ADDIS ABABA UNIVERSITY
GRADUATE STUDIE PROGRAM
COLLEGE OF NATURAL SCIENCES
DEPARTMENT OF STATISTICS

The Statistical Distribution and Some Determinants of Birth Interval for Rural Ethiopia

By

Wondiber Nega

May, 2011
Addis Ababa, Ethiopia

The Statistical Distribution and Some Determinants of Birth Interval for Rural Ethiopia

By

Wondiber Nega

A Thesis Submitted to the Office of Graduate Programs of Addis Ababa University in Partial Fulfillment of the Requirement for the Degree of Master of Science in Statistics.

May 2011
Addis Ababa, Ethiopia

Research Title: *The Statistical Distribution and Some Determinants of Birth Interval for Rural Ethiopia*

Name of Candidate: Wondiber Nega

Approved by the board of examiners

Advisor	Examiner	Examiner
Name: Eshetu Wencheko (prof.)	____________________	____________________
Sign:________________	____________________	____________________
Date:________________	____________________	____________________

Table of Contents

Acknowledgements

My most sincere and heartfelt thanks go to my advisor, Professor Eshetu Wencheko, for his interest, devotion, guidance, generally for his many constructive comments and suggestions throughout the study.

I would also like to thank all the members of my family. My sincere thanks go to my mother, W/o Aynadiss Teshome, she was the source of especial strength towards the successful completion of the study.

Finally, my heartfelt gratitude goes to all who were there with me.

Abstract

Studying the dynamics of spacing of births is important for several reasons including an understanding of completed family size. In this paper the length of birth interval between successive children that is called inter-birth interval in rural community of Ethiopia was considered. The birth-interval approach used to study the tempo of fertility as measured by the transition time for those women who continue reproduction. The data for the study was obtained from Ethiopian Demography and Health Survey data conducted in 2005. The data contains 9647 rural women aged from 15-49 years.

The approach of the study was fitting probability density functions to identify from among the three gamma, lognormal and normal distributions that provide adequate fit. Identifying variables that affect birth intervals was done by parametric survival analysis technique called "accelerated failure time" model.

The result shows that the duration time between successive births is different for different regions, religion groups and education level. Comparatively mothers from the northern region have longer birth interval, while mothers from eastern region have the shortest; and the birth interval for those mothers from central and south regions lie between. The length of birth is short in illiterate mothers than other groups. Muslim mothers have shorter birth interval than followers of "other" religions. Intervals following the death of the previous child (or children) tend to be significantly shorter than intervals where the child survived. The median length of birth intervals in rural community of Ethiopia is 28 months and more than 75% of births occur within 3 years.

Chapter One

1.1 Introduction

The analysis of birth intervals explicitly recognizes the distinct renewable nature of the fertility process. It considers the progression from one birth to the next during a woman's reproductive life. Therefore, birth intervals play a major role on childbearing and receive attention in the study of human fertility. Especially at present, governments of third world countries are aiming at reducing the high rate of fertility. In spite of all-out efforts, the fertility levels of these countries, measured by different indices, continue to be high (World Population Data sheet, 2008). Some of the main factors which affect birth intervals are rooted in social and cultural norms, behaviors of individual women, utilization of different reproductive health services, education level and other personal factors.

Birth history analysis undoubtedly provides useful information on reproduction and family formation. Further, fertility depends not only on the decision of couples but also on socioeconomic, demographic, health-related as well as tradition-related and emotional factors. These factors affecting fertility may have varying effects on child spacing. Thus, birth intervals experienced by women may reveal some insights about their reproduction patterns. Moreover, a detailed analysis of the sequence of steps in the childbearing process could provide a more comprehensive picture of the dynamics of fertility transitions.

Wider birth interval has several benefits. Longer gap between birth allows a mother more time to recover from pregnancy and delivery; the next pregnancy and birth are more likely to be at full conception and growth; there is less competition between existing children for breastfeeding, food, nutrition, the mother's time, and other resources (National Research Council, 1989). According to the World Health Organization (WHO, 2005) report individuals and couples should consider health risks and benefits along with other circumstances such as their age, fecundity, fertility aspirations, access to health services, social and economic circumstances, and personal preferences in making choices for the timing of the next pregnancy.

We know that there are various direct and indirect methods of reducing fertility voluntarily, such as provision of low cost contraceptive pills, delayed marriage, or delayed age at first birth and widening the length of inter-birth intervals. Various studies have been conducted to study the effect of each and all these controlling mechanisms. Indirectly the contribution of birth interval on fertility level has been considered in the present study.

Birth spacing among women is affected by several factors and as a result for some women a birth interval is too short; for some women the interval is too long; and for other women it is just right. In this paper we look at both sides of the coin: what factors contribute to a short birth interval and what are the factors leading to a long birth interval?

Fitting various statistical distributions to birth intervals based on survey data could help to estimate the probability of different fertility levels. And the need for analyzing accelerated failure time model is used to investigating the effects and also relationship between failure time data (inter-birth interval) and important covariates. So far no study had been done on statistical distribution of birth intervals, and accelerated failure time technique was used to know the difference of the average birth intervals for various birth orders in different rural communities of Ethiopia. In this present study an attempt is made in this direction.

1.2 Statement of the Problem

We know that, birth intervals are believed to affect fertility, mortality and population growth. Thus, childbearing or fertility of population has received serious attention. There are several reasons for this. First fertility largely determines a population's age structure: high fertility is responsible for youth dependency mostly in less developed regions. Variable fertility has created major fluctuations in cohort size of population. When we look back at the background of Ethiopia population growth rate there has been a steady increase since 1960. Based on 1984 census data, population growth was estimated at about 2.3 percent for the 1960-70 period, 2.5 percent for the 1970-80 period, and 2.8 percent for the 1980-85 period. Population projections compiled in 1988 by the Central Statistical Authority (CSA) projected a 2.83 percent growth rate for 1985-90 and a 2.96 percent growth rate for 1990-95. According to the 2007 Ethiopia population census, the annual population growth rate within 1994-2007 was estimated as 2.6%. This shows that rate of population growth in previous years was high, and birth intervals are well-thought to affect fertility rate and this has direct relation with population growth.

1.3 Objective of the Study

The major objective of this study is to identify the socio economic and demographic factors, such as religion, region, education level, sex-composition, survival status of index child at a time of infant(before two years), etc that influence birth interval.

The Specific Objectives are:

i. To fit and select the distribution among three probability density functions, namely gamma, lognormal and normal that provides a good fit to different birth intervals.

ii. To select the significant variables under consideration by using accelerated-failure-time model and to calculate probabilities by using the selected probability density function.

iii. To find common patterns in the distribution of birth intervals among social and demographic factors.

Limitations

The data used here, being secondary, may have a number of limitations on the outcome of this study. The first limitation is that there is lack of sequential data which shows a couple's decision to make alternative choice on the age gap between children birth-intervals like matters on the subject of contraceptive use. Again, there is no clear information to know the relative effects for the main constraint variables such as duration of breastfeeding for each child, age of mothers at a time of birth of children, if there is any time gap between spouses, and so on.

Another is that the data used are the 2005 Ethiopian Demographic and Health Survey that includes data over the last 30 years or longer which may not cover the current situation for rural Ethiopia. For instance, if the respondent mother has ten children, the required information which shows the age gap between the first child date of birth and second child date of birth may happen before thirty years.

The other limitation was unavailability of sufficient literature on the subject of the study for rural Ethiopia.

Chapter Two
Review of Literature

For the past half century, several theoretical approaches have been developed and refined to explain variations in fertility (e.g., Easterlin, 1975; Bongaarts, 1978). These models have undoubtedly broadened the perspectives on differentials in fertility. They mostly focus on total or cumulative fertility as the key dependent variable. It is equally important, however, to study the variations in fertility by considering the spacing of births as constraint variable. This is particularly relevant in societies characterized by low levels of contraceptive use where the interval between successive births is a key indicator of total family size like rural part of Ethiopia.

Inter-birth intervals vary widely between populations, even in the absence of deliberate spacing efforts because of differential fecundability and postpartum amenorrhea. These are strongly influenced by coital frequency and breastfeeding, respectively (Wood, 1994). Therefore, it is impossible to draw conclusions about the presence or absence of controlled birth spacing from a simple comparison of mean or median lengths of birth intervals between groups or time periods (David and Mroz, 1989b). David and Mroz (1989b) have turned to multivariate hazard analysis in order to try to control for natural determinants of birth intervals.

Birth intervals are affected by a complex range of factors. Some of which are rooted in social and cultural norms, others in the reproductive histories and behaviors of individual women, utilization of reproductive health services and other personal factors. Group differences in reproductive behavior are usually explained from the characteristics and socio-cultural perspectives (e.g., United Nations, 1987). While the former attributes variations in fertility behavior to socio-economic and demographic differences among groups, the latter assigns a unique role to availabilities of required materials like contraceptive as key to making variations on birth intervals.

Among socio-cultural factors are ethnic-specific practices, norms, and values that capture both observable and unobservable behavioral and cultural factors affect reproductive behavior (Caldwell and Caldwell, 1987)

Studying the dynamics of birth spacing defined as the gap between successive births is interesting for several reasons. First, most of the inferences are based on the view that in the developing countries like Ghana and Kenya by Gyimah (2001); the gap between successive births are short for couples having large families than couples with smaller families. This suggests that the timing of births is inversely related to total or cumulative fertility. Further, the space of births has a significant bearing on maternal and child health through the dynamics of sibling competition, maternal depletion and interval effect hypotheses (Hobcraft et al., 1985 and Pedersen, 2000). According to the competition hypothesis, the birth of each successive child generates competition for insufficient resources among siblings in the household which subsequently leads to a lower quality of care and attention to each child.

Successive births physiologically deplete the mother of energy and nutrition (Gribble, 1993) which may lead to premature births or pregnancy complications, increasing the risk of infant or maternal death, or impair the mother's ability to care for her children. Additionally, it has also been argued that women with closely spaced births may still have very young children and, as such, are less likely to attend prenatal care services. Further, the early arrival of a new child necessitates the premature weaning of the previous child, often exposing the weaned child to malnutrition and increasing their vulnerability to infectious and parasitic diseases. Invariably, longer duration of the inter-birth interval has been found to increase profoundly the probability of infant survival (Pedersen, 2000).

Theoretical perspectives

Ages at first marriage and at first birth are also of tremendous importance in fertility studies because of their inverse relation to the exposure to the risk of conception (see, e.g., Gyimah, 2001). They also represent a number of unmeasured factors that predispose women to differential timing of births and, thus, overall fertility. Women who marry at younger ages or have first births earlier are likely to come from disadvantaged socio-economic backgrounds and are thus more likely to be associated with the higher risks of births than their counterparts whose first marriages or first birth occurs late (Gyimah, 2001). Women who marry late are expected to quicken the pace of the first birth because of the social pressure to prove they have not past their prime reproductive age.

In previous years in Ethiopia, where most marriages start very early (particularly in rural areas), women often stay at their parents' home for several years after getting married. This is likely to prolong the interval between marriage and first birth. Actually, it may not mean that women stay in their parents' home after they have been old enough for sexual intercourse. According to Ghilagaber et al (2005) in rural china results women who married at an older age had a higher fertility rate for the first five years of marriage, while younger married women had a lower initial fertility rate.

Arguably, short birth spacing is not beneficial to the household economy (Alter, 1988). Families with low budgets can be expected to face economic problems, when young children still form an economic burden to the household. This phase of economic hardship lasts depending on birth spacing and on the age at which children could start to earn money. The higher the pace of childbearing creates the harder the economic stress in the household given the level of income and costs (Alter, 1988). The main point is that households are more likely to be under economic stress when a large fraction of the children alive are young and dependent. This may stimulate parents to try to delay the next birth when the proportion of small children is high. For example, more of agricultural laborers directly or indirectly depend on contributions of older children to their family budget. Therefore, agricultural laborers can be expected to be motivated to space births for household economic reasons rather than for limiting their eventual family size.

Other factors, such as the mother's education level has also been correlated with birth spacing although the mechanisms by which these background variable influence birth spacing is less clear. In some settings, maternal education is associated with shorter spacing; in Korea, for example, one study reported that better educated women had shorter second birth intervals than those less educated (Bumpass, Rindfuss, and Palmore, 1986). However, in 38 of 51 countries with DHS data, women with no education were more likely than educated women to have shorter intervals (Setty-Venugopal and Upadhyay, 2002). It can be speculated that better educated women wish to compress childbearing into fewer years and participate in non-childbearing activities and hence have shorter spacing. The pathways through which these happen have been explained through an array of mechanisms including late age at marriage, greater knowledge and access to contraception, high labor force participation and alternative values regarding family size (Martin, 1995; Ware 1984). For example, in Ghana birth interval has been found a positive linear effect of education on the intervals between successive births (Gyimah, 2001).

Also significant in determining the length of the inter-birth interval is the survival status of the index child (Montgomery and Cohen, 1998). Survival status of index of a child refers about information for a child either survived or not up to his/her second birth date. In Ghana, it has been demonstrated that intervals following the death of the index child tend to be significantly shorter than intervals where the child survived, a result of biological and behavioral processes (Gyimah, 2001). We thus expect the death of the index child to be associated with shorter intervals.

Religion may be considered to be an important factor in determining socio-cultural practices. Because religions differed in the amount of personal freedom they allowed vis-à-vis the omnipotence of God. Along with this factor religion-specific practices and values that capture both observable and unobservable factors may affect reproductive behavior. Thus, a stronger religious affiliation is expected to affect the inter-birth intervals or rate of fertility among groups of people.

Sex composition of a family can also influence the age gap between successive children. "Gender" refers to the different role that men and women play in society and also the rights and responsibilities that come with these roles (Centre for Development and Population Activities (CEDPA), 1996). Gender roles are so strong that they are taken for granted. They are reflected in virtually every social institution, including family structures, household responsibilities, labour markets, schools, health care systems, law, and public policies. The influence of gender is similar in strength to that of religion, race, social status, and wealth (Centre for Development and Population Activities, 1996). In all parts of the world, women are facing threats to their lives, health, and well-being as a result of being overburdened with work and of their lack of power and influences.

In many countries, traditional male and female gender roles deter couples from discussing sexual matters, condone risky sexual behavior, and ultimately contribute to poor reproductive health among both men and women (Moser, 2001). Programs can encourage men to adopt positive gender roles, such as being supportive husbands and caring fathers.

Gender roles and gender norms are culturally specific and thus vary tremendously around the world. Almost everywhere, however, men and women differ substantially from each other in power, status and freedom. In virtually all societies' men have more power than women (Moser, 2001). Gender has a powerful influence on reproductive decision making and behavior. Thus, in many developing countries men are the primary decision-makers about sexual activity, fertility, and contraceptive use. These all conditions may have relation with sex preference of a family; and that expected to affect the length of birth intervals.

Generally, differences in the birth intervals can be explained through demographic, socio-economic and socio-cultural factors. On the basis of previous work, relevant variables organized under socio-economic, socio-cultural and demographic controls are selected for this study as influential factors on birth intervals in Ethiopia.

An examination of birth intervals thus provides a detailed understanding of attitudes toward family size as well as differentials in fertility and childhood mortality levels. Even if the technique and way of calculation is differ, related research on identifying of best distribution that fit birth-interval data has been done by Girmay Medhin (1992). This study is an extension of the study by using recent data, and further examines the relative effects of socio-economic and socio-cultural factors on the timing of births by using accelerated failure time model.

Chapter Three

Data and Methodology

3.1 Source of data

Secondary data from Ethiopian Demographic and Health Survey (DHS) conducted by the Central Statistical Agency (CSA) during 2005 are used for this study. This survey was designed to collect information for the purpose of policy decisions, planning, monitoring and evaluating of programs on health in general and reproductive health in particular at both national and regional level. This was the second comprehensive survey designed to provide estimates for the health and demographic variables of interest for the following domains: Ethiopia as a whole, urban and rural areas of Ethiopia and all geographic areas (nine regions and two city administrations), namely: Tigray, Affar, Amhara, Oromiya, Somali, Benishangul-Gumuz, Southern Nations, Nationalities and Peoples (SNNP), Gambela and Harari regional states and two city administrations, that is, Addis Ababa and Dire Dawa.

This survey collected from a nationally representative sample of women and men in the reproductive age groups of 15-49 and 15-59, respectively. The survey was conducted in selected rural and urban areas throughout the country. The sample survey was based on a two stage stratified systematic sample of households. At the first stage of sampling 540 enumeration areas (EAs), 145 in urban areas and 395 in rural areas, were selected using systematic sampling with probability proportional to size (pps).

A complete household listing operation was carried out in the selected EAs to provide a sampling frame for second stage selection of households. At the second-stage of sampling, a systematic sample of 28 households per EA was selected in all the regions and the two city administration to obtain reliable estimates of key demographic and health variables.

The survey was designed to obtain completed questionnaires of interviews from 14,717 women of ages 15-49. In addition, all males of age 15-59 in every second household were interviewed, to obtain a target of 6,778 men interviews. In order to take non-response into account, a total of 14,645 households were selected. Of the 14,645 households selected, interviews were completed for 13,721 households, 14070 women and 6,033 men. Finally, for our purpose we selected 9647 rural women out of all 14070 urban and rural respondents.

3.2 Variables in the Study

The response variable in this study is the age gap between successive children for each mother measured in months that is called "birth interval". The other influential variables/factors included in the study are classified as demographic and socioeconomic. We have selected an array of theoretically relevant variables as likely covariates of birth intervals. These include region, religion, maternal education level, wealth status, survival-index of child and sex (or sex composition) of previous child(ren) of the respondent.

After observing the characteristics of the distribution of each order of birth intervals by region related ones was grouped into the same categories. For instance, Tigray and Amhara region included in one group and Beneshangul Gumuz, Gambela and SNNP incorporated in one region. The way of classifying and coding of the explanatory variables is as follows:

No.	**Factors/ variables**	**Categories**
1.	Region	1 = North $\Rrightarrow \begin{cases} Tigray\ and \\ Amahra \end{cases}$ 2 = East $\Rrightarrow \begin{cases} Affar \\ Dire\ Dawa \\ Harari\ and \\ Somali \end{cases}$ 3 = Central $\Rrightarrow \begin{cases} Addis\ Ababa \\ and\ Oromiya \end{cases}$ 4 = South $\Rrightarrow \begin{cases} \text{Ben} - \text{Gumuz} \\ \text{Gambela and} \\ \text{SNNP} \end{cases}$
2.	Education level	0=illiterate 1=primary 2=secondary & above
3.	Religion group of a woman	1=Coptic Orthodox 2=Protestant 3=Muslim 4=Others
4.	Sex of the 1st two children in a family	1=both males 2=both females 3=one male and one female
5.	Wealth status	1=poor 2=middle 3=rich
6	Survival status of index child (SSIC)	0=died as infant (before two years) 1=Survived

3.3 Statistical Distributions

As one of the objectives of this study the selected probability distributions to fit the inter birth intervals data are Gamma, Lognormal and Normal probability distributions. The choice of these three Continous distributions was dictated by the prevailing literature on birth interval estimates. The best distribution among the three will be identified. Below we briefly introduce the probability density function (pdf) of the above mentioned three distributions.

3.3.1 Normal Distribution

A random variable T is said to be normally distributed with mean μ and standard deviation σ if it has a pdf of

$$f(t) = \frac{1}{\sigma\sqrt{2\Pi}} \exp\left\{\tfrac{-1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right\} \qquad -\infty < t < \infty, -\infty < \mu < \infty \text{ and } \sigma > 0 \qquad (1)$$

This distribution is uniquely determined by its two parameters, namely mean and standard deviation. If a random variable T with mean μ and standard deviation σ has pdf given in equation (1) then the standardized form of T given as $Z = \frac{(T-\mu)}{\sigma}$ is the standard normal random variable with mean zero and standard deviation one.

Most theoretical arguments for the use of this distribution are based on the central limit theorem that identifies sufficient conditions to ensure asymptotic normality. For further discussions of the distribution one may refer to Johnson et al. (1994) and for some important properties Mood et al. (1974) and others may be considered.

The estimates of μ and σ^2 for the normal distribution are

$$\hat{\mu} = \sum_{i=1}^{n} \frac{t_i}{n} \qquad \text{and} \qquad \hat{\sigma}^2 = s^2 = \sum_{i=1}^{n} \frac{(t_i - \hat{\mu})^2}{n}$$

3.3.2. Lognormal Distribution

If T is a random variable such that another random variable Y defined as Y = log(T) is normally distributed with mean μ and standard deviation σ, then T is said to be lognormal distributed random variable and its pdf is given as:

$$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} \exp\left\{\tfrac{-1}{2}\left(\frac{\log(t) - \mu}{\sigma}\right)^2\right\} \quad t > 0\,,\ \mu > 0 \tag{2}$$

and $\sigma > 0$

The mean and the variance of the distribution are given by

$$E(T) = \exp[\mu + \sigma^2/2]$$

$$Var(T) = [\exp(2\mu + \sigma^2)][\exp(\sigma^2) - 1]$$

The distribution is positively skewed to a degree that depends on the standard deviation of log(T). It has been successively fitted to many kinds of duration and size distribution (Bartholomew, 1967). Bartholomew (1967) and others show that the length of human life, age at marriage, the spacing between successive births in a family and others are examples of durational variables of fundamental interest in demography. Lognormal distribution is a good candidate to approximate our variable of interest, that is, birth interval. The two parameters can be estimated by maximum likelihood method. That is given as:

$$\hat{\mu} = \sum_{i=1}^{n} \frac{\log(t_i)}{n} \qquad \text{and} \qquad \hat{\sigma}^2 = s^2 = \sum_{i=1}^{n} \frac{\left(\log(t_i) - \hat{\mu}\right)^2}{n}$$

3.3.3. Gamma Distribution

The gamma model with parameters γ, α *and* β has pdf

$$f(t) = \frac{\beta^{\alpha}(t-\gamma)^{\alpha-1}\exp(-(t-\gamma)\beta)}{\Gamma(\alpha)} \tag{3}$$

where $\alpha > 0, \beta > 0$ *and* $t > \gamma$

In most applications the two parameter form ($\gamma = 0$)

$$f(t) = \frac{\beta^{\alpha} t^{\alpha-1} \exp(-\beta t)}{\Gamma(\alpha)} \quad t > 0, \alpha > 0 \text{ and } \beta > 0 \qquad (4)$$

is used.

The use of gamma distribution to approximate the distribution of quadratic forms (particularly positive definite quadratic forms) in multi-normally distributed variables is well-established and widespread. Gamma distribution also used to make realistic adjustment to exponential distributions in many physical situations to representing lifetimes in life testing situations.

By increasing α the shape of the curve becomes similar to the normal probability curve. Johnson et al (1994) give a good general review of this distribution, including several references to applications in diverse fields. They have also indicated that it is a generalization of the exponential distribution over the positive range.

To estimate parameters for equation (4), the maximum likelihood method may be employed. If $x_1, x_2 \ldots x_n$ are independent random variables each having distribution (4), then equation for the maximum likelihood estimators of α and β are:

$$\frac{1}{n}\sum_{i=1}^{n} \log(xi) = \log(\beta) + \psi(\alpha) \qquad (5)$$

$$\bar{x} = \hat{\alpha}\hat{\beta} \qquad (6)$$

We solve equations (6) for β and use it in equation (5) to get

$$\log(\hat{\alpha}) - \psi(\hat{\alpha}) = \log\left(\frac{\bar{x}}{G}\right)$$

where, $\Psi(\hat{\alpha})$ is digamma function

$\overline{X}$ is arithmetic mean and

G is geometric mean of random variable X.

One possible way to get α is to use table for $\log(\hat{\alpha}) - \psi(\hat{\alpha})$. Other possible is to use one of different approximations that can be made for $\psi(\hat{\alpha})$. In literature different approximation of $\Psi(\hat{\alpha})$ are given and some of these references are given in Johnson et al (1994). One of this approximation is

$$\psi(\hat{\alpha}) = \log(\alpha - 0.5)$$

Using this approximation parameters of equation (4), which the same as equation (3) is with $\Upsilon = 0$ are solved as:

$$\hat{\alpha} = \frac{\bar{x}}{\bar{x}-G} \quad \text{and} \quad \hat{\beta} = \frac{\bar{x}}{\hat{\alpha}}$$

Alternatively, the following highly accurate approximation stressed by Bury (1975) can be used to obtain $\hat{\alpha}$ directly as follows:

$$\hat{\alpha} = \begin{cases} \{0.05001 + 0.1649\mathrm{g} - 0.0544\mathrm{g}^{22}\}\mathrm{g}^{-1} & \text{if} \quad 0 < g < 0.577 \\ (17.80 + 11.978 + \mathrm{g}^2)^{-1}(8.899 + 9.060\mathrm{g} + 0.9775\mathrm{g}^2)\mathrm{g}^{-1} & \\ & \text{if } 0.577 < g < 17 \end{cases}$$

where $\mathrm{g} = \log\left(\frac{\bar{x}}{G}\right)$

Johnson et al (1994) indicted that the error of the formula is less than 0.01% by giving reference and hence in this paper we shall use this approximation for the estimation of gamma parameters.

3.4 *Method of Analysis*

Descriptive analysis is used to describe the median age and average number of children per mother for different socio-demographic and socio-economic variables like region and education level of mothers. The parametric survival analysis which is Lognormal Accelerated Failure Time regression method is employed to estimate the proportion of time for outcome variable (inter-birth intervals) based on some independent variables using time ratios. And also calculating of probabilities of getting a child for a given family in preferable time interval is done based on prominent variables.

Because the range of the dependent variable (inter-birth interval) is wide, we classified it into grouped frequency distributions. The way of constructing grouped frequency distribution is based on Sturges rule $k = 1 + 3.32\log(n)$, where k is number of classes desired to be and n is total number of sample observation.

The statistical analysis was performed by using SPSS 19.0, Stata 10.0, SAS 8.0 and Easyfit softwares. The trial version Easyfit softwar is used to identify the best statistical distribution among gamma, lognormal and normal that fits the given variable birth interval data.

Chapter Four

Technique of model Selection and Parametric Survival Analysis

4.1 Introduction

One purpose in this study is to fit three probability density functions (pdfs), namely the gamma, lognormal and normal probability distributions to birth intervals based on Demography and Health Survey data collected in 2005. Often the hypothesis being tested are statements concerning the unknown probability distribution of the random variable being observed. According to Conover (1980) a test for goodness of fit usually involves examining the random sample from some unknown distribution in order to test the null hypothesis, H_o, that the unknown distribution function is in fact a known, specified function. It has been stated (Gibbons, 1971) that several test statistics for examination of this hypothesis are functions of the deviations between the observed cumulative distribution and the corresponding cumulative probabilities expected under the null hypothesis. The functions of these deviations used to perform a test include the sum of squares, or absolute value, or the maximum deviation. The cumulative distribution function of a random variable is uniquely determined by its probability density (mass) function and vice versa (D'Augostino and Stevens, 1986).

4.2 Goodness of Fit Tests

Once a model has been developed, we would like to know how effective that model is in describing the outcome variable. This is referred to as goodness-of-fit. Thus the goodness-of-fit of a statistical model describes how well it fits a set of observations. The goodness-of-fit test measures the compatibility of a random sample with a theoretical probability distribution function. Measures of goodness of fit typically summarize the discrepancy between observed values and the values expected under the model in question. In other words, these tests show how well the distribution we selected fits to our data (D'Augostino and Stevens, 1986).

The general procedure consists of defining a test statistic which is, some function of the data measuring the distance between the hypothesis and the data, and then calculating the probability of obtaining data which have a still larger value of this test statistic than the value observed assuming the hypothesis is true.

The most commonly used goodness-of-fit tests are Kolmogrov-Smirnov, Anderson-Darling and Chi-squared. Appling only one of these tests is approximately the same, however, they are different in how they are performed (implemented). The software we use to identify the best probability density function is Easyfit which is a data analysis and simulation application software that allows fitting probability distributions to sample data to select a model and apply the analysis results for better decisions. The results are displayed as interactive reports allowing comparing the fitted distributions. The summary report lists the distributions ordered by goodness-of-fit statistics, enabling us to select one or more models which fit to our data. The goodness of fit test which used for our data is Anderson-Darling.

The Anderson-Darling procedure is a general test to compare the fit of an observed cumulative distribution function to gamma, lognormal or normal cumulative distribution function. It is a modification of the Kolmogorov-Smirnov (K-S) test and gives more weight to the tails than does the K-S test. Thus, Anderson-Darling test is an alternative to the chi-square and Kolmogorov-Smirnov goodness-of-fit tests. The test is based on the empirical cumulative distribution function (ECDF). Assume that we have a random sample $t_1, \ldots, t_n$ from some continuous distribution with CDF *F(t)*. The empirical CDF is denoted by

$$F_n(t) = \frac{1}{n}[\text{Number of observation } \leq t]$$

The test problem we consider is given as

H_0: The data follow the specified distribution.

H_A: The data do not follow the specified distribution at a significance α-level.

The test statistic we use in this study is the Anderson-Darling statistic (A^2) defined as

$$A^2 \quad = \quad -n - \frac{-1}{n}\sum_{i=1}^{n}(2i-1).[\ln F(t_i) + \ln(1-F(t_{n-i=1}))]$$

The hypothesis regarding the distributional form is rejected at the chosen significance level (*alpha*) if the computed value of the test statistic, A^2, is greater than the critical value obtained from an appropriate statistical table.

After fitting the pdf's to the observed birth intervals for overall rural Ethiopia, we will select the model that best approximates the observed relative frequency polygon. The same approach was also done by considering socio-economic and demographic determinants like education level, religion of respondents, etc. Further, to select one of the fitted pdf among gamma and lognormal we shall use the rank of the distributions defined in goodness of fit test table.

Various measures-of-goodness of fits of the birth-interval data with its pdf graphically are given in Table4.1 to Table4.6. The first five or six birth intervals are usually the most important and best predictors of subsequent intervals and we put more emphasis on these birth intervals. From the result (Table 4.1to 4.6) we observe that, almost all birth intervals are best approximated by the lognormal distribution than the gamma and normal distribution. Therefore, it seems reasonable to say that higher birth intervals may be approximated by the distribution which estimates the small parities birth interval.

For the purpose of calculating probabilities we shall choose the distribution which proved to be best estimator of the first four or five birth intervals, that is, lognormal distribution.

4.3 Lognormal Accelerated Failure Time Model

In an applied setting, the task of model selection is, to a large extent, based on the goals of the analysis and on the measurement scale of the outcome variable. Regression modeling of the relationship between an outcome variable and independent predictor variable(s) is commonly employed in virtually all fields. The popularity of this approach is due to the fact that plausible models may be easily fit, evaluated and interpreted. Statistically, the specification of a model requires choosing both the systematic and error components. The choice of the systematic component involves an assessment of the relationship between an "average" of the outcome variable and the independent variable(s). This may be guided by an exploratory analysis of the current data and/or past experience.

The accelerated failure time (AFT) model is a linear regression model in which the response variable is the logarithm or a known monotone transformation of a failure time (Kalbfleisch and Prentice, 1980). For instance in studies of mortality, it usually makes sense to think in terms of the hazard and the hazard ratio when making comparisons between groups. In some other situations, it is easier to think in terms of the relative length of time to the event. Thus, AFT model considered comparison intuitively by estimating the effects of covariates in terms of time ratios (TR) by taking one group as reference category (Kalbfleisch and Prentice, 1980).

In most empirical analyses of birth interval data, the focus has been on the quantum of fertility (that is, the rates at which children are born to a defined set of women within a specified unit of time) using proportional hazard models of the type discussed in Cox (1972). There are two commonly used survival models; these are the Cox proportional hazard model and the accelerated failure time model. The widely used Cox model measures causal effect on the hazard (rate) ratio scale, where as the accelerated failure time model measures causal effect on the survival time ratio scale.

The proportional hazards model (Cox, 1972) specifies the intensity of birth as a function of an unspecified time dependent baseline hazard, $\lambda_o(t)$, and the covariates i.e.

$$\lambda(t, z) = \lambda_o(t)\exp(z\beta) \tag{1}$$

As a useful alternative to the Cox model (Cox, 1972), AFT model has an intuitive linear regression interpretation. It means that, even though parametric proportional hazards models are widely used in medical research, AFT models offer a number of advantages. In particular, they provides for a wider variety of shapes of hazard function than parametric proportional hazards models that assume a particular distribution for survival times, since the family includes distributions with bimodal hazard function, such as the lognormal distribution. Moreover, the log-linear formulation of such models emphasizes that the role of regression parameters and dispersion parameters are clearly separated (Hosmer and Lemeshow, 1998). The regression parameters in AFT model are also robust towards neglected covariates, which is not the case for proportional hazards models. In addition, regression parameters in the proportional hazards model are more sensitive to the distributions of the random component.

Suppose we denote by T_o the time (say, birth interval in months) associated with the baseline category corresponding to zero-values for the covariate ($\mathbf{z} = \mathbf{0}$). Such baseline levels may, for example, be women with no education (to be later compared with those having some primary or secondary-level education). Then, the accelerated tempo model specifies that if the vector of covariates had been $\mathbf{z}$ ($\mathbf{z} \neq 0$), the corresponding time is

$$\boldsymbol{T} \;=\; T_o\, exp(\mathbf{Z}\boldsymbol{\beta}) \qquad (2)$$

or, equivalently,

$$ln\boldsymbol{T} \;=\; lnT_o \;+ \boldsymbol{Z\beta} \qquad (2)$$

where, $\boldsymbol{T}$ is the vector of birth intervals (durations), $\boldsymbol{Z}$ is d-dimensional covariate vector, and $\boldsymbol{\beta}$ is a vector of unknown regression parameters. Since covariates alter, by a scale factor, the rate at which an average woman traverses the time axis, (2) may be referred to as the accelerated failure time model. Thus, for proportional hazards model (1), the explanatory variables act multiplicatively on the baseline rate so that their effect is to increase or decrease the rate of birth relative to the baseline rate $\lambda_o(t)$. For accelerated tempo models, on the other hand, the explanatory variables act multiplicatively on time to the event (birth in our case) so that their effect is to accelerate or decelerate transition time to birth relative to that of the baseline category (T_o). In other words, the effect of the covariates in an accelerated failure time model is to change the scale, and not the location, of a baseline distribution of failure times. Usually, the scale function is exp ($\mathbf{Z'\beta}$), where $\mathbf{Z}$ is the vector of covariate values and $\boldsymbol{\beta}$ is a vector of unknown parameters. Thus, if T_o is an event time sampled from the baseline distribution corresponding to values of zero for the covariates, then the accelerated failure time model specifies that, the event time is $T = \exp(Z'\beta)T_o$.

If $y = \log(T)$ and $y_o = \log(T_o)$, then

$$y \;=\; Z'\beta + y_o$$

This is a linear model with y_o as the error term.

Thus the model in (2) is a linear model with $\ln(\mathbf{T}_o)$ playing the role of an error term with an underlying baseline distribution. Usually, an intercept term β_o and a scale parameter δ are allowed in the model to give

$$lnT_i = \beta_o + \boldsymbol{Z_i'\beta} + \delta\varepsilon_i \qquad i \;=\; 1, 2. \ldots, n \qquad (3)$$

where β_o is the intercept, $\beta' = (\beta_1, \beta_2, \ldots, \beta_p)$ is the vector of regression coefficients and ε_i is the error term which assumed follows an unique distribution.

The baseline survival distribution function $S_o(t)$ and the probability density function $f_o(t)$ when **Z** = **0** from equation (10) and (11) are listed for the additive random disturbance. These distributions apply when the log of the response is modeled (this is the default analysis). The chosen baseline functions define the meaning of the intercept, scale, and shape parameters. Thus the distributions supported in the SAS software LIFEREG procedure (we used for analysis) follow μ ~intercept and σ ~ scales in the output.

The other alternative way of writing an accelerated failure time model is

$$T = e^{\beta_o + \beta_1 Z_1 + \beta_2 Z_2 + ... + \beta_p Z_p} \times \varepsilon \tag{4}$$

The error component, ε, follows the exponential distribution with parameter equal to one and its density function is given by:

$$f(\varepsilon) = e^{-\varepsilon}$$

This model, (5), can be "linearized" by taking the natural logarithm of each side of the equation yielding

$$Y = \beta_o + \beta_1 Z_1 + \beta_2 Z_2 + \ldots + \beta_p Z_p + \theta \tag{5}$$

where $Y = \ln(T)$ and $\theta = \ln(\varepsilon)$.

The error component follows an "extreme minimum value" distribution. This distribution is not encountered often outside of applications in survival analysis but plays a central role in models of life-length and is often referred to as the Gumbel distribution.

Note: In probability theory and statistics, the Gumbel distribution (named after Emil Julius Gumbel (1891–1966)) is used to model the distribution of the maximum (or the minimum) of a number of samples of various distributions. For example we would use it to represent the distribution of the maximum level of a river in a particular year if we had the list of maximum values for the past ten years. The potential applicability of the Gumbel distribution to represent the distribution of maxima relates to extreme value theory which indicates that it is likely to be useful if the distribution of the underlying sample data is of the normal or exponential type (Gumbel, 1958).

The cumulative distribution function of the Gumbel distribution is

$$F(\theta; \mu,\beta) = e^{-e^{-(\theta-\mu)/\beta}} \quad \text{for } \mu \text{ and } \beta > 0$$

The standard Gumbel distribution is the case where $\mu = 0$ and $\beta = 1$ with cumulative distribution function

$$F(\theta) = e^{-e^{(-\theta)}} \tag{6}$$

The error term θ in (5), thus, follows a Gumbel distribution with mean 0 and its shape parameter is 1 denoted by $G(0,1)$. The use of the distribution $G(0,1)$ in (5) is somewhat like using the standard normal distribution N(0,1) in linear regression. From practical experience we know that, in linear regression, the errors rarely if ever have variance equal to one. The usual assumption is that the variance is neither a function of the outcome variable nor of the independent variables. It is assumed to be constant and equal to the parameter σ^2 with normal distribution denoted by $N(0, \sigma^2)$. An additional parameter may be introduced into (5) by multiplying θ by σ to yield the model

$$\mathbf{Y} = \beta_o + \beta_1 Z_1 + \beta_2 Z_2 + \ldots + \beta_p Z_p + \sigma \times \theta \tag{7}$$

The distribution of $\sigma \times \theta$ is denoted as $G(0,\sigma)$.

Under this model (7), the survival time follows the exponential distribution when σ =1 and the Weibull distribution when $\sigma \neq 1$. On the other hand when θ is distributed as standard normal, $N\ (0,1)$ or standard logistic, $L(0,1)$, the survival time follows the log-normal distribution or log logistic distribution respectively (Hosmer and Lemeshow, 1998).

Thus, survival time models that can be linearized by taking log-transformation are called accelerated failure time (AFT) models. The reason for this terminology is that the effect of the covariate is multiplicative on the time scale. That is, the effect of the covariate is said to "accelerate" survival time. In terms of the survival function, the accelerated failure-time model states that the survival function of an individual with covariate $\boldsymbol{Z}$ at time t is the same as the survival function of an individual with a baseline survival function at a time $t \times \exp(\boldsymbol{Z}'\boldsymbol{\beta})$.

The main goal for a statistical analysis of survival data is to fit a model that will yield meaningful and interpretable estimates of the effect of covariates on survival time. Therefore, a log normal Accelerated Failure Time model which assumes the log of timing follows a normal distribution will be used in examining the timing of births. The choice of the log-normal distribution derives from the above goodness-of-fit test of birth interval and also its suitability (the birth interval data) in substantive theory (Trussell et al, 1985). Once a distributional model for survival times has been specified in terms of a probability density function, the corresponding survivor and hazard functions can be obtained from the relations

$$S(t) = 1 - \int_0^t f(u)du$$

and

$$\lambda(t) = \frac{f(t)}{S(t)} = \frac{-d\{logS(t)\}}{dt} \qquad (8)$$

where f(t) is the probability density function of the survival time.

Thus, the main target in the analyses of survival (duration) data are: to describe the distribution of the time (duration) variable, to compare the survival experiences (distributions) of different groups of a population, and to investigate explanatory variables that could affect the survival (duration). The survival experience of a given group of individuals is often described by three different but equivalent functions: the density function, denoted by f(t), from which the probability of experiencing the event of interest within a given interval of time is obtained; the hazard function, denoted by λ(t), which is the instantaneous rate at which the event of interest occurs; and the survivor function, denoted by S(t), which is the probability of surviving beyond time t.

Thus the lognormal hazard λ(t), survival S(t), and density f(t) functions of the baseline variable of birth interval data are:

$$\lambda_o(t) = \frac{\frac{1}{t\sigma 2\pi} exp\{-\frac{1}{2\sigma^2}[\ln(t) - \mu]\}}{1 - \Phi\left(\frac{\ln(t)-\mu}{\sigma}\right)} \tag{9}$$

$$S_o(t) = 1 - \Phi\left(\frac{\ln(t)-\mu}{\sigma}\right) \tag{10}$$

$$f_o(t) = \frac{1}{t\sigma 2\pi} exp\{-\frac{1}{2\sigma^2}[\ln(t) - \mu]\} \tag{11}$$

where $\Phi(z)$ is the standard normal cumulative distribution function, σ is the standard deviation and μ is the mean of the normal distribution.

Therefore, under accelerated failure time model, the survivor function for the i^{th} individual corresponding to other explanatory variables is then

$$S_i(t) = P(T_i \geq t) = P\{exp(\beta_o + \boldsymbol{\beta' Z_i} + \sigma\varepsilon_i) \geq t\}$$

Now, $S_i(t)$ can be written in the form

$$S_i(t) = P\{exp(\beta_o + \sigma\varepsilon_i) \geq t/exp(\boldsymbol{\beta' Z_i})\}$$

and the baseline survivor function, $S_o(t)$, the survivor function of an individual for whom **Z = 0**, is

$$S_o(t) = P\{\exp(\beta_o + \sigma\varepsilon_i) \geq t\}$$

It then follows that

$$S_i(t) = S_o\{t/\exp(\boldsymbol{\beta}'\mathbf{Z}_i)\} \tag{12}$$

or

$$S_i(t) = 1 - \Phi\left(\frac{\ln(t) - Zi'\beta - \mu}{\sigma}\right)$$

This is the general form of the survivor function for the i^{th} individual in an accelerated failure time model. The factor exp(β'Z) is called an acceleration factor which tells us how a change in covariate values changes the time scale from the baseline time scale. And the corresponding relationship between the hazard function is obtained using equation (8). The assumption of lognormal distributions which have a hazard rate is that initially the rate increases and then declines after reaching a peak.

After fitting the lognormal accelerated failure time model, the importance of each of the explanatory variables will be assessed by carrying out statistical tests of the significance of the coefficients. The Pearson's Chi-square, the likelihood ratio tests (LRT), Hosmer and Lemeshow Test and the Wald tests are the most commonly used measures of goodness of fits (Hosmer and Lemeshow, 1989).

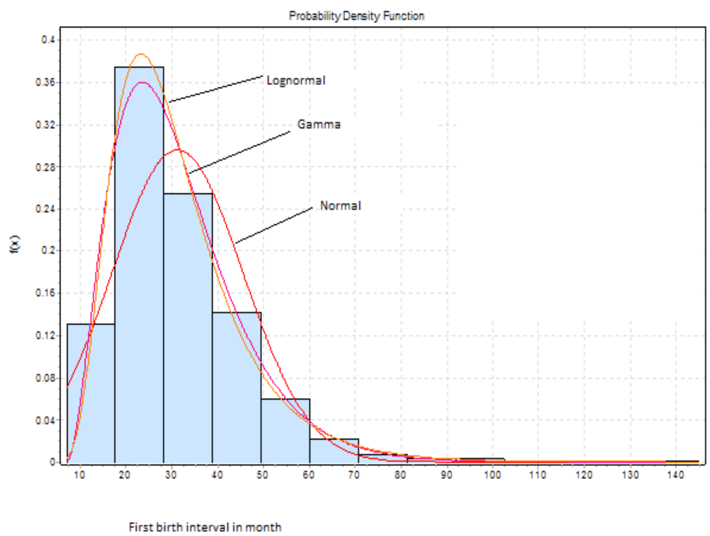
The Wald test is one of a number of ways of testing whether the parameters associated with a group of explanatory variables are zero. If for a particular explanatory variable, or group of explanatory variables, the Wald test is significant, then we would conclude that the parameters associated with these variables are not zero, so that the variables should be included in the model. If the Wald test is not significant then these explanatory variables can be omitted from the model. The Wald test is commonly used to perform multiple degree of freedom tests on sets of dummy variables used to model categorical variables in regression.

The parameter vector $\boldsymbol{\beta}$ can be estimated by using Maximum likelihood method of estimation. The statistical test used as a criterion for comparison is most often the Likelihood ratio (lR) test. However, the 95% CI of β and score test is also frequently used by software Packages.

The focus of parametric models is primarily the timing function. The parameter estimates reflect negative or positive effects on the timing of the event under study. To obtain the average time difference between the reference group and another covariates, the coefficients are suitably transformed by exponentiation (e^{β}), so that they can be interpreted as hazard time ratios. For instance, $\exp(\beta_1)$, $\exp(\beta_2)$, …, $\exp(\beta_p)$ stands for the hazard time ratio covariate of having p +1 categories compared to the hazard of a subject who has a value zero.

For each covariate a time ratio greater than one (or $\beta_i > 0$) indicates a longer duration for the group with the associated characteristic than for the reference group. Conversely, a time ratio less than one (or $\beta i < 0$) works to decelerate the timing of the event compared with the reference category. For example, if in the timing of second birth, the time ratio associated with variable *k* is greater than one, it is expected that women with this attribute would have a longer survival time and thus a lower risk of second birth than women in the reference category (Hosmr and Lemeshow, 1998).

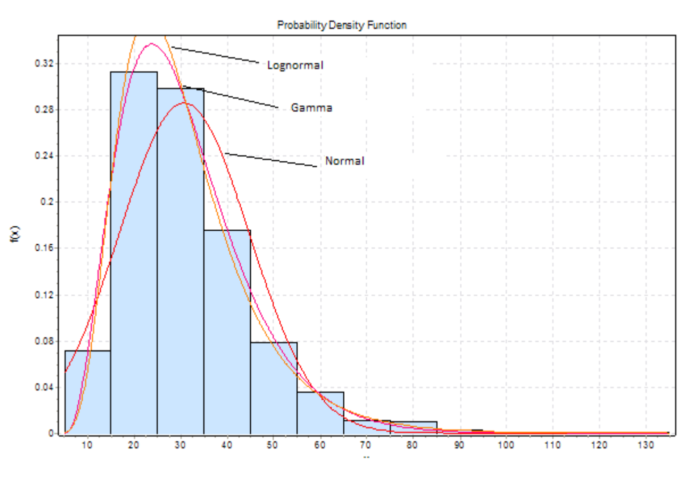
Table 4.1 Goodness of fit of first birth interval



#	Distribution	Parameters
1	Gamma	α=4.7564 β=6.558
2	Lognormal	σ=0.42833 μ=3.3474
3	Normal	σ=14.303 μ=31.193

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	3.8884	2	**Decision (Reject?)**	Yes	Yes	No
2. Lognormal	2.1266	1		No	No	No
3. Normal	103.15	3		Yes	Yes	Yes

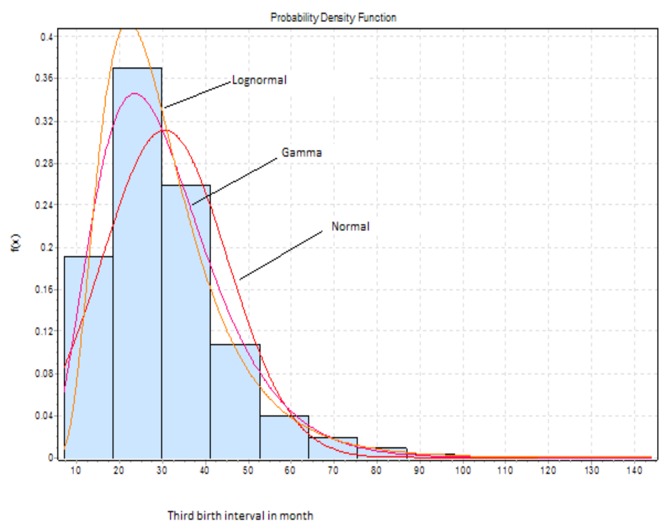
Table 4.2 Goodness of fit of second birth interval



#	Distribution	Parameters
1	Gamma	α=4.8459 β=6.3378
2	Lognormal	σ=0.43463 μ=3.3303
3	Normal	σ=13.952 μ=30.712

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	14.445	2	**Decision (Reject?)**	Yes	Yes	Yes
2. Lognormal	2.1052	1		No	No	No
3. Normal	93.579	3		Yes	Yes	Yes

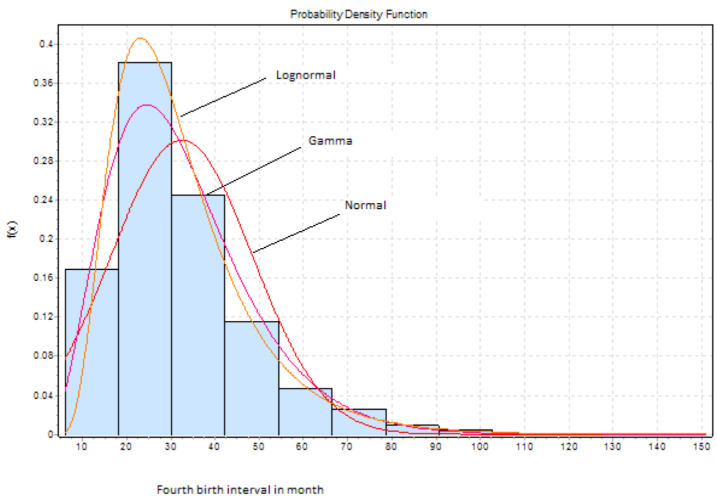
Table 4.3 Goodness-of-fit of third birth interval



#	**Distribution**	**Parameters**
1	Gamma	α=4.3568 β=7.0115
2	Lognormal	σ=0.44916 μ=3.3175
3	Normal	σ=14.635 μ=30.548

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	2.9423	2	**Decision (Reject?)**	Yes	No	No
2. Lognormal	1.9117	1		No	No	No
3. Normal	89.983	3		Yes	Yes	Yes

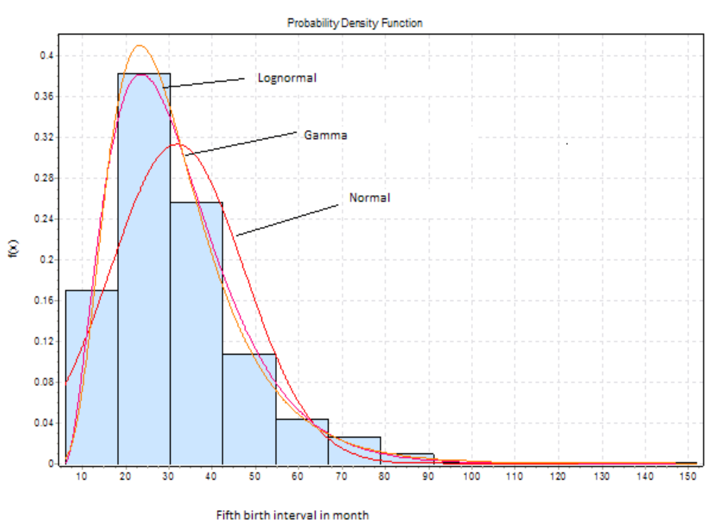
Table 4.4 Goodness-of-fit of fourth Birth Interval



#	Distribution	Parameters
1	Gamma	α=4.2499 β=7.5094
2	Lognormal	σ=0.4568 μ=3.3583
3	Normal	σ=15.481 μ=33.914

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	4.3839	2	**Decision**	Yes	Yes	Yes
2. Lognormal	1.1673	1	**(Reject?)**	No	No	No
3. Normal	77.741	3		Yes	Yes	Yes

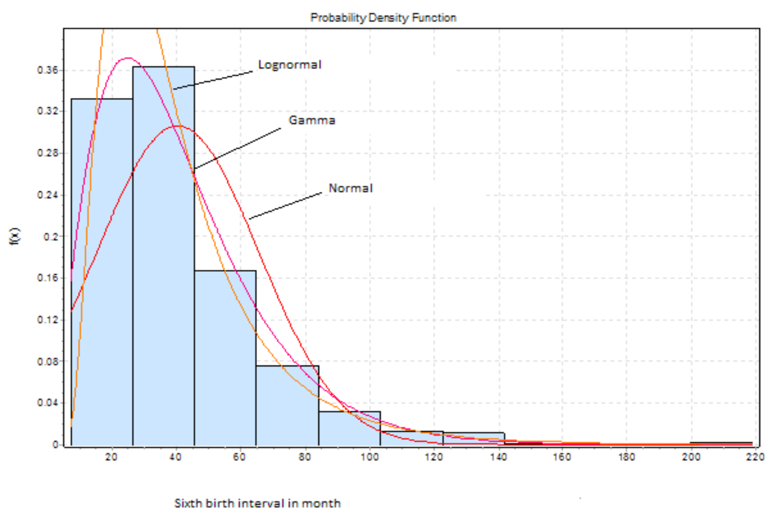
Table 4.5 Goodness-of-fit of fifth birth interval



#	Distribution	Parameters
1	Gamma	α=4.1107 β=7.8898
2	Lognormal	σ=0.45994 μ=3.3721
3	Normal	σ=15.996 μ=32.432

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	2.2915	2	Decision (Reject?)	Yes	Yes	Yes
2. Lognormal	0.90475	1		No	No	No
3. Normal	49.307	3		Yes	Yes	Yes

Table 4.6 Goodness-of-fit of sixth birth interval



#	Distribution	Parameters
1	Gamma	α=2.5781 β=15.632
2	Lognormal	σ=0.56258 μ=3.535
3	Normal	σ=25.1 μ=40.301

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	3.0104	2	Decision (Reject?)	Yes	Yes	Yes
2. Lognormal	1.8166	1		No	No	No
3. Normal	61.098	3		Yes	Yes	Yes

CHATER FIVE

RESULTS

5.1 The Effect of Different Demographic and Socio-economic Covariates on Birth Interval

a. The Effects of variability of Region on Birth Interval

To have an idea about the individual effects of the different explanatory variables on birth interval, we fitted accelerated failure time model for each covariate by taking one category as reference group, which is the approach we have taken here. In Table 5.1a to Table 5.1f, therefore, we have provided a summary of the results for the length of birth intervals in order to facilitate comparisons between different groups of covariates.

The effects of the various covariates on transition to 1st up to 7th births are shown in Tables 5.1a to 5.1f. These tables provide the final log-normal accelerated failure time model results based on the determinants of length of birth interval. The results include the estimates of the regression parameters of the model, the standard error of the coefficients, the standard score(Z), P-values, 95% confidence intervals of the coefficients of predictors of birth interval and the estimated time ratios.

Remember that a time-ratio of one indicates that the estimated risk possibility is equal to that of the reference category; a time-ratio (or relative chance) greater than one indicates that the time of birth interval is higher than that of the reference category; and a time-ratio less than one indicates lower time of birth interval than for the reference category.

First we shall consider the average difference of birth intervals by using lognormal accelerate failure time among different regions for the first sixth birth orders. We get the coefficients of each region from the regression model. The regression equation describes the relationship between region and failure time (birth interval) for the Northern, Central and Southern regions by taking Eastern as reference category. As we can see from the low p-values, the average birth intervals between regions are significantly different at 5 percent significant level.

The literature has often suggested inter-birth intervals vary widely between populations, even in the absence of deliberate spacing efforts, because of differential fecundability and postpartum amenorrhea. These are strongly influenced by variation of coital frequency and breastfeeding in different regions (Wood, 1994) which may or may not be manipulated consciously to space births. But, it is impossible to draw conclusions about the presence or absence of controlled birth spacing from a simple comparison of mean or median lengths of birth intervals between groups or time periods. So, we used method of lognormal accelerated failure time technique in order to know for natural determinants of birth intervals.

In all birth orders the socio demographic factor (region) has influence in the length of birth interval (Table 5.1a – 5.1f). Mothers in the northern region have relatively large mean age gap between their children than other regions. For instance at the first birth order (Table 5.1a), mothers from the northern region on average by 14.60 percent longer than women in eastern region. Or we can say that for the first birth order the median age time (or any percentile) between children in the group with North region ($z = 1$) is $exp(0.1365) = 1.1460$ times the median age time in the group with East region ($z = 0$) . From this the multiplicative covariate (region) effect on the time scale is clearly and easily observed.

Similarly Central and South region mothers elongate the age gap between their first and second children on average by 0.54 and 7.49 percent longer than women in East region, respectively. For example, from Table 5.1b we see that, the time to the third birth is 17.94 percent longer among Northern, 4.70 percent longer among Central, and 10.29 percent longer among Southern region than Eastern region.

Summary results from Lognormal-accelerated failure-time model for the effect of Covariates on Birth Interval

Table 5.1a First Birth Interval

Effect	DF	Wald-Chi-Square	P-value
Wealth	2	2.0134	0.3654
Education	2	22.505	<0.0001
Region	3	54.9922	<0.0001
Religion	3	25.9824	<0.0001
SSIC	1	6.5800	0.0103

Predictor		$\hat{\beta}$	S.E($\hat{\beta}$)	Z-Score	p-value	95.0% CI of β Lower	Upper	exp(β)
Region	Intercept	3.361	0.02697	120.59	0.000	3.308	3.414	28.817
	Eastern(ref)		0.0245					1.1460
	Northern	0.1365	0.0212	5.53	0.000	0.0883	0.1845	1.0054
	Central	0.0054	0.0225	-0.01	0.798	-0.0471	0.0362	1.0749
	Southern	0.0723		3.16	0.001	0.0281	0.1165	
Education level	Illiterate							
	Primary	-0.0305	0.01908	-1.08	0.110	-0.06791	0.0068	0.9699
	Sec.& above	0.249	0.0568	4.25	0.000	0.1376	0.3605	1.2827
Religion	Coptic-Orthodox(ref)							
	Protestant	-0.0160	0.0217	-0.87	0.459	-0.0586	0.0265	0.9841
	Muslim	-0.0836	0.0188	-4.46	0.000	-0.1206	-0.0467	0.9197
	Others	-0.0677	0.0373	-1.21	0.069	-0.1408	0.0054	0.9345
Wealth	Poor(ref)							
	Middle	-0.0256	0.0188	-1.43	0.174	-0.0625	0.0112	0.9747
	Rich	-0.0074	0.0155	-0.73	0.635	-0.0379	0.0231	0.9926
Survival status of index child (SSIC)	Survived(ref)							
	Died as infant	-0.0452	0.0175	-2.58	0.010	-0.0795	-0.0109	0.9558
	Scale	0.4814	0.0043			0.47312	0.48998	

==

Note: All the first groups of categorical covariates are used as reference groups.
exp (β) stands for time ratio.

Table 5.1b Second Birth Interval

Effect	DF	Wald Chi-Square	P-value
Wealth	2	4.3861	0.1116
Education	2	25.1713	<0.0001
Region	3	44.7648	<0.0001
Religion	3	16.5119	0.0024
SSIC	1	6.374	0.0116

Predictor		$\hat{\beta}$	S.E($\hat{\beta}$)	Z-Score	P- value	95.0% CI Lower	Upper	exp($\hat{\beta}$)
	Intercept	3.325	0.029	110.59	0.000	3.26870	3.38319	27.798
Region	Eastern(ref.)							
	Northern	0.165	0.026	6.04	0.000	0.11346	0.21783	1.1794
	Central	0.046	0.022	2.09	0.040	0.00203	0.09169	1.0470
	Southern	0.098	0.024	3.97	0.000	0.05101	0.14662	1.1029
Education level	Illiterate(ref.)							
	Primary	0.0170	0.0212	0.83	0.423	-0.02467	0.05874	1.0170
	Sec.& above	0.3644	0.0730	4.99	0.000	0.2212	0.5077	1.4396
Religion	Coptic-Orthodox(ref.)							
	Protestant	-0.077	0.023	-3.37	0.001	-0.1236	-0.0314	0.9258
	Muslim	-0.069	0.020	-3.67	0.001	-0.1091	-0.0288	0.9333
	Others	-0.027	0.040	-0.71	0.497	-0.1056	0.0512	0.9733
Wealth status	Poor(ref.)							
	Middle	-0.003	0.0202	12.5	0.885	-0.0425	0.0367	0.9970
	Rich	-0.029	0.0168	-0.83	0.076	-0.0627	0.0031	0.9714
Sex-composition	a male &a female(ref.)							
	Males	0.0199	0.0159	-0.49	0.211	-0.011	0.0510	1.020
	Females	-0.0135	0.0164	-2.06	0.409	-0.0457	0.0186	0.986
Survival status of index-child (SSIC)	Survived(ref)							
	Died as infant	-0.0433	0.0174	-2.4	0.013	-0.0779	-0.0093	0.9575
	Scale	0.4753	0.0046			0.4662	0.4845	

Note: All the first groups of categorical covariates are used as reference groups.
exp (β) stands for time ratio.

All value of time ratios indicate that East region has lower mean birth interval in all birth orders than other regions next to Central region. Since we compute the accelerated failure time value by using reference group as eastern region and all corresponding coefficient (β) for each region is greater than zero or the time ratio analogous to each region is almost greater than one. This suggests that women from eastern region have the shortest child spacing.

b. The Effect of Education level on Birth Interval

Martin (1995) showed that the influence of education and other socio-economic factors on number of children ever born, in which, women's education is inversely related to their fertility rate. That is directly related to time space or inter-birth interval between children. In this section we look at the effect of education on birth interval.

Table 5.1c Third Birth Interval

Effect	DF	Wald Chi-Square	P-value
Wealth	2	0.9803	0.6125
Education	2	12.0420	0.0024
Region	3	42.338	<.0001
Religion	3	5.0211	0.2851
Sex-composition	2	0.8967	0.6387
SSIC	1	17.9014	<0.0001

Predictor		$\hat{\beta}$	S.E($\hat{\beta}$)	Z-Score	P-value	95.0% CI for β Lower	Upper	exp($\hat{\beta}$)
	Intercept	3.310	0.0336	97.40	0.000	3.214	3.34690	27.385
Region	Eastern(ref.)							
	Northern	0.1913	0.0298	6.92	0.000	0.1328	0.2498	1.2108
	Central	0.0842	0.0253	3.43	0.001	0.0345	0.1338	1.0878
	Southern	0.1208	0.0270	5.06	0.000	0.0679	0.1738	1.1283

Education level	Illiterate(ref.)							
	Primary	-0.0115	0.02491	0.67	0.644	-0.0603	0.0372	0.9885
	Sec.& above	0.2896	0.0875	3.05	0.001	0.118	0.4611	1.3358
Religion	Coptic-Orthodox(ref)							
	Protestant	-0.0370	0.02636	-1.42	0.160	-0.08871	0.0146	0.9636
	Muslim	-0.0438	0.02291	-1.53	0.056	-0.08872	0.0010	0.9571
	Others	0.0061	0.04489	-0.32	0.891	-0.08182	0.0941	1.0061
Wealth	Poor(ref.)							
	Middle	-0.0090	0.01770	-0.64	0.611	-0.0436	0.0256	0.9910
	Rich	-0.0167	0.01813	-0.75	0.355	-0.0523	0.0187	0.9834
Sex-composition	a male &a female(ref.)							
	Males	-0.009	0.0177	-0.72	0.611	-0.0436	0.0256	0.9910
	Females	-0.016	0.0181	-0.97	0.355	-0.0523	0.0187	0.9841
Survival Status of index child (SSIC)	Survived(ref)							
	Died as infant	-0.0780	0.0184	-4.32	0.000	-0.1142	-0.0418	0.92496
	Scale	0.475	0.0051			0.4651	0.4854	

===

Note: All the first groups of categorical covariates are used as reference groups.
exp (β) stands for time ratio.

For simple method of analysis in this paper we classified the respondents (mothers) into three categories, namely illiterate, those with primary level, and those with secondary level of education and above. Further classification into more than three groups may not give us good proportion of observations for extra analysis because Ethiopian rural community is known as family of farmers whose education level is very low.

The results of regression of lognormal-accelerated failure time in Table 5.1a to Table 5.1d level of education does has a significant effect on birth intervals. For instance, in Table 5.1a the Wald chi-square result corresponding to education level is 22.505, that is significant at 99% confidence interval (since the p-value is less than one percent). Specifically in Table 5.1a, the time ratio of the length of birth interval for primary and secondary educated mothers by taking illiterate mothers as reference category are 0.9699 and 1.2827. This shows that timing of birth interval is not as such different between primary and non-educated mothers; and there is great variation of length of birth interval between mothers with secondary and illiterate mothers. Again, from Table 5.1b, 5.1c, 5.1d we can see that length of birth interval is different between non-educated and secondary level mothers. But in all birth orders there is no significant variation for length of birth interval between primary and illiterate mothers.

The reason behind this might be educated mothers may using contraceptive and other available methods to exhibit voluntary fertility control through child spacing. We know that, relatively educated mothers have insight knowledge about family planning than those who have not educated mothers. Besides better knowledge and use of contraceptives, education also affects reproductive decisions through its influence on a wide spectrum of social psychological orientations in women, including freedom from tradition, heightened aspirations for themselves and their children, and attitudes and response toward smaller families (Gyimah, 2001). And the birth interval approach views the family building process as consisting of a series of transitions, where women move successively from marriage to first birth; from first to second birth; and so on, until they reach their completed family size.

The non significant effect of education on the length of birth at high parities (see Table 5.1e and 5.1f) and also between primary and illiterate respondents may be the data we used (EDHS 2005) includes all respondent mothers whose first childs' age may more than thirty years old. This indicates that the data does not only enclose the current information of rural community. Since, as we recognize in the previous years there were no good supply of contraceptive throughout the country, especially in rural part of Ethiopia.

c. The effect of Religion on Birth Intervals

In Ethiopia the great majority of the populations are followers of Coptic-Orthodox, Muslim and protestant. Our sample data did not contain a large number of Catholic families for consideration of birth interval in model building. As a result we classified the respondents into four categories as Coptic-orthodox, Protestant, Muslim and "others". We note that the religion of the family is taken the same as that of the respondent mother. For method of comparison between different religion groups Coptic-Orthodox is taken randomly as reference category. In realty, a stronger religious affiliation is expected to affect the rate of fertility among groups of people.

The results of regression of lognormal-accelerated failure time (Tables 5.1a, 5.1b and 5.1e) give low p-values (less than 0.05) for Wald chi-square test shows that there is considerable empirical evidence that religious affiliation affects birth interval. This supports the existence of a significant variation of the average age time gap among different religion groups.

Table 5.1a shows p-values of Protestant, Muslim and others are 0.459, 0.000 and 0.069, respectively. These indicate that in the first birth interval (the age gap between first child and second child) there is no significant difference between Coptic-orthodox and protestant families. The p-value of 0.000 associated with Muslim religion shows that there is a considerable difference in birth interval time between the two religions. The sign of the coefficient (β) for Muslim religion is negative (-0.0836) or equivalently the time ratio (0.9197) less than one shows that the average age gap between their first and second child is lower by 8.3 percent than for Coptic-orthodox women. Similarly for other religions, the coefficient value is 0.067, and the time ratio corresponding to that is 0.9345. It is not significant at five percent significant level. The 90% confidence interval shows for the existence of difference in time gap for first birth interval between Coptic-Orthodox and other religions. The large P-value (0.459) corresponding to protestant religion shows the similarity of length of birth-interval for Coptic-orthodox and protestant religion followers.

Similarly for second birth interval (Table 5.1b) the coefficients for protestant, Muslim and others are negative or equivalently the time ratio values are less than one. These results show that comparatively Coptic-orthodox families have on average longer birth interval than others. Likewise in higher birth orders (Tables 5.1c to Table 5.1f) the difference of time gap of children between Coptic-orthodox and Muslim religion seems significant.

Table 5.1d Fourth Birth Interval

Effect	DF	Wald Chi_Square	P_value
Wealth	2	1.2780	0.5278
Education	2	6.5231	0.0383
Region	3	27.5906	<.0001
Religion	4	6.1165	0.1906
Sex-composition	2	1.3042	0.5209
SSIC	1	7.9910	0.0047

Predictor		$\hat{\beta}$	S.E($\hat{\beta}$)	Z-score	P-value	95.0% CI Lower	Upper	exp($\hat{\beta}$)
	Intercept	3.28228	0.03863	85.49	0.000	3.2365	3.3879	26.634
Region	Eastern(ref.)	0.17625	0.03415		0.000	0.10931	0.2431	1.1926
	Northern	0.06930	0.02924	5.45	0.018	0.01199	0.1266	1.0710
	Central	0.09138	0.03096	2.44	0.003	0.03070	0.1520	1.0956
	Southern			3.08				
Education level	Illiterate(ref.)							
	Primary	-0.0161	0.02963	-0.59	0.587	-0.07419	0.0419	0.9840
	Sec.& above	0.274	0.1146	2.01	0.017	0.0495	0.4988	1.31521
Religion	Orthodox(ref.)					-0.1068	0.0124	
	Protestant	-0.04725	0.03043	-1.19	0.121	-0.1079	-0.0054	0.9538
	Muslim	-0.05667	0.02615	-1.84	0.030	-0.1451	0.0601	0.9449
	Others	-0.04247	0.05238	-0.23	0.417			0.9584

Wealth	Poor(ref.)							
	Middle	-0.00134	0.02031	-0.09	0.948	-0.0411	0.0384	0.9986
	Rich	0.02193	0.02089	1.21	0.294	-0.01900	0.0628	1.0221
Sex-composition	A male & a female(ref.)							
	Males	-0.0013	0.0203	-0.67	0.948	-0.0411	0.0384	0.9987
	Females	0.0219	0.0208	-1.24	0.294	-0.0190	0.0628	1.0221
Survival Status of index child (SSIC)	Survived(ref.)							
	Died as infant	-0.0562	0.0199	-2.83	0.005	-0.0952	-0.0172	0.945
	Scale	0.4835	0.0059			0.4720	0.4953	

===

Note: All the first groups of categorical covariates are used as reference groups.
exp (β) stands for time ratio.

d. The effect of Wealth on birth interval

The Wald chi-square results in Table 5.1a to Table 5.1f show that the socio-economic factor wealth status has no significant effect on birth intervals. For instance, in Table 5.1a the Wald chi-square value corresponding to this variable is 2.0134 which is not significance at 95% confidence interval. And the time ratio of the first birth interval for mid range and rich households compared with poor ones are 0.9747 and 0.9926, respectively, and are not significant. Because, the p-value corresponding to middle and rich groups of mothers are 0.174 and 0.635, respectively; that is not significant.

e. Sex Composition and Birth Interval

Another variable associated with birth interval length is sex composition of a family. We know that in rural Ethiopia children are considered as assets. Farmers need help in farm activities and mothers need help in taking care of child and house activities. Rural life is so hard that the livelihood cannot be handled by a single person but co-operatively by hard working. That being the case, we ask whether there is any sex preference that could affect the length of birth interval. To answer this question we considered the sex of the first two children of a family. As a result we classified family groups into three categories. These are: mothers who have no son in the first two parities, mothers who have no daughter in first two birth orders and others who have a son and a daughter. The reference category is therefore taken the third category (those have a son and a daughter).

The Wald chi-square results in Tables 5.1b to 5.1f show that the covariate sex composition has no significant effect on birth interval. In this setting, the sex of the previous child does not seem to influence the length of the interval. For example in Table 5.1c, even if the time ratio for the two categories is less than one (0.9910 for males and 0.9841 for female), the corresponding p-values are extremely large showing that sex composition is not a significant covariate.

Table 5.1e Fifth Birth Interval

Effect	DF	Wald Chi-Square	P-value
Wealth	2	0.0420	0.9792
Education	2	1.8677	0.3930
Region	3	28.974	<.0001
Religion	4	9.7463	0.0449
Sex preference	2	0.2182	0.8966
SSIC	1	0.0025	0.9602

Predictor		$\hat{\beta}$	S.E($\hat{\beta}$)	Z-score	P-value	95.0% CI Lower	Upper	exp($\hat{\beta}$)
	Intercept	3.2658	0.04237	76.18	0.000	3.1828	3.3489	26.201
Region	Eastern(ref.)							
	Northern	0.1899	0.03761	4.93	0.000	0.11622	0.2636	1.2091
	Central	0.0986	0.03195	3.16	0.002	0.03602	0.1612	1.1036
	Southern	0.1533	0.03441	4.33	0.000	0.08595	0.2208	1.1656
Education level	Illiterate (ref.)							
	Primary	0.0466	0.03467	1.47	0.178	-0.02130	0.1145	1.0477
	Sec.& above	0.0408	0.1479	0.28	0.783	-0.2491	0.3307	1.0416

Religion	Orthodox(ref.)							
	Protestant	-0.0908	0.0340	-2.77	0.008	-0.1576	-0.0240	0.9132
	Muslim	-0.0681	0.0288	-2.46	0.018	-0.1246	-0.0116	0.9341
	Others	-0.1133	0.0587	-1.79	0.054	-0.22841	0.0017	0.8928
Wealth status	Poor(ref.)							
	Middle	0.00555	0.02844	-0.43	0.845	-0.05020	0.0613	1.0055
	Rich	0.00387	0.02356	0.12	0.869	-0.04230	0.0500	1.0038
Sex-composition	A male & a female(ref.)							
	Males	-0.0089	0.0224	0.21	0.691	-0.05291	0.0350	0.9911
	Females	0.00215	0.0228	0.22	0.925	-0.0425	0.0468	1.0021
Survival Status of index child (SSIC)	Survived(ref.)							
	Died as infant	0.0010	0.0210	0.05	0.965	-0.0402	0.0423	1.001
	Scale	0.4653	0.0065			0.4526	0.4782	

==

Note: all the first groups of categorical covariates are used as reference groups.
And exp (β) stands for time ratio.

f. The effect of survival status of index child on birth interval

Birth-intervals were estimated using birth history information on birth order, dates of birth, survival status of each child (alive or died before his/her second birth date) and age at death (if it occurs).Thus, comparisons between length of birth-interval is done by taking women whose children are survived as reference group.

There is an expected effect of the survival status of the index child on the length of birth interval. Based on Tables 5.1a to 5.1d the death of the index child significantly reduces the transition time to subsequent births. For instance, in Table 5.1c the time ratio of birth interval when an infant has died is 0.9249. This implies that the interval between the third and the fourth birth is 7.52 percent shorter if the third child died as an infant than if the infant survived. Thus, the results demonstrate that the death of previous children is significantly linked up to fourth birth orders, and behave non significance effect at higher parities (see Table 5.1e and 5.1f).

This finding is consistent with previous research in Ghana and Kenya about the effects of the status of the index child on the risk of subsequent births (Gyimah, 2001). Based on this, women with child loss experience are less likely to use contraception and more likely to discontinue if they are already using contraception. The reason behind this is that couples want make deliberate efforts to bear another child in the hope of replacing the lost one.

Table 5.1f sixth Birth Interval

Effect	**DF**	**Wald Chi-Square**	**P-value**
Region	3	2.8725	<.0001
Education	2	3.9565	0.1383
Religion	4	2.1198	0.7137
Wealth	2	0.6273	0.7308
Sex-composition	2	2.2790	0.3200
SSIC	1	2.6398	0.1042

Predictor		$\hat{\beta}$	S.E($\hat{\beta}$)	Z-Score	P-value	95.0% CI for β Lower	Upper	exp($\hat{\beta}$)
	Intercept	3.2283	0.0516	60.83	0.000	3.1271	3.3295	25.236
Region	Eastern(ref.)	0.1993	0.0453	4.24	0.000	0.1104	0.2882	1.2205
	Northern	0.1510	0.0382	3.93	0.000	0.0760	0.2260	1.1629
	Central	0.1671	0.0417	4.01	0.000	0.0853	0.2489	1.1818
	Southern							
Education level	Illiterate (ref.)							
	Primary	-0.0146	0.0459	-0.22	0.750	-0.1047	0.0754	0.9855
	Sec.& above	0.3829	0.1958	1.99	0.050	-0.000	0.7665	1.4665
Religion	Orthodox(ref.)							
	Protestant	-0.053	0.0412	-1.58	0.198	-0.1338	0.0277	0.9483
	Muslim	-0.0389	0.0346	-1.26	0.262	-0.1069	0.0290	0.9618
	Others	-0.0136	0.0717	-0.52	0.849	-0.1543	0.1269	0.9864
Wealth	Poor(ref.)							
	Middle	-0.0196	0.0348	1.66	0.573	-0.0879	0.0486	0.9805
	Rich	-0.0227	0.0290	0.47	0.434	-0.0797	0.0342	0.9775
Sex-compositio	A male & a female(ref.)							
	Males	0.0406	0.0271	1.66	0.134	-0.0125	0.0937	1.0414
	Females	0.0097	0.0276	0.47	0.724	-0.0443	0.0638	1.0097

Survival Status of index child (SSIC)	Survived(ref.) Died as infant	0.0410	00.0252	-0.43	0.104	-0.0084 0.0905	1.0418
	Scale	0.4770	0.0079	1.62		0.4618 0.4928	

==

Note: All the first groups of categorical covariates are used as reference groups.
exp (β) stands for time ratio.

Table 5.2 shows the median age of inter-birth interval in different regions. This suggests birth intervals tend to become smaller up to third parities and become larger and larger when parities increase. This means on average women tend to desire less birth immediately after marriage and tend to compensate for this at later ages. And also it shows that when mothers become older on average the inter-birth interval between their children becomes longer and longer.

Table 5.2 Median age (in months) of Inter Birth interval in different Regions

Region	Birth order 1	2	3	4	5	6	7	8	Combined mean of eight birth intervals
Eastern	26	25	25	24	24	24	27	28	25.375
Northern	34	33	33	33	32	31	35	36	33.375
Central	29	28	27	27	28	28	29	29	28.125
Southern	30	29	28	28	29	28	29	31	29
Overall	29	29	28	28	28	29	30	30	28.575

Note: Overall refers the aggregate values of Ethiopia

The above result is consistent with two researchers Setty-Venugopal and Upadhyay (2002); older mothers tend to have longer subsequent inter-birth intervals than younger mothers. The reason behind this is that when women get older the fecundability decelerates because successive births physiologically deplete the mother's of energy and thus impair the interest to be pregnant. Therefore, the U-shape of the average birth interval in each region (large median birth interval in low and high birth orders and short median birth intervals in mid birth orders) shows that maternal age at the birth of the index child is associated with birth intervals.

Another result that one observes from Table 5.2 is that the combined mean of birth interval for each region. From these 25.375 months in Eastern region, 33.375 months in Northern region, 28.125 months in central region and 29 months in South region. We can examine that mothers who are in central and eastern regions have short mean birth interval compared to northern and southern regions.

Table 5.3 The average number of children ever-born by region and educational level of mothers

a. Region

Region	Number of children
North	4.56
East	4.65
Central	4.80
South	4.96
Overall	4.62

b. Education level

Education level	Number of children
Illiterate	4.76
Primary	3.66
Secondary	2.66

It may be argued that child spacing may not after all reduce the total fertility rate. In an attempt to verify this we related the overall mean birth interval with the average number of children born per mother in each region. The results are given in Table 5.3. The results suggest that child spacing does reduce the fertility rate as measured by the average number of children born. The average number of children ever-born is 4.76 for illiterate mothers, 3.66 for primary educated mothers and 2.66 for secondary mothers. Even though the average number of children has not been controlled for the age of mothers, there is a temptation to conclude that child spacing is a common means of voluntary birth control in rural communities.

5.2 Graphical view of Duration models

The main goals in the analyses of survival (duration) data are: to describe the distribution of the time (birth interval) variable, to compare the survival experiences (distributions) of different groups of a population, and to identify explanatory variables that could affect the survival (duration). The survival experience of a given group of individuals is often described by three different but equivalent functions: the density function from which the probability of experiencing the event of interest within a given interval of time is obtained; the hazard function, which is the instantaneous rate at which the event of interest occurs; and the survivor function, which is the probability of surviving beyond time *t* (not experiencing the event of interest until time *t*) (Hosmer and Lemeshow, 1998).

The graphs of density functions and the hazard functions are given for our specific data set to see the levels of each in different groups of covariates (Appendix A). The probability density functions clearly show that the first sixth birth intervals under study do not have the same form of distribution in the regions, education level and also religions. Northern region has lower peak (more right skewed), southern region, has the second lower peak of density function and comparatively eastern and central regions have high peaked density functions.

As described earlier mostly hazard-rate models measure the quantum of the event under investigation (child-birth in our present case) and express the rate of occurrence of such event as a function of some covariates. However, in duration models (accelerated failure time models) the interest is about the tempo of the event;

in which the dependent variable (length of birth interval) is the duration of the event. Therefore, we focused on the tempo of birth interval between successive births per mother in each covariate.

In low time intervals, for instance, below combined median birth interval (28 months), eastern and central regions have high hazard rates for birth compared with northern and southern regions. In high time intervals, for example above combined median (28 months), on average central and northern regions have high hazard rate for birth in all parities. Thus, southern region is found within the graph of other hazard curves. But it does not mean that, the hazard rate of birth interval in southern region is low. This suggests that, in southern region there is approximately equal proportion of births below and above the combined median time of inter-birth interval (Appendix A).

The shape of the density functions of birth interval for different educational levels is not also the same. The density function of secondary level and higher educated mothers' birth interval is skewed to the right than those for primary and illiterate groups. Therefore, the given hazard rate for birth interval of each group of covariate give us direction to infer about the magnitude of probability of bearing a child in an interval length.

5.3 Probabilities of Birth Interval Length

As we noted in the above section(s) the appropriateness of a model for a given set of duration data depends on the specific data set at hand. Parallel to this, we have also noted that inferences concerning covariate effects on a given time variable depend on model choice. Thus, for our data set, the log-normal-accelerated failure time model is a parsimonious model to identify significant variables for durations between successive births (see Table 5.1a to Table 5.1f).

The other model that is used for comparisons of birth interval between groups is calculated probabilities. Birth-interval probabilities, p, represent the probability that a woman will give birth between time t and $t + h$, $h > 0$, after a prior birth, conditional on her surviving up to time h. Thus, the probability of giving birth to a child for any group of family gives direction for large or short length of birth interval, which has a direct correlation with rate of fertility. To identify the statistical distribution which fits birth interval, we fitted three statistical distributions. The result shows that almost all birth intervals are approximated very well by lognormal distribution (see in Appendix of Table B). Therefore, the computation of probability for significant variables is done using the lognormal distribution.

The observed covariates that influence (see in Section 5.1.1) length of birth interval are religion, region, survival status index of child and mother's education level. The calculation of probabilities is based on classifying the interval length into eight classes. The way of classifying is Sturges procedure $k = 1 + 3.32\log(n)$, where k is number of classes desired to be and n is total number of sample observation.

5.4.1 Region

The fit based on region is given in Appendix B, where we observe that the lognormal distribution was found more appropriate. Estimates of the probabilities for the first five birth orders for each region are shown in Tables 5.4a – 5.4c. Some interesting and consistent results seem to prevail.

In each of the four regions and each of the first five birth intervals, the common birth interval length seems to be in the range of 24 – 31 months. Taking the class mid-point as representative of the interval length, the common birth interval length is about 28 months or 2.3 years. Hence, one can say that on average in rural Ethiopia women give birth to child with in a period of 2.3 years. The probability of having a child within this interval, almost in all parities, is highest in the East and lowest in the North with the probability for South and central region in between.

Probability of Having Child in a given Birth Interval of a Family

Table 5.4a North and East regions

Birth Interval in month	North region					East region				
	Birth order					Birth order				
	1	2	3	4	5	1	2	3	4	5
<=15	0.058	0.66	0.072	0.068	0.059	0.160	0.158	0.190	0.172	0.203
16 – 23	0.157	0.152	0.159	0.156	0.186	0.251	0.279	0.278	0.259	0.251
24 – 31	0.244	0.245	0.238	0.241	0.238	0.290	0.279	0.259	0.288	0.261
32 – 39	0.238	0.243	0.256	0.228	0.226	0.119	0.128	0.102	0.090	0.125
40 – 47	0.132	0.132	0.137	0.123	0.141	0.057	0.065	0.053	0.062	0.068
48 – 55	0.064	0.062	0.055	0.071	0.064	0.039	0.025	0.042	0.051	0.041
56 – 63	0.041	0.040	0.031	0.047	0.031	0.026	0.023	0.024	0.024	0.025
64 +	0.066	0.060	0.052	0.065	0.054	0.059	0.043	0.053	0.054	0.027

The probability of bearing a second child for each region below 31 months are: 0.459 of northern, 0.701 of eastern, 0.632 of central and 0.546 of southern regions. That means on average 45.9% of northern, 70.1% of eastern, 63.2% of central and 54.6% of southern mothers give birth to their second child below 31 months. The implication is that comparatively the length of a birth interval is short in eastern region, large in northern region. When we look at the more probable interval length (greater than 64 months), northern region exhibits large probabilities compared to the other three regions which is happens at lower and also higher parities.

Results that support Table 5.2 are also observed here. That is, there is a higher probability in eastern and central regions for shorter birth intervals, less than about 31 months, and small probability for large intervals compared with North and South regions. Hence, it seems reasonable to say that eastern region exhibits a high fertility pattern and the North has low fertility rate with South and Central part in between relative to each other.

Table 5.4b Central and South regions

Birth Interval in month	Central region					South region				
	Birth order					Birth order				
	1	2	3	4	5	1	2	3	4	5
<=15	0.132	0.120	0.116	0.121	0.117	0.101	0.118	0.109	0.105	0.108
16 – 23	0.234	0.217	0.251	0.218	0.230	0.205	0.208	0.218	0.223	0.217
24 – 31	0.266	0.271	0.256	0.281	0.278	0.240	0.241	0.254	0.249	0.245
32 – 39	0.169	0.193	0.165	0.154	0.183	0.184	0.187	0.172	0.163	0.182
40 – 47	0.095	0.096	0.104	0.093	0.087	0.111	0.114	0.110	0.125	0.112
48 – 55	0.035	0.045	0.054	0.052	0.039	0.054	0.051	0.057	0.054	0.055
56 – 63	0.030	0.019	0.027	0.028	0.024	0.040	0.030	0.030	0.032	0.035
64 +	0.039	0.039	0.027	0.052	0.043	0.064	0.050	0.050	0.050	0.045

Table 5.4c Overall Ethiopia

Birth Interval in Month	Birth Order					
	1	2	3	4	5	6
<=15	0.107	0.112	0.116	0.111	0.114	0.118
16 – 23	0.207	0.209	0.221	0.210	0.217	0.222
24 – 31	0.257	0.256	0.252	0.261	0.253	0.250
32 – 39	0.184	0.192	0.179	0.166	0.184	0.173
40 – 47	0.102	0.105	0.104	0.106	0.107	0.105
48 – 55	0.049	0.048	0.053	0.057	0.052	0.055
56 – 63	0.035	0.029	0.028	0.034	0.030	0.032
64^{+}	0.059	0.049	0.046	0.054	0.044	0.044

Table 5.4c shows an overall fertility pattern. One can easily see the probability of having a child within a given interval among first seven children. The most common birth interval with large probabilities seems to be about two to three years. Hence, the previous conclusions based on median age time in month seem to be consistent.

5.4.2 Education

The approximating distributions for birth interval in all categories of families are given in Table 2a, 2b, and 2c of Appendix B. Thus, each group of birth interval data follows the lognormal distribution with different parameters; we calculated the probabilities for a given range of birth intervals. Secondary level educated mothers have relatively lower probabilities in lower birth intervals than primary and non-educated mothers. From this one can conclude that education level of the mother has little effect on fertility pattern in rural community.

Probability of Having Child in a given Birth Interval by Education Level

Table 5.5a Non-educated

Birth Interval in Month	Birth Order 1	2	3	4	5
<=15	0.107	0.117	0.114	0.111	0.116
16 – 23	0.207	0.213	0.234	0.219	0.219
24 – 31	0.258	0.259	0.251	0.261	0.255
32 – 39	0.188	0.195	0.182	0.177	0.193
40 – 47	0.106	0.110	0.108	0.109	0.109
48 – 55	0.049	0.048	0.052	0.058	0.054
56 – 63	0.036	0.030	0.029	0.033	0.028
64 +	0.049	0.028	0.030	0.032	0.027

Table 5.5a and Table 5.5b present probabilities of bearing a child within a given interval length by considering education of mothers. It is clear from Table 5.5a and 5.5b that 24-31month is the more probable interval length of having a child for all categories.

Table 5.5b Primary and secondary

Birth Interval (month)	Primary					Secondary			
	Birth order					Birth order			
	1	2	3	4	5	1	2	3	4
<=15	0.112	0.093	0.147	0.116	0.083	0.086	0.047	0.167	0.114
16 – 23	0.21.8	0.200	0.160	0.190	0.244	0.200	0.279	0.267	0.234
24 – 31	0.255	0.272	0.298	0.294	0.288	0.300	0.186	0.133	0.251
32 – 39	0.177	0.220	0.207	0.165	0.161	0.200	0.256	0.267	0.182
40 – 47	0.102	0.115	0.093	0.129	0.107	0.100	0.070	0.100	0.108
48 – 55	0.057	0.046	0.060	0.055	0.029	0.071	0.093	0.067	0.052
56 – 63	0.029	0.024	0.016	0.029	0.054	0.029	0.023	0.033	0.029
64 +	0.050	0.029	0.019	0.022	0.034	0.014	0.047	0.113	0.030

5.4.3 Religion

The birth interval data that are grouped based on religion affiliation also follows a lognormal distribution (Table 3 in Appendix B); the probabilities calculated are given in Table 5.11a and 5.11b. When we look at if we see the cumulative probability of getting a second child below 31 months in different religion groups, we have: 0.487 for Coptic-orthodox, 0.553 for protestant, 0.668 for Muslim and 0.631 for "other" religions (see Table 5.6). This means that 48.7 % of orthodox, 55.3% of protestant, 66.8 % of Muslim and 63.1% of other religion mothers given birth to their second child within less than 31 months. These suggests that the median birth interval for Muslim is short, large for Coptic-orthodox religion followers, and the interval for protestant and other religion are between the two.

Probability of bearing a child in a given interval length by Religion

Table 5.6a Orthodox and Protestant families

Birth Interval (months)	Coptic Orthodox					Protestant				
	Birth order					Birth order				
	1	2	3	4	5	1	2	3	4	5
<=15	0.075	0.078	0.08	0.073	0.073	0.093	0.119	0.110	0.114	0.177
16 – 23	0.166	0.164	0.182	0.176	0.198	0.214	0.229	0.244	0.225	0.247
24 – 31	0.246	0.253	0.245	0.243	0.270	0.246	0.269	0.269	0.270	0.255
32 – 39	0.227	0.237	0.256	0.242	0.207	0.203	0.191	0.140	0.166	0.139
40 – 47	0.127	0.138	0.122	0.123	0.122	0.109	0.096	0.124	0.121	0.074
48 – 55	0.063	0.059	0.049	0.069	0.060	0.046	0.051	0.069	0.052	0.044
56 – 63	0.044	0.038	0.031	0.041	0.026	0.038	0.021	0,025	0.032	0.029
64 +	0.052	0.070	0.033	0.033	0.048	0.051	0.025	0.019	0.018	0.035

Table 5.6b Muslim and other religion family

Birth Interval	Muslim					Others				
	Birth order					Birth order				
	1	2	3	4	5	1	2	3	4	5
<=15	0.146	0.146	0.153	0.146	0.142	0.121	0.124	0.102	0.130	0.142
16 – 23	0.245	0.251	0.261	0.250	0.238	0.232	0.207	0.219	0.225	0.256
24 – 31	0.277	0.270	0.260	0.278	0.252	0.278	0.207	0.270	0.0268	0.276
32 – 39	0.143	0.157	0.139	0.126	0.166	0.152	0.260	0.153	0.109	0.142
40 – 47	0.081	0.091	0.083	0.094	0.092	0.091	0.101	0.109	0.080	0.066
48 – 55	0.039	0.036	0.049	0.047	0.045	0.045	0.024	0.073	0.087	0057
56 – 63	0.026	0.023	0.024	0.025	0.027	0.025	0.041	0.053	0.036	0.028
64 +	0.053	0.026	0.031	0.034	0.038	0.055	0.036	0.022	0.065	0.038

Chapter Six

Discussion conclusion and recommendation

Discussion

Table 5.1a to Table 5.1f present the covariates associated with the first six-birth interval in a multivariate context and lognormal accelerated failure time technique. Starting with the interval between first birth and the second birth, survival analysis indicates that the median duration of inter-birth interval is 28 months (Table 5.1). That is not similar in different region, education level or religion groups. For example, Table 5.1a suggests that about 50 percent of women with non-education had a second birth within 28 months after first birth compared with about 36 percent of those with secondary education. Results had been shown for all intervals; that is, secondary educated women space births more widely than their counterpart non-educated women, the effect being largest at higher parities. Thus, it is consistent with theoretical expectation (eg. Girmay, 1992), maternal education significantly relate with the later timing of births.

Maternal education can influence birth spacing through a number of pathways. Mothers space births more widely because they may have better knowledge of contraception and are more likely to use more efficient methods. As demonstrated in the Ethiopian 2005 DHS published report, about 52.6 and 23.4 per cent of women with secondary and primary education respectively were current users of contraception compared with only 10.0 percent of those with no education.

Variations of the length of birth interval by region difference are also noticeable. The time to the second birth is 14.60 percent longer among Northern, 0.54 percent longer among Central, and 7.49 percent longer among Southern regions than Eastern region(Table 5.1a). The explanation should rather be sought through ethnic norms and practices such as duration of breastfeeding and post-partum sexual abstinence that traditionally lengthen intervals between successive births.

There is an expected effect of the survival status of the index child on the transition to subsequent births. The death of the index child significantly reduces the transition time to subsequent births. For instance, in Table 5.1a the interval between the first and the second birth is 4.42 percent shorter if the first child died as an infant than if the child survived. The theoretical pathways through which childhood mortality affects fertility have been identified as physiological and behavioral replacement effects (e.g., Montgomery and Cohen, 1998). The physiological effect mainly operates through premature weaning (accustom) and shorten of the inter-birth interval following an infant death. In its usual interpretation, the behavioral hypothesis refers to the deliberate efforts a couple makes to bear another child in the hope of replacing the lost one. Again, within the framework of the behavioral effect, women with child loss experience are less likely to use contraception and, more likely to discontinue if they are already using contraception.

Further, some significant religions differences are also visible. Although not always significant Coptic-orthodox, Protestant, and Others like catholic and traditionalists tend to space births more distantly than Muslims. That means comparatively in all birth orders Muslim religion followers have short birth interval than followers of other religions. A number of explanations can be advanced to account for the observed pattern. For instance in Table 5.1a about 50 percent of Coptic-orthodox women had a second birth within 28 months after first birth compared with about 54 percent of Muslims. At all parities there is no significant difference of length of birth-interval between successive children by sex-composition and wealth variation.

Conclusions

This study has highlighted the relative significance of socio-economic, demographic and socio-cultural factors on the duration of birth intervals in rural Ethiopia. For all birth intervals, the log likelihood ratios suggest that region, education level, survival status of index child and religion affiliation explain more of the variance in the dependent variable than the two other variables, namely sex composition and wealth variation. Although the modernizing effects of the education seem to be powerful influence on the timing of births variation, demographic background or difference in region are also supported by these findings.

Child spacing has been found to play an important role in fertility transition as a result of its inverse relationship with family size. Groups with longer birth spacing tended to have a significantly fewer number of children ever-born. For example, in Table 5.3 we presented the mean number of children ever-born by selected characteristics. The mean number of children ever-born to women with secondary education is 2.66 compared with 4.76 among those with no education.

Recommendation

We know that making the spacing of birth intervals large has its own advantages. In addition to maternal and child health through the dynamics of sibling competition and maternal depletion hypotheses (theoretical view), large birth intervals reduce the total number of children in a family, which has a direct relation with fertility rate. According to 2007 Ethiopia population census, the annual population growth rate is very high (2.7%); and the fertility rate of the younger group among whom the concept of birth spacing is not yet popular has, in fact, shown a decreasing trend. Even if that happens, lengthening the median age gap to more than 28 month is desirable. According to many literatures the preferable time length for inter-birth is more than two years.

Comparatively, as we have shown in East and Central part of rural Ethiopia the length of successive birth is too short than in southern and northern regions. The other influential variable "education" has also a positive effect on birth intervals. Against these informations, it is logical to wonder the target of reducing fertility rate, especially given the prevailing information for advantage of birth spacing at the policy level. In respect of population growth, the government policy is to promote the attainment of small family size on a voluntary basis and a reduced population growth aligned with replacement fertility level. While efforts to make available the knowledge, means and opportunity to practice family planning by making large age gap between children should be intensified.

REFERNCE

Alter, G. (1988). Family and the Female Life Course: The Women of Verviers Belgium, 1849-1880: University of Wisconsin Press, Madison, Wisconsin.

Bartholomew, D.J. (1967). Stochastic Models for Social Processes (3rd Ed.). John Wiley and Sons, Inc.: New York.

Bongaarts, J. (1978). A framework for analyzing the proximate determinants of fertility: Population and Development Review 4(1):105-132.

Bumpass, L., Rindfuss, R. and Palmore J. (1986). Determinants of Korean birth intervals: Population Studies, 40(3): 403-423.

Bury, K.V. (1975). Statistical Models in Applied science. John Wiley and Sons, Inc.: New York.

Caldwell, J. C. and Caldwell, P. (1987). The cultural context of high fertility in sub-Saharan Africa. Population and Development Review 13(3):409-437.

Centre for Development and Population Activities (CEDPA) (1996). Gender Equity, Concepts and Tools for Development. Washington, DC.

Conover, W.J. (1980). Practical Nonparametric Statistics (2nd). John Wiley and Sons, Inc.: New York.

Cox, D. R. (1972). Regression models and life-tables (with discussion). Journal of the Royal Statistical Society - Series B 34, 187-220.

David, P. A. and T. A. Mroz (1989b). Evidence of fertility regulation among rural French villagers, 1749-1789. A sequential econometric model of birth-spacing behavior (Part 2), European Journal of Population 5(2). 2, 173-206.

D′Augostino, R.B. and Stevens, M.A. (1986). Goodness-of-Fit Techniques. Marcel Dekker.

Esterlin, R. (1975). A framework for fertility analysis. Studies in Family Planning 6:54-63.

Ethiopian Demographic and Health Survey (EDHS) (2005). Central Statistical Agency, Addis Ababa, Ethiopia and ORC Macro, Calverton, Maryland, USA.

Ghilagaber, G. Gyimah, S., Lofstedt, P. and Johansson, A. (2005). The Use of Flexible Parametric Duration Functions in Modeling the Tempo of Fertility: Applications to the Analysis of Birth Intervals in Rural China. Population Review 44(1): 11-35.

Gibbons, J.D. (1971). Nonparametric Statistical Inference. Mcgraw-Hill, Inc.: London.

Girmay Medhin (1992). Statistical distribution of birth intervals. M.sc thesis, Addis Ababa University, Ethiopia.

Gribble, J.N. (1993). Birth intervals, gestational age and low birth weight: Population Studies 47:133-146.

Gumbel, E. J. (1958). Statistics of Extreme. Columbia University Press, New York.

Gyimah, S. O. (2001). Childhood Mortality and Reproductive Behavior in sub-Saharan with Emphasis on Ghana and Kenya. A paper presented at the Population Association of America Meetings, Atlanta Hilton, GA May 8-11, 2002.

Hosmer D.W and Lemeshow S. (1998). Applied Survival Analysis Regression Modeling of Time to Event Data. John Wiley and Sons, Inc. New York.

Hobcraft, J., McDonald, J. and Rustein, S. (1985). Demographic determinants of early infant and early child mortality: A comparative analysis. Population Studies 39:363-385.

Kalbfleisch, J. D. and Prentice, R. L. (1980). The Statistical Analysis of Failure Time Data. John Wiley & Sons, Inc.: New York.

Johnson, N.L. and Kotz, S. and Balakrishnan, N. (1994). Distribution in Statistics: Continous Univariate Distributions (2nd Ed., Vol-1). John Wiley and Sons, Inc.: New York.

Martin, T. (1995). Women's education and fertility: Results from 26 Demographic and Health Surveys. Studies in Family Planning 26:187-202.

Mood, A.M., Graybill, F. A. and Boes D.C. (1974). Introduction to the theory of Statistics (3rd Ed.). McGraw–Hill: New York.

Montgomery, M and Cohen, B. (1998). From Death to Birth: Mortality Decline and Reproductive Change. Washington, D.C.: National Academy Press.

Moser, R. M. (2001). Reproductive health issue for refugees in Latin America. Personal Communication April 22 (2001).

National Research Council (1989). Contraception and Reproduction: Health Consequences for Women and Children in the Developing World. Washington, DC: National Academy Press.

Pederson, J. (2000). Determinants of infant and child mortality in the West Bank and Gaza strip. Journal of Biosocial Science 32 (4): 527-546.

Setty-Venugopal V. and Upadhyay, U.D. (2002). Birth Spacing: Three to Five Saves Lives. Population Reports, Series L, No. 13. Baltimore: Johns Hopkins Bloomberg School of Public Health, Population Information Program.

Trussell, J., Martin L., Feldman, R. and Palmore, J. A. (1985). Determinants of birth interval length in the Philippines: A Hazard Model Analysis. Demography 22 (2):145-168.

United Nations (1987). Fertility Behavior in the Context of Development: Evidence from the World Fertility Surveys. New York: United Nations.

Ware, H. (1984). Effects of maternal education, women's roles and child care on child mortality. The Population Council: New York.

Wood, J. W. (1994). Dynamics of Human Reproduction. Aldine De Gruyter: New York.

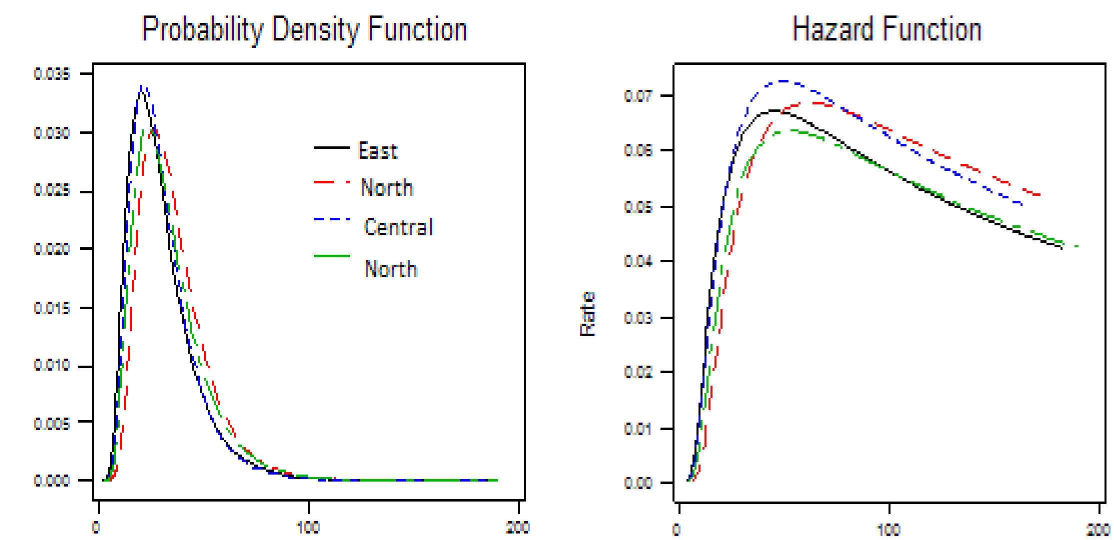
World Population Data sheet (2008). World Population High lights: The impact of Regional Patterns of Fertility on World Population Growth. Accessed from http://www.prb.org/pdf08/08wpds_eng.pdf

World Health Report (WHR) (2005). Make Every Mother and Child Count. World Health Organization, Geneva, Switzerland.

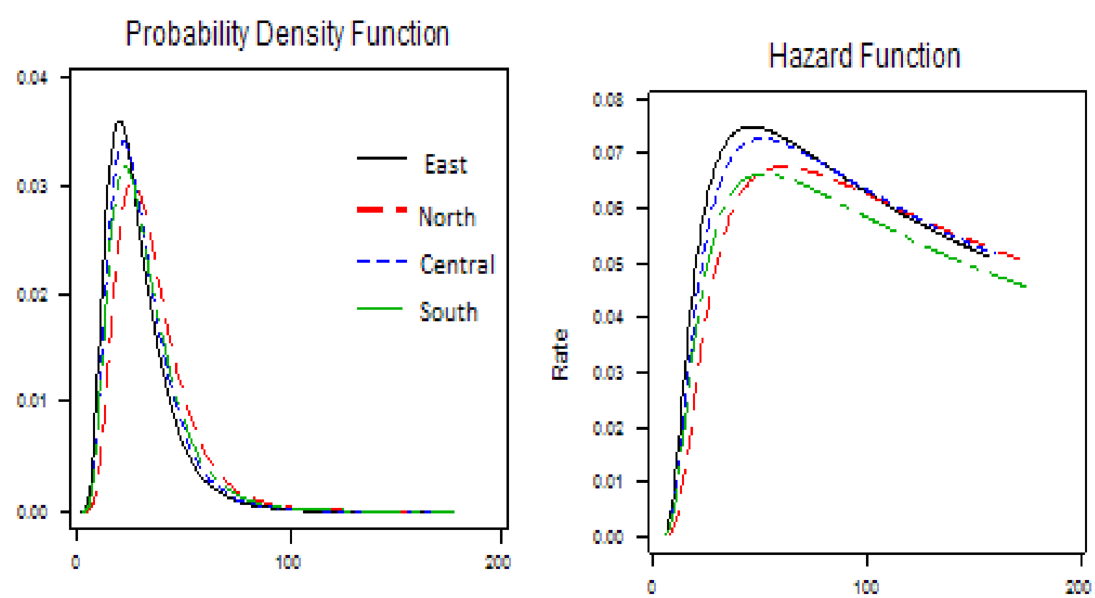
Appendix A

1) Probability density and hazard function by Region

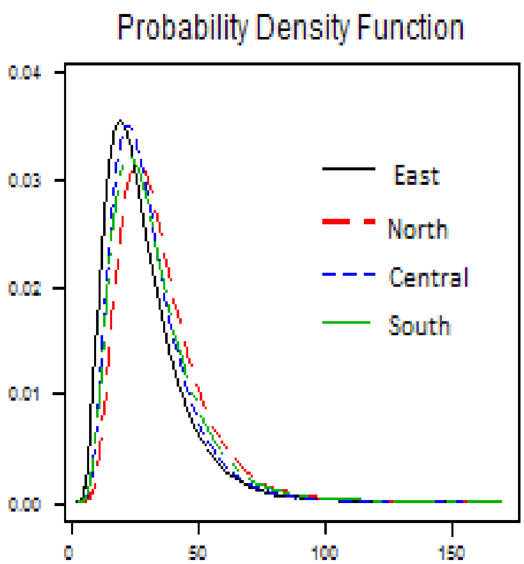
a. First birth-interval

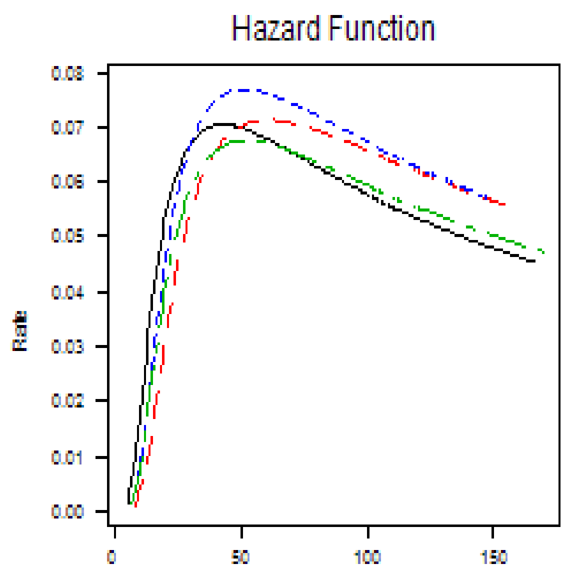


b. Second birth-interval

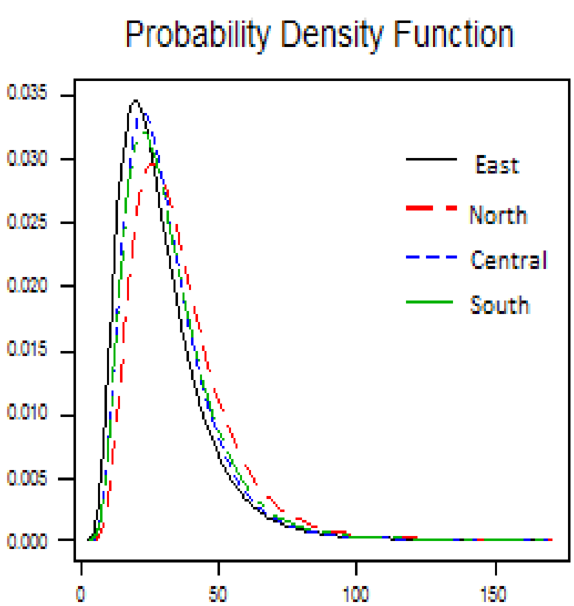


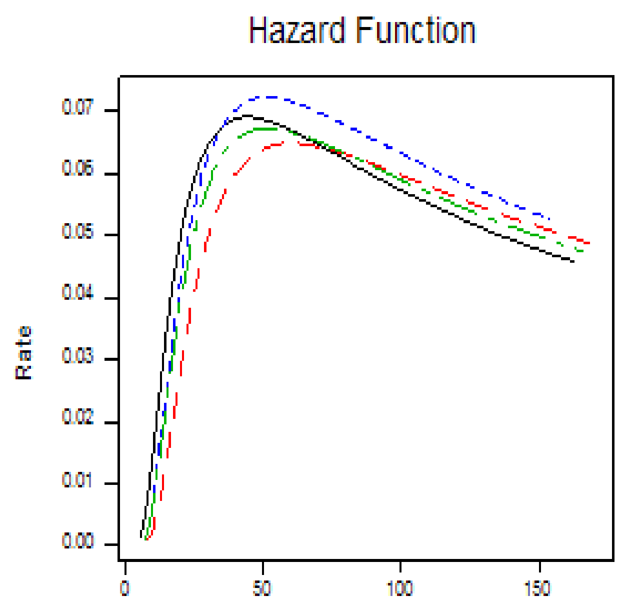
c. Third birth-interval



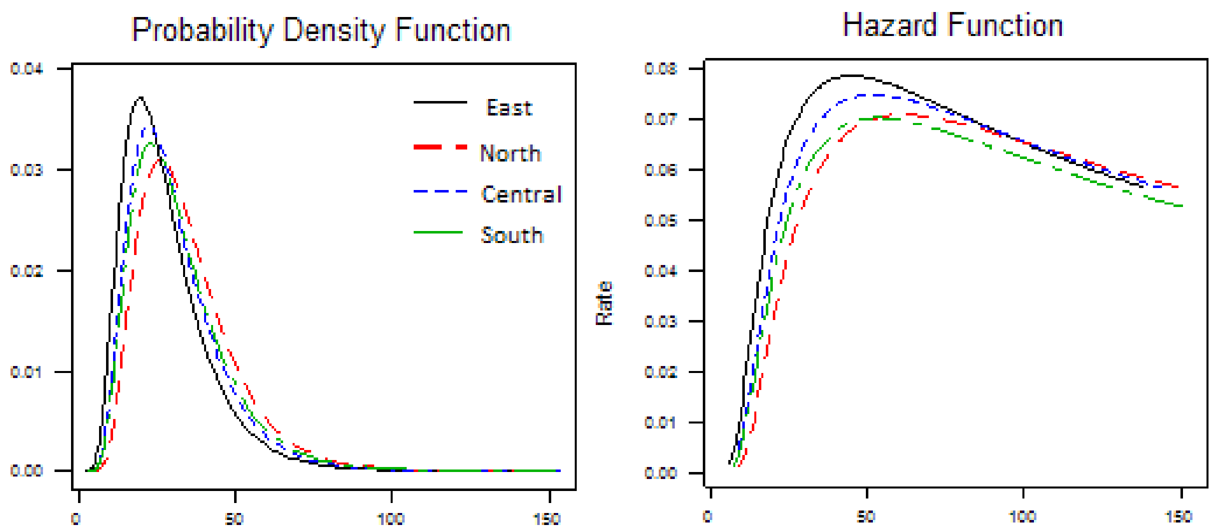


d. Fourth birth-interval

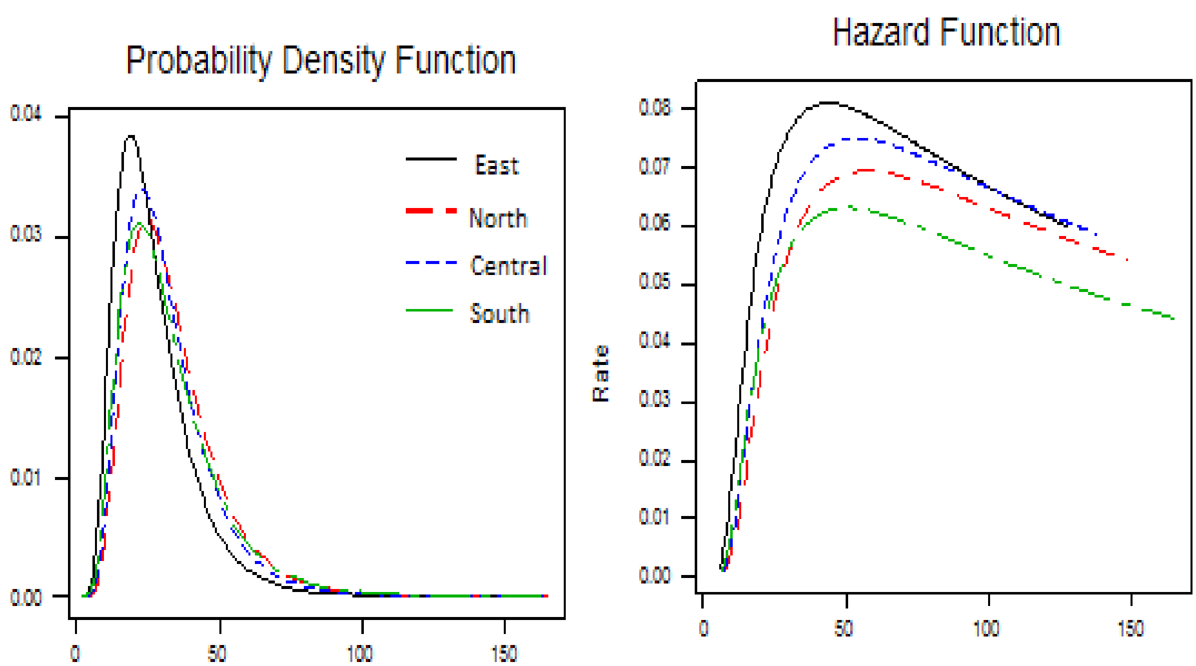




e. Fifth birth-interval

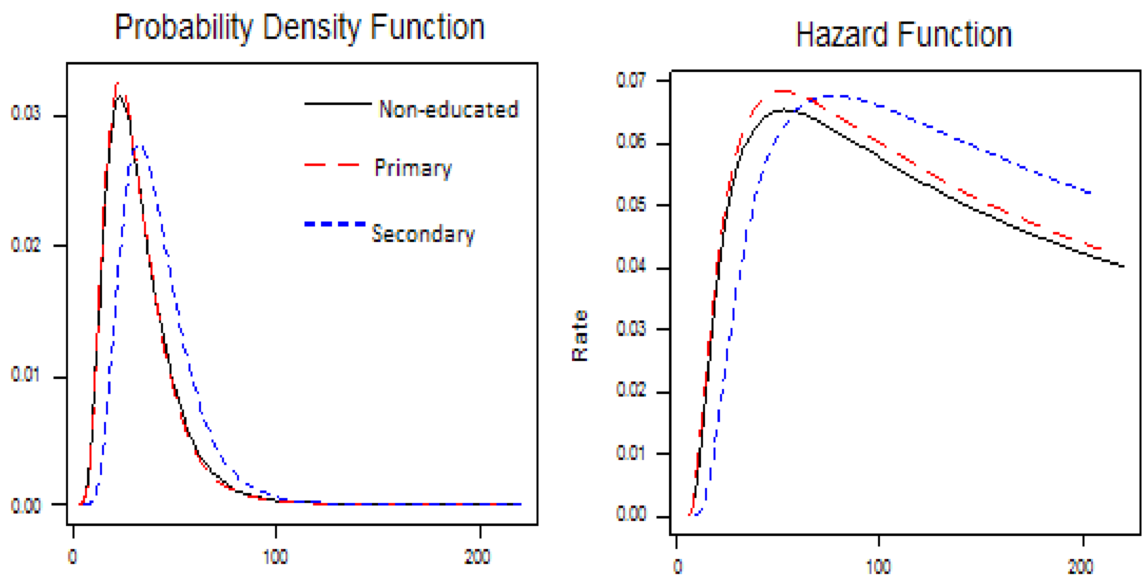


f. Sixth birth-interval

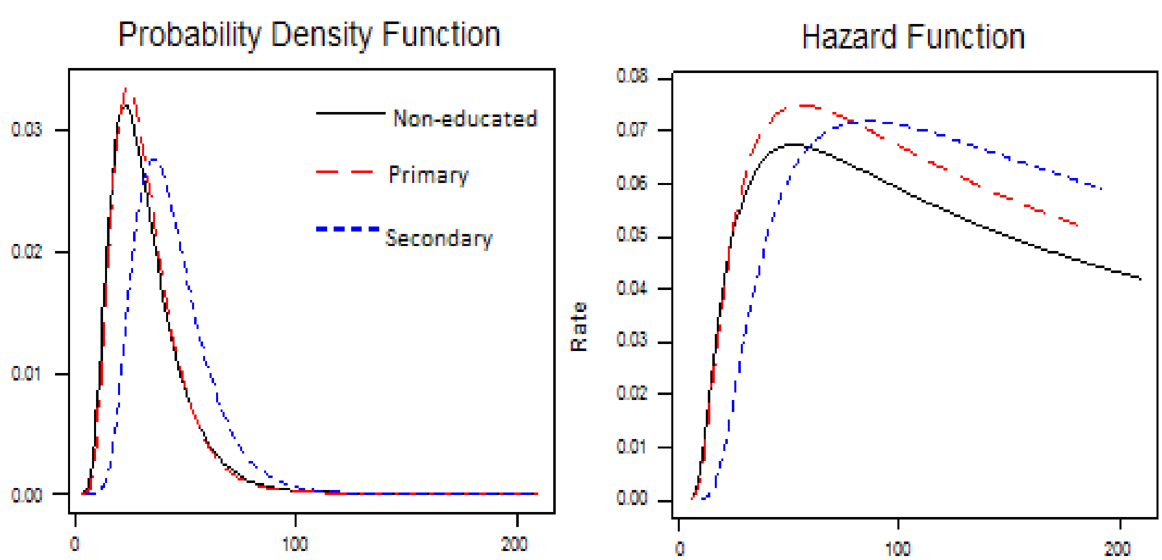


2) Probability density and hazard function by education level

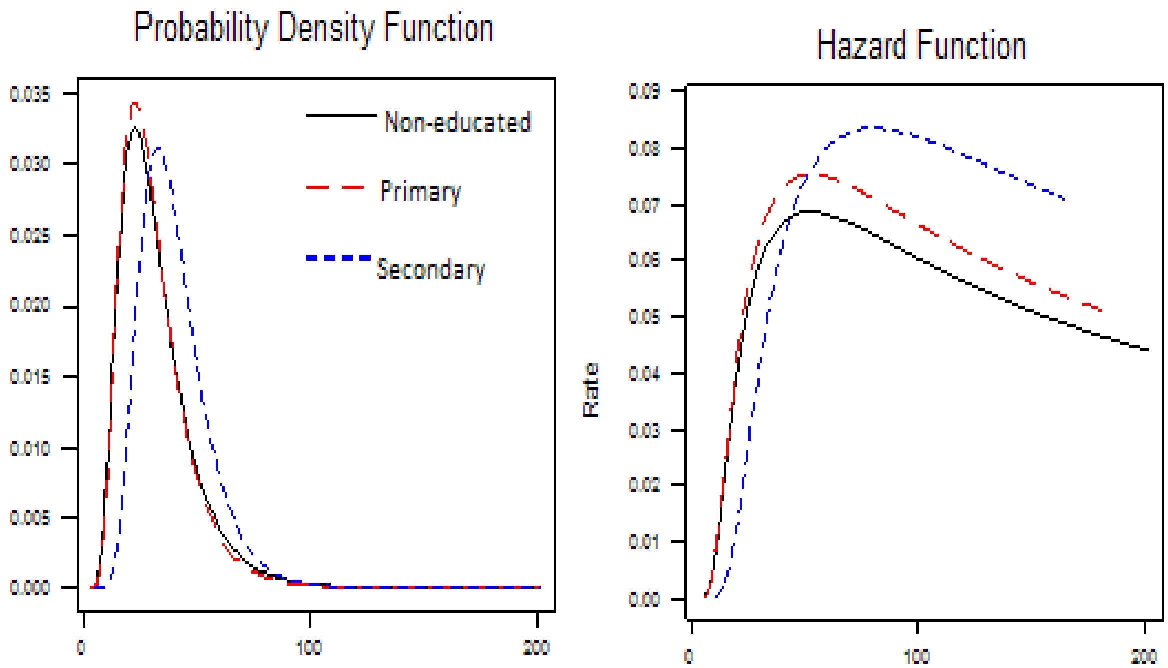
a. First birth-interval



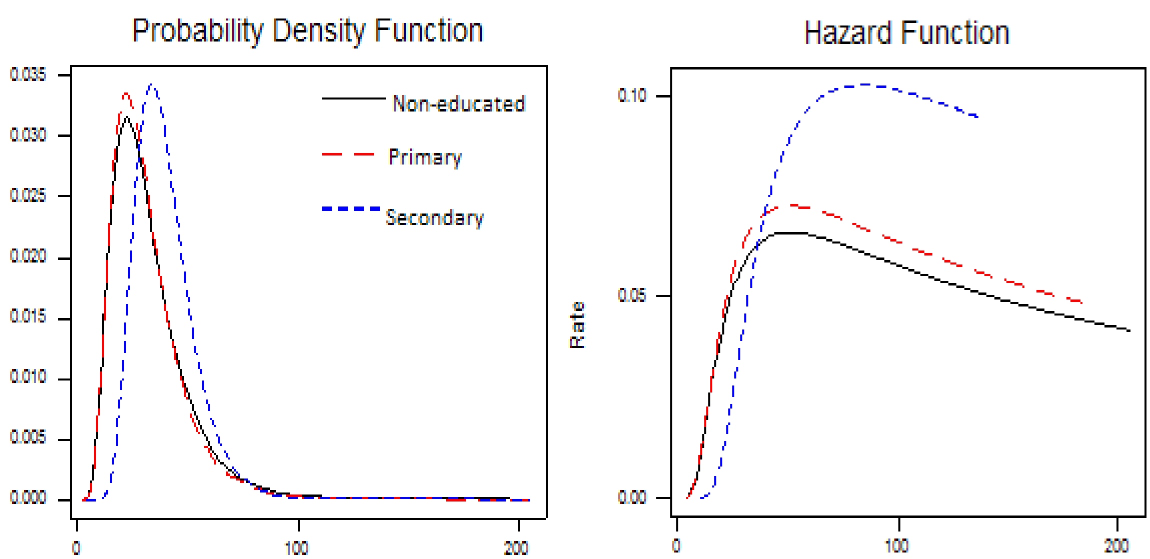
b. Second birth-interval



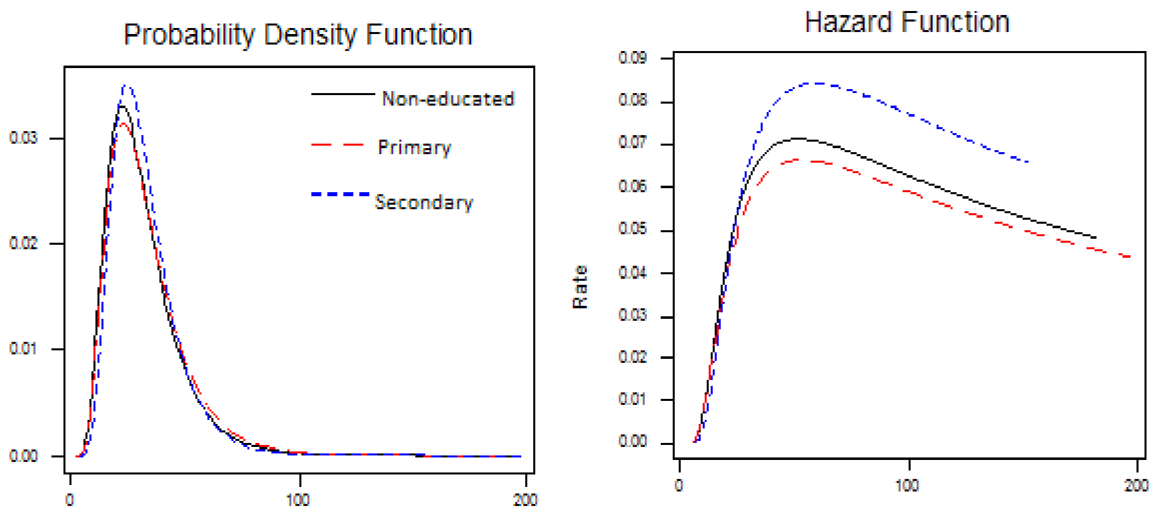
c. Third birth-interval



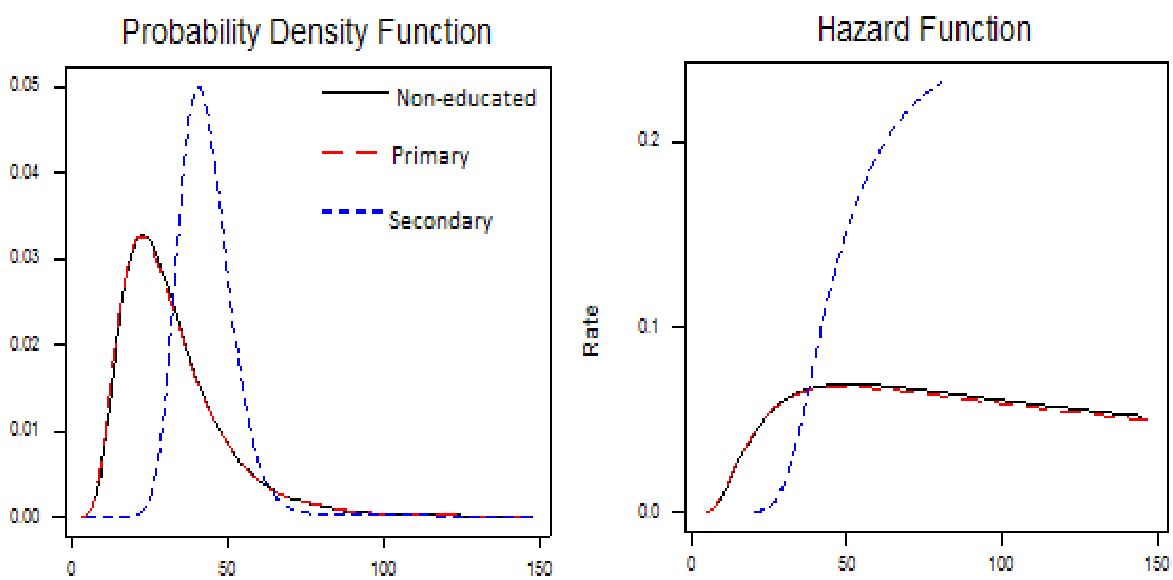
d. Fourth birth-interval



e. Fifth birth-interval

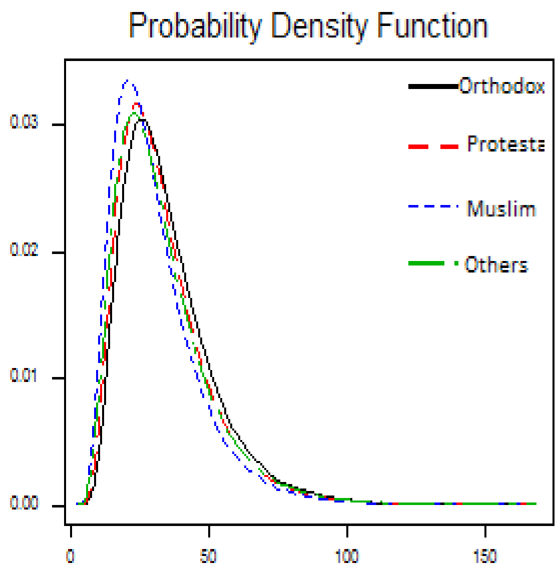


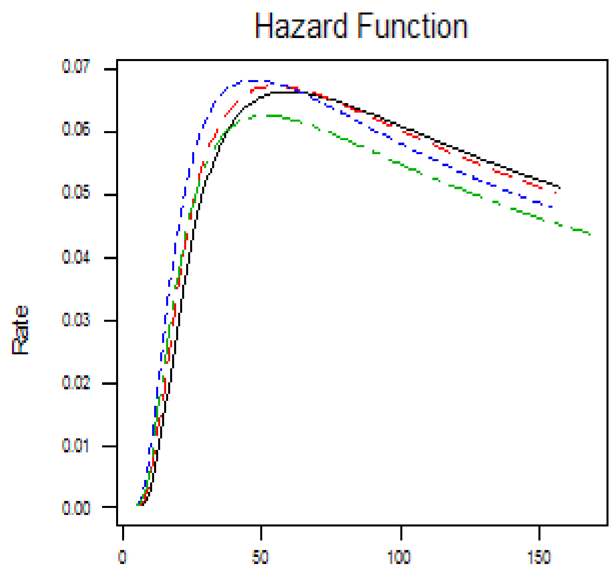
f. Sixth birth-interval



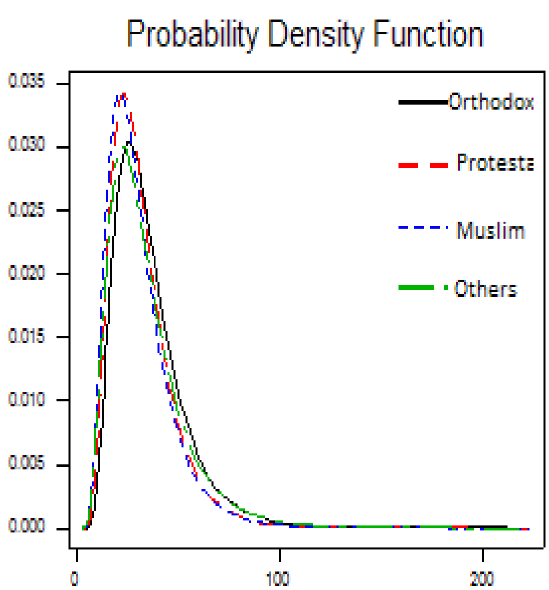
3) Probability density and hazard function by Religion groups

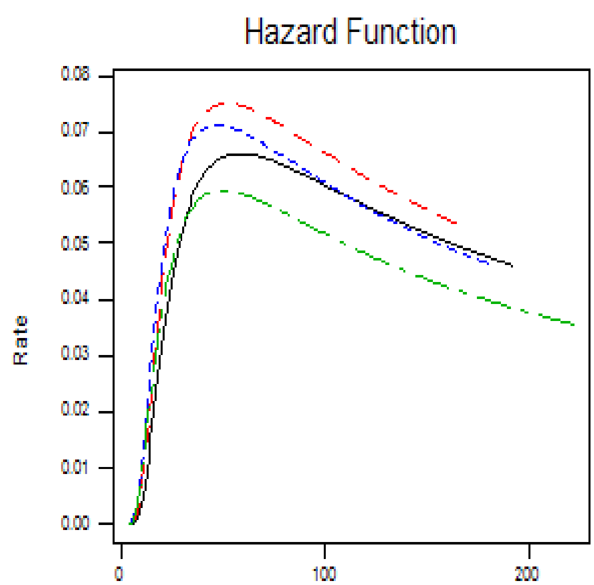
a. First birth-interval



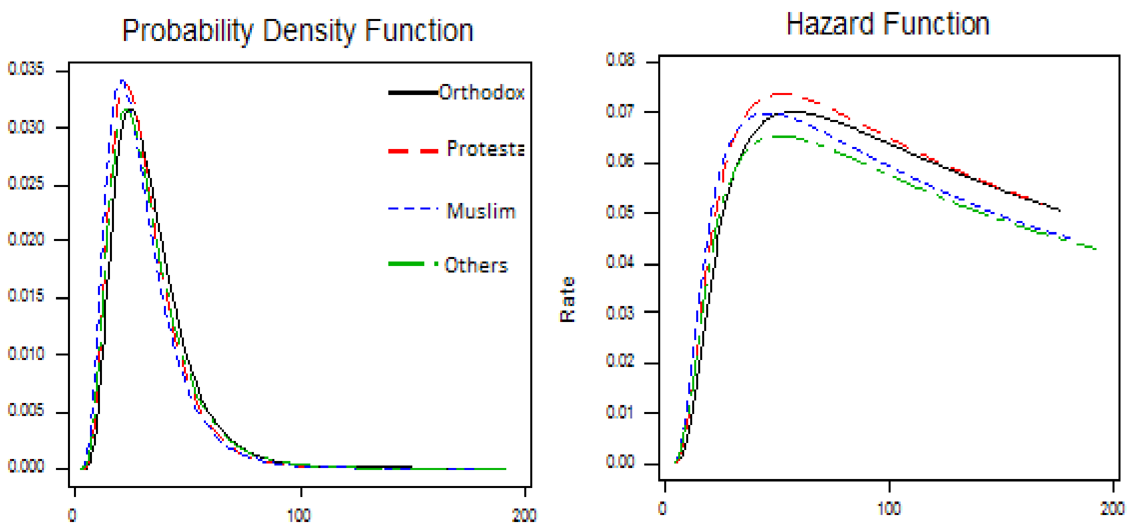


b. Second birth-interval

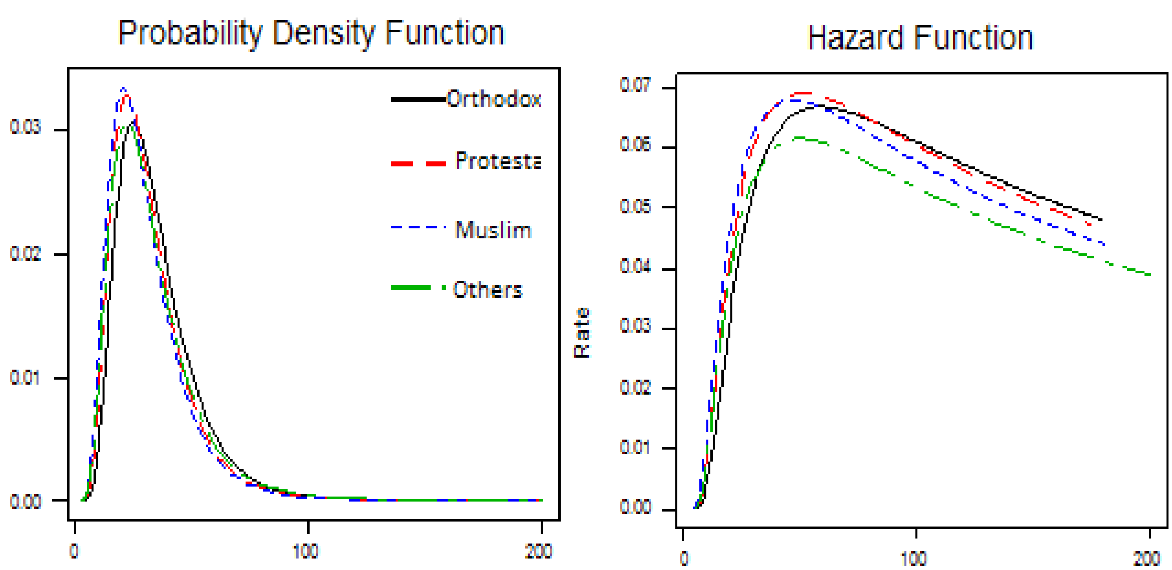




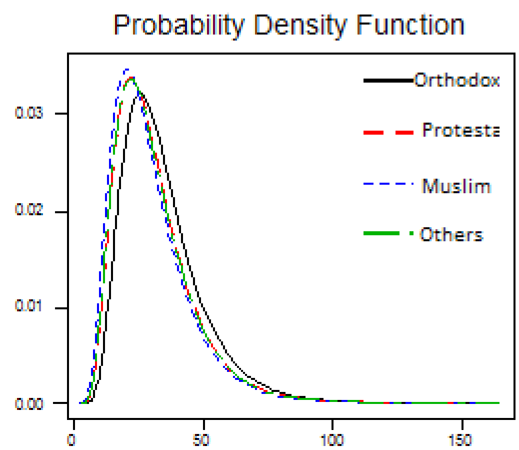
c. Third birth-interval

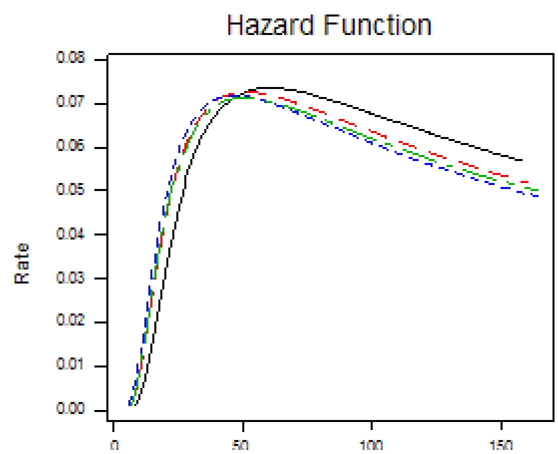


d. Fourth birth-interval

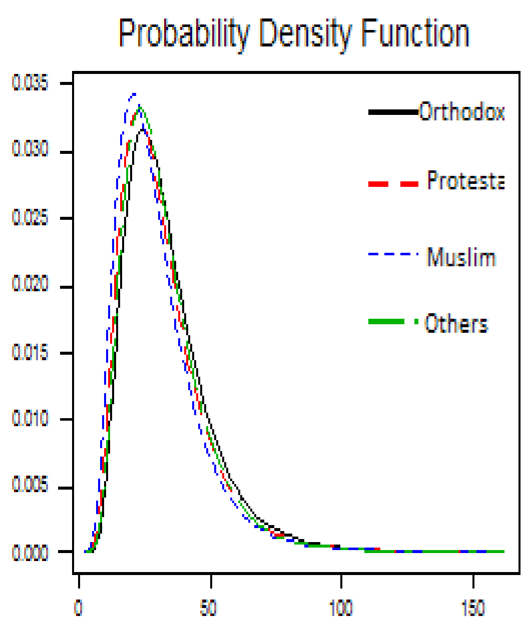


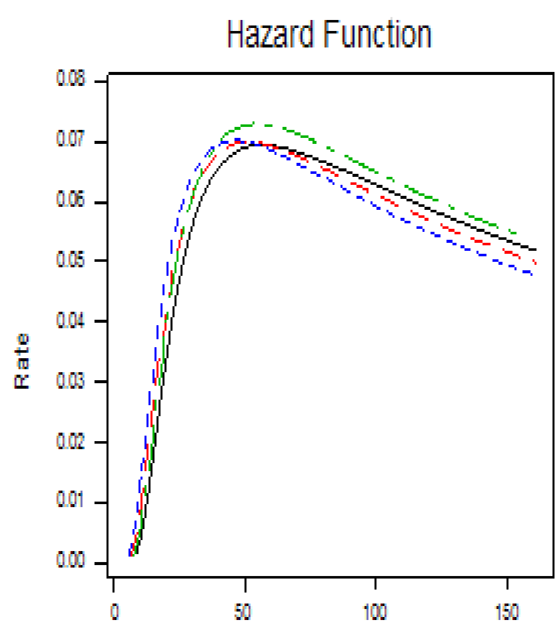
e. Fifth birth-interval





f. Sixth birth-interval

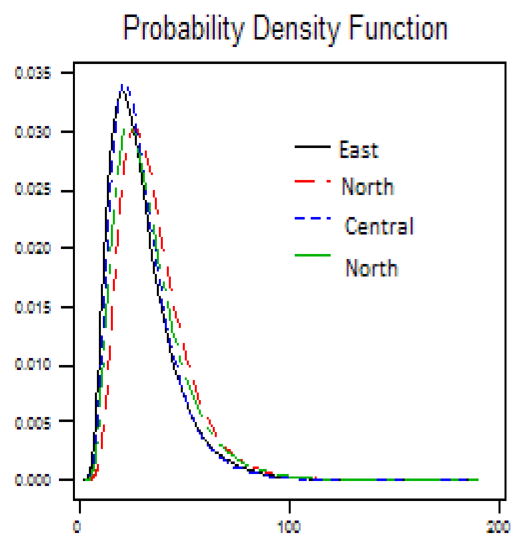


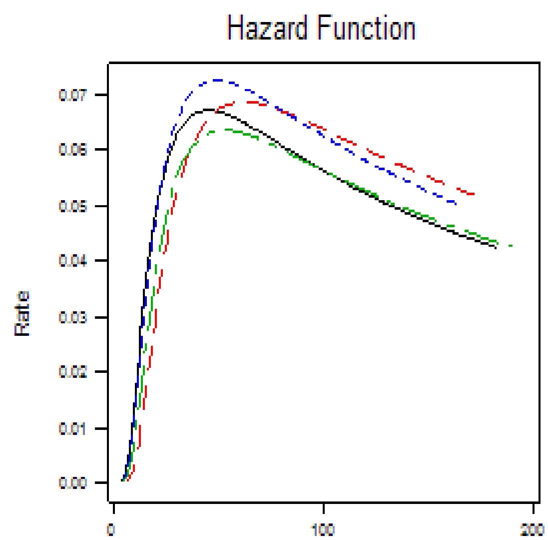


Appendix A

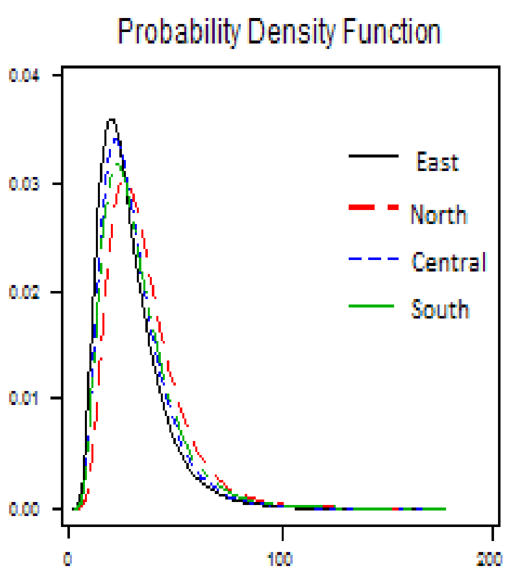
1) Probability density and hazard function by Region

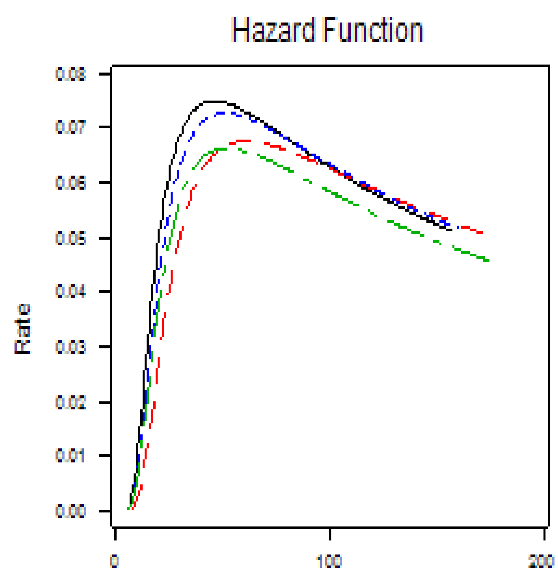
a. First birth-interval



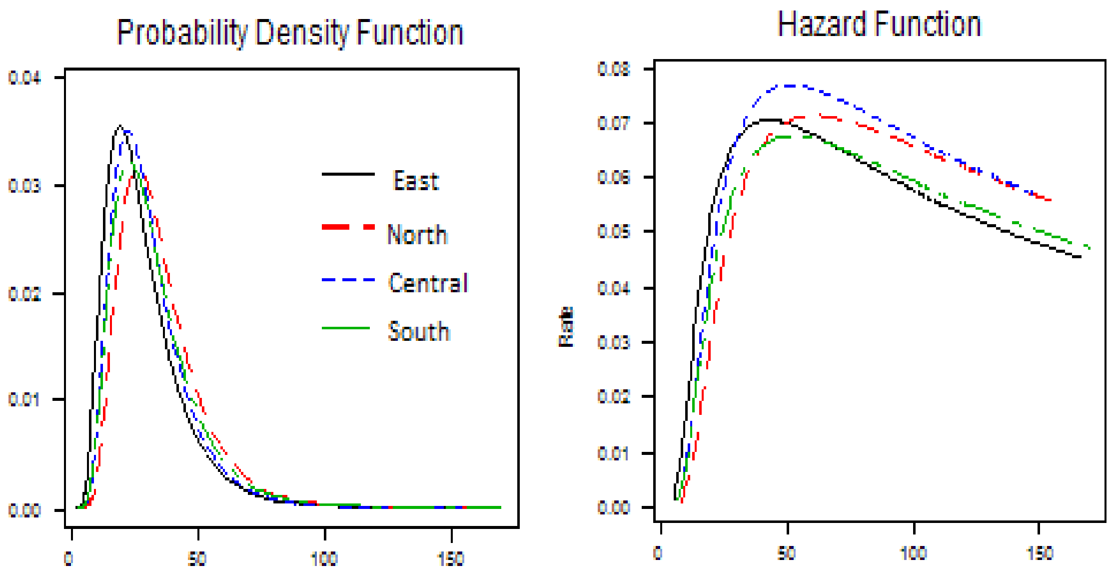


b. Second birth-interval

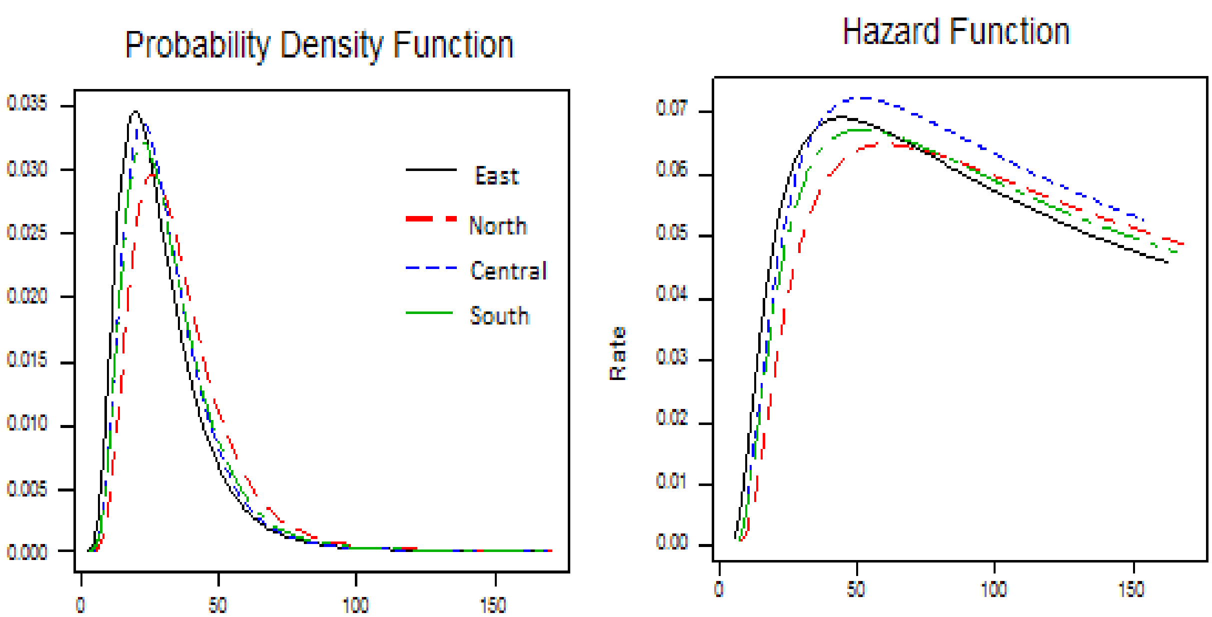




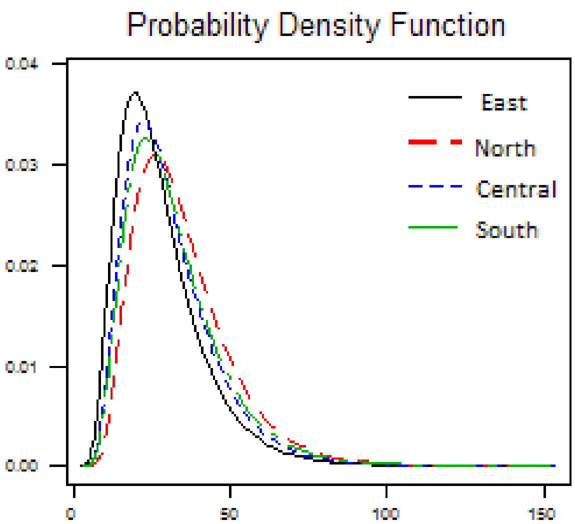
c. Third birth-interval

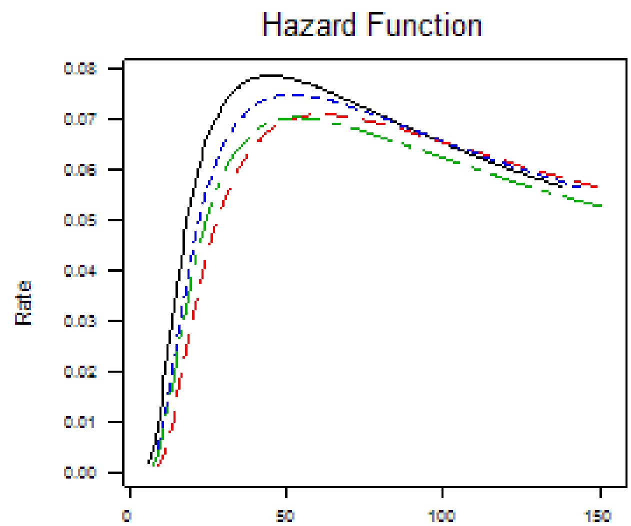


d. Fourth birth-interval

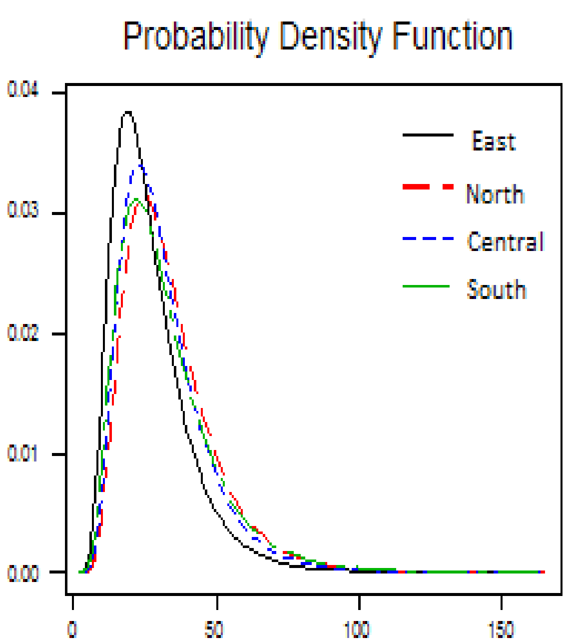


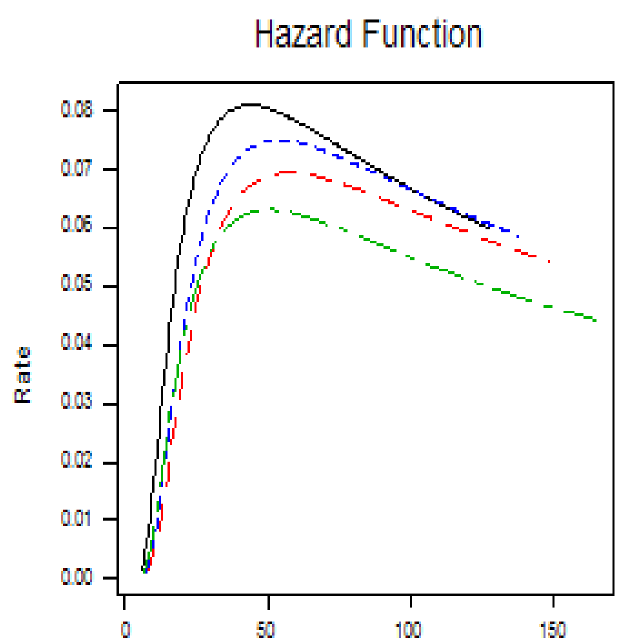
e. Fifth birth-interval





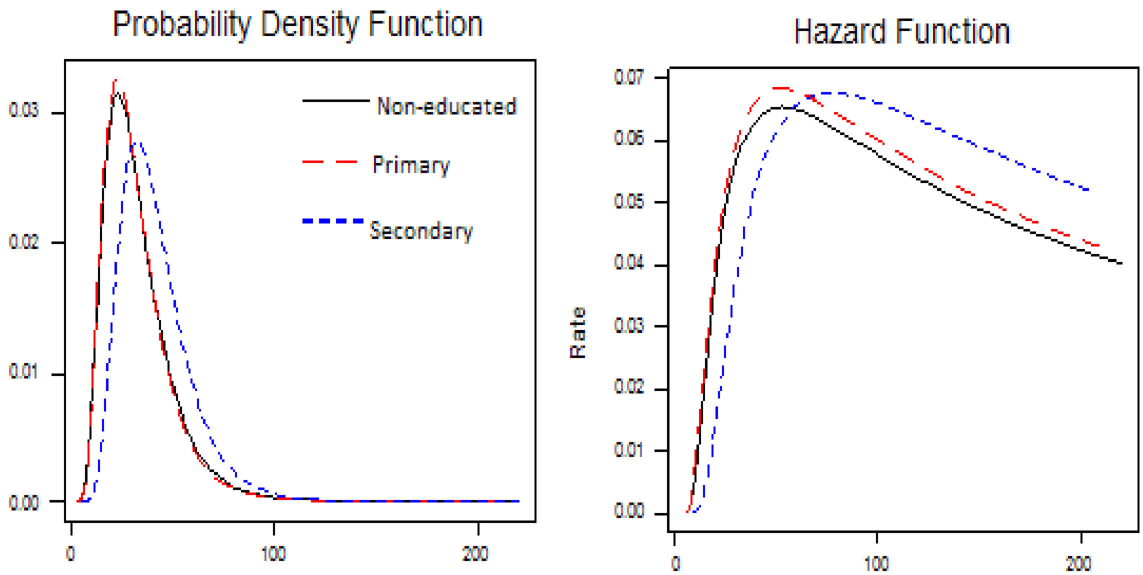
f. Sixth birth-interval



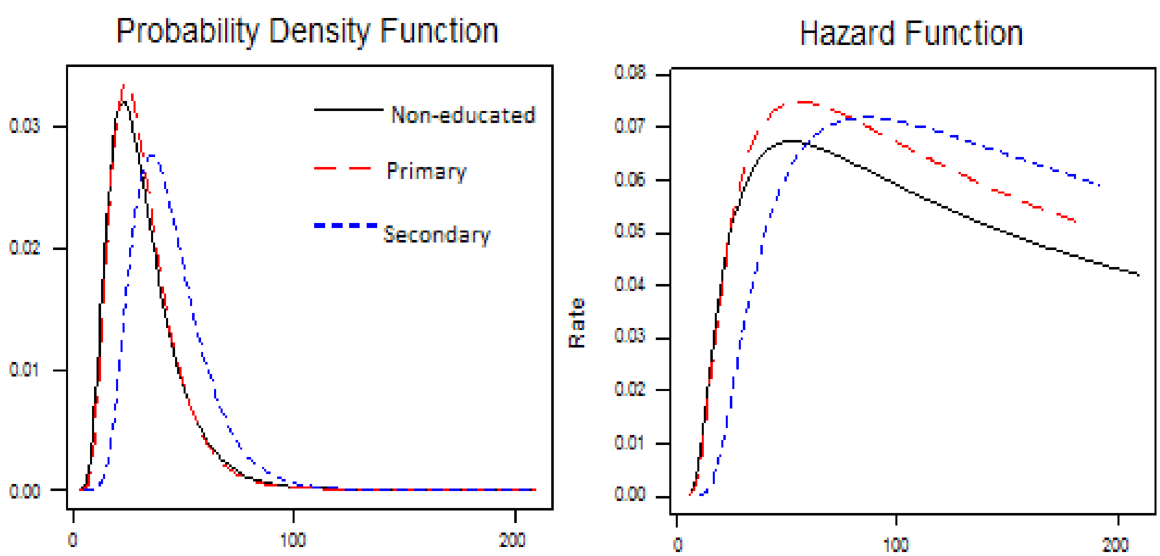


2) Probability density and hazard function by education level

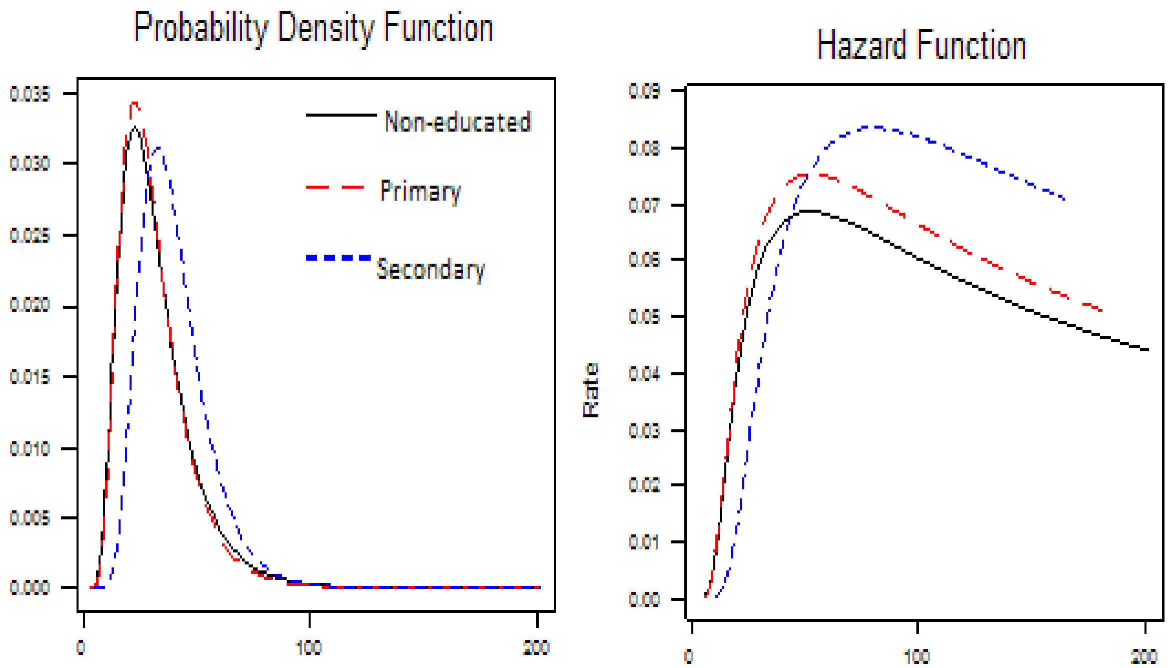
a. First birth-interval



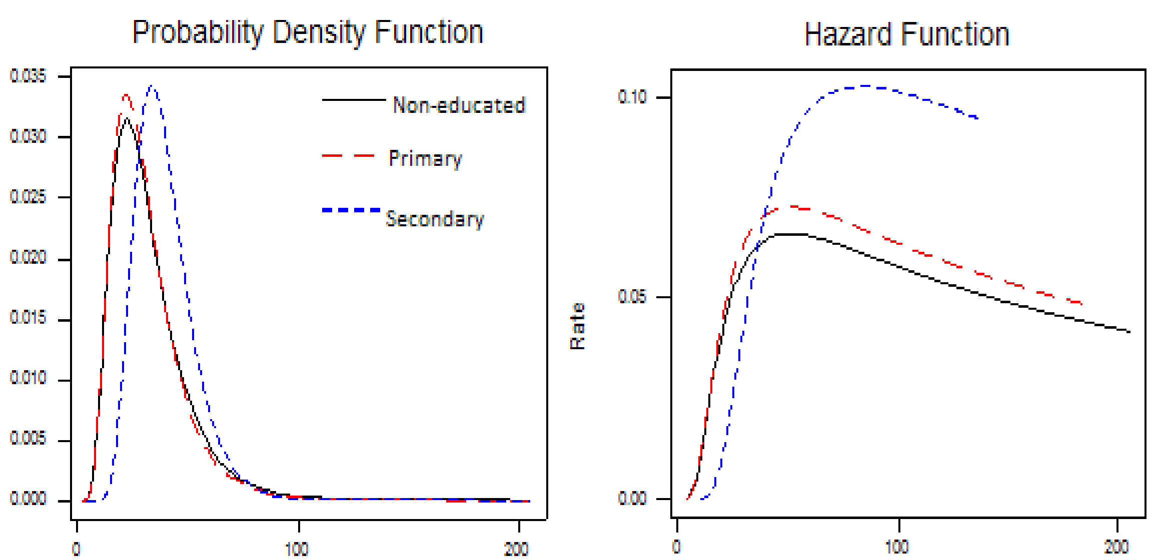
b. Second birth-interval



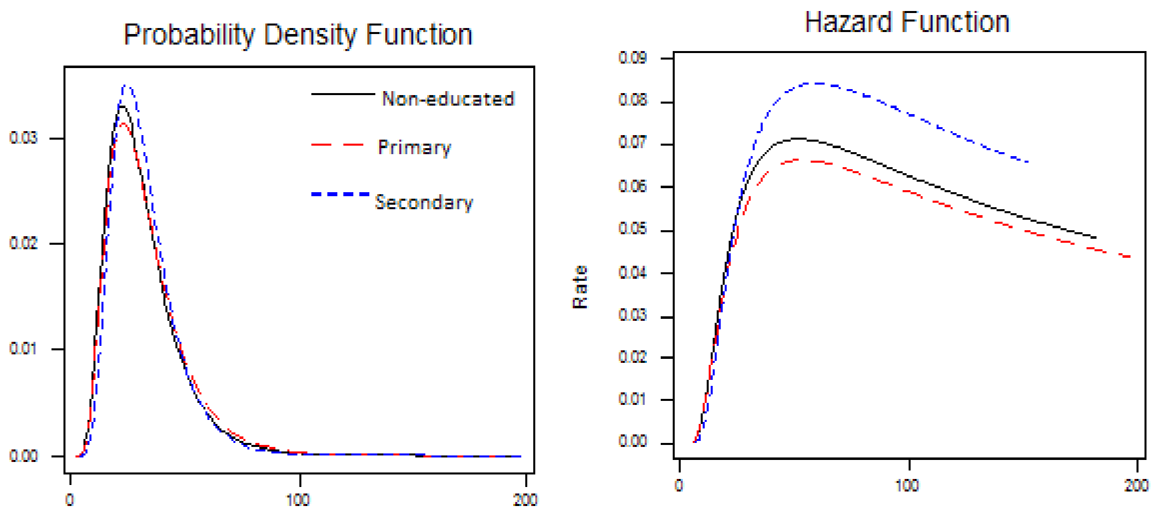
c. Third birth-interval



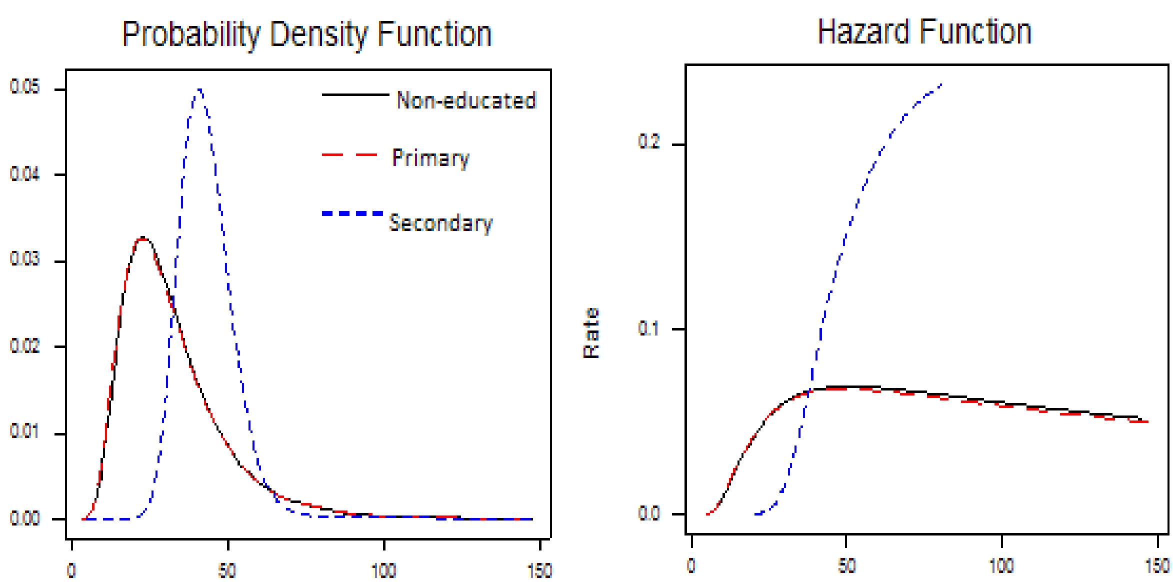
d. Fourth birth-interval



e. Fifth birth-interval

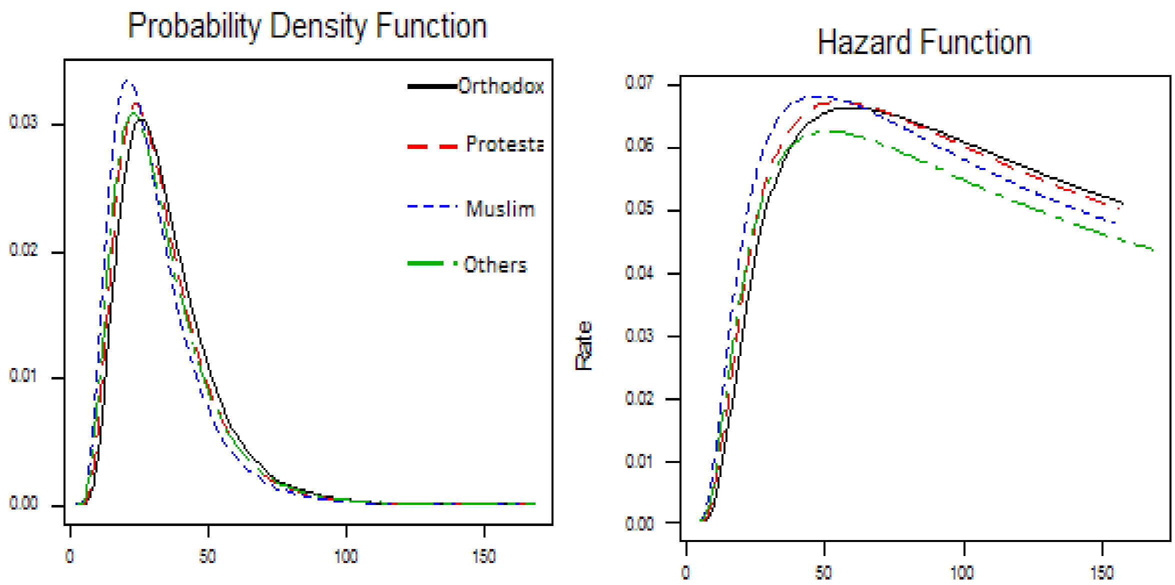


f. Sixth birth-interval

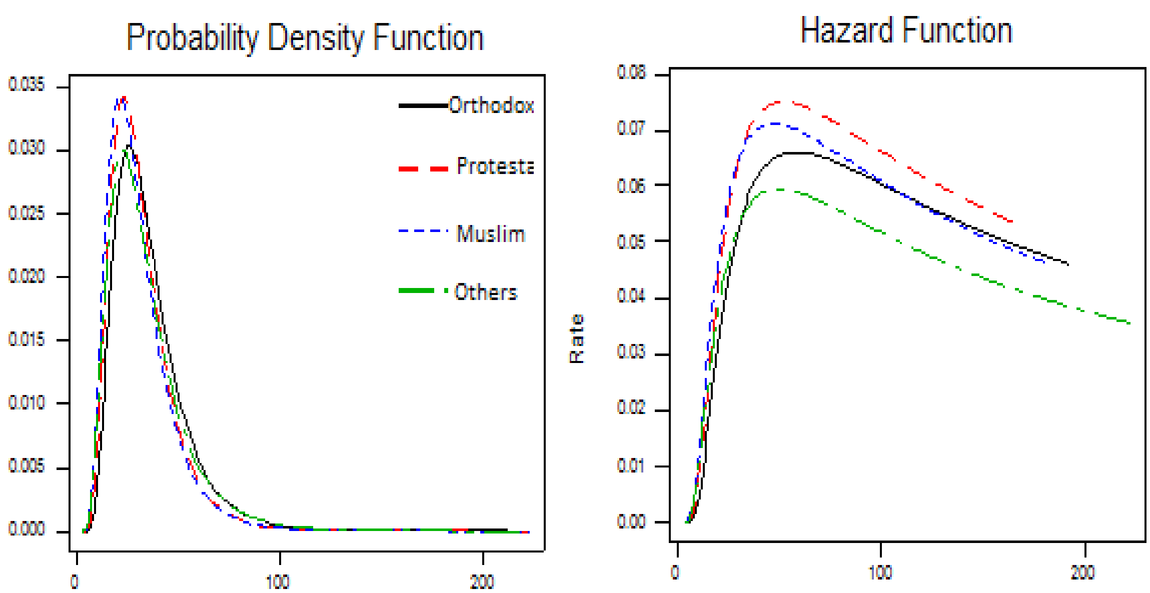


3) Probability density and hazard function by Religion groups

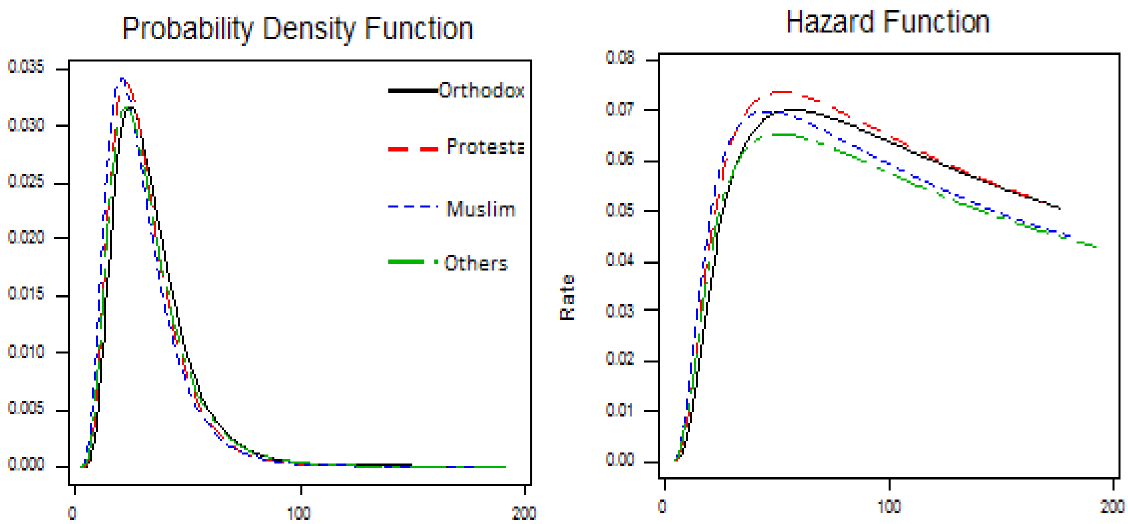
a. First birth-interval



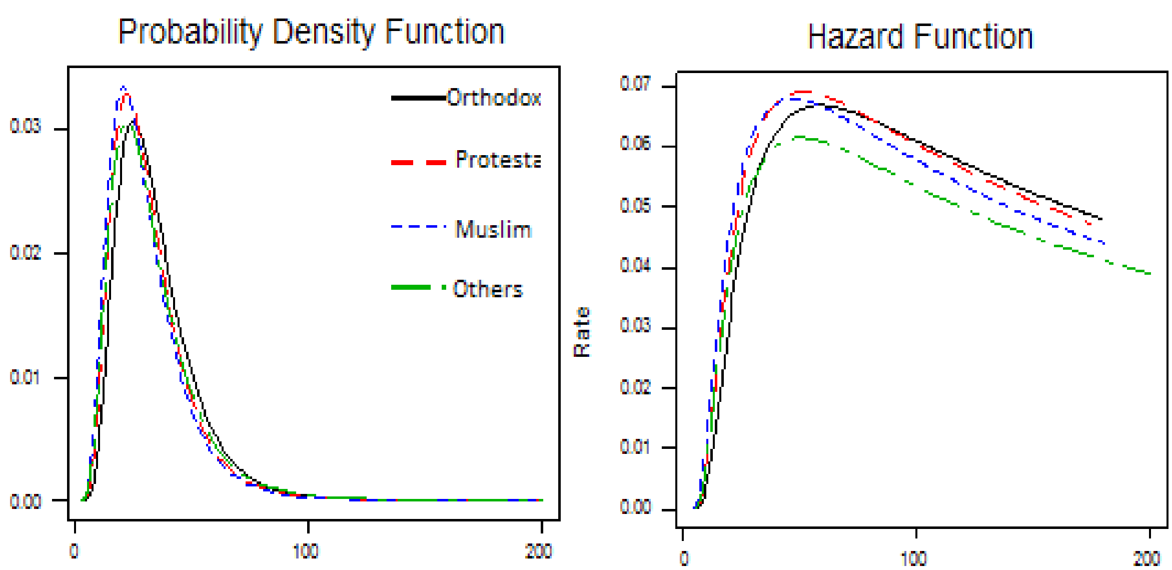
b. Second birth-interval



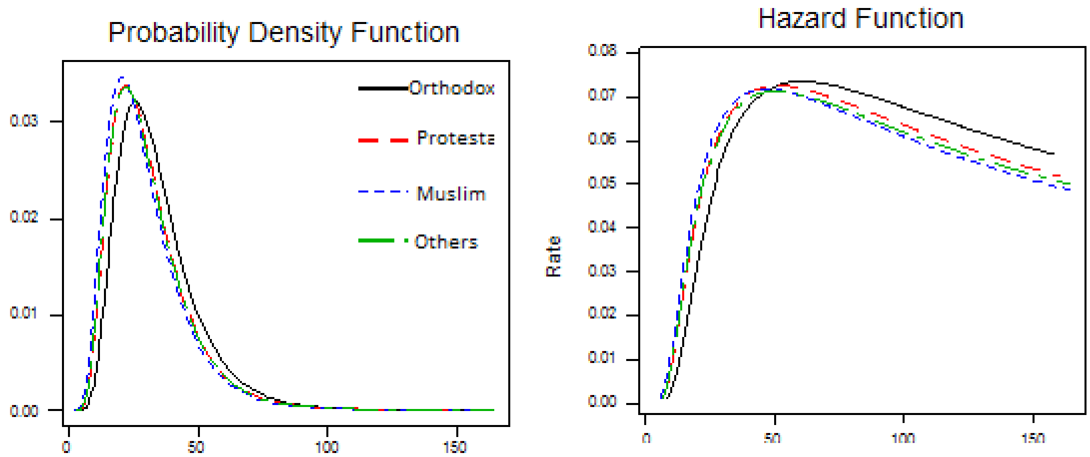
c. Third birth-interval



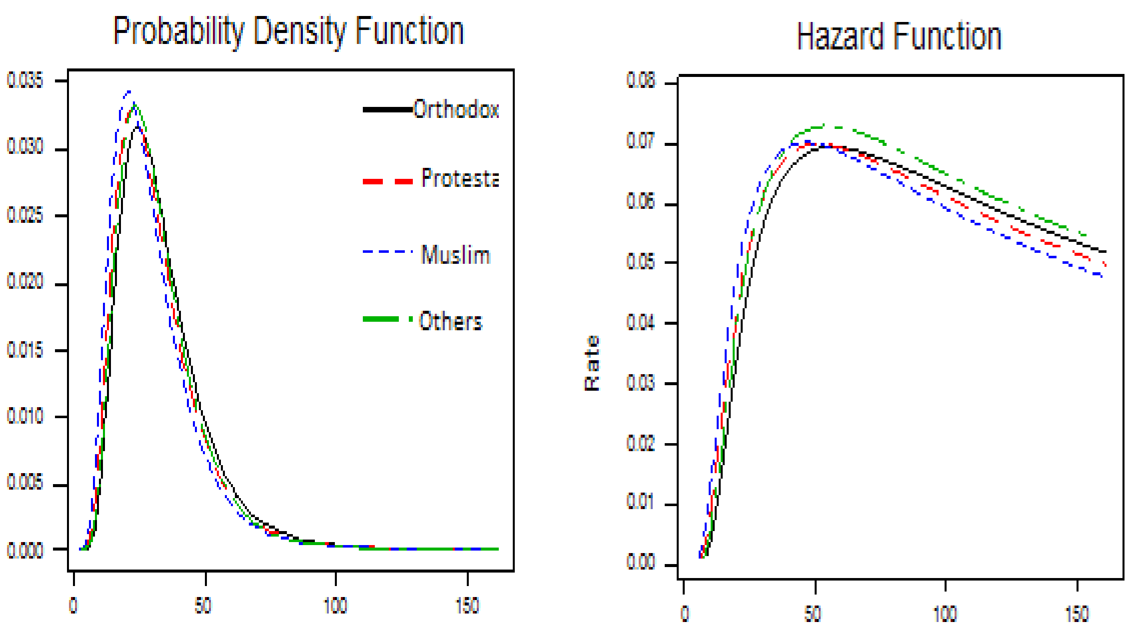
d. Fourth birth-interval



e. Fifth birth-interval



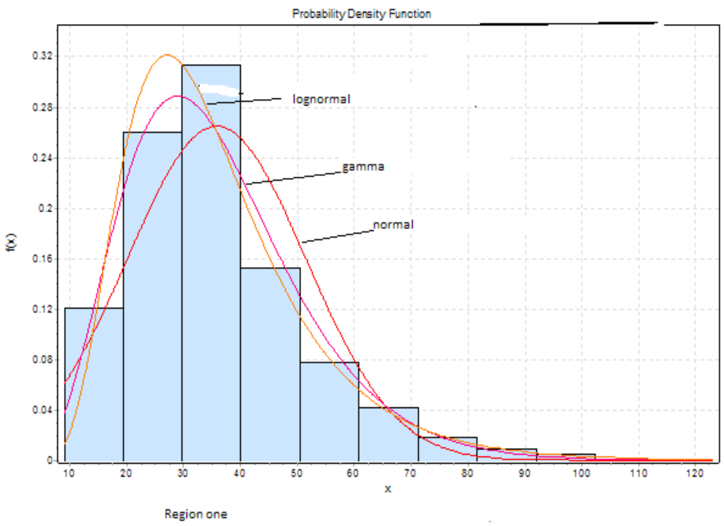
f. Sixth birth-interval



Appendix B Goodness of Fit Results

Table 1.a First birth interval by region

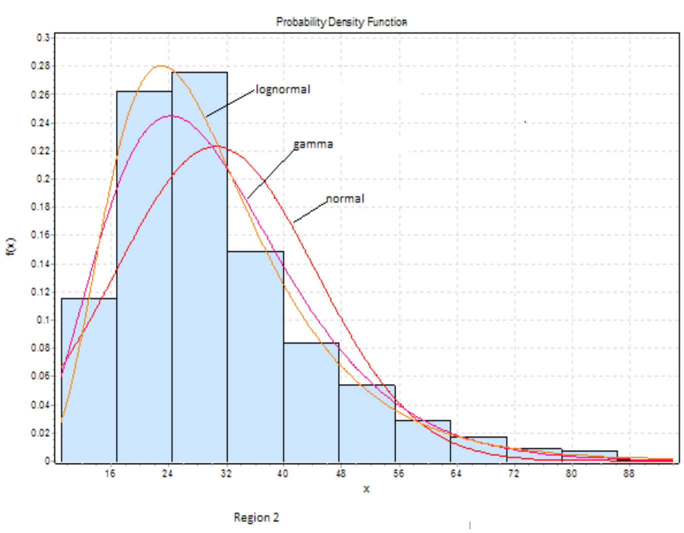
i. Northern



#	Distribution	Parameters
1	Gamma	α=4.8725 β=6.2646
2	Lognormal	σ=0.43538 μ=3.3238
3	Normal	σ=13.828 μ=30.524

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	3.1944	2	**Decision** **(Reject?)**	Yes	No	No
2. Lognormal	2.3335	1		No	No	No
3. Normal	126.693	3		Yes	Yes	Yes

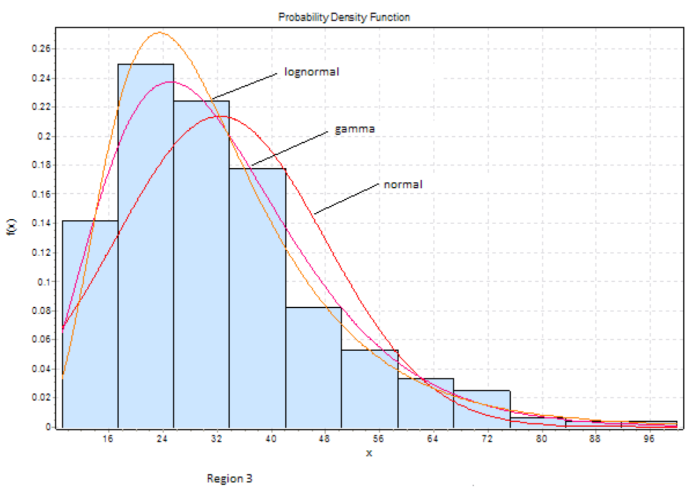
ii. East Region



#	Distribution	Parameters
1	Gamma	α=5.2458 β=6.8209
2	Lognormal	σ=0.43343 μ=3.4858
3	Normal	σ=15.623 μ=35.782

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	4.4124	2	**Decision**	Yes	Yes	Yes
2. Lognormal	1.1348	1	**(Reject?)**	No	No	No
3. Normal	22.745	3		Yes	Yes	Yes

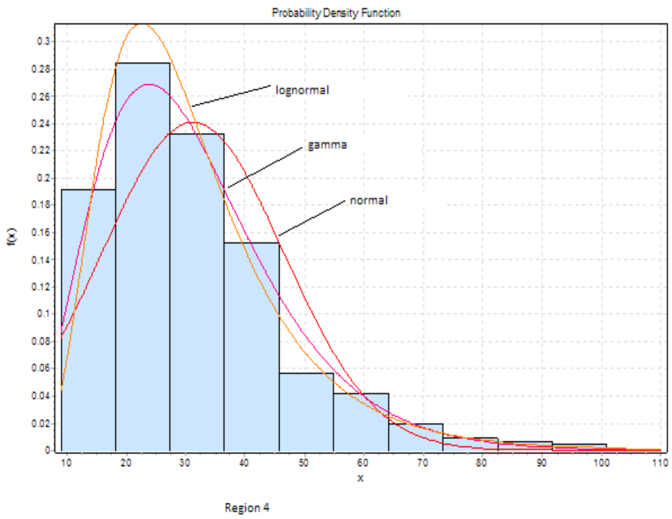
iii. Central



#	**Distribution**	**Parameters**
1	Gamma	α=4.3997 β=7.3653
2	Lognormal	σ=0.46677 μ=3.3709
3	Normal	σ=15.449 μ=32.405

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	3.2933	2		Yes	Yes	No
2. Lognormal	0.96795	1	**Reject?**	No	No	No
3. Normal	26.975	3		Yes	Yes	Yes

iv. South Region

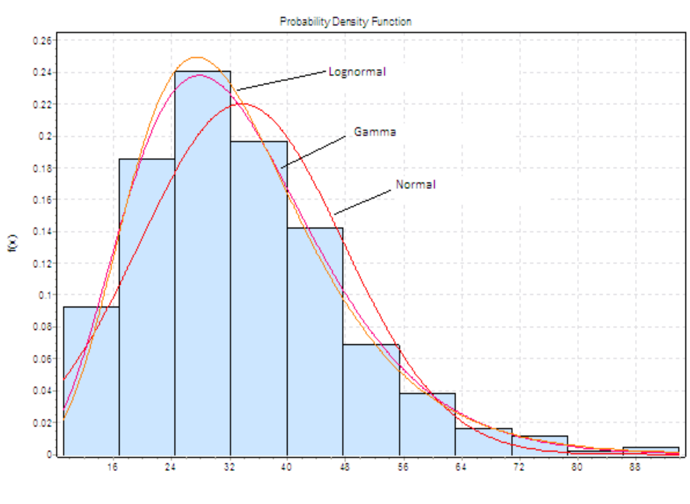


#	Distribution	Parameters
1	Gamma	α=4.2132 β=7.4091
2	Lognormal	σ=0.46383 μ=3.3331
3	Normal	σ=15.208 μ=31.216

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	1.768	2	**Decision (Reject?)**	No	No	No
2. Lognormal	1.2062	1		No	No	No
3. Normal	43.771	3		Yes	Yes	Yes

Table 1.b Second birth by Region

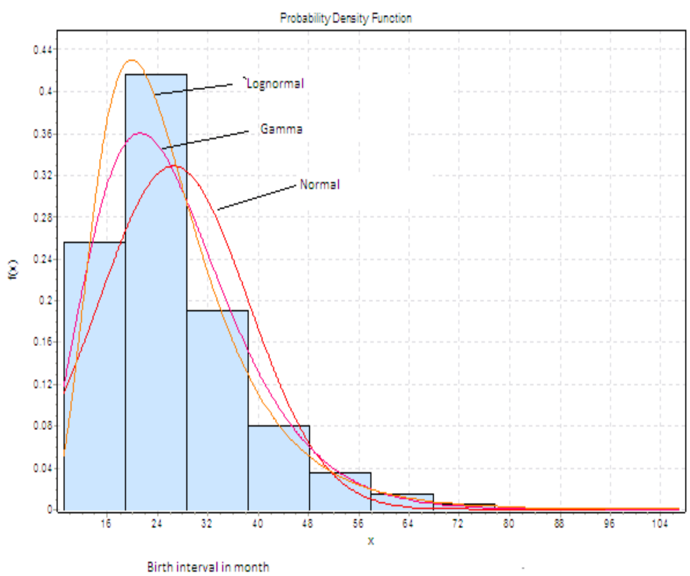
i. North Region



#	Distribution	Parameters
1	Gamma	α=5.7678 β=5.8331
2	Lognormal	σ=0.42522 μ=3.4291
3	Normal	σ=14.009 μ=33.644

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	1.1288	1	**Decision**	No	No	No
2. Lognormal	1.1943	2	**(Reject?)**	No	No	No
3. Normal	13.79	3		Yes	Yes	Yes

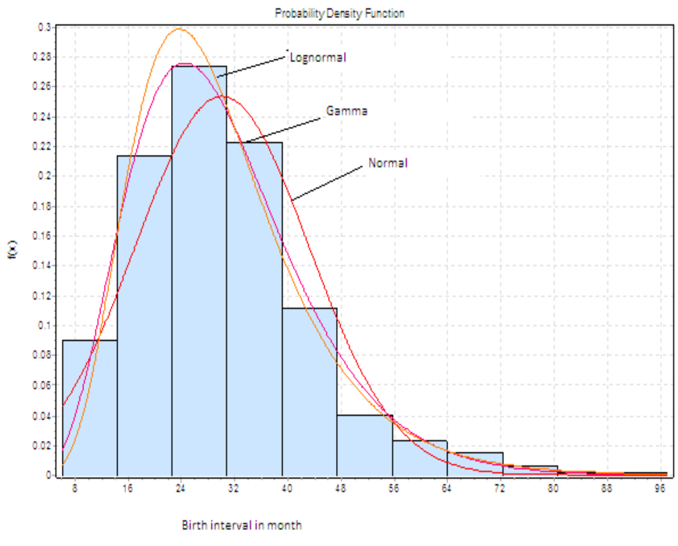
ii. East Region



#	**Distribution**	**Parameters**
1	Gamma	α=4.982 β=5.3266
2	Lognormal	σ=0.41726 μ=3.19
3	Normal	σ=11.889 μ=26.537

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	1.625	2	**Decision**	No	No	No
2. Lognormal	0.94196	1	**(Reject?)**	No	No	No
3. Normal	20.457	3		Yes	Yes	Yes

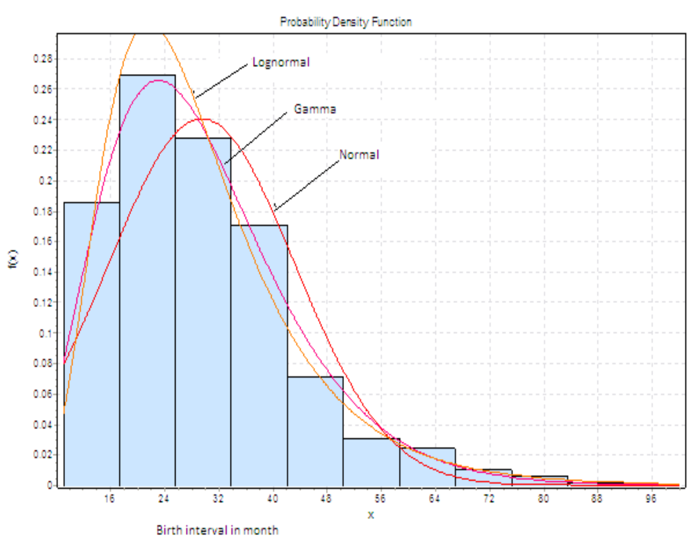
iii. Central Region



#	Distribution	Parameters
1	Gamma	α=5.328 β=5.6405
2	Lognormal	σ=0.43322 μ=3.3117
3	Normal	σ=13.02 μ=30.052

Distribution	Statistic	Rank	**α-value and critical values**			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	1.2909	2	**Decision (Reject?)**	No	No	No
2. Lognormal	1.2707	1		No	No	No
3. Normal	13.421	3		Yes	Yes	Yes

iv. South Region

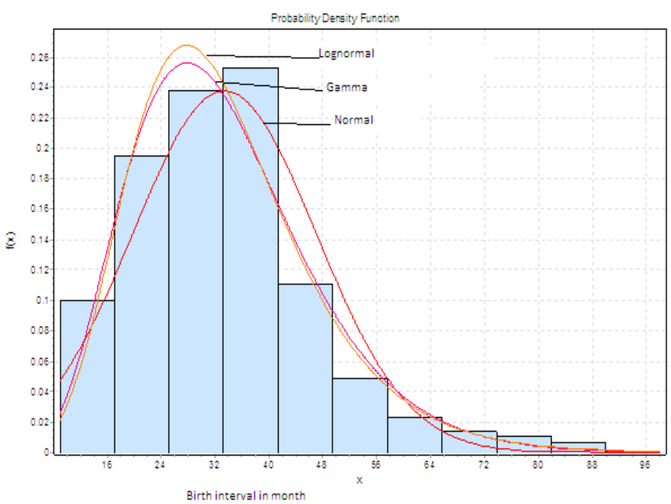


#	Distribution	Parameters
1	Gamma	α=4.6027 β=6.394
2	Lognormal	σ=0.45557 μ=3.2796
3	Normal	σ=13.718 μ=29.429

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	3.0801	2	Decision (Reject?)	Yes	No	No
2. Lognormal	1.7674	1		No	No	No
3. Normal	26.815	3		Yes	Yes	Yes

Table 1.c Third Birth interval in Region

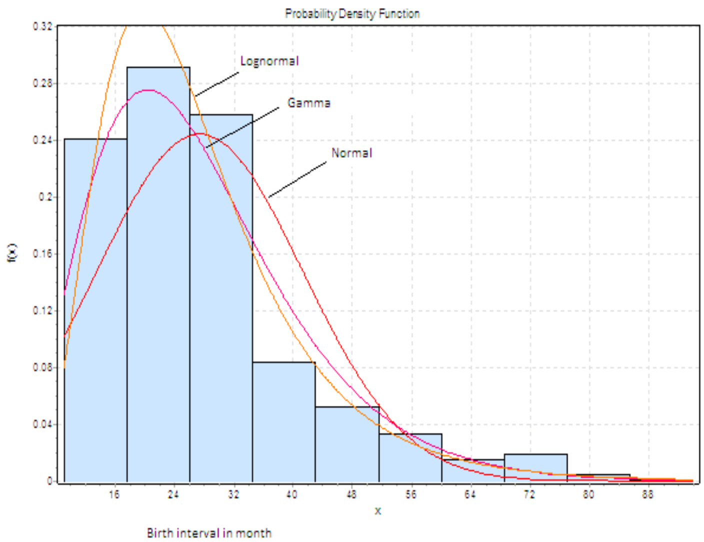
i. **North Region**



#	Distribution	Parameters
1	Gamma	α=6.0504 β=5.5203
2	Lognormal	σ=0.41399 μ=3.4264
3	Normal	σ=13.579 μ=33.4

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	3.2144	2	Decision (Reject?)	Yes	No	No
2. Lognormal	2.7593	1		Yes	No	No
3. Normal	11.756	3		Yes	Yes	Yes

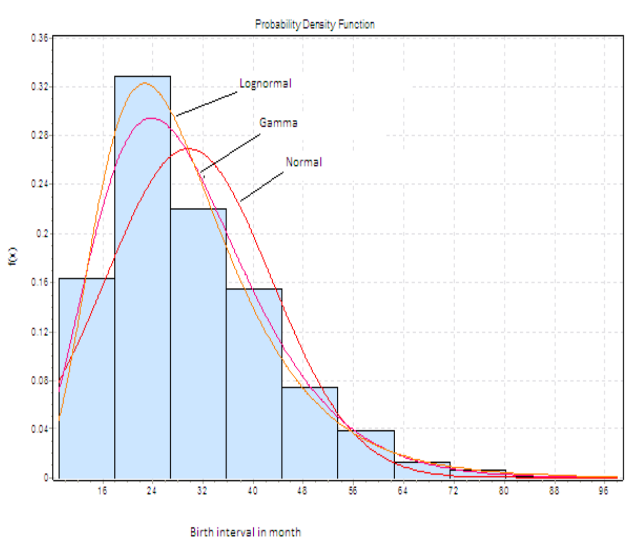
ii. East Region



#	Distribution	Parameters
1	Gamma	α=3.8904 β=7.0367
2	Lognormal	σ=0.4647 μ=3.1985
3	Normal	σ=13.879 μ=27.376

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	2.4581	2	**Decision (Reject?)**	No	No	No
2. Lognormal	2.1399	1		No	No	No
3. Normal	23.764	3		Yes	Yes	Yes

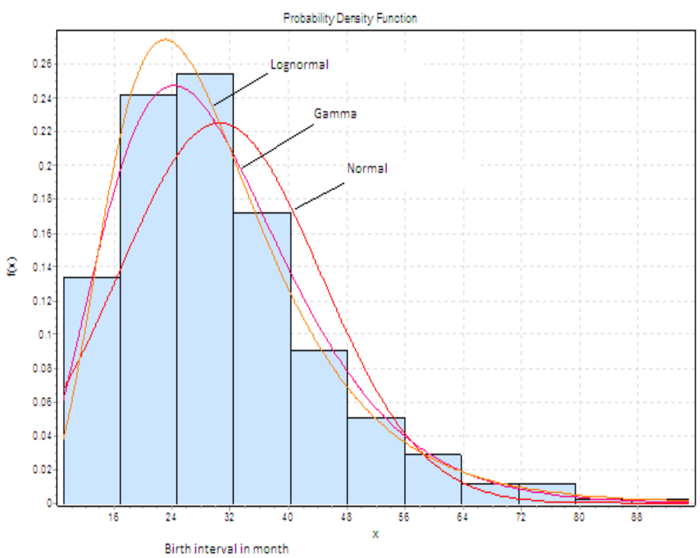
iii. Central Region



#	Distribution	Parameters
1	Gamma	α=5.078 β=5.8473
2	Lognormal	σ=0.44343 μ=3.295
3	Normal	σ=13.177 μ=29.693

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	1.2704	2	Decision (Reject?)	No	No	No
2. Lognormal	1.2421	1		No	No	No
3. Normal	14.024	3		Yes	Yes	Yes

iv. South Region

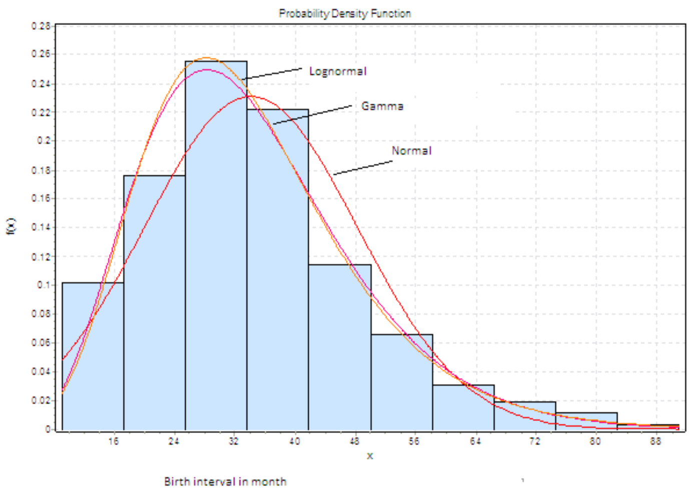


#	Distribution	Parameters
1	Gamma	α=4.854 β=6.2833
2	Lognormal	σ=0.4498 μ=3.3187
3	Normal	σ=13.843 μ=30.499

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	1.3487	1	Decision (Reject?)	No	No	No
2. Lognormal	1.5082	2		No	No	No
3. Normal	20.216	3		Yes	Yes	Yes

Table 1.d Fourth Birth interval in Region

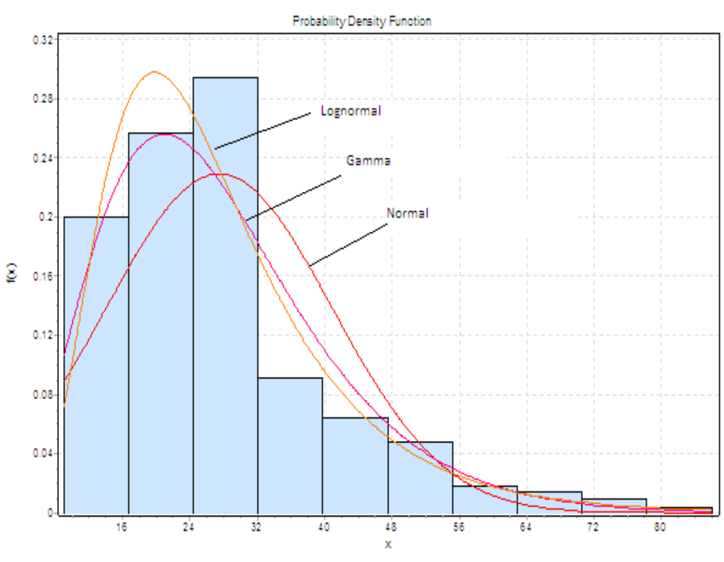
i. North Region



#	Distribution	Parameters
1	Gamma	α=5.8093 β=5.8744
2	Lognormal	σ=0.42961 μ=3.4425
3	Normal	σ=14.159 μ=34.126

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	1.0892	2	**Decision (Reject?)**	No	No	No
2. Lognormal	0.89961	1		No	No	No
3. Normal	9.5143	3		Yes	Yes	Yes

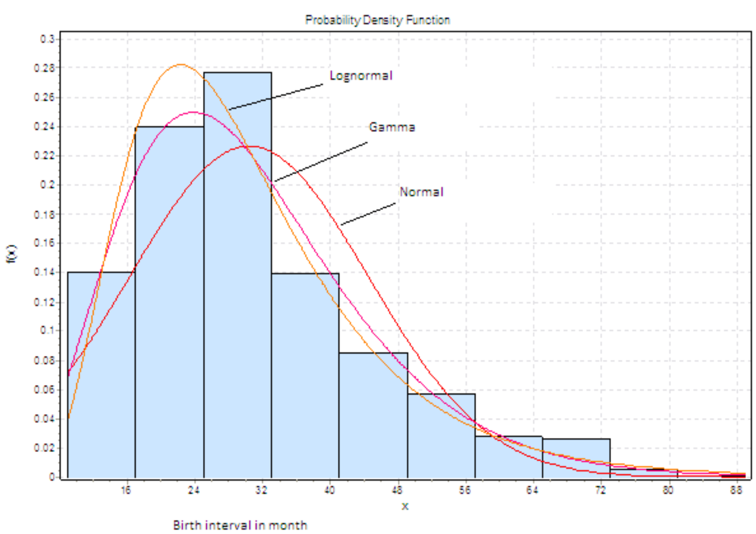
ii. East Region



#	Distribution	Parameters
1	Gamma	α=4.2039 β=6.5309
2	Lognormal	σ=0.46744 μ=3.2035
3	Normal	σ=13.391 μ=27.455

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	2.356	2	**Decision** **(Reject?)**	No	No	No
2. Lognormal	1.6694	1		No	No	No
3. Normal	13.678	3		Yes	Yes	Yes

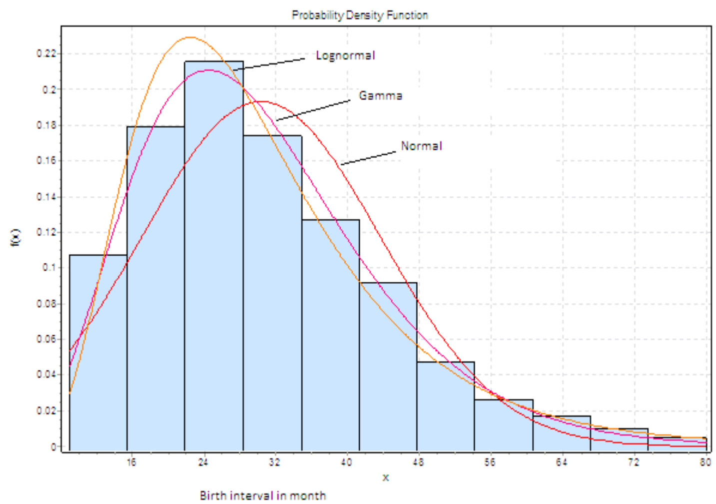
iii. Central Region



#	Distribution	Parameters
1	Gamma	α=4.653 β=6.5309
2	Lognormal	σ=0.45718 μ=3.3113
3	Normal	σ=14.088 μ=30.389

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	2.4252	2	**Decision** **(Reject?)**	No	No	No
2. Lognormal	1.5595	1		No	No	No
3. Normal	14.756	3		Yes	Yes	Yes

iv. South Region



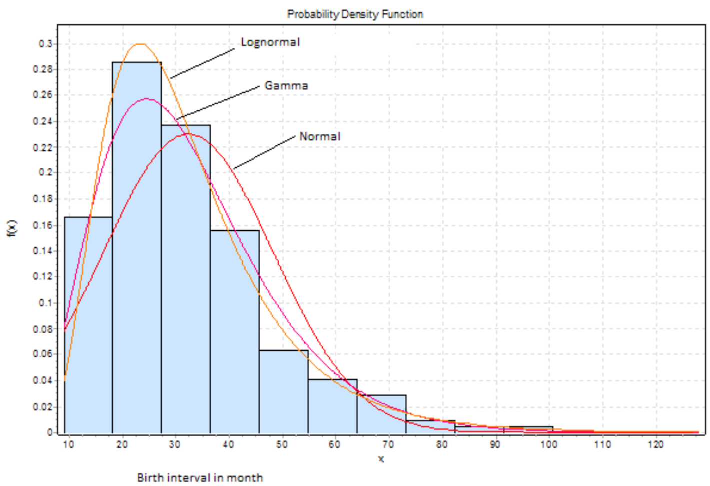
#	**Distribution**	**Parameters**
1	Gamma	α=5.1914 β=5.8477
2	Lognormal	σ=0.45155 μ=3.3154
3	Normal	σ=13.324 μ=30.358

Distribution	Statistic	Rank	**α-value and critical values**			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	0.81708	1	**Decision**	No	No	No
2. Lognormal	2.3082	2	**(Reject?)**	No	No	No
3. Normal	10.623	3		Yes	Yes	Yes

Goodness of fit test result of birth Interval data by Education Level

Table 2.a First birth Interval

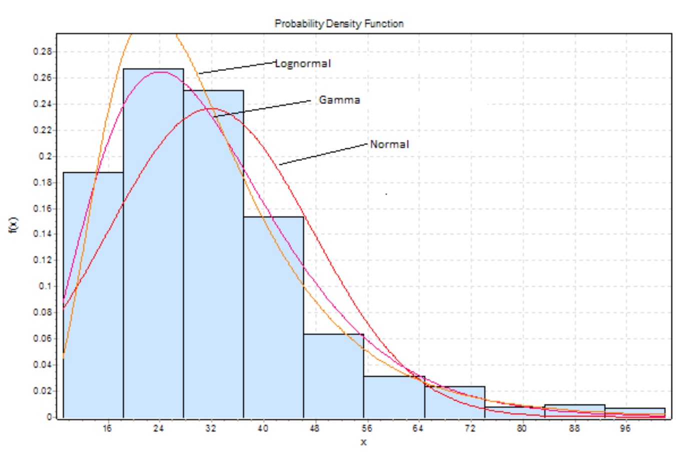
i. **Non-educated**



#	Distribution	Parameters
1	Gamma	α=4.1417 β=7.7971
2	Lognormal	σ=0.47037 μ=3.3649
3	Normal	σ=15.868 μ=32.293

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	8.2354	1	**Decision (Reject?)**	Yes	Yes	Yes
2. Lognormal	5.2591	2		Yes	Yes	Yes
3. Normal	14.756	3		Yes	Yes	Yes

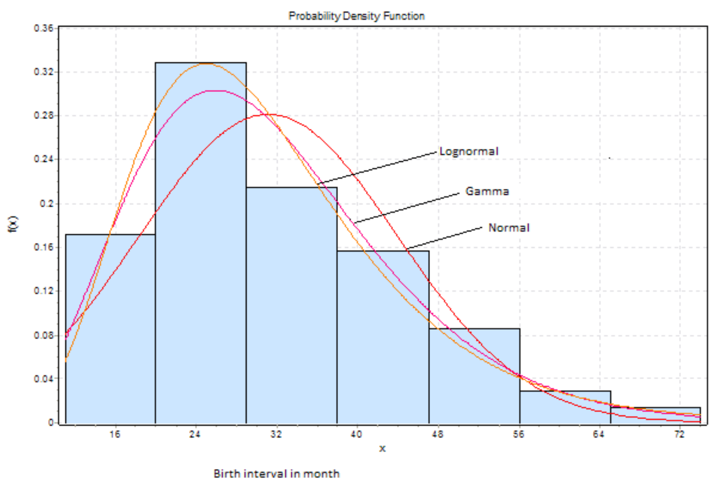
ii. **Primary**



#	Distribution	Parameters
1	Gamma	α=4.0972 β=7.7615
2	Lognormal	σ=0.47318 μ=3.3482
3	Normal	σ=15.71 μ=31.801

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	1.0951	3	**Decision** **(Reject?)**	No	No	No
2. Lognormal	0.74808	1		No	No	No
3. Normal	0.91869	2		No	No	No

iii. Secondary

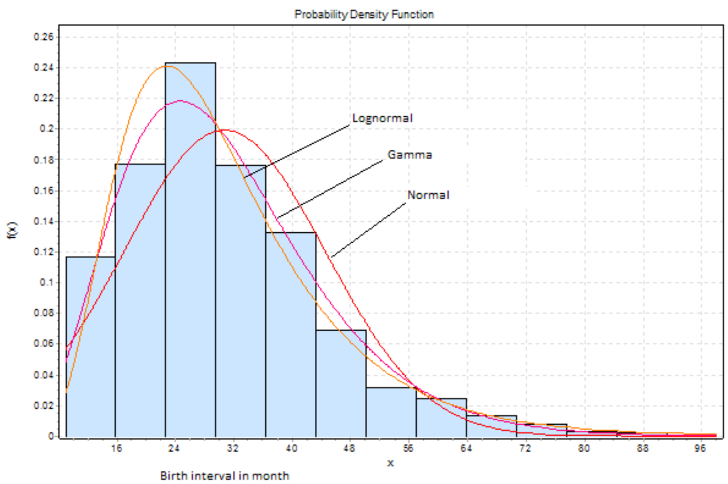


#	Distribution	Parameters
1	Gamma	α=5.9578 β=5.2272
2	Lognormal	σ=0.4088 μ=3.3571
3	Normal	σ=12.759 μ=31.143

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	2.19695	1	**Decision** **(Reject?)**	No	No	No
2. Lognormal	2.3299	2		No	No	No
3. Normal	16.011	3		Yes	Yes	Yes

Table 2.b Second Birth Interval in Month

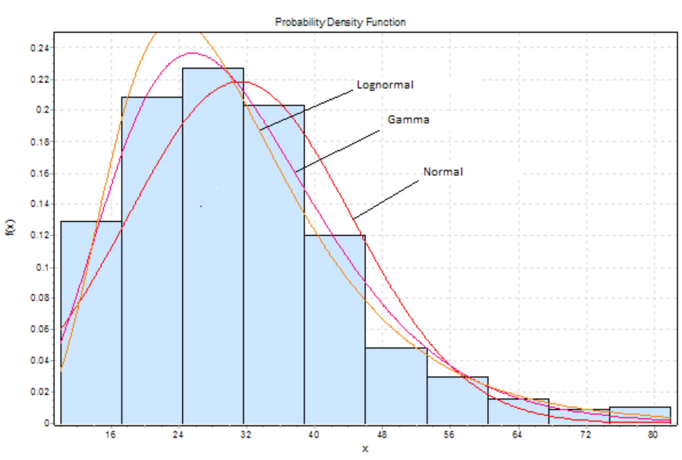
i. Non-educated



#	Distribution	Parameters
1	Gamma	α=5.0336 β=6.1072
2	Lognormal	σ=0.44907 μ=3.3281
3	Normal	σ=13.702 μ=30.741

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	4.9186	1	**Decision (Reject?)**	Yes	Yes	Yes
2. Lognormal	5.1892	2		Yes	Yes	Yes
3. Normal	55.081	3		Yes	Yes	Yes

ii. Primary

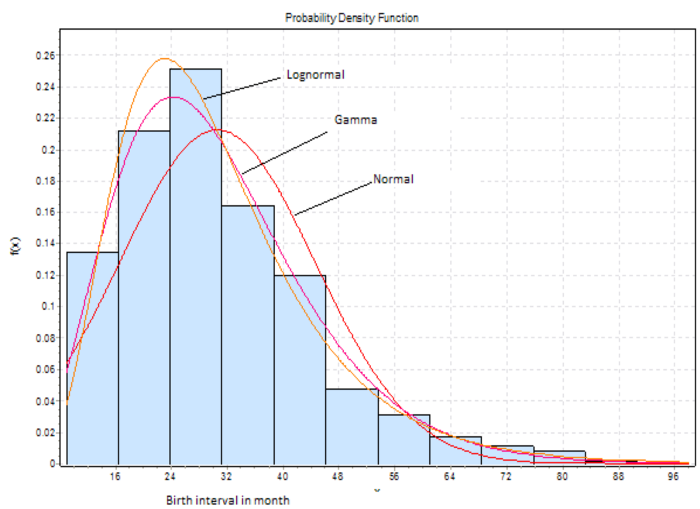


#	Distribution	Parameters
1	Gamma	α=5.6168 β=5.5495
2	Lognormal	σ=0.42772 μ=3.3512
3	Normal	σ=13.152 μ=31.17

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	0.65462	1	**Decision (Reject?)**	No	No	No
2. Lognormal	0.75665	2		No	No	No
3. Normal	5.8602	3		Yes	Yes	Yes

Table 2.c Third birth interval by education Level

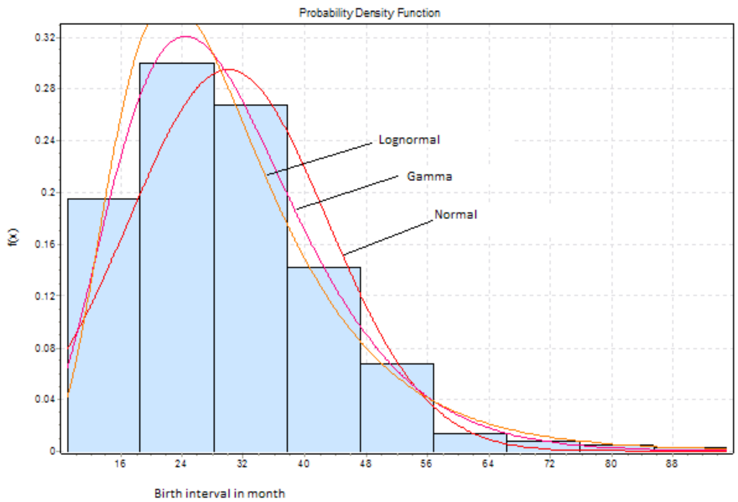
i. Non-educated



#	**Distribution**	**Parameters**
1	Gamma	α=4.825 β=6.3384
2	Lognormal	σ=0.45266 μ=3.3204
3	Normal	σ=13.923 μ=30.583

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	3.9701	2	**Decision (Reject?)**	Yes	Yes	Yes
2. Lognormal	3.4547	1		Yes	Yes	No
3. Normal	54.319	3		Yes	Yes	Yes

ii. Primary

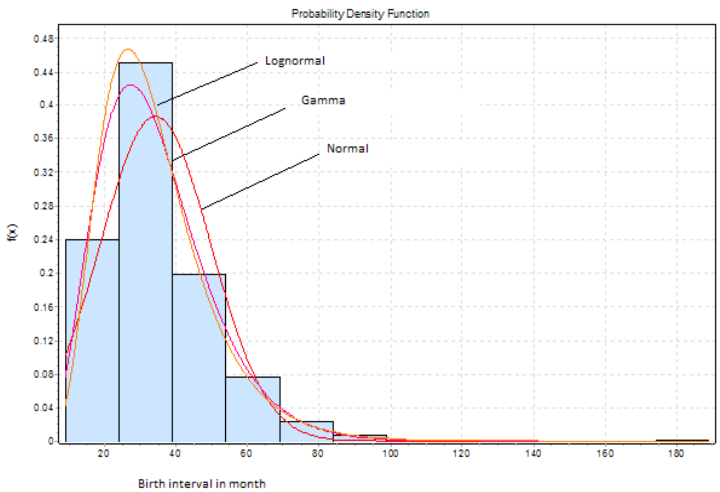


#	**Distribution**	**Parameters**
1	Gamma	α=5.3812 β=5.5737
2	Lognormal	σ=0.44259 μ=3.3076
3	Normal	σ=12.93 μ=29.993

Distribution	Statistic	Rank	**α-value and critical values**			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	1.3857	2	**Decision (Reject?)**	No	No	No
2. Lognormal	1.6001	1		No	No	No
3. Normal	3.7367	3		Yes	Yes	No

Table 3.a First Birth Interval in Religion

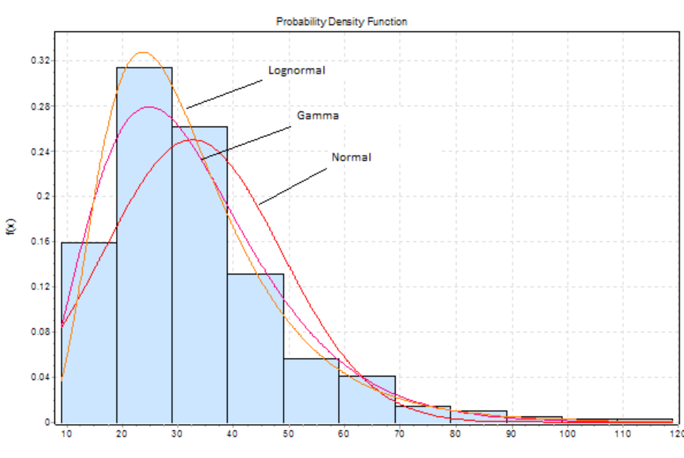
i. Orthodox



#	Distribution	Parameters
1	Gamma	α=4.9096 β=6.9777
2	Lognormal	σ=0.44453 μ=3.4376
3	Normal	σ=15.461 μ=34.258

Distribution	Statistic	Rank	**α-value and critical values**			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	3.4905	2	**Decision (Reject?)**	Yes	Yes	No
2. Lognormal	2.5842	1		Yes	No	No
3. Normal	37.869	3		Yes	Yes	Yes

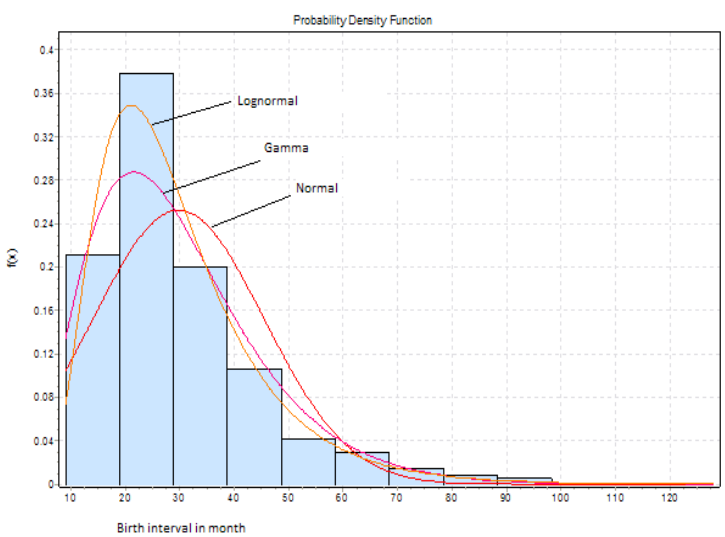
ii. Protestant



#	Distribution	Parameters
1	Gamma	α=4.1878 β=7.7923
2	Lognormal	σ=0.46114 μ=3.3786
3	Normal	σ=15.946 μ=32.632

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	1.9388	2	**Decision**	No	No	No
2. Lognormal	1.1822	1	**(Reject?)**	No	No	No
3. Normal	26.131	3		Yes	Yes	Yes

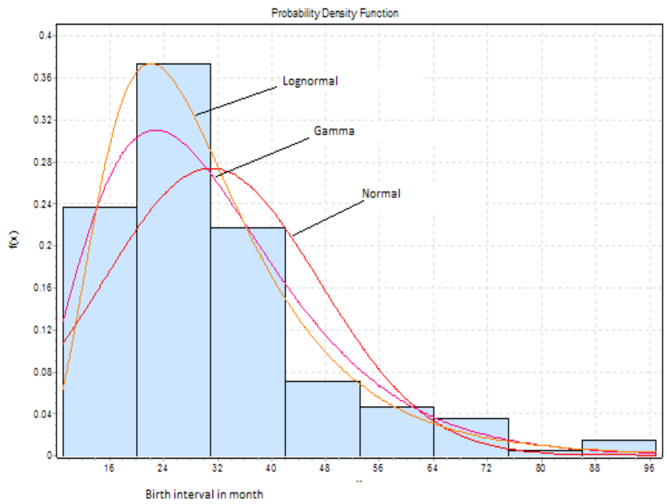
iii. Muslim



#	Distribution	Parameters
1	Gamma	α=3.6147 β=8.2338
3	Lognormal	σ=0.48026 μ=3.2754
5	Normal	σ=15.654 μ=29.762

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	8.339	2	**Decision**	Yes	Yes	Yes
2. Lognormal	4.7167	1	**(Reject?)**	Yes	Yes	Yes
3. Normal	75.318	3		Yes	Yes	Yes

iv. Other Religions

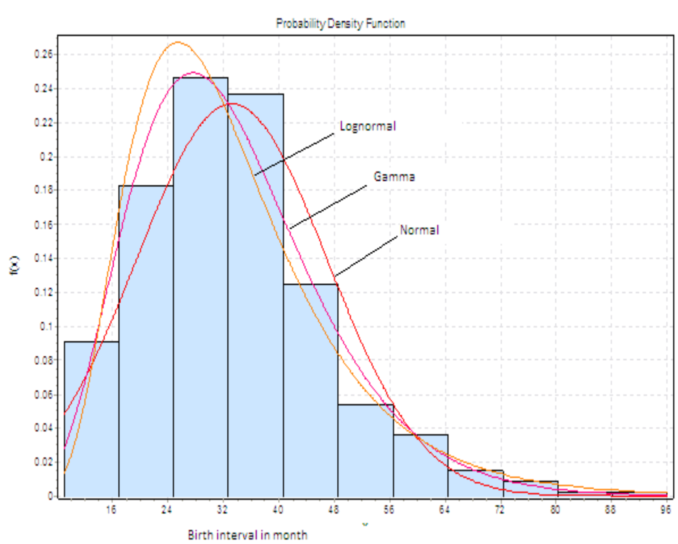


#	Distribution	Parameters
1	Gamma	α=3.7573 β=8.2627
2	Lognormal	σ=0.47474 μ=3.3205
3	Normal	σ=16.016 μ=31.045

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	0.58611	2	**Decision** **(Reject?)**	No	No	No
2. Lognormal	0.35611	1		No	No	No
3. Normal	6.2503	3		Yes	Yes	Yes

Table 3.b Second Birth Interval by Religion

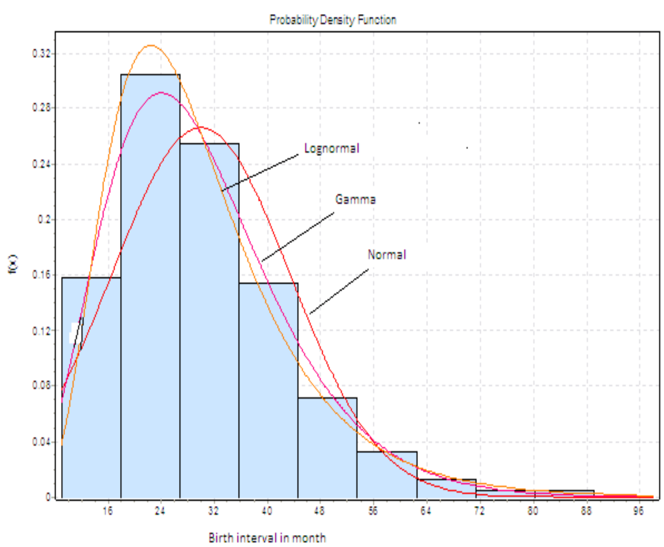
i. Orthodox



#	Distribution	Parameters
1	Gamma	α=5.9016 β=5.6266
2	Lognormal	σ=0.42416 μ=3.4171
3	Normal	σ=13.669 μ=33.206

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	2.4675	2	**Decision**	No	No	No
2. Lognormal	2.3505	1	**(Reject?)**	No	No	No
3. Normal	17.338	3		Yes	Yes	Yes

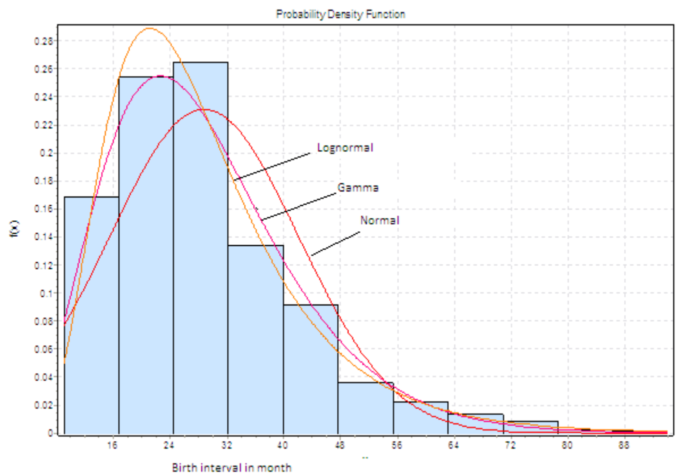
ii. Protestant



#	Distribution	Parameters
1	Gamma	α=5.0471 β=5.9368
2	Lognormal	σ=0.44109 μ=3.305
3	Normal	σ=13.338 μ=29.964

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	1.3481	1	**Decision**	No	No	No
2. Lognormal	1.6103	2	**(Reject?)**	No	No	No
3. Normal	12.439	3		Yes	Yes	Yes

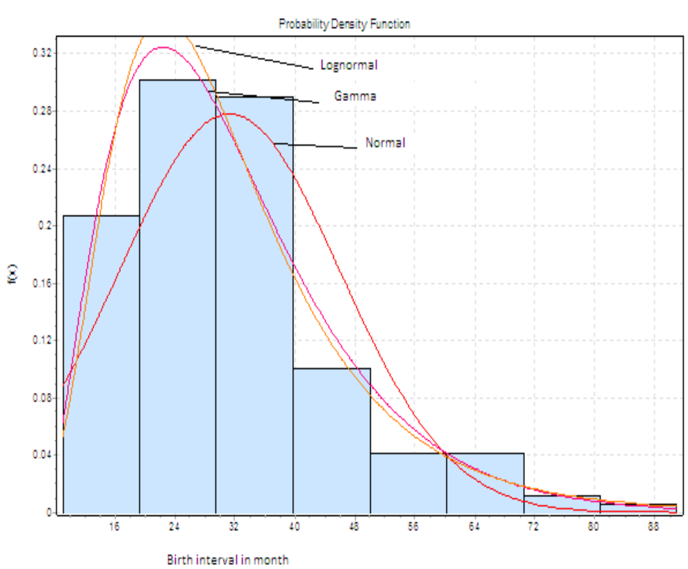
iii. Muslim



#	Distribution	Parameters
1	Gamma	α=4.6601 β=6.1705
2	Lognormal	σ=0.45454 μ=3.2572
3	Normal	σ=13.32 μ=28.755

Distribution	Statistic	Rank	**α-value and critical values**			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	4.9812	2	**Decision (Reject?)**	Yes	Yes	Yes
2. Lognormal	3.4417	1		Yes	Yes	No
3. Normal	34.608	3		Yes	Yes	Yes

iv. Others



#	Distribution	Parameters
1	Gamma	α=4.5176 β=6.9237
2	Lognormal	σ=0.47304 μ=3.3352
3	Normal	σ=14.716 μ=31.278

Distribution	Statistic	Rank	α-value and critical values			
			α-value	0.05	0.02	0.01
			critical values	2.5018	3.2892	3.9074
1. Gamma	0.47238	1	**Decision (Reject?)**	No	No	No
2. Lognormal	0.77299	2		No	No	No
3. Normal	2.4738	3		No	No	No