. and Andrews, D. W.K. (1992), Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis, *Journal of Business & Economic Statistics*, 10(3), 251–270.

Appendix A

TABLE 4.1 Criteria for assessing goodness of fit for occurrence of maternal mortality

Criteria For Assessing Goodness Of Fit			
Criterion	**DF**	**Value**	**Value/DF**
Deviance	121	119.0390	0.9838
Scaled Deviance	121	119.0390	0.9838
Pearson Chi-Square	121	116.9710	0.9667
Scaled Pearson X2	121	116.9710	0.9667
Log Likelihood		1446.3300	

Table 4.3 Contrast results of occurrence of maternal deaths

Contrast Results				
Contrast	**DF**	**Chi-Square**	**Pr > ChiSq**	**Type**
year, 2000 vs 2005	1	0.81	0.3684	LR
year, 2005 vs 2010	1	0.00	0.9999	LR

TABLE 4.4 Analysis of parameter estimates for occurrence of maternal mortality with intercepts

Analysis Of Parameter Estimates								
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald 95% Confidence Limits**		**Chi-Square**	**Pr > ChiSq**
Intercept		1	2.2246	0.0949	2.0386	2.4107	549.33	<.0001
YEAR	**2000**	1	-0.1246	0.1386	-0.3962	0.1471	0.81	0.3688
YEAR	**2001**	1	-0.0461	0.1358	-0.3123	0.2201	0.12	0.7343
YEAR	**2002**	1	-0.0090	0.1345	-0.2727	0.2546	0.00	0.9464
YEAR	**2003**	1	-0.1044	0.1379	-0.3746	0.1659	0.57	0.4491
YEAR	**2004**	1	0.0000	0.1342	-0.2631	0.2631	0.00	1.0000
YEAR	**2005**	1	0.0354	0.1331	-0.2254	0.2962	0.07	0.7902
YEAR	**2006**	1	-0.1662	0.1402	-0.4410	0.1085	1.41	0.2356
YEAR	**2007**	1	-0.1144	0.1382	-0.3854	0.1565	0.68	0.4079
YEAR	**2008**	1	0.2177	0.1275	-0.0322	0.4676	2.92	0.0877
YEAR	**2009**	1	0.0267	0.1333	-0.2347	0.2880	0.04	0.8415
YEAR	**2010**	0	0.0000	0.0000	0.0000	0.0000	.	.
Scale		0	1.0000	0.0000	1.0000	1.0000		

TABLE 4.5 Analysis of parameter estimates for occurrence for continuous time.

Analysis Of Parameter Estimates							
Parameter	**DF**	**Estimate**	**Standard Error**	**Wald 95% Confidence Limits**		**Chi-Square**	**Pr > ChiSq**
Intercept	1	2.1217	0.0631	1.9980	2.2455	1129.06	<.0001
tme	1	0.0136	0.0091	-0.0044	0.0315	2.19	0.1386
Scale	0	1.0000	0.0000	1.0000	1.0000		

TABLE 4.6 Criteria for assessing goodness of fit for incidence of maternal mortality

Criteria For Assessing Goodness Of Fit			
Criterion	**DF**	**Value**	**Value/DF**
Deviance	121	124.9686	1.0328
Scaled Deviance	121	124.9686	1.0328
Pearson Chi-Square	121	125.4037	1.0364
Scaled Pearson X2	121	125.4037	1.0364
Log Likelihood		1443.3651	

TABLE 4.8 Analysis of parameter estimates for occurrence for continuous time

Analysis Of Parameter Estimates							
Parameter	**DF**	**Estimate**	**Standard Error**	**Wald 95% Confidence Limits**		**Chi-Square**	**Pr > ChiSq**
Intercept	1	-4.5797	0.0622	-4.7016	-4.4578	5423.84	<.0001
tme	1	-0.0116	0.0090	-0.0292	0.0060	1.67	0.1963
Scale	0	1.0000	0.0000	1.0000	1.0000		

Appendix B

1. **Quarterly Maternal Mortality Ratios**

	Qtr1	Qtr2	Qtr3	Qtr4
2000	1152.9	848.3	837.0	858.2
2001	1017.4	1002.0	1203.0	770.4
2002	1231.1	905.1	1127.7	864.4
2003	950.2	1008.7	1152.7	828.0
2004	983.2	880.9	1351.4	977.6
2005	1331.2	1211.3	969.0	963.7
2006	923.2	835.7	923.4	1050.1
2007	936.3	939.1	780.7	574.5

2008	951.3	995.8	1373.0	969.3
2009	928.5	1133.0	841.8	835.1
2010	724.0	711.1	833.1	896.1

2. Poisson data

Deaths	Deliveries	Year
9	774	2000
9	743	2000
9	832	2000
8	987	2000
9	1113	2000
10	1112	2000
5	945	2000
10	632	2000
6	751	2000
9	998	2000
5	934	2000
9	763	2000
8	747	2001
6	743	2001
10	878	2001
13	967	2001
9	1081	2001
8	968	2001
9	906	2001
13	799	2001
9	887	2001
5	949	2001
9	974	2001
7	842	2001
7	861	2002
13	758	2002
11	911	2002
7	966	2002
12	1327	2002
10	914	2002
11	918	2002
7	774	2002
10	812	2002
13	835	2002
3	949	2002
6	816	2002

7	755	2003
3	669	2003
11	815	2003
7	937	2003
12	980	2003
10	971	2003
11	800	2003
7	814	2003
10	845	2003
13	929	2003
3	950	2003
6	850	2003
5	721	2004
10	779	2004
9	977	2004
9	1076	2004
11	1042	2004
7	966	2004
7	868	2004
15	832	2004
11	698	2004
12	1048	2004
9	1067	2004
6	659	2004
17	870	2005
10	795	2005
7	923	2005
12	1043	2005
13	1061	2005
11	924	2005
9	888	2005
9	868	2005
6	727	2005
7	1022	2005
6	889	2005
8	733	2005
9	773	2006
4	739	2006
9	843	2006
4	763	2006
9	1126	2006
5	281	2006
4	818	2006
9	865	2006
11	937	2006
8	1042	2006

10	959	2006
12	869	2006
9	853	2007
7	828	2007
9	941	2007
15	1102	2007
13	1341	2007
5	1089	2007
7	894	2007
7	891	2007
8	1029	2007
5	1226	2007
8	1118	2007
6	1006	2007
9	989	2008
11	875	2008
6	862	2008
12	1152	2008
12	1147	2008
9	1021	2008
16	909	2008
19	1038	2008
9	1195	2008
8	1279	2008
8	1305	2008
19	1035	2008
13	1003	2009
8	833	2009
6	1058	2009
9	1139	2009
11	773	2009
14	1095	2009
4	984	2009
14	1020	2009
7	932	2009
7	1114	2009
10	1236	2009
11	979	2009
7	1052	2010
6	990	2010
10	1130	2010
9	1196	2010
8	1302	2010
10	1254	2010
9	1234	2010
10	1145	2010

10	1126	2010
9	1304	2010
13	1110	2010
10	1171	2010