TRANSITIONAL PROBABILITY AND TIMING OF FIRST SEXUAL INTERCOURSE AND FIRST BIRTH AMONG NIGERIAN WOMEN

BY

OMISILE, KEHINDE OLUGBENGA

B.Sc. Mathematics (FUNAAB)

MATRIC NUMBER: 182374

SUBMITTED TO THE DEPARTMENT OF EPIDEMIOLOGY AND MEDICAL STATISTICS, FACULTY OF PUBLIC HEALTH, COLLEGE OF MEDICINE, UNIVERSITY OF IBADAN, IBADAN, NIGERIA

IN PARTIAL FULFILMENT OF THE REQUIREMENT FOR THE AWARD OF MASTERS DEGREE IN BIOSTATISTICS OF THE UNIVERSITY OF IBADAN

APRIL 2015

Dedication

I dedicate this work to the Lord Jesus Christ of Nazareth, and to my mother Victoria Omobonike Agbeke Ayobami Omisile.

Abstract

Age-at-first sex and age-at-first-birth have been identified as important markers of changes in social and health needs especially in the lives of women and an understanding of these important reproductive health events can provide clear understanding of how individual progress throughout these stages. Age-at-first-sex is an important indicator of risk of pregnancy, as well as STIs and HIV/AIDS among young adults. The interval between these reproductive events has also called for consideration. The wider the interval between first sex and first marriage, for instance, exposes a woman to having multiple sexual partners. The interval between first sex and first birth also suggests possible use of contraceptives and intentional delay in childbearing among the couples given several personal and community level factors.

These reproductive health events have survival times. Although many studies settle for the Cox proportional hazard model for estimating survival times but parametric methods have been described as having better estimates. Both parametric and non-parametric methods have been suggested for fitting these reproductive health events and even for the parametric methods, graphical and analytical methods have been used.

The main objective of this study was to model the baseline hazard of age at first sex, age at first birth and the interval between them using the Weibull distribution. Data for the study was obtained the Nigeria DHS 2013 for 38948 women of which 31075 (84.4%) ever had sex among whom 25370 (81.6%) ever gave birth. Exploratory data analysis, curve fitting, and model fitting were employed for the analyses. Kaplan-Meier estimates of the survival function was estimated and substituted into the Log cumulative hazard plot to check the suitability of Weibull distribution to model the hazard function. stgenreg procedure was used for analytical fitting of the hazard functions through maximum likelihood estimation. Fixed and random effect models for interval between first sex and first birth were fitted using streg and stmixed procedures respectively.

The results showed that the age at first sex, age at first birth and their interval changed with birth cohort. Specifically, older women had wider interval of sex debut and first birth compared to younger women. Weibull distribution was appropriate for the baseline hazard of age at first sex, age at first birth and the interval between them. The median age at first sex obtained from the log-cumulative hazard plot was 16.61 years (95% CI: 16.43–16.80) which was slightly higher than the Kaplan-Meier estimate of 16 years (95% CI: 15.94-16.05). The estimate of median age at first sex from the MLE of Weibull parameters was 16.97 years (95% CI: 14.97-19.23). The median age at first

birth for both graphical and MLE estimates were 19.47 years (95% CI: 19.23-19.72) and 19.60 years (95% CI: 17.02–22.56) respectively, which are comparable to the KM estimate of 19 years (95% CI: 18.94-19.06). The KM estimate for first sex-first birth interval was 2 years (95% CI: 1.97-2.03), which is also comparable to the median estimate from the Weibull estimation of 2.12 years (95% CI: 2.06 – 2.18).

Random effects Weibull model showed that age at first sex, education, religion, women's birth cohort, contraceptive use, marital status, and ever terminated a pregnancy were significant factors of the interval between first sex and first birth.

Word Count = 532.

Acknowledgements

This work is a combination of the effort of several persons. First is my supervisor Dr. Joshua O. Akinyemi who would always listen to me and tell me what to do every time I managed to show up. I acknowledge also my lecturers – Prof. O. Ayeni, Dr. Onoja M. Akpa, Dr. Bidemi O. Yusuf (Sub-Dean Faculty of Public Health), Dr. Adeniyi F. Fagbamigbe, Dr. Ayo S. Adebowale, Mr. Babatunde M. Gbadebo, Dr. Ibitola O. Fawole (HOD, Department of EMS), Mr. Rotimi F. Afolabi, Dr. Babatunde O. Adedokun, as well as Mr. Nathaniel, who took particular interest in my work and have been very supportive. I do not take your kindness for granted.

Funmilola Asa, thank you for staying by me throughout this task. My Moma and my siblings – Taiwo and Idowu were quite supportive and never complained when I had to do my thing. I became stronger during this period and I can always look up to them.

I appreciate the contribution of the members of staff of Association for Reproductive and Family Health, especially Mrs. Anne Taiwo and Ukeni Mercy, who encouraged me throughout the period; Temilolu Kumapayi and Damilola Osakile, who first told me about Biostatistics; and the Executives of the organization for their support during trying times.

I must also mention all my colleagues in the Biostatistics Class of 2013 who kept giving me the push to keep on – Bolu (first on the list), Idowu Philippe, Mr. Adetunji, Femi Alo (Class Rep), Tunrayo, Linda, Layo Ogunnika, Babatunde, and Wole. I acknowledge the prayers of my childhood friends – Ope & Ayo Famakinde, Tade & Lanre Afolabi, Wumi & Femi Akanmu, Opeyemi and Kayode Bababunmi, Eyiyemi and Adewole Ogunmodede, Abigael Olawoore and her husband, Folafemi & Kunle Korede. Thank you everyone.

Certification

I certify that this work was carried out by OMISILE, Kehinde Olugbenga in the Department of Epidemiology and Medical Statistics, Faculty of Public Health, University of Ibadan, Ibadan, Nigeria.

.... 07/05/15 ....

Dr. Joshua O. Akinyemi Date

Department of Epidemiology and Medical Statistics

Faculty of Public Health, College of Medicine

University of Ibadan, Ibadan, Nigeria.

Table of Contents

List of Tables

List of Figures

Chapter One

Introduction

The growing world population poses great threat to the future of mankind given the monotonous increase in population growth rate across the nations of the world, and sub-Saharan Africa especially Nigeria, having its fair share of the burden. Resources are shrinking and there is little time available to adequately plan for the inevitable challenge. In view of this, donor agencies, government and non-governmental organizations across the nations of the world have spent the last few decades investing in birth control commodities and their distribution, campaigns on need for child spacing and limiting, as well as behavioural change communication with regards to fertility issues at all levels – from community levels to policy advocacy. The Nigeria experience shows that contraceptives prevalence has stagnated at 15% between 2008 and 2013 after an initial 2% increase from 2003 rate (NPC and ICF International, 2014). These are in spite of the fact that 85% of women and 95% of men aged 15-49 years in Nigeria know at least one method of family planning. Nevertheless, on the one hand are fertility issues and on the other hand are reproductive health issues, including HIV and AIDS.

Since year 2005, the proportion of young people aged 15-24 years in Nigeria that had experienced sexual debut before age 15 years has almost doubled from 10% in 2005 to 16% in 2012 (Nigeria GARPR, 2014) while sexual experience is almost twice as much among females (37%) as among males (20%) aged 15-19 years. Of course, this would include consensual and coercive sexual intercourse. However, Nigeria has also seen a steady increase in condom use among sexually active individuals with multiple sexual partners within the period aforementioned with a rise from 56% in 2005 to 65% in 2012 (Nigeria GARPR, 2014). Study done among female sex workers in Nigeria shows that more than three-quarter had sexual debut before age 19 years

(IBBSS 2010, 2010). Early sexual debut does not only put one at risk of sexually transmitted infections and unwanted pregnancy but can also influence behavior, including multiple sexual partners, either as a sex worker or not (Eric & Eleanor, 2008).

Age-at-first-sex is an important indicator of risk of unwanted pregnancy (Atuhaire, 2011), sexually transmitted infections and HIV/AIDS in the lives of young people, and it could also be an important indicator for age-at-first-marriage (Basia, et al., 2002). On the other hand, some authors have used age-at-first-marriage as a proxy for measuring women's exposure to pregnancy in fertility studies (Basia, et al., 2002), but Stover (1998) and Ngalinda (1998), in view of societies with common premarital sexual activity, have suggested that age-at-first-sex be used as a better proxy. Eric & Eleanor (2008) have also noted that intiation of sexual intercourse at an early age exposes adolescents to multiple sexual partners and longer period of sexual activity before they eventually marry or "form long-term monogamous relationships". However, in Northern Nigeria, early marriage is more common compared to other places in the country (Olalekan, 2008); and early marriage has been linked with early childbearing.

There are two important questions that come to mind given this background. First, we would want to know what is the probability that a certain individual would have sex at a paticular age given his or her socio-demographic characteristics and other important charactersitics. Second, we would want to know what background characenistics of an individual predisposes him or her to sexual intercourse at an early age. These questions also apply to age-at-first-marriage and age-at-first-birth as well as the factors that determine the length of the interval between the events. Some researchers have argued that socioeconomic environment in most sub-Saharan countries encourage early sexual activity especially among females (Okonufua, 2000; Luke, 2003). Hence, there is need to account for community-level factors that influence sexual activities among the individuals. A number of authors, some of a number of whose literatures were reviewed, have worked on these reproductive health events or milestones using both parametric, semi-parametric and multi-level modeling, but borrowing from the perception of Finer & Philbin (2014), the recent release of the 2013 Nigeria DHS implies that we have new data to update the past published figures.

This study investigates transitional probability of timing of first sexual intercourse and first birth among women in Nigeria, as well as their paired intervals, before fitting a parametric model to determine the individual- and community-level factors that influence interval of first sex and first birth.

1.1. Problem statement

Despite that much researches have been carried out on age-at-first-sex, age-at-first-marriage and age-at-first-birth, none of the studies provided useful information about the probability distribution of their baseline hazards. Hazard here refers to the probability of initiating sex, for instance, given a specific age. This is so because the variables are usually fitted into the Cox proportional hazard models (Oyedokun & Odumegwu, 2013; Chandrasekhar, 2010; Lofstedt, et al., 2005), which does not assume a probability distribution for baseline hazard.

Also these variables have been fitted separately into proportional hazard models without accounting for community-level influences on the variables and where the community-level influences were accounted for, the time-to-event properties were expunged (Olalekan, 2008; Ngalinda, 1998). Subaiya & Johnson (2008) assumed a normal distribution for the length of time between age-at-first-sex and age-at-first-marriage by fitting an ordinary least square (OLS) linear regression but OLS estimation does not account for the multi-level structure in data collected and produces less precise estimates of the standard error (Angeles & Mroz, 2001). Angeles & Mroz individuals. A number of authors, some of a number of whose literatures were reviewed, have worked on these reproductive health events or milestones using both parametric, semi-parametric and multi-level modeling, but borrowing from the perception of Finer & Philbin (2014), the recent release of the 2013 Nigeria DHS implies that we have new data to update the past published figures.

This study investigates transitional probability of timing of first sexual intercourse and first birth among women in Nigeria, as well as their paired intervals, before fitting a parametric model to determine the individual- and community-level factors that influence interval of first sex and first birth.

1.1. Problem statement

Despite that much researches have been carried out on age-at-first-sex, age-at-first-marriage and age-at-first-birth, none of the studies provided useful information about the probability distribution of their baseline hazards. Hazard here refers to the probability of initiating sex, for instance, given a specific age. This is so because the variables are usually fitted into the Cox proportional hazard models (Oyedokun & Odumegwu, 2013; Chandrasekhar, 2010; Lofstedt, et al., 2005), which does not assume a probability distribution for baseline hazard.

Also these variables have been fitted separately into proportional hazard models without accounting for community-level influences on the variables and where the community-level influences were accounted for, the time-to-event properties were expunged (Olalekan, 2008; Ngalinda, 1998). Subaiya & Johnson (2008) assumed a normal distribution for the length of time between age-at-first-sex and age-at-first-marriage by fitting an ordinary least square (OLS) linear regression but OLS estimation does not account for the multi-level structure in data collected and produces less precise estimates of the standard error (Angeles & Mroz, 2001). Angeles & Mroz

(2001) also noted that "no other unbiased estimator can provide more efficient estimates than the maximum likelihood procedure". Furthermore, inferences based on assumption of a particular probability distribution, even for skewed variables, are more precise and have the advantage of in terms of prediction, extrapolation and quantification, if the assumption is valid (Crowther, et al., 2014; Collet, 1994).

1.2. Purpose of the study

The purpose of this study is to provide information regarding probability distribution of the hazard of age-at-first-sex and age-at-first-birth and employ a multilevel parametric proportional hazard model to investigating what factors determine the length of first sex to first birth.

1.3. Objective of the study

The main objective of the study is to investigate the probability distribution of baseline hazard of age-at-first sex and age at first birth.

1.4. Specific objectives

Objectives of this study include:

1. To investigate the probability distribution of baseline hazard of age-at-first-sex and age-at-first-birth.
2. To explore the probability distribution of the intervals between pairs of age-at-first-sex and age-at-first-birth.
3. To examine the individual- and community-level factors that determine the interval between age at first sex and age at first birth using parametric survival analysis.

1.5. The research questions

1. What are probability distribution of hazard functions of age-at-first-sex and age-at-first-birth among women in Nigeria?

2. What are the probability distribution of hazard functions of the intervals between pairs of age-at-first-sex and age-at-first-birth among women in Nigeria?
3. What are the individual- and state-level factors that influence interval between first sex and first birth among women in Nigeria?

1.6. Limitations of the study

Time to event data from surveys, especially nationally representative survey such as Nigeria DHS, depend on recall information and could introduce recall bias into the RH milestones. There were observed inconsistencies in the reporting of age at reproductive events and these were explicitly identified in the methods section. The computer resources – hardware and software – available were not capable of running more than two-level mixed effect parametric models hence the entire hierarchical structure of the data could not be fully accounted for. However, all the studies reviewed showed that researchers generally do not go beyond the second level of the hierarchy.

Chapter Two

Literature Review

2.0. Introduction

Five reproductive health events or milestones (as used by Finer and Philbin (2014)) stand out in the lives of women and they mark obvious change in all aspects of a woman's life – physically, emotionally, socially and intellectually. These events are menarche, first sexual intercourse, marriage or first cohabitation, first birth, and menopause. In their study of these RH milestones among men and women in the United States, using the National Survey and Family Growth (NSFG) data, Finer and Philbin (2014) also included age at first contraceptive use as one of the milestones in a woman's reproductive health issues. The age at which these RH events occur are influenced by norms and values at community levels and have called for regulatory policies. Hence, a lot of work have been done in this area in medical, demographical, mathematical, statistical, and social sciences, as well as policy interventions. This literature review is broadly divided into two parts. The first part focuses on the reproductive health events, their nature, burden and determinants, and the second part focuses on the statistical methods of survival analysis for analyzing the RH events.

2.1. The burden of these RH milestones

The magnitude of sexually transmitted infections and HIV/AIDS in sub-Saharan African countries have necessitated several studies on the nature and predictors of sexual activities among individuals, especially young people; and, sexual behavior and maternal health remain issues of important public health concern. Nigeria, being a sub-Saharan African nation has also had its fair share of the scourge of STIs and HIV/AIDS with current prevalence of 3.4% (NARHS Plus, 2013), and 35% of the Nigerian States having prevalence higher than that of the national estimate. Almost two-fifth of the males and one-fifth of females aged 15-19years in Nigeria have ever had sex (NARHS Plus, 2013), with median age at first sex of 17 years for both genders among those aged 15-24 years, which is slightly less than that of Nigeria DHS 2013 (NPC [Nigeria] and ICF International, 2014), estimated at 17.6 years among women and 21.1 years among men aged 25-49 years. Oladokun, et al. (2007) also found that male adolscents are three times more likely to be sexually exposed than the female adolescents, with a median age at sex debut of 15 years for both gender.

Furthermore, apart from STIs and HIV/AIDS, of serious reproductive health and demographic concern is fertility. Fertility depends not only on marital unions but also on unwanted pregnancies as a result of early sexual initiation, unprotected sexual intercourse, age at first sex (Akpa & Ikpotokin, 2012) and frequency of sexual intercourse (NPC [Nigeria] and ICF International, 2014). Fertility rate in Nigeria is currently approximately 6 births per woman (NPC [Nigeria] and ICF International, 2014). Changes in age at marriages have implications for fertility, as an increase in the age at marriage reduces the number of years available for childbearing (Chandrasekhar, 2010) and Adebowale et al. (2012) also pointed out that age at first marriage has implication for the health of the mother and of their under-five children.

Worldwide there are 60 million women aged 20-24 years who were married before the age of 18 years (Chandrasekhar, 2010), and while median age at first marriage among women in Nigeria is estimated at 18.1 years and among men at 27.2 years, median age at first birth among Nigerian women is 19 years (NARHS Plus, 2013), which is slightly less than 20.2 years estimated by Nigeria DHS 2013 (NPC [Nigeria] and ICF International, 2014) among women aged 25-49 years. Median age at first marriage among women aged 25-49 years in urban areas of Pakistan and

India are 19.7 years and 18.7 years respectively (Chandrasekhar, 2010), which are less than what holds in Nigieria.

2.2. Reproductive events

Age-at-first sex, age-at-first marriage and age-at-first-birth have been identified as important markers of changes in social and health needs especially in the lives of women (Finer & Philbin, 2014) and the median ages at which these events occur can be helpful for understanding how individuals progress through these stages.

2.2.1. Age at first sex

While age-at-first-sex is an important indicator of risk of pregnancy, as well as sexually transmitted infections and HIV/AIDS among young adults (Finer & Philbin, 2014; Atuhaire, 2011; Kaestle, et al., 2005; Zaba, et al., 2004) it could also be an important indicator for age-at-first-marriage (Adebowale, et al., 2012; Basia, et al., 2002). Eric & Eleanor (2008) have also noted that intiation of sexual intercourse at an early age exposes adolscents to multiple sexual partners and longer period of sexual activity before they eventually marry or "form long-term monogamous relationships", and early childbearing and its risks of maternal mortality and morbidity (Subaiya & Johnson, 2008). Multiple sexual partners is a mediator variable between risk of human papillomavirus (HPV) infection and age at first sex, the latter which has been found to be strongly associated with HPV infection by several authors (Khan, et al., 2002). Akpa & Ikpotokin (2012) have identified age at first sex as a strong proximate determinant of number of children a woman ever gives birth to before age fifty.

Age at first sex was defined by Measure Evaluation DHS as "age at which an individual initiates sexual intercourse" (NPC [Nigeria] and ICF International, 2014). Many programs targeted at adolescents and youths on sexual reproductive health (SRH) issues encourage delay of initiation of sexual intercourse, especially through the Minimum Preventive Package Interventions (MPPI) that provides both in-school and out-of-school adolescents with the ABC of HIV. The ABC refers to abstinence, being faithful to one's partner, and consistent and correct condom use for prevention of STIs and HIV & AIDS. In Nigeria, for instance, in-school adolescents are only provided with information about the A – abstinence till marriage; with life building skills to help them make important decisions to delay sexual unions till marriage.

2.2.1.1. Determinants of age at first sex

Zaba *et al.* (2004) have noted that in several countries prevalence of HIV among young pregnant women attending antenatal clinics have reduced and this they associated with changes in age at first sex, rates of partner change, sexual mixing patterns, and condom use. Specifically, Cremin *et al.* (2009) noted that delay in sexual debut coincided with declines in HIV prevalence in Uganda and Zimbabwe, although these changes are not likely to contribute substantially to lifetime risk of infection or HIV incidence (Hallett, et al., 2007). A number of factors have also been associated with age-at-first-sex. In the study of adolescents and young adults aged 14-22 years in South Africa, Oyedokun & Odumegwu (2013) found that being older, being female, racial origin, having secondary or tertiary level of education demonstrated a delaying effect on the age at first sex. Specifically, secondary education has been noted as a strong determinant of age at first sex (Zaba, et al., 2004). This result also identified differentials between gender which is similar to findings from Nigeria DHS 2013 that showed that 16% of females compared to 3% of males had had sex before age 15 years (NPC [Nigeria] and ICF International, 2014).

Comparing trends of age at first sex for six African countries, Zaba, et al. (2004) found that median age at first sexual intercourse has considerably declined in Ghana, Kenya, and Uganda than other countries such as Tanzania, Zambia and Zimbabwe. Median age at first sexual intercourse in Nigeria has slightly increased from 17.7 years among women and 20.6 years among men in 2008 to about 18.1 years among women and about 21 years among men in 2013 (NPC [Nigeria] and ICF International, 2014; NPC [Nigeria] and ICF Macro, 2009).

As a variable, age at first sex is a quantitative variable depicting number of years that had elapsed before an individual commences or experiences first sexual intercourse, either consensual or coercive. This, in turn, makes the age at first sexual intercourse a time-to-event variable that should be analyzed using appropriate censored data statistical methods. This would be further discussed in subsequent sections.

It should, however, be noted that in most cross-sectional survey that explore age at sexual debut and other event histories the data obtained generally depend on recall information or retrospective interview (NPC [Nigeria] and ICF International, 2014; Gayawan & Adebayo, 2014; NARHS Plus, 2013; Mensch & Singh, 2005; Hogan, et al., 2000). For instance, ever-married individuals are assumed to have initiated sex and hence are asked at what age they initiated sex (recall data) while adolescents are asked if they had ever had sex (current status) and what age they initiated sex (recall data) if they had had sex (Zaba, et al., 2002). This poses major challenge in the analysis of such data due to inconsistencies that could be observed and reservation about such information due to religious, social, and cultural beliefs.

2.2.2. Age at first marriage

"Marriage is one of the primary events during the transition to adulthood" (Copen, et al., 2012; Jensen & Thornton, 2003), but recent studies have showed that many young people now defer first marriage, which might be linked to staggering socioeconomic statuses and need to complete higher education – a major determinant of age at first sex (Chandrasekhar, 2010; Garenne, 2004).

Deferring first marriage has also been found to be influenced by the practice of cohabitation, a growing trend in western culture (Finer & Philbin, 2014; Copen, et al., 2012).

Age at first marriage is the age, in years, at which an ever-married person started living with their first spouse/partner (NPC [Nigeria] and ICF International, 2014; Gayawan & Adebayo, 2014). This definition may not be all encompassing, as it does not particularly indentify couples in 'consesual unions', which is widely occuring and acknowledged form of union in Latin America (Mensch & Singh, 2005). Chandrasekhar (2010) noted that "Worldwide there are 60 million women aged 20-24 years who were married before the age of 18 years". Median age at first marriage, just as median age at first sex, varies from country to country and from culture to culture (Adebowale, et al., 2012). For instance, in a study carried out among ever-marrid women in Awka Town, Anambra State, Nigeria, Agbim & Ikyernum (2013, p. 50) found the average age at first marriage among the study population to be between 20-24 years, which is on the average four years above the national average reported later in the NDHS 2013. In the United States, median age at first marriage is 25.8 years for women and 28.3 years for men, which are also as a result of the fact that many of them cohabit (11%) before they marry (Copen, et al., 2012). Studies conducted on age at first marriage among 32 sub-Saharan African countries revealed that age at first marriage is still low, in the 15-19 years band for most countries, except Botswana, Namibia and South Africa (Garenne, 2004). Erulkar & Bello (as cited in Gayawan & Adebayo, 2013) have stated that 76% of sexually active women in Nigeria are married, similar to earlier findings of Singh & Samara (1996), who noted hat 20-50% of women in developieng countries are married by age 18 years and 40-70% are married before age 20 years. Trend in age at first marriage in Nigeria shows that the median age at first marriage for women aged 25--49 years was 17.7 years in 1990, 17.9 years in 1999, 16.6 years in 2003, 18.3 years in 2008, and 18.1 years in 2013

(Gayawan & Adebayo, 2014; NPC [Nigeria] and ICF International, 2014; Feyisetan & Bankole, 2002).

Age-at-first-marriage has been used as a proxy to measuring a woman's exposure to pregnancy in fertility studies (NPC [Nigeria] and ICF International, 2014; Shayan, et al., 2014; Garenne, 2004; Basia, et al., 2002; Ngalinda, 1998), but Stover (1998), in view of societies with common premarital sexual activity, has suggested that age-at-first-sex be used as a better proxy. Also, age at which a woman commences childbearing influences the total number of children she would bear throughout her reproductive period (Gayawan & Adebayo, 2013; Chandrasekhar, 2010; Akpa & Ikpotokin, 2012; Garenne, 2004). Lofstedt *et al.* (2005) and Shayan et al. (2014) have also showed that late age at first marriage and/or longer intervals between age at marriage and first birth slows down the rate of population growth. On other hand, studies by Bongaart *et al.* (as cited in Subaiya & Johnson, 2008) found that later marriages have their own challenges including increased risks of HIV/AIDS and higher rates of premarital births. Other factors influenced by age at first marriage include health, social status life cycle, dynamics of family formation, exposure to multiple sexual partners and risk of STIs (Garenne, 2004). In particlar, early marriages (15-19 years), with their pros and cons, have been linked to higher fertility, larger families, lower numbers of sexual partners and fewer STDs (Kumchulesi, et al., 2011; Garenne, 2004).

Nations of the world have spent considerable time enacting and reviewing policies to guide age at marriage. This becomes necessary in order to have a good grasp on and slow the pace of fertility and population growth, and ultimately improve quality of life. Nigeria too has such policies with the Child Right Act passed in 2003 that imposed the minimum age of marriage to be 18 years for girls (Gayawan & Adebayo, 2014), similar to the National Population Policy 2000, National Youth Policy 2003, National Policy for the Empowerment of Women 2001, and Prohibition of Child Marriage Act 2006 all in India (Chandrasekhar, 2010). "United Nations' 1962 Convention on the 'Consent to Marriage, Minimum Age for Marriage and Registration and an Associated Recommendation' calls on member states to establish a minimum age of marriage of no less than 15 years" (Jensen & Thornton, 2003). However, in Northern Nigeria, child marriage is more common compared to other places in the country (Olalekan, 2008).

2.2.2.1. Determinants of age at first marriage

Several studies have been carried out to determine the socioeconomic and demogrpahic factors that influence age at first marriage; and this inludes studies by Copen, et al. (2012) who noted that premarital cohabitation contributes significantly to delay in first marriage for both men and women. Age at first marriage has also been found to be positively associated with prolonged education, urbanization, new attitudes towards partnership, and across religion and ethnic divide (Gayawan & Adebayo, 2014; Adebowale, et al., 2012; Kumchulesi, et al., 2011; Garenne, 2004). Specifically, Nigerian women with primary education are at higher risk of early marriage than women with at least secondary education (Singh & Samara, 1996); and women that live in urban areas, with high exposure to higher education and wealth, tend to delay marriage (Gayawan & Adebayo, 2014). Furthermore, Gayawan & Adebayo (2014) also found that Moslems and Catholics marry earlier than respective comparative groups and women from Hausa ethnic groups are at higher risk of early marriage than other non-Hausa groups in Nigeria.

On the othe hand, Mensch & Singh (2005) have argued that education attainement, though has an association with marriage, as most developing countries have experienced both a rise in educational attainment and a rise in age of marriage, the association is weaker than might be expected given the determining power often attributed to educational change. Their study was a review of association between educational attainment and marriage from data from 49 DHS countries. They further argued that educational attainment may be endogenous, as those who intend to marry may leave school earlier while those who marry later, for whatever reason, may stay in school longer. Shretha & Shretha (2008) put it much clearly that on one extreme is the fact that marriage is a response to social norms and on the other extreme, marriage happens when rational individuals decide the time will maximize their utility function subject to constraints. Also, they considered the heterogeneity of individuals. They said marriage could either be considered that all individuals belonging to the same cohort or population are subject to the same type of influence or it can be thought that individuals, as being heterogeneous in their choice of when to marry according to some unobservable characteristics (Shretha & Shretha, 2008).

2.2.3. Age at first birth

As metioned in the preceeding section, age at first birth is determined not only by age at first marriage but also by age at first sexual intercourse (Agbim & Ikyernum, 2013), given the common premarital sexual debut, contraceptive use, and abortion among young people (Lofstedt, et al., 2005). It has been shown also that about 10% of women in Nigeria are likely to have their first birth before their 15th birthday (Akpa & Ikpotokin, 2012) and that women who have their first birth before age 15 years have 18% and 61% higher fertility rate than women who had their first birth between age 15-30 years and above age 30 years respectively. The finding of Akpa & Ikpotokin (2012) is comparable to that of Nigeria DHS 2013 where it was stated that 8% of women aged 25-49 years have given birth by age 15 years (NPC [Nigeria] and ICF International, 2014) and 23% of adolescent women aged 15-19 years are already mothers or pregnant with their first child. Timing of first birth has social, medical and economc consequences (Hogan, et al., 2000; Hirschman & Rindfuss, 1980)

Age at first birth has been defined as the age in years at which a woman commences childbearing and "is an important determinant of the overall level of fertility as well as the health and welfare of the mother and child" (NPC [Nigeria] and ICF International, 2014). Insight into age at first birth provides could provide a baseline for also measuring use of contraceptives and fertility. In studying the association between age at first birth and knowledge of contraceptive methods among Yoruba Women in Nigeria, Oyefara (2012) found that knowledge of contraceptive methods was significantly higher among older mothers than adolescent mothers. However, one of the challenges with his finding was to completely attribute the knowledge of family planning methods to whether a mother was older or adolescent without accounting for other sources of variation such as urban-rural divide, educational attainment, religion, and his work did not also capture sources of information about family planning. On the other hand, Gbaroro & Igbafe (2002) have found that older women have better attitude and attend ANC better than the teenage mothers. Low ANC attendance could be as a result of the fact that the teenage mothers were trying to avoid stigmatization due to early pregnancy and/or pregnancy occuring earlier than expected. On the other hand, women that gave birth later than 35 years have been found to be at high risk of cervical incompetence at delivery, malaria, preterm labour, and fibroid (Lisonkova, et al., 2010; Gharoro & Igbafe, 2002).

Age at first birth not only influences a womans cummulative fertility or children-ever-born (Gayawan & Adebayo, 2013), but also has broader implications for women's roles and social change in general (Hirschman & Rindfuss, 1980), as well as biological, medical and economic factors (Feyisetan & Bankole, 2002; Hogan, et al., 2000). Several parts of the world have experienced a rise in timing of first marriage and, hence, there is expected shortened timing of first marital birth (MacQuarrie, 2014). As mentioned earlier, later timing of first birth due to postponed first marriage, prolonged educational attainment, career puisuit in lieu of family life and so on are more of individual level factors that affect age at first birth with little or no community level effect. This is supported by the findings of Gayawan & Adebayo (2013), who found that "variation in age at first birth in Nigeria is determined more by individual household than by community, although substantian geographical variations in timing of first birth also exist". Hence, Hirschman & Rindfuss (1980) rightly noted that a woman's fertility is actualy an outcome of a series of behaviours, decisions, and events that may span the childbearing period of the woman.

Median age at first birth in Nigeria has changed significantly over time. Median age at first birth among women of reproductive age was 20.4 years in 2008 and 20.2 years in 2013 (NPC [Nigeria] and ICF International, 2014; NPC [Nigeria] and ICF Macro, 2009).

2.2.3.1. Determinants of age at first birth

Prominent among the determinants of age at first birth is age at first marriage (Agbim & Ikyemum, 2013). It has been observed that delayed marriage is also a way to delay childbearing especially in cultures where non-marital childbearing is rare, such as India (MacQuarrie, 2014; Amin & Bajracharya, 2011). From Nigeria DHS 2013, it was observed that median age at first birth increased with level of education and women with no education had their first birth four years earlier than women with secondary education (NPC [Nigeria] and ICF International, 2014).

2.2.4. First sex-first marriage interval

Subaiya & Johnson (2008) have argued that age at first marriage should not be the focus of analysis on reproductive health concerns but rather the gap in years between sexual debut and marriage. Understanding what happens in the lives of women between sexual debut and marriage could provide insight into women's reproductive health as well as what happens to these women in the interim. Subaiya & Johnson (2008) hypothesized that women that have longer interval between sexual debut and marriage are at risk of multiple sexual partners. For communities where premarital sexual intercourse is not prescribed the interval between sexual debut and marriage may be negligible, as age at sexual debut is cultural-specific and not easily influence by external factors (Subaiya & Johnson, 2008).

Subaiya & Johnson (2008) in studying factors associated with the interval between age at first sex and first marriage assumed a normal distribution for the interval and, hence, fitted it to an ordinaly least square regression (OLS). No attempt was made to fit the variable into a probability distribution before multivariate analysis. It has been earlier pointed out that age variables from homogenous groups are usually skewed. Their study further showed that the factors that affect the interval between first sex and first marriage include having a birth prior to first union, initiating sexual activity while still in school, number of lifetime partners, terminating a pregnancy, and practicing contraception, which are associated with longer intervals of sex debut and marriage, and hence longer interval of premarital sexual activity.

2.2.5. First sex-first birth interval

There is dearth of literature that focuses on the interval between first sex and first birth. This is so because most studies focused much on the interval between first marriage and first birth, perhaps assuming that most birth take place within the neighborhood of marital union or other long term cohabitation. Gayawan & Adebayo (2013) have described age at first birth and age at first sex as better indicators of the onset of sexual exposure than age at first marriage given the presence of pre-marital sexual intercourse which is becoming increasingly more common in the Nigerian culture. For instance, they mentioned the study of Orubuloye, Caldwell and Caldwell (in Gayawan & Adebayo, 2013) which was carried out in Ekiti State, Southwest Nigeria, where is was found that four-fifth and two-thirds of women from urban and rural areas, respectively, did not marry as virgins. Finer and Philbin (2014), in presenting their work on the intervals between successive marital sexual intercourse is not prescribed the interval between sexual debut and marriage may be negligible, as age at sexual debut is cultural-specific and not easily influence by external factors (Subaiya & Johnson, 2008).

Subaiya & Johnson (2008) in studying factors associated with the interval between age at first sex and first marriage assumed a normal distribution for the interval and, hence, fitted it to an ordinaly least square regression (OLS). No attempt was made to fit the variable into a probability distribution before multivariate analysis. It has been earlier pointed out that age variables from homogenous groups are usually skewed. Their study further showed that the factors that affect the interval between first sex and first marriage include having a birth prior to first union, initiating sexual activity while still in school, number of lifetime partners, terminating a pregnancy, and practicing contraception, which are associated with longer intervals of sex debut and marriage, and hence longer interval of premarital sexual activity.

2.2.5. First sex-first birth interval

There is dearth of literature that focuses on the interval between first sex and first birth. This is so because most studies focused much on the interval between first marriage and first birth, perhaps assuming that most birth take place within the neighborhood of marital union or other long term cohabitation. Gayawan & Adebayo (2013) have described age at first birth and age at first sex as better indicators of the onset of sexual exposure than age at first marriage given the presence of pre-marital sexual intercourse which is becoming increasingly more common in the Nigerian culture. For instance, they mentioned the study of Orubuloye, Caldwell and Caldwell (in Gayawan & Adebayo, 2013) which was carried out in Ekiti State, Southwest Nigeria, where is was found that four-fifth and two-thirds of women from urban and rural areas, respectively, did not marry as virgins. Finer and Philbin (2014), in presenting their work on the intervals between successive reproductive health milestones provided insight into the interval between age at first sex and age at first contraception and not exactly first birth.

For the purpose of this study, the interval between sexual debut and first birth will be defined as the number of years that had elapsed between the two events. Measuring the event in years as a result of the fact that first, the study focuses on the events more as a public health concern than reproductive health issue. Second, because the data that used for the study does not measure capture the interval in months and third, because of the possibility of pre-marital births as well as delayed birth due to deliberate contraception or any other physiological issue. Hence, the interval between first sex and first birth for married women that had their first sex and first birth same year as first marriage will be measured as zero while others will be measured as positive whole numbers. More insight into this is provided in Section 3.6 of methods.

2.2.6. First marriage-first birth interval

The interval between first marriage and first birth has been defined as "length of time in months between an individual's time of first marriage and the timing of the birth of their first child" (Amin & Bajracharya, 2011).

Several studies have been carried out on marriage and first birth interval but there is dearth of information on the modelling of marriage and first birth (Shayan, et al., 2014), as a lot of studies have focused mainly on birth spacing (Amusan & Mohd, 2014), which is also of utmost public health and demographic concern. The definition of marriage and first birth interval precludes births that are less than 7 months into marriage (Amusan & Mohd, 2014), as this would not imply conception before marriage.

Study by Amusan and Mohd (2014) showed that the interval between marriage and first birth is best modelled using parametric distributions. In their study, inverse Gaussian, Log-logistic, Weibull, and Burr Type XII were used to model marriage and first birth interval and it was observed that Inverse Gaussian had the best fit for the interval. Shayan *et al.* (2014) used Log-logistic and Generalized Gamma to model marriage and first birth interval.

2.2.7. Importance and relevance of the intervals

Finer & Philbin (2014) have said that women's and men's lives contain a number of milestones that mark changing life stages with different health and social needs. They further showed that "specific events, such as the initiation of sexual intercourse and the completion of childbearing, often mark the beginning or end of stages during which different needs predominate. For example, menarche marks the beginning of the biological capacity to bear children, and first intercourse (coitarche) marks the beginning of the period of risk for pregnancy and sexually ~~transmitted~~ diseases. First marriage closes a period at which an individual is at risk for a nonmarital birth, although risk for sexually transmitted infections still exists if individuals have other partners."

MacQuarrie (2014) noted that "as women's age at marriage increases, prior research directs us to expect a shorter first marital birth interval". Subaiya & Johnson (2008) also observed that "while the factors associated with age at first marriage have been well-researched, what is lacking is an understanding of what factors are associated with longer intervals and what the implications of those factors might be for women's health and welfare". Finer & Philbin (2014) further indicated that "ordering the median ages at which these events occur can be helpful for understanding how individuals progress through these stages, and also enables us to compare various subgroups on the events of interest. Generational differences in the time between these events, their order, and their frequency are important markers of social change that allow us to track and predict demographic trends, economic and societal inequality, and changes in social norms". Long average interval between sexual debut and first marriage have been hypothesized as key factor driving the spread of HIV (Johnson, et al., 2009), especially in countries of Southern Africa, where marriage tends to occur relatively late. Although the underlying assumption is that mariage is protective of spread of HIV. Finer & Philbin (2014) have also suggested that having insight into the gap between first sex and first use of contraception, as it may represent a period of heightened vulnerability to unintended pregnancy.

In describing the relationship between the health markers or reproductive events and their intervals, Finer & Philbin (2014) used the diagram below:

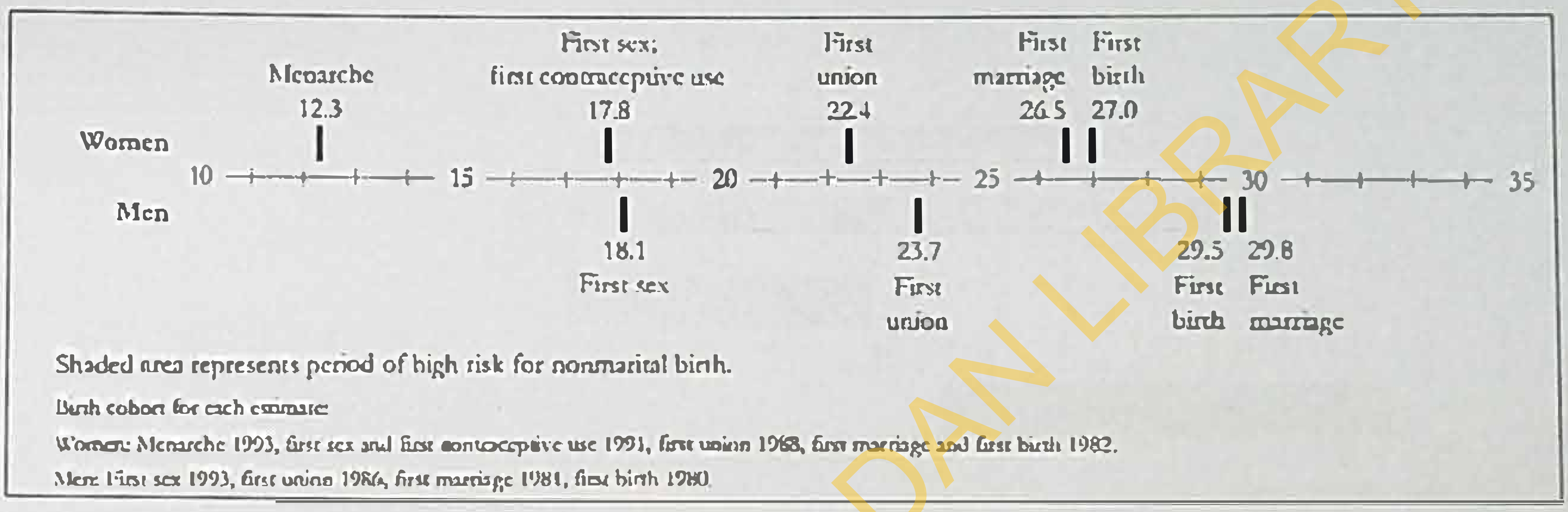


Figure 2.2-1: "Current" median ages at reproductive events (i.e. median age at event for most recent cohort for whom data are available)

Source: (Finer & Philbin, 2014)

spread of HIV. Finer & Philbin (2014) have also suggested that having insight into the gap between first sex and first use of contraception, as it may represent a period of heightened vulnerability to unintended pregnancy.

In describing the relationship between the health markers or reproductive events and their intervals, Finer & Philbin (2014) used the diagram below:

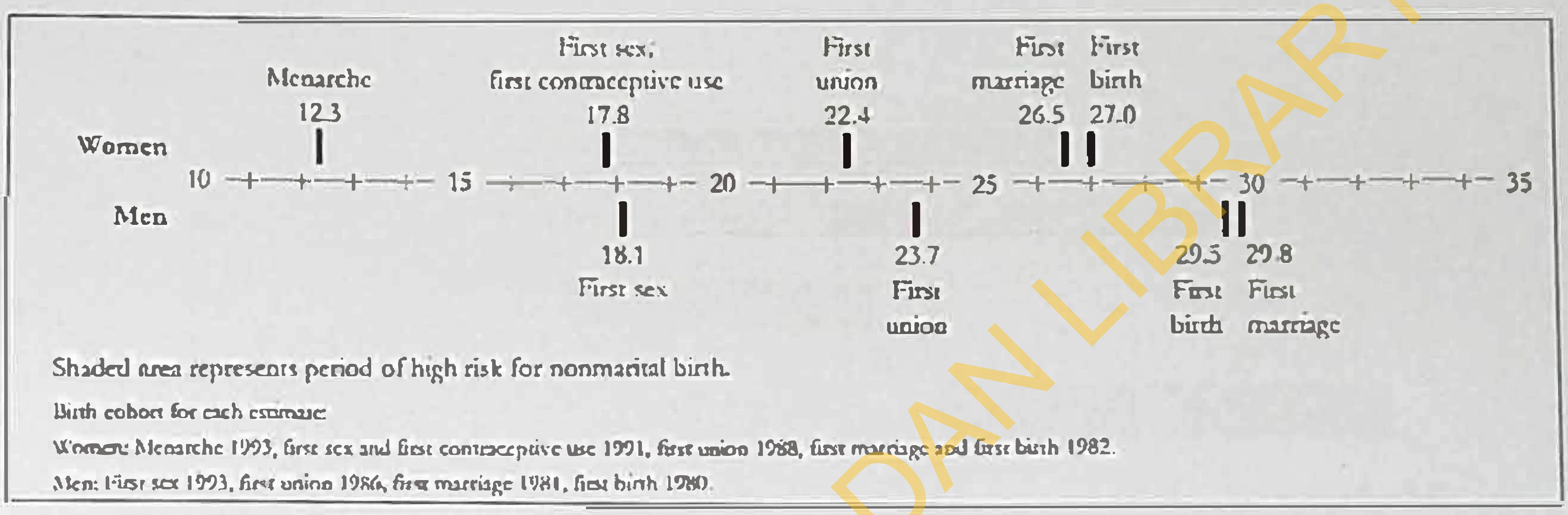
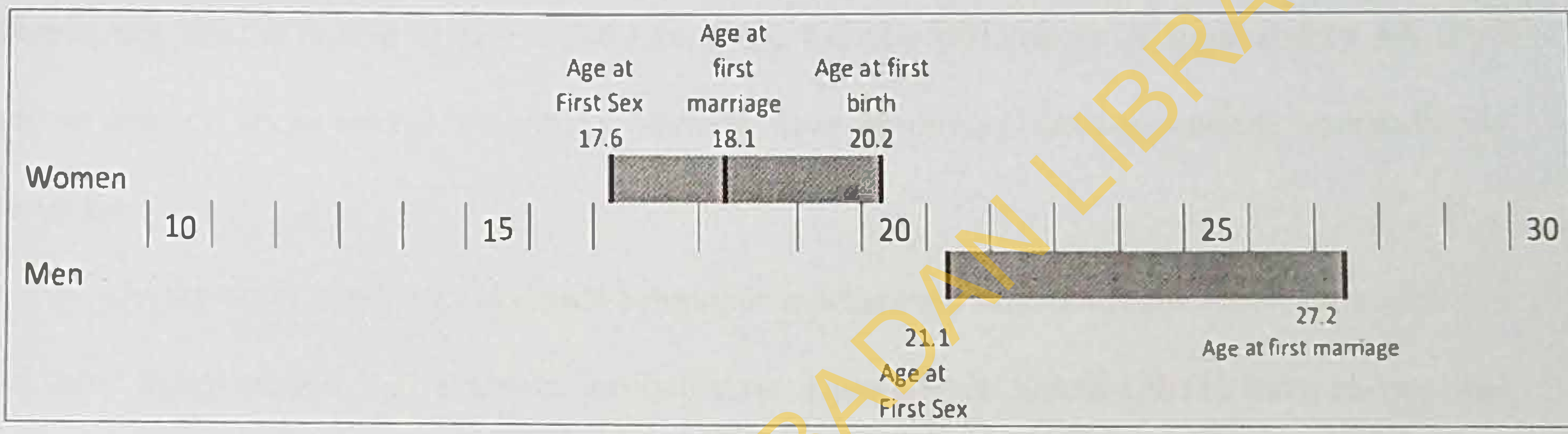


Figure 2.2-1: "Current" median ages at reproductive events (i.e. median age at event for most recent cohort for whom data are available)
Source: (Finer & Philbin, 2014)

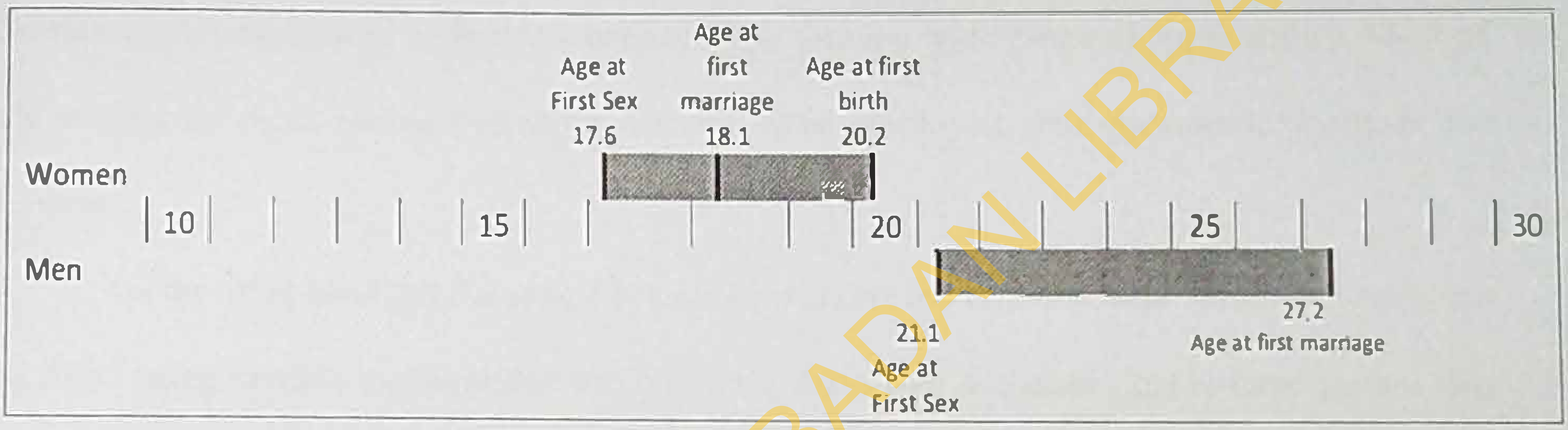
Using NDHS 2013 data, median age of reproductive events is as presented in Figure 2.2-2. Since the margin between sexual debut and first marriage is on the average 6 years for men, there is an increased risk of multiple sexual partners and sexual encounters, STIs and HIV/AIDS, and unwanted pregnancies. This is not too different for the women, as the wider the interval between sexual debut and first birth, the indication exists for contraception, multiple sexual partners and sexual encounters and unwanted pregnancies.



Shaded area shows interval between successive events

Figure 2.2-2: Median age at reproductive events
Data Source: (NDHS, 2013)

Using NDHS 2013 data, median age of reproductive events is as presented in Figure 2.2-2. Since the margin between sexual debut and first marriage is on the average 6 years for men, there is an increased risk of multiple sexual partners and sexual encounters, STIs and HIV/AIDS, and unwanted pregnancies. This is not too different for the women, as the wider the interval between sexual debut and first birth, the indication exists for contraception, multiple sexual partners and sexual encounters and unwanted pregnancies.



Shaded area shows interval between successive events

Figure 2.2-2: Median age at reproductive events
Data Source: (NDHS, 2013)

2.3. Previous studies on the sexual behaviour markers

Numerous studies have been conducted on age at first sex, age at first birth, and age at first marriage; and few studies have also considered the intervals between them. However it is almost rare to find studies that have shown much interest in the instantaneous probability of this event occuring given a specific age i.e. distribution of the baseline hazard, since it is easy eschew this assumtpion once the Cox proportional hazard model is used for analyzing such data. Cox proportional hazard model is a semi-parametric statistical method that assumes no probability distribution for the hazard of debut; and hence, has found a wide range of applicability. Most of the studies on these sexual behaviour markers have employed semi-parametric methods for analyses.

On the other hand, all the sexual behaviour markers are time-to-event variables which can be fitted using models that evaluate survival time. Dätwyler & Stucki (2011) have shown that survival time follows a distribution and several authors have shown that time-to-event variables follow five main statistical distributions namely: exponential, log-normal, Weibull, generalized gamma, and log-logistic distributions (Shayan, et al., 2014; Atuhaire, 2011; Dätwyler & Stucki, 2011; Rodríguez, 2010). Provided the assumption of a probability distribution for survival data is valid, parameters estimation is more precise and inferences based on such assumptions will be more reliable (Amusan & Mohd, 2014; Shayan, et al., 2014; Collet, 1994). Shayan et al. (2014) also pointed out that with the parametric models, one can plot a description of the hazard rate of live birth over time and we can also measure the direct effect of the explanatory variables on the survival time instead of hazard.

2.3.1. Modelling sexual behaviour markers

It was suggested by Zaba *et. al.* (2002) that the most appropriate method for estimating median age at first sex from censored observation is survival analysis, when no assumption is made about the probability distribution for the variable while maximum likelihood estmation has been used when assumption of probability distribution is valid. They further indicated that among several measures that could be used to summarize the distribution of age at first sex, the median and interquartile range are possibly the best simple measure that capture location and shape of the curve. This also applies quite logically to age at first marriage and age at first birth, the three, being time-to-event variables with possibility of censoring.

2.4. Survival Analysis and Survival Functions

Survival analysis is a branch of statistical methods that deals with analyzing data that correspond to a well-defined time origin to the occurece of an event of interest. This branch of statistics is quite unique and not amenable to standard statistical procedures basically because of the nature of the data which ususally have the following special features:

1. Not symmetrically distributed: survival times obtained from similar individuals with similar characteristics tend to be positively skewed and as a consequence, statistical methods that assume a normal distribution, such as the ordinary least squares and methods for comparing means, cannot be used to analyze such data.
2. Censoring: The survival time of an individual is said to be censored if the event of interest had not been observed for that individual, which could be as a result of the individual dropping out of a study, or the event occurred before the study commenced of after the study had ended in a logitudinal study. Most standard statistical procedures that may account for skewness may not account for censoring.

Survival analysis finds applicability in a very wide range of human endeavour including epidemiology and biology (surival times), engineering (failure time), medicine (treatment effect or drug efficacy) (Weinken, 2007), quality control (lifetime of a component), credit risk modelling in finance (defalt time of a firm) (Fan & Jiang, 2008), as well as in demographic studies including fertility (age at marriage, duration of marriage, interval between successive births), mortality (length of life) and migration (duration of stay on a job or a city) (Rodriguez, 2010).

2.4.1. Statistical functions of the survival data

Let T be a non-negative random variable representing the waiting time unto the occurrence of an event. Without loss of generality, we shall refer to the waiting time as survival time and the event of occurrence shall be referred to as "debut". Suppose T has a probability density function (pdf) $f(t)$ and cumulative distribution function (cdf) $F(t) = \Pr\{T<t\}$, provided that debut has occurred by duration of t. $F(t)$ is the probability that debut occurs any time before or at time t.

2.4.1.1. *The survival function*

The probability that debut has occurred after the time t, then we have the probability that debut has not occurred up to time t as the survival function given as:

$$S(t) = \Pr\{T \geq t\} = 1 - F(t) = \int_t^\infty f(x)dx$$ **Equation 2.4-1**

2.4.1.2. *Hazard function*

The hazard function is defined as the probability that an individual debuts as time t provided that s(he) has survived until that time. It is the instantaneous debut rate for an individual having survived to time t, written as:

$$h(t) = \lim_{dt \to 0} \frac{\Pr\{t \leq T \leq t + dt | T \geq t\}}{dt}$$ **Equation 2.4-2**

The numerator of the equation is the conditional probability that debut occurs in the interval $[t, t+dt)$ given that it had not occurred before while the denominator is the width of the interval. Taking limits as $dt \to 0$, we have the instantaneous rate of occurrence.

Writing the hazard function in terms of a conditional probability, we obtain:

$$h(t) = \frac{\Pr\{t \leq T < t + dt\}}{\Pr\{T \geq t\}},$$

and from the definitions of the cumulative probability distribution function of T, $F(t)$, and survival function, $S(t)$, we have that

$$h(t) = \frac{F(t + dt) - F(t)}{S(t)}$$

Now, taking limits on the numerator, we have the derivative of the probability density function, $f(t)$, hence, we have the relationship:

$$h(t) = \frac{f(t)}{S(t)}$$ **Equation 2.4-3**

This goes on to shows that "the rate of occurrence of the event at duration t equals the density of events at t divided by the probability of surviving to that duration without experiencing the event" (Rodríguez, 2010). The equation implies that $f(t) = S(t)h(t)$. This relationship forms the basis for evaluating survival data, as one can easily obtain the other functions given at least one of the functions. Also, the cumulative hazard function is given by:

$$H(t) = \int_0^t h(u)du = -\log\{S(t)\}$$ **Equation 2.4-4**

The survival function and hazard function are estimated from the observed survival times. There are both parametric and non-parametric methods of estimating the survival functions.

Commonly used non-parametric methods are the life-table approach and the Kaplan-Meier method but this current study focuses on the parametric methods, which are discussed forthwith.

2.4.2. Parametric survival distributions

Parametric models have been suggested as standard tools for the analysis of survival times because they are flexible for obtaining relative effects, absolute measures of riks, and for extrapolation (Crowther & Lambert, 2013). Survival data are assumed to follow a number of distributions including exponential, log-normal, Weibull, Inverse Gaussian, Log-Logistic, Gompertz-Makeham, Gamma, Generalized Gamma, Generalized F and the Coale-McNeil Model for fitting age at first marriage (Rodríguez, 2010); and four of these distributions - exponential, Weibull, lognormal, and log-logistic distributions, which have been identified as very relevant to modelling age variables (Atuhaire, 2011; Weinken, 2007) are discussed here.

2.4.2.1. Exponential distribution

The exponential distribution has been described as the simplest of the parametric survival models that assumes constant risk of death over time (Weinken, 2007). The probability density function of the exponential distribution is of the form:

$$f(t|\lambda) = \lambda e^{-\lambda t} \qquad \textbf{Equation 2.4-5}$$

For $0 \leq t \leq \infty$

It assumes that the hazard of death is constant over time hence, the hazard function is of the form:

$$h(t) = \lambda$$

and the survivor function is of the form:

$$S(t) = e^{-\lambda t}$$

Estimate of the pth percentile is of the form:

$$\hat{t}(p) = \frac{1}{\hat{\lambda}} \log\left(\frac{100}{100 - p}\right)$$

A major limitation of the exponential function is the assumption of constant hazard rate and hence, does not account for previous knowledge of survivorship if a person had survived up to a given time, say t_0.

2.4.2.2. *Weibull distribution*

The Weibull distribution, introduced by Waloddi Weibull in 1939 (Weinken, 2007), is a generalization of the exponential distribution with probability density function of the form:

$$f(t|\lambda, \gamma) = \lambda \gamma t^{\gamma-1} e^{-\lambda t^{\gamma}} \quad \text{Equation 2.4-6}$$

Where λ is the scale parameter and γ is the shape parameter.

For $0 \leq t \leq \infty$

The hazard function and survivor functions are of the form:

$$h(t) = \lambda \gamma t^{\gamma-1}$$

$$S(t) = e^{-\lambda t^{\gamma}}$$

Estimate of the pth percentile is of the form:

$$\hat{t}(p) = \left[\frac{1}{\hat{\lambda}} \log\left(\frac{100}{100 - p}\right)\right]^{1/\hat{\gamma}}$$

The shape of the Weibull distribution depends on the value of γ, called the shape parameter while λ is the scale parameter. In particular case where $\gamma = 1$, the hazard function takes a constant value and the survival times have an exponential distribution. The Weibull hazard function is monotonic and, therefore, cannot account for situations in which the hazard function changes. Assumptions of a Weibull proportional hazards model includes a monotonically increasing or decreasing baseline hazard rate (Crowther & Lambert, 2013).

2.4.2.3. *Lognormal distribution*

A random variable T is said to have a lognormal distribution with scale parameters μ, shape parameter σ, and location parameter θ, if $\log T$ has a normal distribution with mean μ and variance σ^2. The pdf of a lognormal distribution is given by (NIST/SEMATECH e-Handbook of Statistical Methods, 2013):

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} t^{-1} \exp\left\{-\frac{(\log(t-\theta)-\mu)^2}{2\sigma^2}\right\} \qquad \text{Equation 2.4-7}$$

For $0 \leq \theta \leq t \leq \infty;\ \mu, \sigma > 0$.

The hazard function and survivor functions, respectively, are given by:

$$h(t) = \frac{\left(\frac{1}{t\sigma}\right)\phi\left(\frac{\log t}{\sigma}\right)}{\Phi\left(\frac{-\log t}{\sigma}\right)}$$

and

$$S(t) = 1 - \Phi\left(\frac{-\log t}{\sigma}\right)$$

Where ϕ and Φ are the pdf and cdf of the normal distribution respectively.

Estimate of the p^{th} percentile of the lognormal distribution is obtained from (Krishnamoorthy, et al., 2014):

$$t(p) = \exp(\hat{\mu} + z_p\hat{\sigma})$$

Where z_p is the p^{th} percentile of the standard normal distribution.

2.4.2.4. *Log-logistic distribution*

A random variable T is said to have a log-logistic distribution with parameters θ and κ if $\log T$ has a logistic distribution. The pdf of the log-logistic distribution is symmetric distribution and similar to that of the normal distribution. The probability density function of the log-logistic distribution is the function:

$$f(t|\theta,\kappa) = \frac{e^{-\theta}\kappa t^{\kappa-1}}{(1+e^{-\theta}t^{\kappa})^2}$$ Equation 2.4-8

With parameters θ and κ; $t \geq 0$

The hazard function and survivor function are given by:

$$h(t) = \frac{e^{\theta}\kappa t^{\kappa-1}}{1+e^{-\theta}t^{\kappa}}$$

$$S(t) = \left(1 + e^{\theta}t^{\kappa}\right)^{-1}$$

Estimate of the *pth* percentile is of the form:

$$\hat{t}(p) = \left(\frac{100-p}{pe^{\hat{\theta}}}\right)^{1/\hat{\kappa}}$$

Collet (1994) was quick to point out that the hazard and survivor functions of the gamma and lognormal distributions can only be expressed in terms of integrals which gives them considerable disadvantage and hence, unsuitable for general use. On the other hand, in view of the similarity between the normal and logistic distributions, the lognormal distribution will tend to be similar to the log-logistic distribution. Similarly, the gamma and Weibull distributions will generally lead to similar results. The works of Leemis and McQuestion (2008), Song (2005), Abdelkader & Al-Marzouq (2010) and Cook (2014) provide very comprehensive insight into the reparameterization and relationships among probability distributions. Hence, the exponential and Weibull distributions have found more applicability among statisticians than the other distributions, especially because of the ease of the interpretability of their hazard functions (Weinken, 2007).

2.5. Fitting the Survival Distribution

In order to investigate that a given data fits a probability distribution, two methods have been proposed which are graphical methods and analytical methods (Al-Fawzan, 2000).

2.5.1. Graphic methods of estimating distribution estimates

Graphical methods include probability plotting, hazard plotting, and cumulative hazard plots.

2.5.1.1. Probability plotting technique

A probability plot has been described as a graphical method of investigating if sample data conform to a distribution (Simeon & Khalid, 2014). Given the probability density function *f(t)* of a given censored data with cumulative distribution function *F(t)*. The aim of the probability plotting method is to fit the cumulative function into a straight line graph using double-logarithmic transformation. For instance, the cumulative distribution function of the two-parameter Weibull distribution if given as:

$$F(t) = 1 - e^{-\left(\frac{t}{\alpha}\right)^{\beta}}$$ Equation 2.5-1

Equation 2.5-1 can be re-written as:

$$\frac{1}{1 - F(t)} = e^{\left(\frac{t}{\alpha}\right)^{\beta}}$$ Equation 2.5-2

Taking double logarithmic transformation of the Equation 2.5-2 we have the equation of a straight line:

$$\ln \ln \left[\frac{1}{1 - F(t)}\right] = \beta \ln \alpha - \beta \ln t$$ **Equation 2.5-3**

To plot $F(t)$ versus t, we apply the following procedure:

1. Rank failure times in ascending order
2. Estimate $F(t_i)$ of the i^{th} failure
3. Plot $F(t_i)$ versus t in the Weibull probability paper.

To estimate $F(t_i)$ in (2 and 3) above, we may use one of the following methods presented in Table 2.5-1.

Table 2.5-1: Methods of estimating cumulative density function of a Weibull distribution using probability plotting

Method	$F(t_i)$
Mean rank	$\frac{i}{n+i}$
Median rank	$\frac{i-0.3}{n+0.4}$
Symmetrical CDF	$\frac{i-0.5}{n}$

Source: (Al-Fawzan, 2000).

2.5.1.2. *Hazard plotting technique*

This estimation procedure for the Weibull parameters is done by plotting hazard function $H(t)$ against failure times on a hazard paper or a simple log-log paper (Al-Fawzan, 2000). The hazard function for a two-parameter Weibull distribution is given by:

$$h(t) = \frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1}$$ **Equation 2.5-4**

Where α is the scale parameter and β is the shape parameter and t is failure time; and they are all non-zero. The cumulative hazard function is given by:

$$H(t) = \int_0^t h(x)dx = \int_0^t \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} dx = \left(\frac{t}{\alpha}\right)^{\beta}$$ **Equation 2.5-5**

Equation 2.5-5 is transformed by taking logarithms as follows

$$\ln H(t) = \beta\{\ln t - \ln \alpha\}$$

$$\ln t = \frac{1}{\beta}\ln H(t) + \ln \alpha$$ **Equation 2.5-6**

From Equation 2.5-6 we can plot ln$H(t)$ versus ln$t$ using the following procedure:

1. Rank the failure times in ascending order
2. For each failure, calculate $\Delta H_i = \frac{1}{(n+1)-1}$
3. For each failure, calculate $H_i = \Delta H_1 + \Delta H_2 + \cdots + \Delta H_i$
4. Plot $\ln H$ versus $\ln t$.
5. Fit a straight line.

Upon completing the plotting, the estimated parameters will be as follows:

$$\beta = \frac{t}{y} = \frac{1}{slope}$$

At $H = 1$, $\alpha = t$

2.5.1.3. *Log cumulative hazard plotting*

The work of Collet (1994, p. 200) is discussed here. Given the survival function of a Weibull distribution with scale parameter, β, and shape parameter, α, we have

$$S(t) = e^{-\beta t^{\alpha}}$$ Equation 2.5-7

Taking double logarithmic transformation of Equation 2.5-7, we have

$$\log\{-\log S(t)\} = \log \beta + \alpha \log t$$ Equation 2.5-8

Equation 2.5-8 is linear equation of a straight line of log{-log $S(t)$} against log t, with intercept log β and slope α. Recall that, the survival function and the cumulative hazard function are related thus: $\log S(t) = -H(t)$, hence, Equation 2.5-8 can be re-written as:

$$\log H(t) = \log \beta + \alpha \log t \qquad \text{Equation 2.5-9}$$

Which is the log cumulative hazard function equation. Plotting log $H(t)$ against log (t) could be used to estimate the Weibull parameters. The estimates could be read off the log cumulative hazard plot or fit the model to an ordinary least squares linear regression equation of a straight line.

2.5.2. Analytical methods

There is much chance for error when reading off parameter estimates from graphs using the graphical methods for parameters estimation, hence, the analytical methods are preferable (Al-Fawzan, 2000; Collet, 1994). Analytical methods for estimating Weibull parameters include method of moments (MOM), method of ordinary least squares (OLS), method of maximum likelihood estimation (MLE) (Crowther & Lambert, 2013; Bartkute & Sakalauskas, 2008; Lei, 2008; Al-Fawzan, 2000) and the iterative generalized least squares (IGLS) implemented in MLwiN and penalized quasi likelihood (PQL) (Stewart, 2010). The method of maximum likelihood estimation is the most common method used for estimating censored observations (Lei, 2008; Al-Fawzan, 2000) and has the advantage of handling unbalanced data and produces positive variance estimates always (Albright & Marinova, 2010) and "is asymptotically the most efficient method" (Lei, 2008). Although Lei (2008), who compared the estimation of the three aforementioned methods of parameter estimation for three-parameter Weibull distribution for the Chinese pine, concluded that the MOM is the best method of estimating the Weibull parameters based on the method with lower means square error (MSE), which was corroborated by the simulation study of Al-Fawzan (2000); however, Al-Fawzan, went further to show that MLE has lower computing time compared to MOM. Hence, MLE is described in the next section.

2.5.2.1. *Method of maximum likelihood estimation (MLE) for estimating Weibull parameters*

We shall refer to the works of Crowther & Lambert (2013) in discussing this section. The log-likelihood contribution of the ith respondent, allowing for right censoring and delayed entry (left truncation), using the relation $f(t) = h(t)S(t)$, can be written as

$$l_i = \log\left[h(t_i)^{d_i}\left(\frac{S(t_i)}{S(t_{0i})}\right)^{1-d_i}\right]$$

$$= d_i \log\{h(t_i)\} + (1-d_i)\log\{S(t_i)\} - (1-d_i)\log\{S(t_{0i})\}$$ **Equation 2.5-10**

By substituting Equation 2.4-4, this becomes

$$l_i = d_i \log\{h(t_i) - \int_{t_{0i}}^{t_i} h(u)du$$ **Equation 2.5-11**

Hence, from Equation 2.5-11, the task is to maximize the likelihood function given the hazard function. The maximization is done using numerical quadrature to evaluate the cumulative hazard (Crowther & Lambert, 2013).

2.6. Some gaps observed in literature

There have been several attempts by several authors to describe age at first sex, age at first marriage and age at first birth as well as their intervals using different statistical techniques but there is paucity of information on the distribution of the baseline hazards of these variables. Specifically, none of these studies provided information on their respective instantaneous probability of occurrence given a person's age or time interval. This is so because the variables are usually fitted into the Cox proportional hazard models (Oyedokun & Odumegwu, 2013; Adebowale, et al., 2012;

Chandrasekhar, 2010; Lofstedt, et al., 2005), which does not assume a probability distribution for baseline hazard. Cox proportional hazard models assume that time is measured on a continuous scale and based on proportional hazards (Gayawan & Adebayo, 2013), therefore, the Cox PH models have been identified as the most popular survival model (Crowther, et al., 2014). Amusan and Mohd (2014) used four parametric distributions, including Weibull and log-logistic, to fit first marriage and first birth interval with an emphasis on the inverse Gaussian providing the best fit. However, their work did not provide focus on the hazard function of the interval and Leemis and McQuestion (2008, p. 47) have showed that the Weibull distribution is amenable to the inverse Gaussian through indirect parameterization.

Also these variables have been fitted separately into proportional hazard models without accounting for community-level influences on the variables and where the community-level influences were accounted for, the time-to-event properties were expunged (Olalekan, 2008; Ngalinda, 1998). Subaiya & Johnson (2008) assumed a normal distribution for the length of time between age-at-first-sex and age-at-first-marriage by fitting an ordinary least square (OLS) linear regression but OLS estimation does not account for the multi-level structure in data collected and produces less precise estimates of the standard error (Angeles & Mroz, 2001). Angeles & Mroz (2001) also noted that described the MLE procedure as the most unbiased estimator. Furthermore, inferences based on assumption of a particular probability distribution, even for skewed variables, are more precise if the assumption is valid (Collet, 1994).

When the survival data is assumed to be influenced by covariates, then parametric and nonparametric or semi-paramteric methods have been proposed to model such data. "An important problem in survival analysis is how to model well the conditional hazard rate of failure times given certain covariates, because it involves frequently asked questions about whether or not certain independent variables are correlated with the survival or failure times" (Fan & Jiang, 2008). Rodríguez (2010) proposed four models which included parametric families, accelerated life model, proportional hazards model, and proportional odds. Hogan, Sun, & Conrwell (2000) used the Weibull hazards regression model to estimate the timing of first sexual intercourse and timing of first pregnancy.

Amusan & Mohd (2014) fitted paramtric models to first birth interval among women of reproductive age in Nigeria using Inverse Gaussian, Log-Logistic, Weibull and Burr Type XII models. Non-parametric method was also considered for comparison and Inverse Guassian was favoured in the end, using the lowest Akaike Information Criteria for model selection and fitted with Maximum Likelihood estimation method. However, the gap in their work is that there is no information with regards to the values of the parameter estimates using the assumed distributions and neither did they compare the compare the normal distribution, which would further confirm or question the work of Subaiya and Johnson (2008) for fitting the interval between two skewed distributions.

2.7. Types of data used for sexual behaviour markers

Nationally representative data sources have always been used to study age at first sex, age at first marriage, age at first birth, as well as their intervals. For instance, several authors have used data from demographic and health surveys for different countries including Cameroun (Subaiya & Johnson, 2008), Nigeria (Gayawan & Adebayo, 2013; Aremu, 2013; Adebowale, et al., 2012; Olalekan, 2008), Tanzania (Ngalinda, 1998), six African countries (Zaba, et al., 2004), Malawi (Ueyama & Yamauchi, 2009); and other kinds of nationally representative data such as National Family Health Survey in India (Chandrasekhar, 2010), National Survey of Family Growth in the United States (Finer & Philbin, 2014), and India Human Developemnt Survey (MacQuarrie, 2014).

independent variables are correlated with the survival or failure times" (Fan & Jiang, 2008). Rodriguez (2010) proposed four models which included parametric families, accelerated life model, proportional hazards model, and proportional odds. Hogan, Sun, & Conrwell (2000) used the Weibull hazards regression model to estimate the timing of first sexual intercourse and timing of first pregnancy.

Amusan & Mohd (2014) fitted paramtric models to first birth interval among women of reproductive age in Nigeria using Inverse Gaussian, Log-Logistic, Weibull and Burr Type XII models. Non-parametric method was also considered for comparison and Inverse Guassian was favoured in the end, using the lowest Akaike Information Criteria for model selection and fitted with Maximum Likelihood estimation method. However, the gap in their work is that there is no information with regards to the values of the parameter estimates using the assumed distributions and neither did they compare the compare the normal distribution, which would further confirm or question the work of Subaiya and Johnson (2008) for fitting the interval between two skewed distributions.

2.7. Types of data used for sexual behaviour markers

Nationally representative data sources have always been used to study age at first sex, age at first marriage, age at first birth, as well as their intervals. For instance, several authors have used data from demographic and health surveys for different countries including Cameroun (Subaiya & Johnson, 2008), Nigeria (Gayawan & Adebayo, 2013; Aremu, 2013; Adebowale, et al., 2012; Olalekan, 2008), Tanzania (Ngalinda, 1998), six African countries (Zaba, et al., 2004), Malawi (Ueyama & Yamauchi, 2009); and other kinds of nationally representative data such as National Family Health Survey in India (Chandrasekhar, 2010), National Survey of Family Growth in the United States (Finer & Philbin, 2014), and India Human Developemnt Survey (MacQuarrie, 2014).

2.8. Summary

Literature shows that extensive work has been done on the reproductive health/fertility markers and several attempts have been made to model the variables and their probabilities using both parametric, semi-parametric and non-parametric approaches for analysis of time-to-event data. It was also shown through literature that parametric estimation of distribution estimates are favoured over other methods because they are judged as being more precise. This work, hence, focuses on the distribution of the hazard function of the RH events, age at first sex and age at first birth as well as their interval, using MLE to estimate Weibull parameters for the distribution of the baseline hazard.

Chapter Three

Methodology

3.0. Introduction

Nigeria is located in Western Africa on the Gulf of Guinea and has a total area of 923,768 kilometer squared making it the world's 32nd largest country (after Tanzania). Nigeria is the most populous country in Africa and represents 2.35 percent of the world's total population (Trading Economics, 2014). The United Nations estimates Nigeria's population in 2004 was at 131,530,000, with the population distributed as 48.3% urban and 51.7% rural and population density at 139 people per sq km. Fulani/Hausa, Yoruba, and Igbo are the largest ethnic groups and account for 68% of Nigeria's population (Olalekan, 2008), Edo, Ijaw, Kanuri, Ibibio, Ebira, Nupe, Tiv and other minorities comprise 32%. Current population estimate is placed at 170 million with total fertility rate 9 (TFR) of 5.6 births per woman compared to overall of 5.2 births per woman in Africa (Amusan & Mohd, 2014).

3.1. Research design

Cross-sectional and population-based study using data from the 2013 Nigeria Demographic and Health Survey (NDHS).

3.2. Data collection techniques

Secondary data from Nigeria Demographic and Health Survey was used for the study. Three questionnaires were used in NDHS 2013 – household questionnaire, women's questionnaire, and men's questionnaire, and the questionnaires were administered in four languages, which were English, Hausa, Igbo, and Yoruba. The household questionnaire listed and collected information about usual members of and visitors to the selected households, as well as and basic socio-demographic information about each person. The Household Questionnaire also collected information about the household's dwelling unit such as basic amenities, types of materials used for building, roofing, flooring, ownership of various durable goods, ownership of livestock and farmland, and ownership of long-lasting insecticide nets (LLINs).

Women's and Men's Questionnaires collected information from those aged 15-49 years who had been identified by the Household Questionnaires. Information garnered included:

- Background characteristics (age, religion, education, literacy, media exposure, etc.)
- Reproductive history and childhood mortality
- Knowledge, source, and use of family planning methods
- Fertility preferences
- Antenatal, delivery, and postnatal care
- Breastfeeding and infant feeding practices
- Child immunization and childhood illnesses
- Marriage and sexual activity
- Women's work and husbands' background characteristics
- Malaria prevention and treatment
- Women's decision-making
- Awareness of AIDS and other sexually transmitted infections
- Maternal mortality
- Domestic violence

However, the Men's Questionnaire did not include such sections as reproductive history or questions on maternal and child health or nutrition.

3.3. Sampling techniques

The method of data collection is as described in Nigeria Demographic and Health Survey 2013 (NPC and ICF International, 2014). The 2013 NDHS sample was selected using a stratified three-stage cluster design consisting of 904 clusters, which are the primary sampling units (PSU) defined on the basis of the 2006 EA census frame (NPC [Nigeria] and ICF International, 2014). Hence, it allowed for indicators to be calculated for each of the six geo-political zones, 36 states and the federal capital territory. The sampling was nationally representative and covered the entire nation, unlike what obtained in all NDHS before 2008. The survey sampling frame of the list of enumeration areas (EAs) prepared for the 2006 Population Census was provided by the National Population Commission.

Nigeria is administratively divided into states, which are subdivided by local government areas (LGAs) and then communities. For the first stage, the nation was stratified by separating each state into urban and rural areas, which provided 893 localities selected with probability proportional to size and with independent selection in each sampling stratum. One EA was then randomly selected from each identified localities, with equal probability of selection, which meant that more than one EA was selected in few larger localities. This also provided 904 EAs.

Household listing was then carried out for the selected EAs using location maps, detailed sketch map, and recording the household listings. The resulting household listing served as the sampling frame for the selection of the households in the third stage.

In the third and final stage, a fixed number of 45 households were selected in every urban and rural cluster using equal probability systemic sampling method.

3.4. Study population

The study population was women of reproductive age between 15-49 years, who had ever had sex. Studies on age at first sex was done for the entire study population while studies on age at first birth and the interval between first sex and first birth was done for women that ever had a birth only. Women that ever had sex but never had sex were considered left censored at their current age.

3.5. Data sources

Data for the study was obtained from NDHS 2013, which is available free from the website of Nigeria Population Commission. Data was collected from 38,522 households, 38,948 individual women, 17,359 individual men, and 8,658 couples (NPC [Nigeria] and ICF International, 2014). For the purpose of this study, only data from the Woman's Questionnaire relating to sexual behavior, marriage and births was used. The following variables will be used for the analyses:

3.5.1. Identification variables:

Variable Name	Variable Label
V000	Country code and phase
V001	Cluster number
V002	Household number
V003	Respondent's line number
V004	Ultimate area unit
V005	Women's individual sample weight (6 decimals)
V010	Respondent's year of birth
V011	Date of birth (CMC)
V012	Respondent's current age
V013	Age in 5-year groups

3.5.2. Sampling variables

Variable Name	Variable Label
V021	Primary sampling unit
V022	Sample strata for sampling errors
V023	Stratification used in sample design

3.5.3. Background information about the women

Variable Name	Variable Label
V101	Region
V102	Type of place of residence
V106	Highest educational level
V107	Highest year of education
V130	Religion
V131	Ethnicity
V133	Education in single years
V149	Educational attainment
V150	Relationship to household head
V151	Sex of household head
V152	Age of household head
V153	Household has: telephone (land-line)
V155	Literacy
V190	Wealth index
V191	Wealth index factor score (5 decimals)

3.5.4. Sex, marriage and birth variables

Variable Name	Variable Label
V211	Date of first birth (CMC)
V212	Age of respondent at 1st birth
V221	Marriage to first birth interval (months)
V302A	Ever used anything or tried to delay or avoid getting pregnant
V511	Age at first cohabitation
V512	Years since first cohabitation
V513	Cohabitation duration (grouped)
V525	Age at first sex
V527	Time since last sex
V528	Time since last sex (in days)
V529	Time since last sex (in months)
V530	Flag for V529
V531	Age at first sex (imputed)

Variable Name	Variable Label
V532	Flag for V531
V535	Ever been married or in union
V536	Recent sexual activity
V537	Months of abstinence
V538	How previous marriage or union ended

3.6. Definitions of key terms, concepts and variables

- **Age-at-first-sex (AAFX):** Age at which an individual initiates sexual intercourse [Age at first sex – V525]. Collected on NDHS 2013 as age-at-first-sexual Intercourse.
- **Age-at-first-birth (AAFB):** Age at which childbearing commences. This information is collected on NDHS 2013 as Age a First Birth. [Age of respondent at 1st birth – V212].
- **First sex-first birth interval (FXFB):** Number of years between onset of sexual intercourse and age at which childbearing commences. It is calculated as:

$$FXFB = AAFX - AAFB$$

However, this variable can be very subjective because use of contraceptives and irregular menstrual cycle may delay birth. Also, time of first sex may not correspond with timing of first birth. Hence, we would have to scenarios

1. First birth before marriage, and
2. First birth after marriage

The second scenario is well taken care of, as a married woman is naturally exposed to sexual intercourse and marriage-first birth interval would take care of this. For the first scenario, it would only imply that we are trying to measure fertility rate of the women, and we only want to focus on women of reproductive age (16-49 years) in the study.

3.7. Data analysis and interpretation

Age-at-first-sex and age-at-first-birth are time-to-event variables and of essence, follow a distribution (Dätwyler & Stucki, 2011). Specifically, as a person grows older, the susceptibility to having sex after a certain age, say 15 years, becomes more imminent. Using the age 15 years as cut-off just as it has been assumed by Zaba *et al* (2004) and Olalekan (2008). Exploratory data analysis and inferential statistical methods were used in the study. Three methods were proposed by Lei (2008) and Al-Fawzan (2000) to investigate the survival data distributions namely – maximum likelihood estimation (MLE), method of moments estimation (MOM), and least squares estimation (LSE) method; but MLE was used for this study.

3.8. Ethical considerations

Data obtained from NDHS 2013 is entirely de-identified and does not require any ethical consideration at this level, as all necessary ethical consideration for the data collection had been built into the data collection procedure (NPC [Nigeria] and ICF International, 2014). This also involved informed consent from the participants and confidentiality

3.9. Steps in Data Analysis

3.9.1. Description of Population

Starting with age at first sex, there are 38948 observations of which 46.67% (19,345) were missing because they were at their first union. 6 (0.02%) don't know and 95 (0.24%) were missing. Hence another variable was used in lieu which was age at first sex imputed.

The following were excluded and declared missing because they were (i) inconsistent, 2027 (5.2%), don't know, 6 (0.02%), and missing, 95, (0.24%). Hence, a total of 36820 cases were available for analysis. Comparing the values for the original values and imputed variables and the original variable is as shown on Table 3.9-1:

Table 3.9-1: Description of original and imputed age at first sex variables

Variable	Observations	Minimum	Maximum
Age at first sex	19502	0	45
Age at first sex imputed	36820	0	46

Note: 0 implies that the respondent had never had sex.

The variable ever had sex was further recoded into a dummy variable to show persons that had never had sex (0) and those that ever had sex (1). Hence, there were 31075 (84.40%) women that ever had sex in the data set and 5,745 (15.6%) that had never had sex. This work focused on the 31075 women that ever had sex.

3.9.2. Exploratory data analysis

Exploratory data analysis including frequencies, measures of central tendencies and dispersion, histograms and density probability plots were used to describe each of the quantitative outcome variables.

This section provides a description of the key variables for this study. As a preamble, some modifications were made to the data before the analysis commenced. For instance, Table 3.9-2 shows the distribution of missing values observed in the dataset. The respondent's current age had no missing values. Age at first sex was imputed using Multiple Imputation (MI) procedure from the Nigeria DHS 2013 data source hence had no missing values. There are 5705 (18.4%) missing values in Age of respondent at first birth which would correspond to the number of women who had never given birth.

The interval between first sex and first birth had 5706 (18.4%) missing values, which included the 5705 that had never had a birth and an additional respondent that had inconsistent records (i.e. negative interval). Dummy variables were created for each variable to filter them at the level of analysis. Analysis was carried by excluding missing values pairwise in order not to lose relevant information.

Table 3.9-2: Distribution of missing values in the women's data

Variable	Non-missing		Missing		Total	
	Number of women	Percent	Number of women	Percent	Number of women	Percent
Respondent's current age (years)	31075	100.0	0	0.0	31075	100.0
Age at first birth (years)	25370	81.6	5705	18.4	31075	100.0
Age at first sex (years)	31075	100.0	0	0.0	31075	100.0
Age at first sex and first birth interval (years)	25369	81.6	5706	18.4	31075	100.0

3.9.3. Fitting Cumulative Hazard Function of the Weibull Distribution

This section describes how the estimates of the Weibull distribution in describing the hazard function were obtained. First we explore the cumulative hazard plotting method and then the analytical method.

3.9.4. Checking the suitability of the Weibull distribution using the log cumulative hazard plot

The graphical method is used to evaluate the suitability of the hazard function to Weibull distribution. If the log cumulative hazard plot produces a straight line graph, then the data is suitable to be fitted to a Weibull distribution. First step was to obtain the Kaplan-Meier estimate for the survivorship function, $\hat{S}(t)$ and then substitute in the equation

$$\log\{-\log S(t)\} = \log\beta + \alpha\log t \qquad \textbf{Equation 3.9-1}$$

Where α is the shape parameter, β is the scale parameter, and t is the time of debut. If we can assume a Weibull distribution for the data then the survivorship function estimate, $\hat{S}(t)$, will be close to the Kaplan-Meier estimate and a plot of log{-log $S(t)$} against log t will be approximately a straight line. Recall that the expression –log $S(t)$ is the cumulative hazard function, $H(t)$. The parameters were estimated using ordinary least square linear regression and the percentiles were estimated using the using the Weibull estimates

$$\hat{t}(p) = \left[\frac{1}{\hat{\beta}}\log\frac{100}{100-p}\right]^{1/\hat{\alpha}}$$ Equation 3.9-2

Where $p = 50$ for the median.

3.9.5. Maximum likelihood estimation of Weibull parameters using stgenreg

Having established the suitability of the hazard function for a Weibull distribution, Stata stgenreg macro (Crowther & Lambert, 2013) is a General parametric survival model used to estimate the Weibull parameters. The following is from the release note of the macro

> "stgenreg fits parametric survival models using any user-defined [log] baseline hazard function and uses Gaussian quadrature to evaluate the cumulative hazard function and consequently the survival function allowing the estimation of a parametric survival model with almost any form. The [log] hazard function must be written in Mata code using colon operators. Each parameter defined in a [log] hazard function can include components, whereby each component can contain variables, user-defined functions of time, restricted cubic spline functions and/or fractional polynomial functions. Time-dependent effects can be included in any component as well fitting as relative survival models."

The stgenreg macro enables the user to fit general parametric models by specifying any baseline hazard function which can be written in a standard analytical form, and implemented through numerical integration of the user-defined hazard function (Crowther & Lambert, 2013). The command used for the procedure is provided below:

```
stset V531 /// Set V531 – age at first sex as survival time variable
stgenreg, loghazard([ln_lambda] :+ [ln_gamma] :+ (exp([ln_gamma])-1):*log(#t))
ln_lambda(NHX) /// estimate Weibull parameters
nlcom (theta:exp([ln_gamma]_b[_cons])) (beta:1/exp(-[ln_lambda]_b[_cons]))
///nonlinear combination of estimators
estat vce /// variance-covariance matrix
estat ic ///estimated Information criteria
predict haz_1, hazard ci ///predict the baseline hazard with condifence intervals
```

The estimators obtained were then used to obtain estimates of the hazard function and the survival function, which were plotted against times to in order to estimate the median and other percentiles. However, the median was also estimated analytically using Equation 3.9-2. The standard error of the estimated percentiles (p) were also estimated using

$$s.e.\{t(p)\} = \frac{\hat{t}(p)}{\hat{\beta}\hat{\alpha}^2}\{\hat{\alpha}^2 \operatorname{var}(\hat{\beta}) + \hat{\beta}^2(c_p - \log\hat{\beta})^2 \operatorname{var}(\hat{\alpha}) + 2\hat{\alpha}\hat{\beta}(c_p - \log\hat{\beta}) \operatorname{cov}(\hat{\alpha}\hat{\beta})\}^{1/2}$$

Where

$$c_p = \log\log\left(\frac{100}{100-p}\right)$$

The variances and covariance of $\hat{\alpha}$ and $\hat{\beta}$ are obtained from the variance-covariance matrix.

3.9.6. Multilevel modelling of interval between first sex and first birth

Multi-level modelling is adopted to investigate the individual- and community- level variables that determine the interval between age at sexual debut and first. Multi-level modelling is adopted because of the hierarchical nature of the data as described in Section 3.3. Fixed effect model was analyzed using the streg procedure of Stata while multilevel mixed effects parametric survival analysis [stmixed] Stata Procedure (Crowther, et al., 2014) was used for the random effects modeling. The release note for the procedure is described as follows:

> stmixed fits multilevel mixed effects parametric survival models using maximum likelihood assuming a Gaussian distribution of the random effects. Adaptive or non-adaptive Gauss-Hermite quadrature is used to evaluate the likelihood. Parametric survival models available include the exponential, Weibull, and Gompertz proportional hazards models, log-logistic, log-normal, and generalized gamma accelerated failure time models, and the Royston-Parmar flexible parametric survival model. The model currently supports only two level models. The random effects are included in the linear predictor on the log hazard scale for proportional hazards models, and the log time scale for accelerated failure time models. In particular, stmixed provides normally distributed frailties as an alternative to those implemented in streg (gamma and inverse normal), as well as allowing random slopes.

The proportional hazard mixed effects model is given by:

$$h_{ij}(t) = h_0(t)\exp\left(X_{ij}^T\beta + Z_i^T b_i\right) \qquad \text{Equation 3.9-3}$$

Where $i = 1,\ldots,m$ regions; $j = 1,\ldots,n$ observations; T_{ij} is the true survival time; X_{ij}^T is the design matrix for the fixed (Level 1) effect β; Z_i^T is the design matrix for the random

(Level 2) effects b_i; h_0 is the baseline hazard. The stata command for the multilevel model is:

```
xi: stmixed V531 V212 i.V013 V102 i.V130 i.V149 V228 V302A V502 || V021: , dist(weibull) nolog
```

The description of the variable names is provided in Sections 3.5.1 to 3.5.4 above. The interaction expansion command (xi:) was included to account for the factor variables prefixed with "i.". Other Stata commands for modeling are in the Appendices.

3.9.7. Level-1 and Level-2 Variables

As in previous studies on the determinants of age at first birth and age at first sex, some of these factors that influenced both variables were included in the study. Individual-level variables (fixed effect) included were women's highest educational attainment, religion, age at first sexual intercourse, age at first birth, woman's year of birth (birth cohort), contraceptive use, ever terminated a pregnancy, place of residence and ever been in a union while the primary sampling unit was included as the community-level variable (mixed effect). The PSU is expected to account for the geopolitical zone and state-level variation.

3.10. Software consideration

Descriptive and exploratory data analysis were done using IBM SPSS Statistics Version 22 (IBM Corporation, 2013), while the model fitting and multi-level analyses were done using Stata 12 (StataCorp LP, 2012). Kaplan Meier estimation was done with the SPSS KM procedure (see syntax in the appendices). Microsoft® Office Excel® 2013 (Microsoft Corporation, 2012) was used to evaluate the percentiles and produce some of the charts. All SPSS syntaxes and Stata commands used for this research are included in the appendices to encourage reproducible research.

Chapter Four

Results

4.1. Introduction

Exploratory data analysis including frequencies, measures of central tendencies and dispersion, histograms and density probability plots were used to describe each of the quantitative outcome variables. This section also provides results obtained from model fitting into the probability distributions using method of Maximum Likelihood Estimate (MLE). Single-level Weibull PH model was fitted and compared with single-level Cox PH. The single level models do no account for the fact that the data is hierarchically structured, consisting of three levels, households nested within clusters and nested within states. Finally, results obtained from multi-level model fitting of interval between age at first sex and age at first birth using parametric proportional hazard models was presented.

4.1.1. Socio-demographic characteristics

This section presents the socio-demographic characteristic of the women in terms of age, location, marital status, region, and fertility.

4.1.1.1. Age

Table 4.1-1 shows the description of the current ages of the women. Mean age was 30.6 (SD=9.1) years, which is slightly higher than the median age of 30 years that is same as the modal age. The minimum age was 15 years and the maximum age was 49 years. A quarter of the women were below 23 years and a quarter were above 38 years.

Table 4.1-1: Description of current ages of the women

Statistics		Unweighted
Mean		30.61
Median		30.00
Mode		30.00
Std. Deviation		9.10
Minimum		15.00
Maximum		49.00
Percentiles	25th	23.00
	50th	30.00
	75th	38.00

4.1.1.2. Geographic distribution and religion

Table 4.1-2 shows the demographic distribution of the women. More than a quarter (27%) were from the Northwest part of Nigeria, 18% from North East, 16% from South South, 15% each from South West and North Central respectively while 10% were from South East. More than two-fifth (62%) lived in rural areas while 38% lived in urban areas. Half of the women were Muslims, almost two-fifth (39%) were other Christians, one-tenth (19%) were Catholic and less than one percent were Traditionalists and a neglible percentage belonged to other religions.

Table 4.1-1: Description of current ages of the women

Statistics		Unweighted
Mean		30.61
Median		30.00
Mode		30.00
Std. Deviation		9.10
Minimum		15.00
Maximum		49.00
Percentiles	25th	23.00
	50th	30.00
	75th	38.00

4.1.1.2. Geographic distribution and religion

Table 4.1-2 shows the demographic distribution of the women. More than a quarter (27%) were from the Northwest part of Nigeria, 18% from North East, 16% from South South, 15% each from South West and North Central respectively while 10% were from South East. More than two-fifth (62%) lived in rural areas while 38% lived in urban areas. Half of the women were Muslims, almost two-fifth (39%) were other Christians, one-tenth (19%) were Catholic and less than one percent were Traditionalists and a neglible percentage belonged to other religions.

Table 4.1-2: Distribution demographic characteristics of the women

Demographic Characteristics	Number of Women	Percentage
Region		
North Central	4609	14.8
North East	5505	17.7
North West	8322	26.8
South East	3215	10.3
South South	4803	15.5
South West	4621	14.9
Type of place of residence		
Urban	11701	37.7
Rural	19374	62.3
Religion		
Catholic	3048	9.9
Other Christian	12091	39.1
Islam	15509	50.1
Traditionalist	280	0.9
Other	11	0.0

4.1.1.3. Education

With regards to educational attainment, two-fifth (40%) had no formal education, 18% had completed secondary school, 14% had completed primary education, and 10% had higher education (see Table 4.1-3). The women were further required to provide information about their ability to read and write; half of the women (50%) cannot read at all, more than two-fifth (43%) could read a whole sentence and 7% were able to read only parts of a sentence while less than 1% each had no card with required language or blind/visually impaired. However, 151 (0.5%) of the women had missing records on literacy.

Table 4.1-3: Distribution of education and literacy level of the women

Demographic Characteristics	Number of Women	Percentage
Educational attainment		
No education	12370	39.8
Incomplete primary	1700	5.5
Complete primary	4271	13.7
Incomplete secondary	3955	12.7
Complete secondary	5676	18.3
Higher	3103	10.0
Literacy		
Cannot read at all	15306	49.5
Able to read only parts of sentence	2174	7.0
Able to read whole sentence	13330	43.1
No card with required language	96	0.3
Blind/visually impaired	18	0.1

4.1.1.4. *Number of years schooling*

Of those that had ever attended school, average period of schooling was 10 (SD=4) years. Those that had incomplete primary education spent an average of 4 (SD=1) years schooling and those that had incomplete secondary education had spent average of 9 (SD=1) years schooling, as shown on Table 4.1-4.

Table 4.1-4: Number of years in schooling among women in Nigeria

Educational Attainment	Number of Women	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower	Upper		
Incomplete primary	1685	3.54	1.17	3.49	3.60	1	5
Complete primary	4271	6.00	0.00	6.00	6.00	6	6
Incomplete secondary	3944	9.22	1.24	9.18	9.26	6	11
Complete secondary	5676	12.00	0.00	12.00	12.00	12	12
Higher	3101	15.05	1.15	15.01	15.09	12	20
Total	18677	9.78	3.69	9.73	9.84	1	20

4.1.1.4. *Number of years schooling*

Of those that had ever attended school, average period of schooling was 10 (SD=4) years. Those that had incomplete primary education spent an average of 4 (SD=1) years schooling and those that had incomplete secondary education had spent average of 9 (SD=1) years schooling, as shown on Table 4.1-4.

Table 4.1-4: Number of years in schooling among women in Nigeria

Educational Attainment	Number of Women	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower	Upper		
Incomplete primary	1685	3.54	1.17	3.49	3.60	1	5
Complete primary	4271	6.00	0.00	6.00	6.00	6	6
Incomplete secondary	3944	9.22	1.24	9.18	9.26	6	11
Complete secondary	5676	12.00	0.00	12.00	12.00	12	12
Higher	3101	15.05	1.15	15.01	15.09	12	20
Total	18677	9.78	3.69	9.73	9.84	1	20

4.1.1.5. Marital status

Table 4.1-5 shows the distribution of marital status and unions the women ever had. Almost four-fifth (79%) of the women were currently married at the time of the survey, more than a tenth (13%) were never in a union, 3% were living with a partner, 3% were widowed, and one percent each were either divorced or no longer together/separated from their partners. Sixty percent of the teenagers (15-19 years) were married; and 39% of same cohort had never been in a union, followed by 20-24 years age cohort of which 26% had never been in a union, which are higher than other age cohorts.

More than four-fifth (82%) of the women were currently in a union or living with a man while 5% were formerly in a union/living with a man. Similar to the current marital status, 60% of the 15-19 years age cohort were currently living with a man or married. Also more than six percent (6%) of those 35 years or older had never been in a union.

With regards to number of unions 87% of the women had ever had one union and 11% had ever had more than one union. More than one union increased with the age of the women, which implies that as a woman grows so is her propensity for more than one union. Multiple unions grew from 3% among the 15-19 years cohort to reach 18% among the 45-49 years cohort.

Table 4.1-5: Distribution of marital status of the women

Distribution of background characteristics of current marital status, ever in a union, and number of unions ever had.

Background Characteristics	Age in 5-year groups 15-19	20-24	25-29	30-34	35-39	40-44	45-49	Total	Number of Women
Current marital status									
Never in union	38.6	26.2	13.0	6.0	2.4	1.4	0.7	12.9	4021
Married	57.9	66.8	80.3	87.0	89.6	87.5	84.3	79.1	24577
Living with partner	1.8	4.2	3.3	2.6	1.9	1.6	1.1	2.6	799
Widowed	0.3	0.3	0.9	1.7	3.3	6.7	10.4	2.8	873
Divorced	1.2	1.5	1.4	1.1	1.3	1.3	1.5	1.3	410
No longer living together/separated	0.3	1.0	1.2	1.6	1.5	1.6	1.9	1.3	395
Currently/formerly/never in union									
Never in union	38.6	26.2	13.0	6.0	2.4	1.4	0.7	12.9	4021
Currently in union/living with a man	59.6	71.0	83.5	89.6	91.5	89.1	85.4	81.7	25376
Formerly in union/living with a man	1.8	2.8	3.4	4.3	6.1	9.5	13.9	5.4	1678
Number of unions									
Once	97.2	93.4	91.2	87.5	85.5	84.7	82.4	88.6	23710
More than once	2.8	6.6	8.8	12.5	14.5	15.3	17.6	11.4	3047

4.1.1.6. Pregnancy termination

More than one in ten (12%) of the women had ever terminated a pregnancy, as shown on Table 4.1-6. This included miscarriages, abortions and still births. Termination of pregnancy increased with increase in age of the respondents from 5% among the 15-19 years cohort and peaked at 17% among the 40-44 year cohort. This could be attributed to the fact that multiple unions also increased with age of the women.

With regards to their level of education, there was no particular pattern in the incidence of termination of abortion, although the proportion was highest among those that had incomplete primary education, followed by those with higher education.

However, with regards to marital status, pregnancy termination was more common among those that were formerly in a union/living with a man (15%) than the other groups; while pregnancy termination was 6% among those that were never in a union. Almost a fifth (19%) of those that had had more than one union had terminated pregnancy compared to 13% of those that had a union once.

Table 4.1-6: Percent distribution of termination of pregnancy according to selected background characteristics of the respondents

Background characteristics	Ever had a terminated pregnancy*			Number of
	No	Yes	Total	women
All women	87.6	12.4	100.0	31042
Age in 5-year groups				
15-19	95.4	4.6	100.0	3274
20-24	91.8	8.2	100.0	5524
25-29	88.1	11.9	100.0	6395
30-34	85.6	14.4	100.0	4931
35-39	83.2	16.8	100.0	4355
40-44	82.8	17.2	100.0	3341
45-49	85.1	14.9	100.0	3222
Educational attainment				
No education	89.4	10.6	100.0	12354
Incomplete primary	84.6	15.4	100.0	1695
Complete primary	85.9	14.1	100.0	4270
Incomplete secondary	87.3	12.7	100.0	3952
Complete secondary	86.9	13.1	100.0	5670
Higher	85.7	14.3	100.0	3101
Currently/formerly/never in union				
Never in union	94.4	5.6	100.0	4018
Currently in union/living with a man	86.7	13.3	100.0	25348
Formerly in union/living with a man	84.8	15.2	100.0	1676
Number of unions				
Once	87.3	12.7	100.0	23683
More than once	80.7	19.3	100.0	3045

**33 responses were missing **297 responses were missing*

4.1.1.7. Employment status

Current employment status of the women was defined as "if the woman had done any work in the past 12 months preceding survey" (NPC [Nigeria] and ICF International, 2014) and this is presented in Table 4.1-7. Employment status is a measure of a woman's wealth. Almost seven in ten (68%) of the women were currently working, which was higher than 62% for all women reported by Nigeria DHS 2013. Current working status increased monotonically with age of the respondents while the highest working cohort with respect to educational attainment were those that had completed primary school (81%) while women with no education had the least working status (61%). Also, women who had never married had the least working status (49%) while women that were formerly in a union had the highest working status (82%).

Table 4.1-7: Percent distribution of employment status of the women

	Woman currently working			Number of women
	No	Yes	Total	
All women	31.9	68.1	100.0	30934
Age in 5-year groups				
15-19	64.6	35.4	100.0	3261
20-24	47.3	52.7	100.0	5490
25-29	33.0	67.0	100.0	6372
30-34	24.5	75.5	100.0	4915
35-39	18.7	81.3	100.0	4343
40-44	16.7	83.3	100.0	3336
45-49	15.6	84.4	100.0	3217
Educational attainment				
No education	39.2	60.8	100.0	12318
Incomplete primary	23.7	76.3	100.0	1693
Complete primary	18.6	81.4	100.0	4261
Incomplete secondary	35.3	64.7	100.0	3941
Complete secondary	27.4	72.6	100.0	5640
Higher	29.9	70.1	100.0	3081
Currently/formerly/never in union				
Never in union	51.5	48.5	100.0	3989
Currently in union/living with a man	29.8	70.2	100.0	25275
Formerly in union/living with a man	17.9	82.1	100.0	1670

4.1.1.8. *Birth history*

24,665 (79%) of the women had ever given birth to a child while 6,410 (21%) had never given birth to a child, as presented on Table 4.1-8. Ever given birth increased with the age of the women with 32% of the adolescent women being mothers. Comparing locations, birth history is relatively higher in the rural areas than the urban areas. Never having had a birth is more than twice as high in South-South and South East as in the North East and North West.

Nine in ten of the women that had no education had ever given birth compared to 60% of those that had higher education. Birth history reduced with increased education. Furthermore, all the women who had never been in a union had indicated that they had never had a birth (this is a subset of 92% of all women never in a union from the Nigeria DHS 2013). More than nine in ten of the women that ever had or currently in a union had given birth.

Table 4.1-8: Percentage distribution of women that had given birth to a child according to background characteristics

Background characteristics	Ever had a birth		Total	Number of Women
	Never had a birth	Ever had a birth		
All women	18.4	81.6	100.0	31075
Age in 5-year groups				
15-19	63.0	37.0	100.0	3278
20-24	31.7	68.3	100.0	5530
25-29	16.2	83.8	100.0	6401
30-34	8.7	91.3	100.0	4932
35-39	4.5	95.5	100.0	4361
40-44	3.8	96.2	100.0	3347
45-49	3.0	97.0	100.0	3226
Region				
North Central	19.5	80.5	100.0	4609
North East	13.7	86.3	100.0	5505
North West	12.0	88.0	100.0	8322
South East	27.5	72.5	100.0	3215
South South	26.6	73.4	100.0	4803
South West	19.3	80.7	100.0	4621
Type of place of residence				
Urban	22.9	77.1	100.0	11701
Rural	15.6	84.4	100.0	19374
Religion				
Catholic	28.1	71.9	100.0	3048
Other Christian	23.9	76.1	100.0	12091
Islam	12.3	87.7	100.0	15509
Traditionalist	8.9	91.1	100.0	280
Other	9.1	90.9	100.0	11
Educational attainment				
No education	9.8	90.2	100.0	12370
Incomplete primary	9.6	90.4	100.0	1700
Complete primary	8.9	91.1	100.0	4271
Incomplete secondary	28.0	72.0	100.0	3955
Complete secondary	29.3	70.7	100.0	5676
Higher	37.9	62.1	100.0	3103

Table 4.1-8: Percentage distribution of women that had given birth to a child according to background characteristics

Background characteristics	Ever had a birth			
	Never had a birth	Ever had a birth	Total	Number of Women
All women	18.4	81.6	100.0	31075
Age in 5-year groups				
15-19	63.0	37.0	100.0	3278
20-24	31.7	68.3	100.0	5530
25-29	16.2	83.8	100.0	6401
30-34	8.7	91.3	100.0	4932
35-39	4.5	95.5	100.0	4361
40-44	3.8	96.2	100.0	3347
45-49	3.0	97.0	100.0	3226
Region				
North Central	19.5	80.5	100.0	4609
North East	13.7	86.3	100.0	5505
North West	12.0	88.0	100.0	8322
South East	27.5	72.5	100.0	3215
South South	26.6	73.4	100.0	4803
South West	19.3	80.7	100.0	4621
Type of place of residence				
Urban	22.9	77.1	100.0	11701
Rural	15.6	84.4	100.0	19374
Religion				
Catholic	28.1	71.9	100.0	3048
Other Christian	23.9	76.1	100.0	12091
Islam	12.3	87.7	100.0	15509
Traditionalist	8.9	91.1	100.0	280
Other	9.1	90.9	100.0	11
Educational attainment				
No education	9.8	90.2	100.0	12370
Incomplete primary	9.6	90.4	100.0	1700
Complete primary	8.9	91.1	100.0	4271
Incomplete secondary	28.0	72.0	100.0	3955
Complete secondary	29.3	70.7	100.0	5676
Higher	37.9	62.1	100.0	3103

Background characteristics	Ever had a birth			
	Never had a birth	Ever had a birth	Total	Number of Women
Currently/formerly/never in union				
Never in union	82.5	17.5	100.0	4021
Currently in union/living with a man	8.9	91.1	100.0	25376
Formerly in union/living with a man	7.9	92.1	100.0	1678

4.1.1.9. *Contraceptive use*

Table 4.1-9 shows percent distribution of the women that had ever used contraceptives among the women that ever had sex in Nigeria. 8825 (28%) of the women had ever used contraceptives which increased with age cohort with 32% among the 30-39 years cohort and 20% among the 15-19 years cohort. Across the region, contraceptive use was almost ten times as high in South West (56%) compared to the North West (6%) and five time as high in the South South (50%) as in the North East (10%). Contraceptive use was more than twice as high in the urban (44%) as in the rural areas (19%). Contraceptive use was comparable between Catholic and other Christians but quite low (11%) among Muslims.

With regards to educational attainment, contraceptive use increased with more education with 6% among women with no education and 60% among women with higher education. Women that were never in a union used contraceptives two times more than women that had ever been in any union. With regards to birth history and contraceptive use, women who had never had a contraceptive use was higher among women that never had a birth (38%) than women that ever had a birth (26%).

Table 4.1-9: Percentage distribution of women that had ever used contraceptives according to background characteristics

Background characteristics	Ever used anything or tried to delay or avoid getting pregnant			Number of women
	No	Yes	Total	
All women	71.6	28.4	100.0	31075
Age in 5-year groups				
15-19	80.3	19.7	100.0	3278
20-24	70.4	29.6	100.0	5530
25-29	70.5	29.5	100.0	6401
30-34	68.9	31.1	100.0	4932
35-39	68.0	32.0	100.0	4361
40-44	69.6	30.4	100.0	3347
45-49	78.2	21.8	100.0	3226
Region				
North Central	69.1	30.9	100.0	4609
North East	90.1	9.9	100.0	5505
North West	93.8	6.2	100.0	8322
South East	58.3	41.7	100.0	3215
South South	50.2	49.8	100.0	4803
South West	43.6	56.4	100.0	4621
Type of place of residence				
Urban	56.4	43.6	100.0	11701
Rural	80.8	19.2	100.0	19374
Religion				
Catholic	54.5	45.5	100.0	3048
Other Christian	53.6	46.4	100.0	12091
Islam	88.6	11.4	100.0	15509
Traditionalist	87.9	12.1	100.0	280
Other	81.8	18.2	100.0	11
Educational attainment				
No education	94.5	5.5	100.0	12370
Incomplete primary	76.6	23.4	100.0	1700
Complete primary	67.3	32.7	100.0	4271
Incomplete secondary	58.4	41.6	100.0	3955
Complete secondary	49.9	50.1	100.0	5676
Higher	40.0	60.0	100.0	3103

Background characteristics	Ever used anything or tried to delay or avoid getting pregnant			Number of women
	No	Yes	Total	
Currently/formerly/never in union				
Never in union	43.7	56.3	100.0	4021
Currently in union/living with a man	75.9	24.1	100.0	25376
Fonnerly in union/living with a man	73.2	26.8	100.0	1678
Ever had a birth				
Never had a birth	61.7	38.3	100.0	5705
Ever had a birth	73.8	26.2	100.0	25370

4.1.2. Description of median age at first sex, age at first birth and mean of their intervals

Table 4.1-10 describes the median age at the first sex and first birth as well as the mean of their paired intervals. For all the women, median age at first sex was 16 years which showed an inverse relationship with birth cohort. Younger women tend to give birth earlier than older women. Figure 4.1-4 provides a clearer description of this changing trend. Median age at first sex was highest in North Central, South East and South West at 18 years while North West and North East had the lowest at 15 years. Women in the rural areas experienced sex debut 3 years ahead of those from urban centers and women with no education experienced sexual debut 3 years ahead of women that completed secondary education and 5 years ahead of women with higher education. Also, women who had never been in a union experienced sexual debut 2 years later than those that ever had any union.

Median age at first birth was 19 years for all women and comparable for all age cohorts except for 15-19 years cohort with a median of 16 years. On the average, women in the North East give birth a year later than women in the North West while women from South East and South West give birth most latest at 21 years. Women in urban areas give birth to their first child 2 years later than women from rural areas and women with no education give birth to their first child 3 years ahead of women that had completed secondary education and 7 years ahead of women with higher education. Median age at first birth is same (19 years) for women that ever had union or not.

The interval between first sex and first birth is 2.8 years for all women, which is about 1.7 years for the 15-19 years age cohort and approximately 3 years for women above 25 years. Mean interval between first sex and first birth is highest in the South West and South East at 3.08 years and lowest in North Central at 2.46 years. Number of years between sexual debut and first birth is

4.1.2. Description of median age at first sex, age at first birth and mean of their intervals

Table 4.1-10 describes the median age at the first sex and first birth as well as the mean of their paired intervals. For all the women, median age at first sex was 16 years which showed an inverse relationship with birth cohort. Younger women tend to give birth earlier than older women. Figure 4.1-4 provides a clearer description of this changing trend. Median age at first sex was highest in North Central, South East and South West at 18 years while North West and North East had the lowest at 15 years. Women in the rural areas experienced sex debut 3 years ahead of those from urban centers and women with no education experienced sexual debut 3 years ahead of women that completed secondary education and 5 years ahead of women with higher education. Also, women who had never been in a union experienced sexual debut 2 years later than those that ever had any union.

Median age at first birth was 19 years for all women and comparable for all age cohorts except for 15-19 years cohort with a median of 16 years. On the average, women in the North East give birth a year later than women in the North West while women from South East and South West give birth most latest at 21 years. Women in urban areas give birth to their first child 2 years later than women from rural areas and women with no education give birth to their first child 3 years ahead of women that had completed secondary education and 7 years ahead of women with higher education. Median age at first birth is same (19 years) for women that ever had union or not.

The interval between first sex and first birth is 2.8 years for all women, which is about 1.7 years for the 15-19 years age cohort and approximately 3 years for women above 25 years. Mean interval between first sex and first birth is highest in the South West and South East at 3.08 years and lowest in North Central at 2.46 years. Number of years between sexual debut and first birth is comparable between urban and rural areas. While this value is approximately three years for women with no education and secondary education completed and about 4 years for women with higher education. Average interval between first sex and first birth was slightly higher among women currently living in a union/living with a man (2.8 years) than among those formerly in a union (2.7 years) and those never in a union (2.6 years).

Table 4.1-10: Median age at first sex, first birth, and mean of their intervals

Background characteristics	Median Age at first sex (Years)	Median Age at first birth (Years)	Mean Age at first sex and first birth interval (Years)
All women	16	19	2.82
Age in 5-year groups			
15-19	15	16	1.7
20-24	17	18	2.04
25-29	17	19	2.66
30-34	17	19	3.03
35-39	17	20	3.26
40-44	17	20	3.26
45-49	16	19	3.12
Region			
North Central	18	20	2.46
North East	15	18	2.56
North West	15	17	2.82
South East	18	21	3.08
South South	17	19	3.1
South West	18	21	3.08
Type of place of residence			
Urban	18	20	2.92
Rural	15	18	2.76
Educational attainment			
No education	15	17	2.74
Incomplete primary	16	18	2.49
Complete primary	16	19	2.56
Incomplete secondary	17	19	2.66
Complete secondary	18	21	3.03
Higher	20	24	3.86
Ever in a union			
Never in union	18	19	2.55
Currently in union/living with a man	16	19	2.83
Formerly in union/living with a man	16	19	2.7

4.1.3. Descriptive statistics of age at first sex and age at first birth, and their intervals

This section provides description of the variables using measure of dispersion, measures of central tendencies, measures of skewness and graphical representation. The descriptive statistics are as shown on

Table 4.1-11.

4.1.3.1. Age at first sex

Minimum and maximum ages at first sex are 8 years and 46 years respectively with median age at first sex of 16 years. The 25th and 75th percentiles are also given at 14 years and 19 years respectively, which imply that 25% of the respondents had had sex by age 14 years and 75% of the respondents had had sex by age 19 years. The data is positively skewed (Skewness = 1.14, SE = 0.01). Figure 4.1-1 displays the distribution of age at first sex on a histogram, highlighting the median age at first sex. It should be noted that 17872 (58%) of the women had their first sex in their first union.

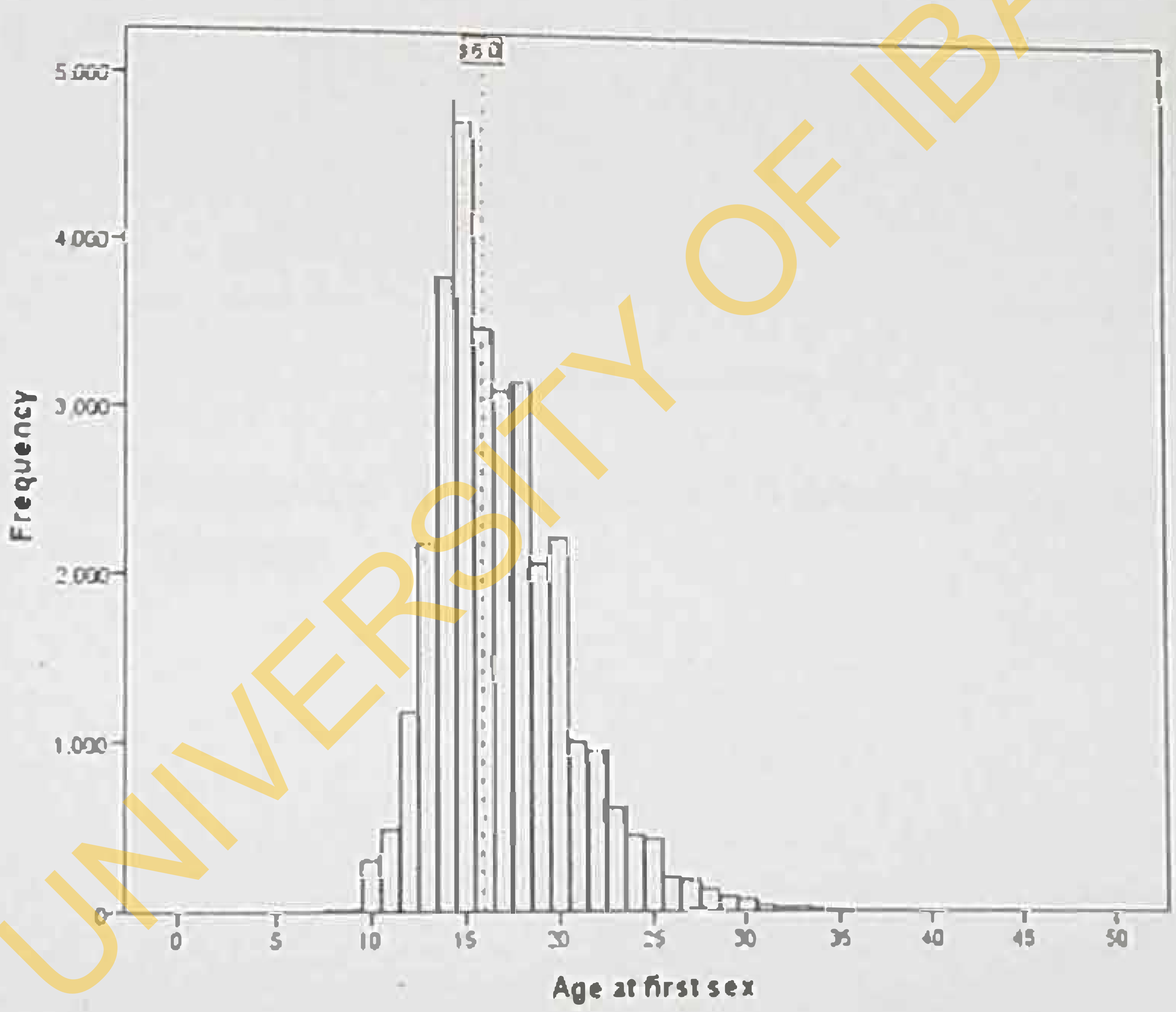


Figure 4.1-1: Histogram showing the distribution of age at first sex

4.1.3.2. *Age at first birth*

Minimum and maximum ages at first sex are 12 years and 45 years respectively with median of 19 years, as shown on Table 4.1-11. The 25th and 75th percentiles are also given at 16 years and 22 years respectively, which imply that 25% of the respondents had had sex by age 16 years and 75% of the respondents had had sex by age 22 years. The data is positively skewed (Skewness= 1.01, SE = 0.02). Figure 4.1-2 displays the distribution of age at first birth on a histogram, highlighting the median age.

Figure 4.1-2: Histogram showing the distribution of age at first birth among women in Nigeria

4.1.3.3. *Interval between age at first sex and age at first birth*

Minimum and maximum interval were 0.0 years and 26 years respectively with median of 2 years, as shown on Table 4.1-11. The 25th and 75th percentiles are also given at 1 year and 4 years respectively, which imply that 25% of those that ever had sex had a child about a year later and 75% by age 4 years later. The data is positively skewed (Skewness = 2.30, SE = 0.02). Figure 4.1-3 displays the distribution of the interval on a histogram, highlighting the mean and median ages.

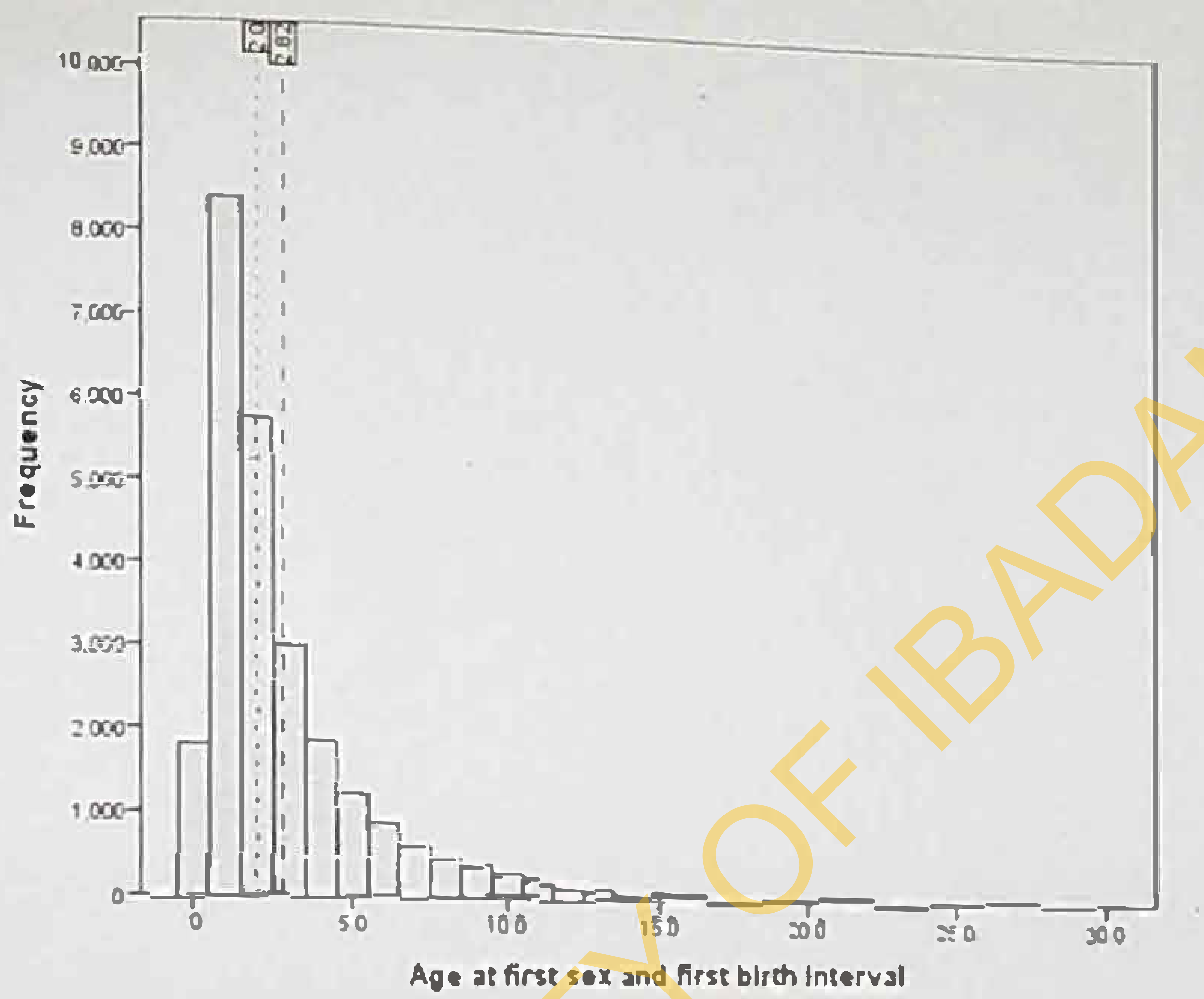


Figure 4.1-3: Histogram showing the distribution of interval between first sex and first marriage among women in Nigeria

Table 4.1-11: Descriptive statistics of age at first sex, age at first birth, age at first marriage and their intervals

Statistics		Age at first sex	Age at first birth	Age at first sex and first birth interval
Mean		16.95	19.53	2.82
Median		16.00	19.00	2.00
Mode		15.00	17.00	1.00
Std. Deviation		3.66	4.36	2.82
Skewness		1.14	1.01	2.30
Std. Error of Skewness		0.01	0.02	0.02
Kurtosis		2.58	1.21	7.10
Std. Error of Kurtosis		0.03	0.03	0.03
Minimum		8.00	12.00	0.00
Maximum		46.00	45.00	26.00
Percentiles	25th	14.00	16.00	1.00
	50th	16.00	19.00	2.00
	75th	19.00	22.00	4.00

Figure 4.1-4 shows the trend in the reproductive health event by birth cohorts of the women since 1963. The vertical bars have been added to provide easy tracing of the ages with the cohorts and to emphasize the interval between the first sex and first birth. There is a general decline in age at first sex and age at first birth and the width of the interval is shrinking. Specifically the median age at first sex undulated briefly for almost all women born before the 1988 and after a sharp increase in 1989, there had been a general decline in the median ages at both RH events with considerable shrinking interval. The interval is much narrower for the younger generation compared to the older women. This cannot be excused from the recent trend in exposure to sexually explicit materials, increased availability of reversible contraceptives, increasing practice of cohabitation, which had been foreign to most Nigerian culture until recently and generally

changing culture. Hence, the birth cohort, was included in the multivariate analysis to provide insight into changing trend by age cohort.

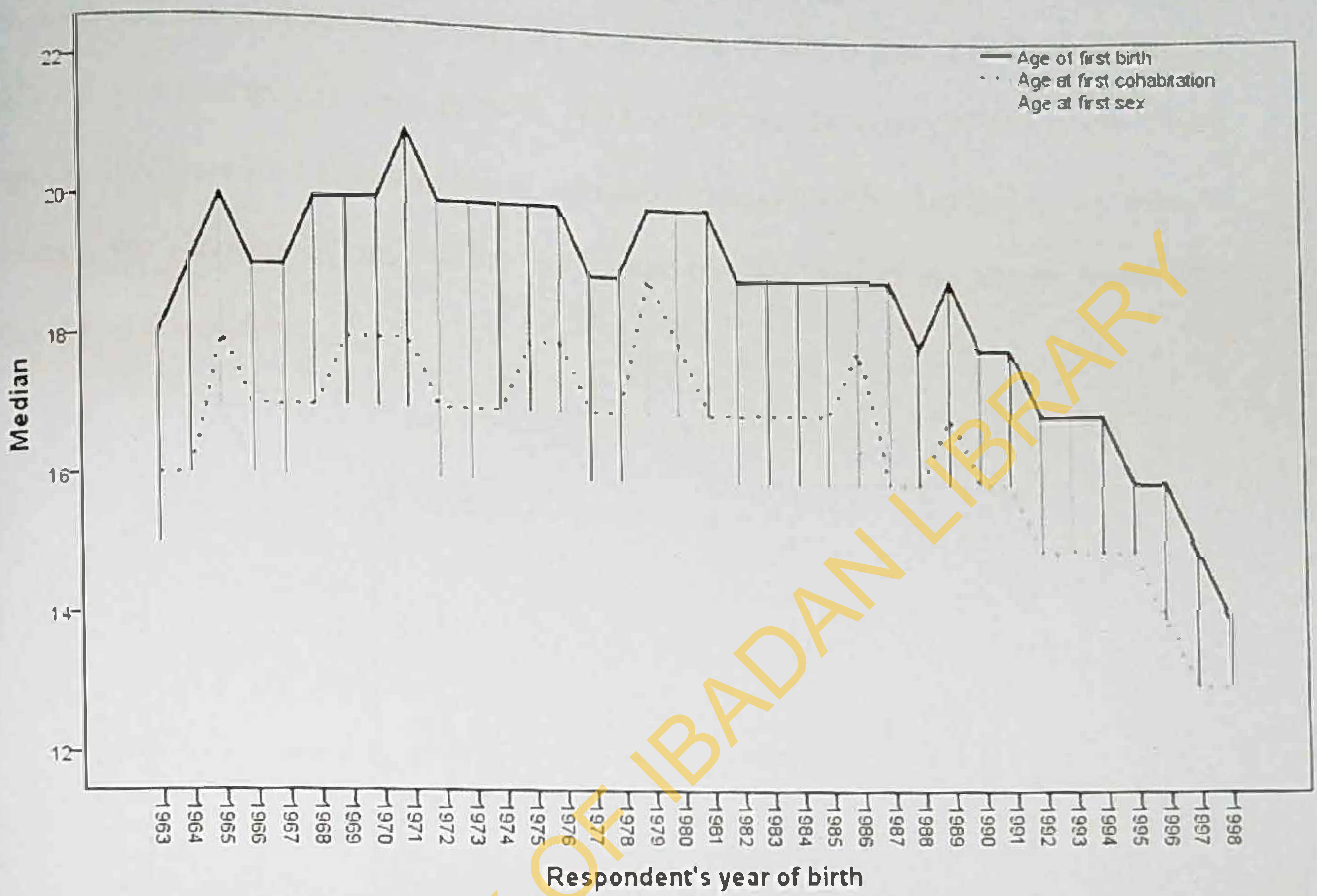


Figure 4.1-4: Nigerian women's median ages at reproductive events

changing culture. Hence, the birth cohort, was included in the multivariate analysis to provide insight into changing trend by age cohort.

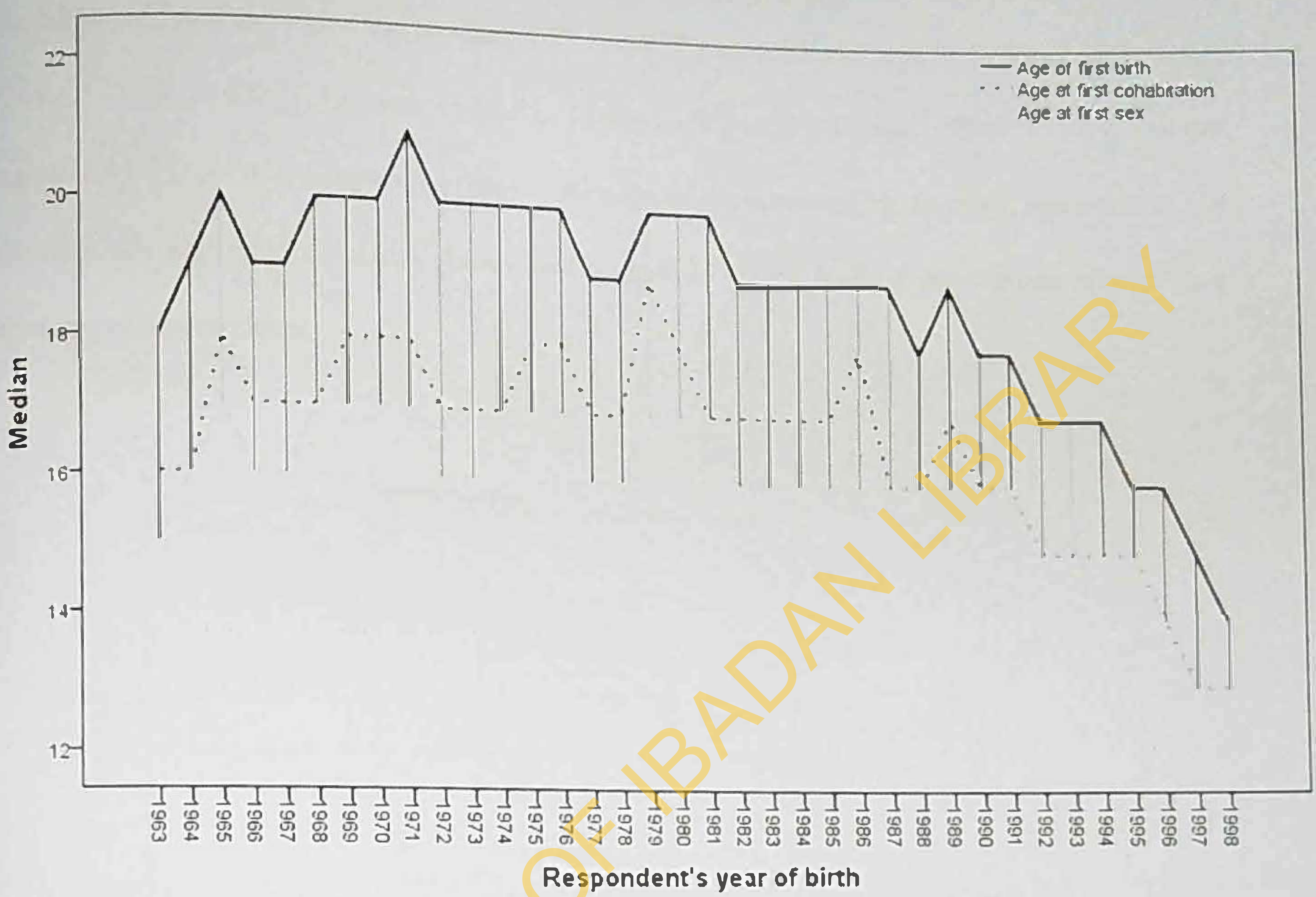


Figure 4.1-4: Nigerian women's median ages at reproductive events

4.2. Model fitting

4.2.1. Fitting the log-cumulative hazard plot for age at first sex

The Kaplan-Meier estimates the median, 25th and 75th percentiles of age at first sex as 16 years (95% CI: 15.94–16.05), 19 years (95%CI: 18.94–19.06), and 14 years (95%CI: 13.96–14.04) respectively. Figure 4.2-1 shows the plot of estimates of the survivorship function against time. As expected, the survivorship curve shows that by age 16 years, half of the women would have experienced sexual debut.

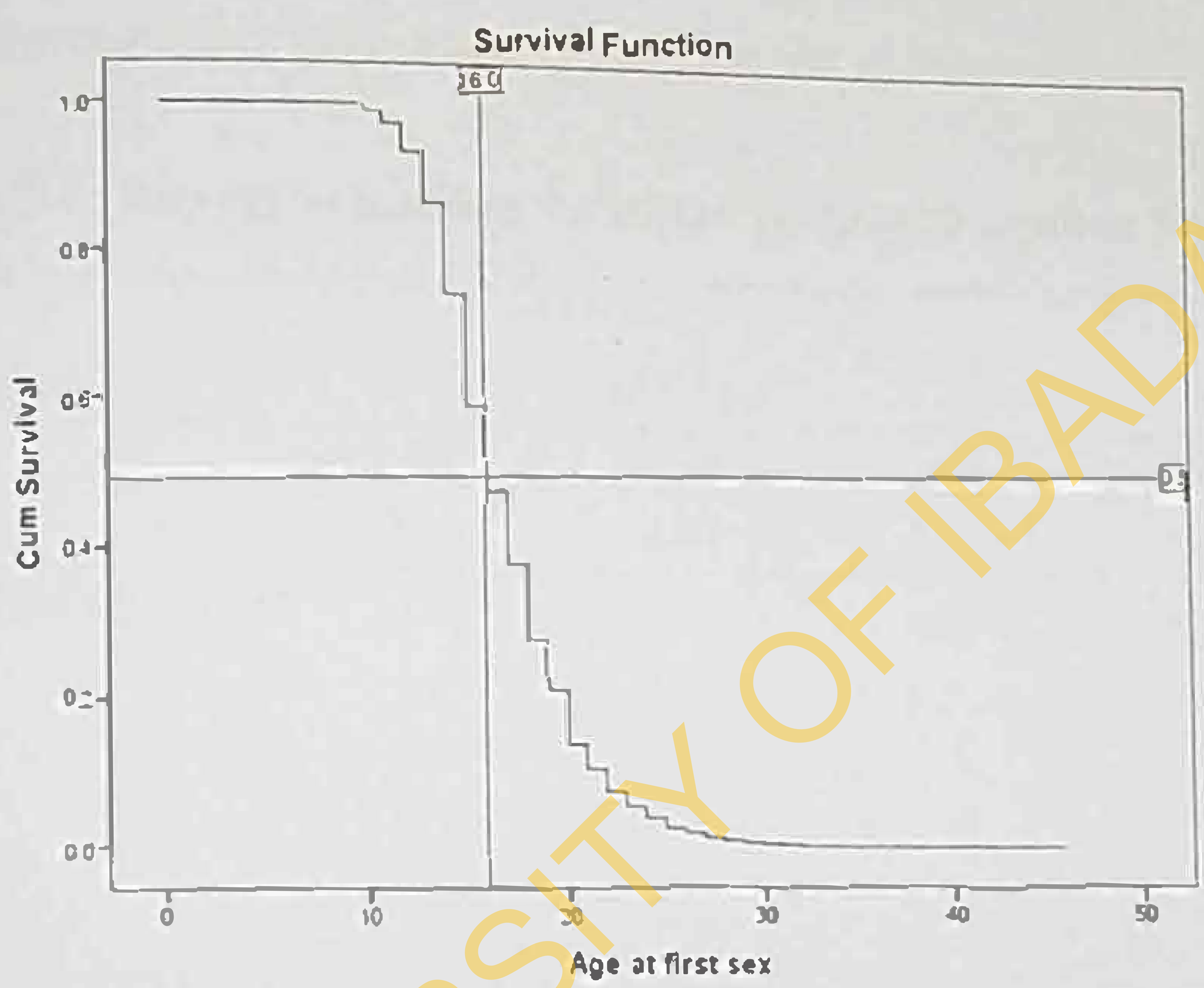


Figure 4.2-1: Cumulative survival function for age at first sex using Kaplan-Meier estimate

Figure 4.2-2 shows the log-cumulative hazard plot for age at first sex which visually defies an assumption of a straight line. The relationship between the two variables was tested with ordinary least squares regression and to obtain estimates of the shape and scale parameters, as shown on Table 4.2-1. The linear model was significant (F=287243.2, p<0.001, $R^2 = 0.902$). The values of the shape parameter, $\alpha = 5.37$ (SE = 0.01) and value of the scale parameter for the Weibull distribution is $\beta = \exp\{-15.456\} = 1.939\times10^{-07}$ (SE=0.028). Since the value of the scale parameter is not close to unity, the distribution cannot assume an exponential distribution and hence favors a Weibull distribution. Furthermore, the model fitting shows that the hazard function takes on a value approximately zero when age at first sex is zero.

Table 4.2-1: Regression of log-cumulative hazard against log of age at first sex

Parameters	B	Std. Error	t	Sig.	Lower 95% CI	Upper 95% CI
log β	-15.46	0.03	-547.74	0.00	-15.51	-15.40
α	5.37	0.01	535.95	0.00	5.35	5.39

RMSE = 0.36

Figure 4.2-2 shows the log-cumulative hazard plot for age at first sex which visually defies an assumption of a straight line. The relationship between the two variables was tested with ordinary least squares regression and to obtain estimates of the shape and scale parameters, as shown on Table 4.2-1. The linear model was significant (F=287243.2, p<0.001, R^2 = 0.902). The values of the shape parameter, $\alpha = 5.37$ (SE = 0.01) and value of the scale parameter for the Weibull distribution is $\beta = \exp\{-15.456\} = 1.939\times10^{-07}$ (SE=0.028). Since the value of the scale parameter is not close to unity, the distribution cannot assume an exponential distribution and hence favors a Weibull distribution. Furthermore, the model fitting shows that the hazard function takes on a value approximately zero when age at first sex is zero.

Table 4.2-1: Regression of log-cumulative hazard against log of age at first sex

Parameters	B	Std. Error	t	Sig.	Lower 95% CI	Upper 95% CI
log β	-15.46	0.03	-547.74	0.00	-15.51	-15.40
α	5.37	0.01	535.95	0.00	5.35	5.39

RMSE = 0.36

Figure 4.2-2: Log-cumulative hazard plot of age at first sex

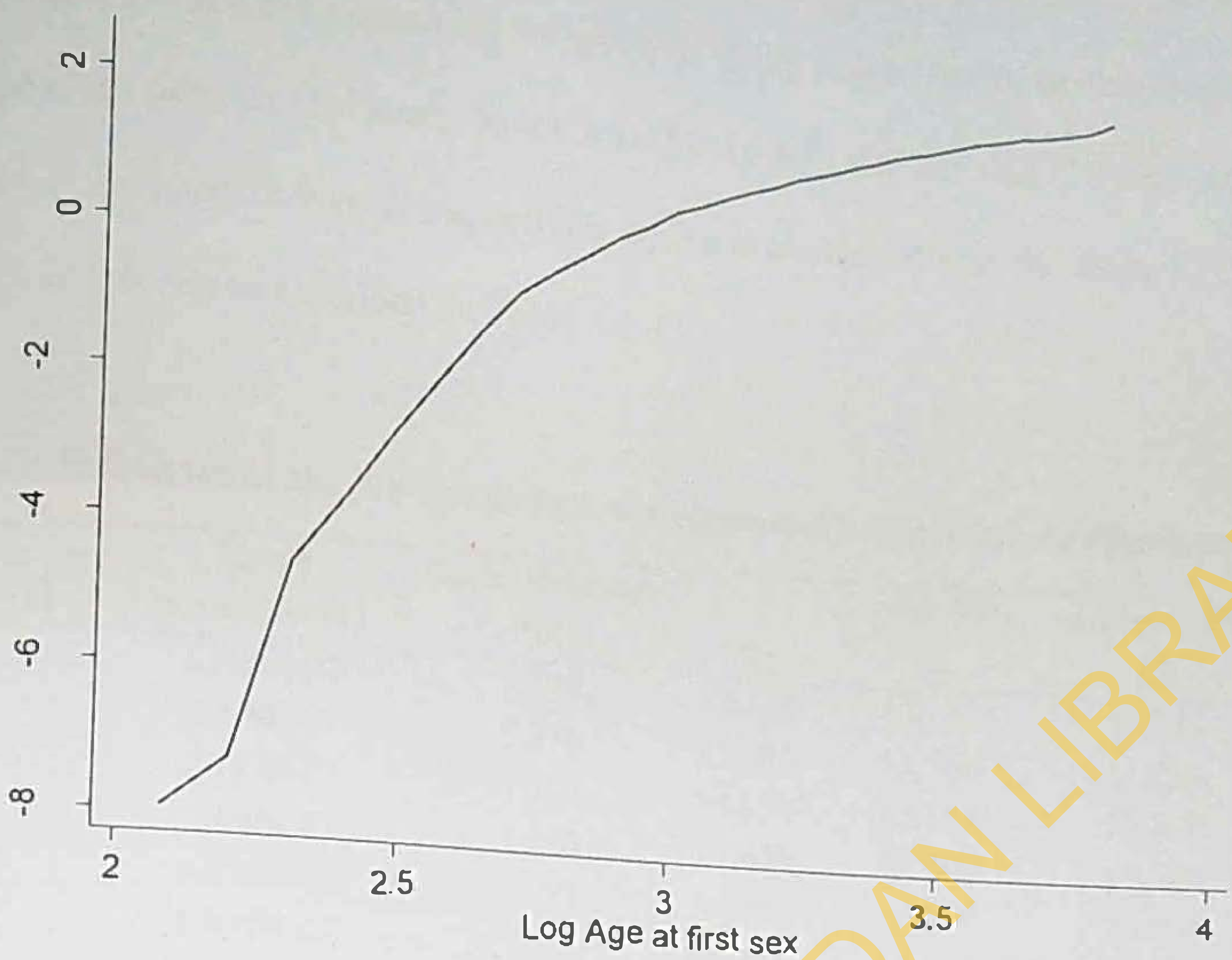


Figure 4.2-2: Log-cumulative hazard plot of age at first sex

Estimates of the pth percentiles are given in Table 4.2-2. The table shows an estimate of median age at sex debut is 16.6 years (95% CI: 16.43–16.80). The 25th and 75th percentiles are also given as 14.1 years and 18.9 years respectively which is consistent with the range of the expected median age at first sex as described on Table 4.1-11.

Table 4.2-2: Estimates of the pth age at first sex estimates using Weibull distribution

Percentile (p)	Shape parameter (a)	Scale parameter (b)	t(p)	Lower 95% CI	Upper 95% CI
5	1.939E-07	5.370	10.228	10.115	10.342
10	1.939E-07	5.370	11.695	11.566	11.826
15	1.939E-07	5.370	12.678	12.538	12.820
20	1.939E-07	5.370	13.449	13.300	13.599
25	1.939E-07	5.370	14.100	13.944	14.258
30	1.939E-07	5.370	14.676	14.514	14.840
35	1.939E-07	5.370	15.201	15.033	15.371
40	1.939E-07	5.370	15.692	15.518	15.867
45	1.939E-07	5.370	16.158	15.979	16.339
50	***1.939E-07***	***5.370***	***16.609***	***16.426***	***16.795***
55	1.939E-07	5.370	17.053	16.864	17.243
60	1.939E-07	5.370	17.495	17.302	17.691
65	1.939E-07	5.370	17.944	17.746	18.145
70	1.939E-07	5.370	18.408	18.204	18.614
75	1.939E-07	5.370	18.898	18.689	19.109
80	1.939E-07	5.370	19.430	19.215	19.647
85	1.939E-07	5.370	20.034	19.813	20.258
90	1.939E-07	5.370	20.770	20.541	21.003
95	1.939E-07	5.370	21.814	21.572	22.057

4.2.2. Fitting the log-cumulative hazard plot for age at first birth

The Kaplan-Meier estimates the median age at first birth as 19 years (95% CI: 18.94–19.06), with the 25th and 75th percentiles estimated as 16 years (95% CI: 15.95–16.05) and 22 years (95% CI: 21.91–22.09). Figure 4.2-3 shows the plot of estimates of the survivorship function and cumulative hazard function respectively against time. As expected, the survivorship curve shows that by age 19 years, half of the women that ever had sex would have given birth.

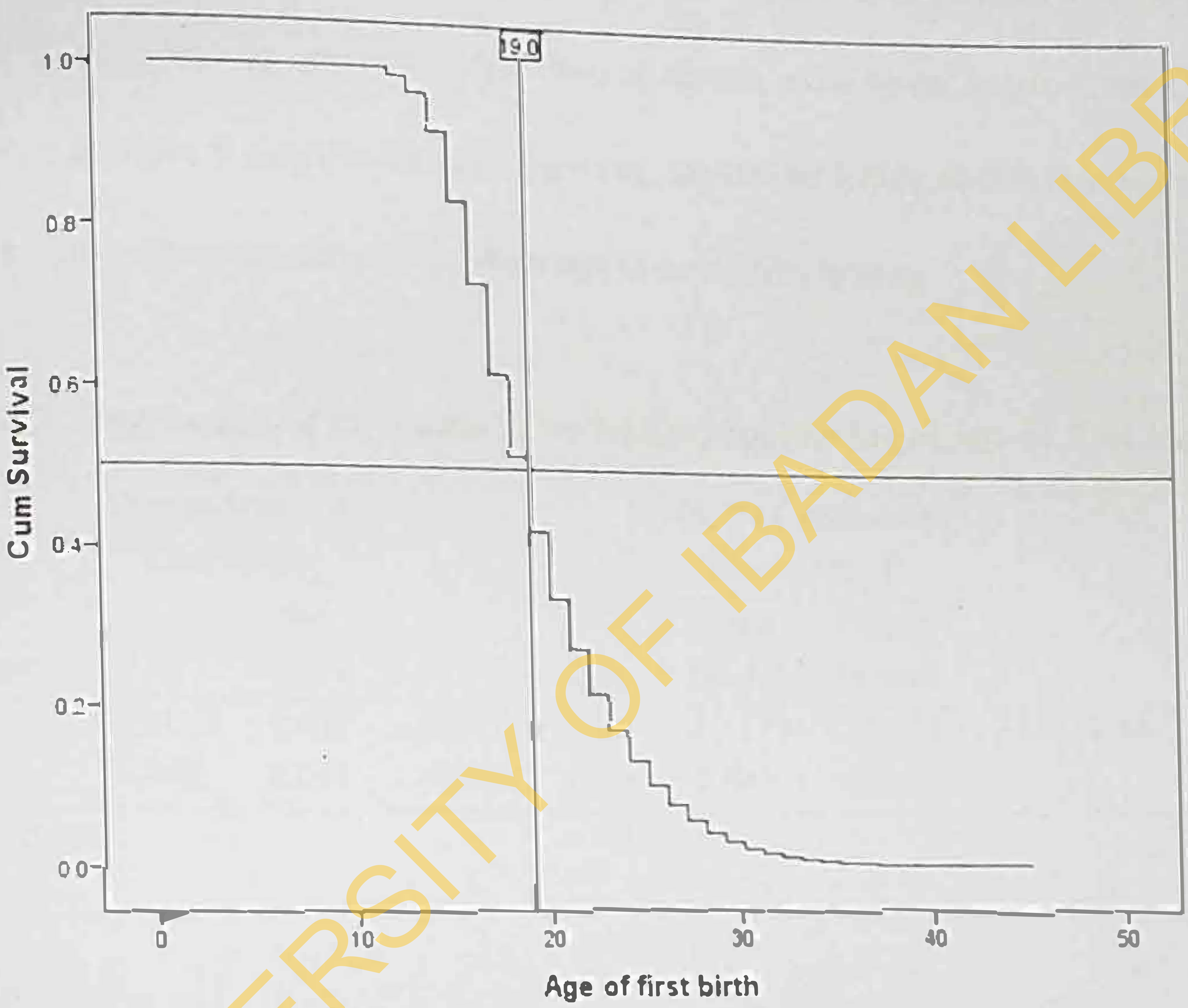


Figure 4.2-3: Cumulative survival function for age at birth sex using Kaplan-Meier estimate

Figure 4.2-4 shows the log-cumulative hazard plot for age at first birth which also visually defies an assumption of a straight line. The relationship between the two variables was tested with ordinary least squares regression and to obtain estimates of the shape and scale parameters were given, as shown on Table 4.2-3. The linear model was significant with an r-squared value of 0.894. The values of the shape parameter, $\alpha = 4.97$ (SE = 0.01) and value of the scale parameter for the Weibull distribution is $\beta = \exp\{-15.12\} = 2.72\times10^{-07}$ (SE = 0.03). Since the value of the scale parameter is not close to unity, the distribution cannot assume an exponential distribution and hence favors a Weibull distribution. Furthermore, the model fitting shows that the hazard function takes on a value approximately zero when age at first birth is zero.

Table 4.2-3: Regression of log-cumulative hazard against log of age at first birth

Parameters	Unstandardized Coefficient B	Std. Error	t	Sig.	95.0% Confidence Interval for B Lower Bound	Upper Bound	F	p-value	R-squared
log β	-15.117	0.032	-476.57	0.00	-15.179	-15.054	214419.44	<0.0001	0.894
α	4.968	0.011	463.05	0.00	4.947	4.989			

RMSE = 0.362

Figure 4.2-4: Log cumulative hazard function curve of age at first birth

Table 4.2-4 shows estimates of the pth percentile using the Weibull parameters for age at first birth as estimates. The median age at first birth from the model is 19.5 years (95% CI: 19.23–19.72) which is comparable to the raw median of 19 years. The 25th and 75th percentiles are 16.3 (95% CI: 16.12–16.52) years and 22.4 (95 CI: 22.12–22.67) respectively.

Table 4.2-4: Estimates of the *p*th age at first birth estimates using Weibull distribution

Percentile (p)	Shape parameter (a)	Scale parameter (b)	t(p)	Lower 95% CI	Upper 95% CI
5	2.721E-07	4.968	11.531	11.390	11.673
10	2.721E-07	4.968	13.328	13.165	13.493
15	2.721E-07	4.968	14.543	14.366	14.724
20	2.721E-07	4.968	15.502	15.312	15.694
25	2.721E-07	4.968	16.315	16.116	16.517
30	2.721E-07	4.968	17.037	16.828	17.247
35	2.721E-07	4.968	17.696	17.480	17.915
40	2.721E-07	4.968	18.314	18.090	18.541
45	2.721E-07	4.968	18.903	18.672	19.137
50	***2.721E-07***	***4.968***	***19.474***	***19.236***	***19.715***
55	2.721E-07	4.968	20.037	19.792	20.285
60	2.721E-07	4.968	20.600	20.348	20.855
65	2.721E-07	4.968	21.172	20.913	21.434
70	2.721E-07	4.968	21.764	21.497	22.033
75	2.721E-07	4.968	22.390	22.116	22.667
80	2.721E-07	4.968	23.073	22.791	23.359
85	2.721E-07	4.968	23.850	23.558	24.145
90	2.721E-07	4.968	24.798	24.495	25.105
95	2.721E-07	4.968	26.147	25.827	26.471

4.2.3. Fitting log-cumulative hazard curve for interval between first sex and first birth

The Kaplan-Meier estimates the median age at first sex as 2 years (95% CI: 1.97–2.03). Figure 4.2-5 shows the plot of estimates of the survivorship function time. As expected, the survivorship curve shows half of the women would have given birth within two years of their first sex.

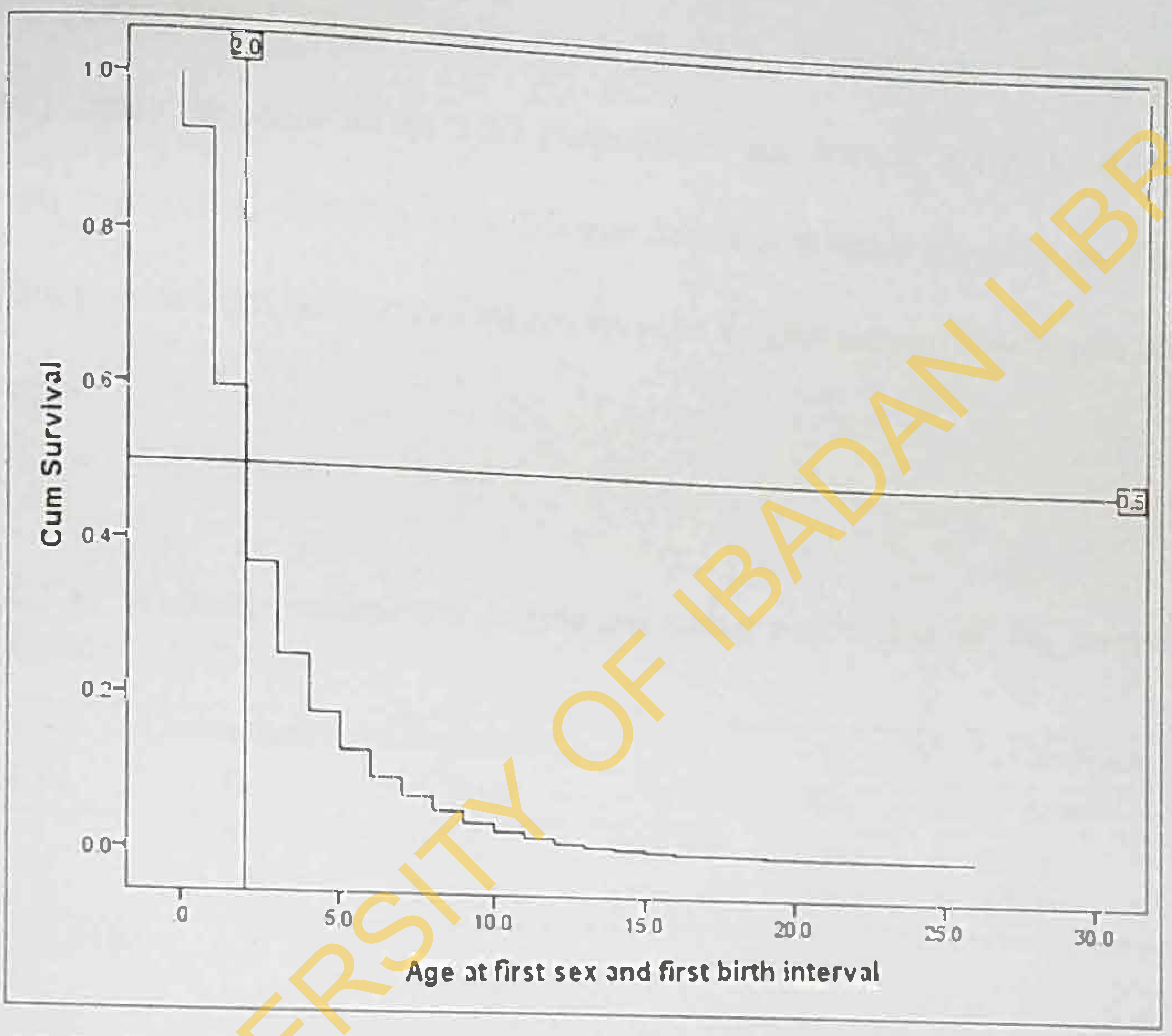


Figure 4.2-5: Cumulative survival function for interval between age at first sex and age at first birth using Kaplan-Meier estimate

Figure 4.2-4 shows the log-cumulative hazard plot for interval between age at sex debut and age at first birth which assumes straight line. The relationship between the two variables was tested with ordinary least squares regression and to obtain estimates of the shape and scale parameters were given, as shown on Table 4.2-5. The linear model was significant, F = 7713903.71, $p < 0.001$, $R^2 = 0.997$, RMSE = 0.0384. The values of the shape parameter, α = -1.00 (SE = 0.07) and value of the scale parameter for the Weibull distribution is $\beta = \exp\{0.928\} = 2.53$ (SE=0.02). Since the value of the scale parameter is not close to unity, the distribution cannot assume an exponential distribution and hence favors a Weibull distribution. Furthermore, the model fitting shows that the hazard function takes on a value approximately zero when age at first birth is zero.

Table 4.2-5: Weibull parameters estimation using regression of log cumulative hazard against log of time

	Unstandardized Coefficient				95.0% Confidence Interval for B	
Parameters	B	Std. Error	t	Sig.	Lower Bound	Upper Bound
log β	0.928	0.023	2779.590	0.000	0.927	0.929
α	-1.001	0.069	-2734.440	0.000	-1.002	-1.001

RMSE = 0.0384

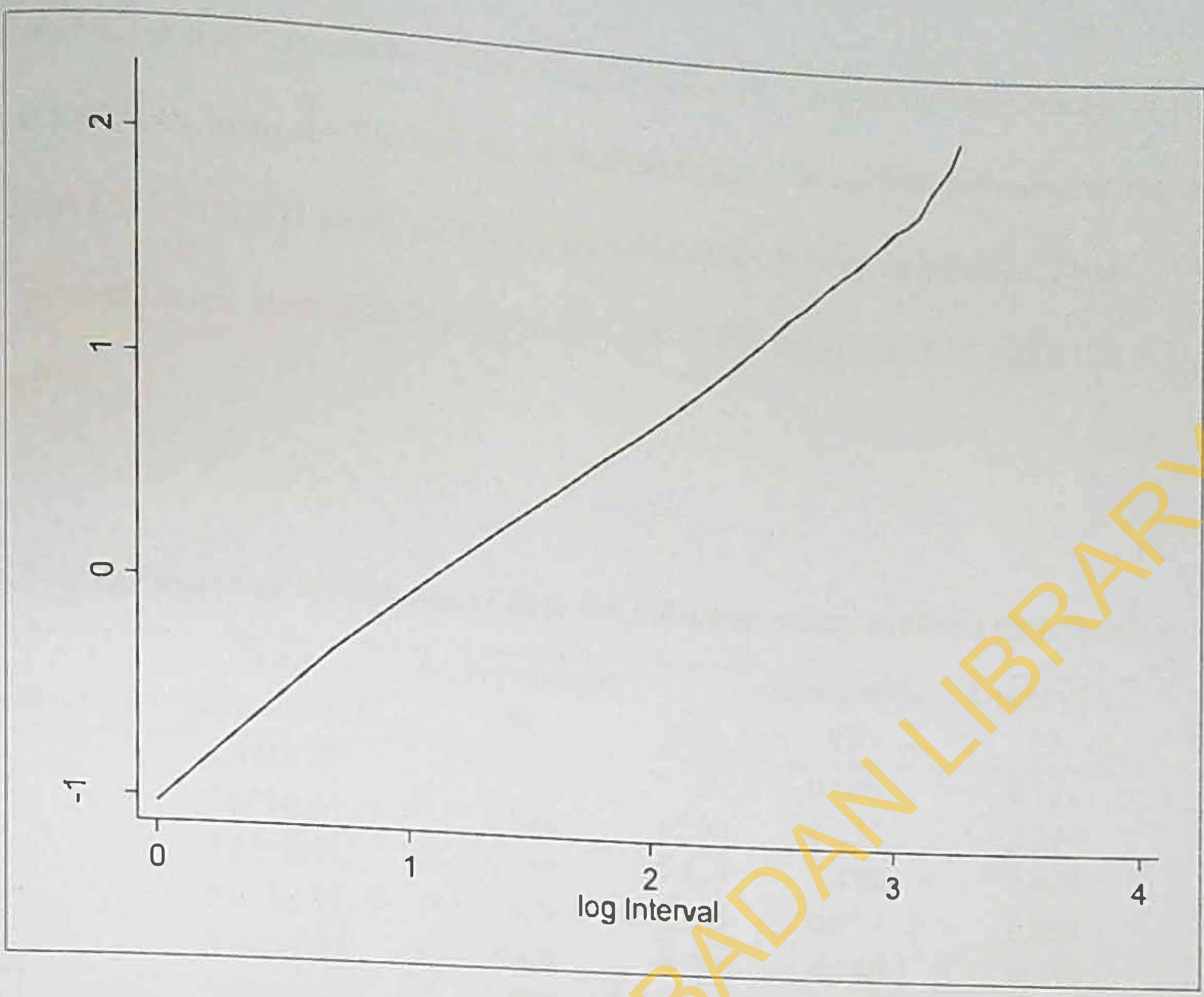


Figure 4.2-6: Log cumulative hazard function of interval between age at first sex and age at first birth

Table 4.2-6 shows estimates of the *p*th percentiles of the interval between age at first sex and age at first birth, using the Weibull parameter estimates. The median estimate for the interval is 1.98 (95% CI: 1.96–2.01) years, which approximates the observed median of 2 years. The 25th and 75th percentiles are given as 0.768 (95% CI: 0.76–0.78) years and 4.18 (95% CI: 4.13–4.24) respectively.

Table 4.2-6: Estimates of the *p*th age at first sex estimates using Weibull distribution

Percentile (p)	Shape parameter (a)	Scale parameter (b)	t(p)	Lower 95% CI	Upper 95% CI
5	3.675E-01	0.928	0.120	0.118	0.121
10	3.675E-01	0.928	0.260	0.257	0.264
15	3.675E-01	0.928	0.415	0.410	0.420
20	3.675E-01	0.928	0.584	0.577	0.592
25	***3.675E-01***	***0.928***	***0.768***	***0.758***	***0.778***
30	3.675E-01	0.928	0.968	0.956	0.981
35	3.675E-01	0.928	1.187	1.172	1.202
40	3.675E-01	0.928	1.426	1.408	1.444
45	3.675E-01	0.928	1.689	1.668	1.711
50	***3.675E-01***	***0.928***	***1.981***	***1.956***	***2.007***
55	3.675E-01	0.928	2.308	2.278	2.337
60	3.675E-01	0.928	2.676	2.642	2.711
65	3.675E-01	0.928	3.099	3.060	3.139
70	3.675E-01	0.928	3.592	3.546	3.638
75	***3.675E-01***	***0.928***	***4.181***	***4.128***	***4.235***
80	3.675E-01	0.928	4.911	4.849	4.974
85	3.675E-01	0.928	5.863	5.789	5.938
90	3.675E-01	0.928	7.224	7.132	7.317
95	3.675E-01	0.928	9.592	9.471	9.716

4.3. Maximum Likelihood Estimation of Weibull Parameters

4.3.1. Age at first sex

The estimates of the scale and shape parameters of the Weibull model were 2.55×10^{-4} (95% CI: $1.98-2.84\times10^{-4}$) and 4.42 (95% CI: 4.09–4.75) respectively, with a log likelihood of 612.44 (AIC = -1220.886). Figure 4.3-1 shows the survival curve of age at first sex highlighting the estimate of the median age at first sex which is approximately 17 years. The confidence interval of the shape parameter does not contain unity and this shows that the hazard function for age at first sex cannot be modelled with exponential distribution.

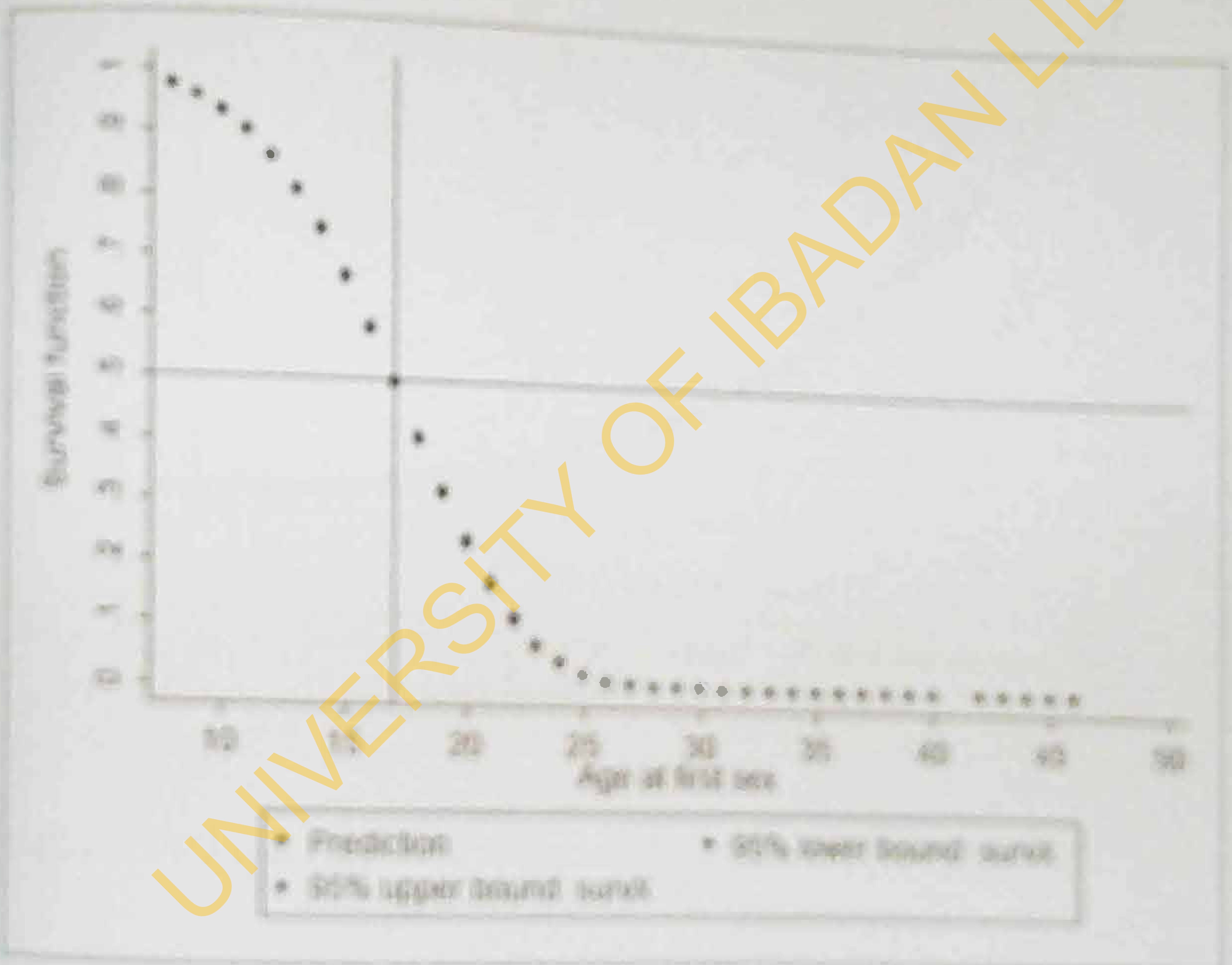


Figure 4.3-1: Survival function of age at first sex highlighting estimate of median age at first sex using Weibull estimates

Table 4.4-1 shows the estimates of the percentiles. The estimated median age at first sex is 16.97 years (95% CI: 14.97–19.23). Hence, 95% of the time, age at first sex is expected to fall within the range of about 15 years to 19 years. The interval is less precise compared to the estimate median from Kaplan-Meier estimates but the results go to show that age at first sex could be estimated using the Weibull distribution. It is observed that width of the 95% CI increased with the percentiles.

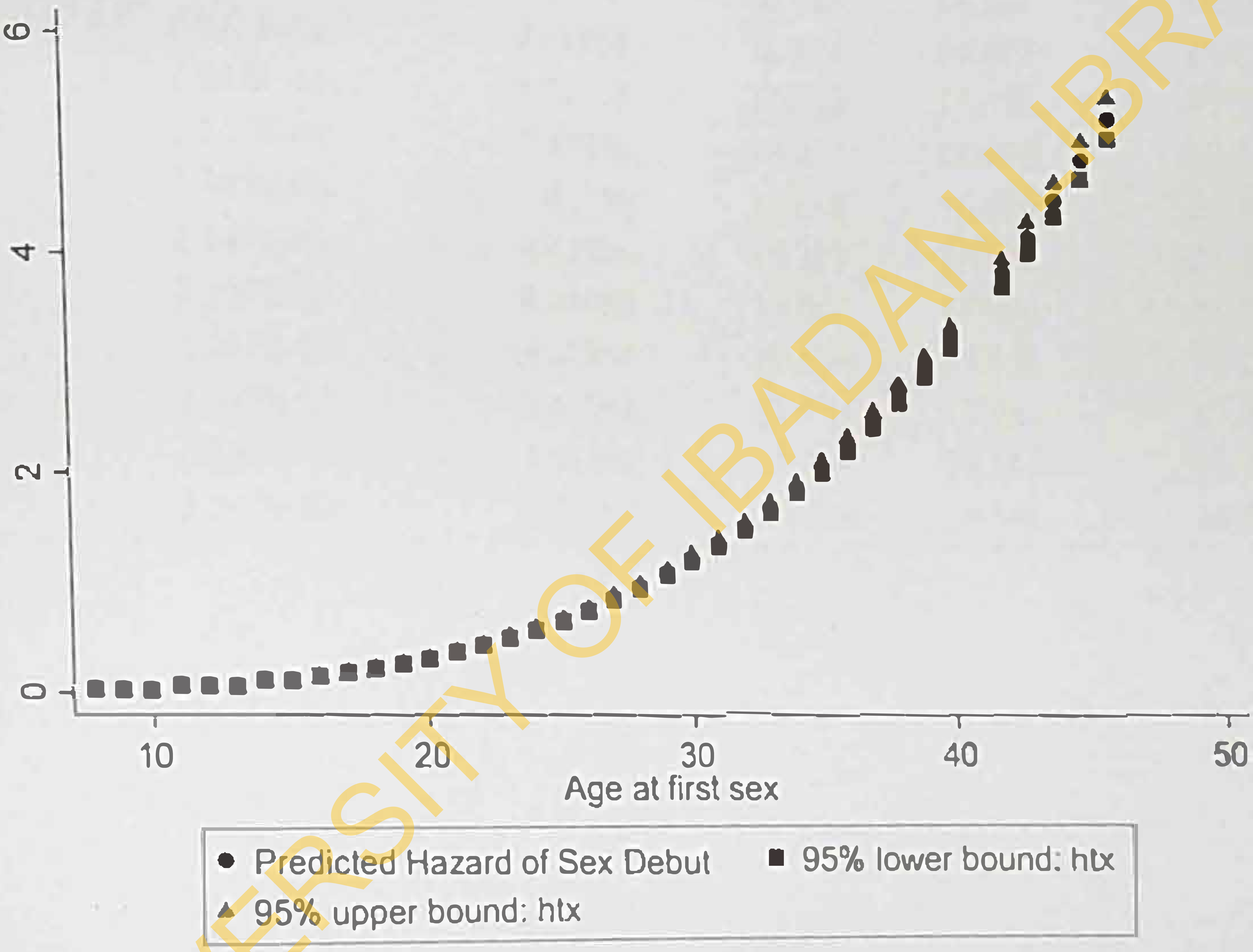


Figure 4.3-2: Predicted hazard function of sex debut

Table 4.4-1 shows the estimates of the percentiles. The estimated median age at first sex is 16.97 years (95% CI: 14.97–19.23). Hence, 95% of the time, age at first sex is expected to fall within the range of about 15 years to 19 years. The interval is less precise compared to the estimate median from Kaplan-Meier estimates but the results go to show that age at first sex could be estimated using the Weibull distribution. It is observed that width of the 95% CI increased with the percentiles.

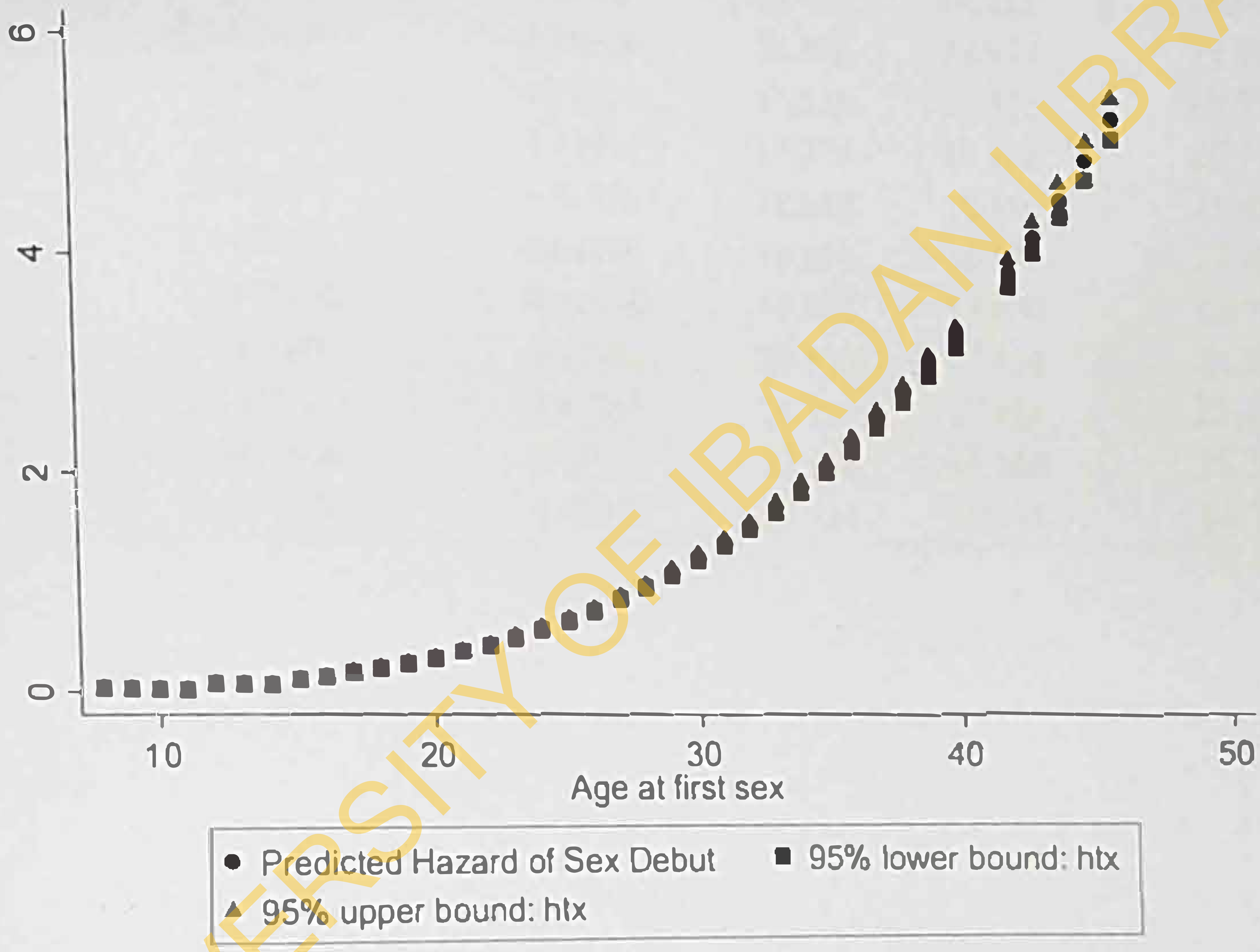


Figure 4.3-2: Predicted hazard function of sex debut

Table 4.3-1: Maximum Likelihood estimates of Weibull parameters for age at first sex

Percentile (p)	Shape parameter (a)	Scale parameter (b)	t(p)	Lower 95% CI	Upper 95% CI
5	2.547E-06	4.41956	9.416	8.895	9.967
10	2.547E-06	4.41956	11.081	10.326	11.891
15	2.547E-06	4.41956	12.223	11.278	13.246
20	2.547E-06	4.41956	13.132	12.019	14.347
25	2.547E-06	4.41956	13.909	12.640	15.304
30	2.547E-06	4.41956	14.602	13.185	16.171
35	2.547E-06	4.41956	15.239	13.677	16.980
40	2.547E-06	4.41956	15.838	14.133	17.750
45	2.547E-06	4.41956	16.412	14.562	18.497
50	***2.547E-06***	***4.41956***	***16.971***	***14.974***	***19.233***
55	2.547E-06	4.41956	17.523	15.376	19.970
60	2.547E-06	4.41956	18.077	15.772	20.719
65	2.547E-06	4.41956	18.642	16.170	21.492
70	2.547E-06	4.41956	19.229	16.576	22.306
75	2.547E-06	4.41956	19.852	17.000	23.183
80	2.547E-06	4.41956	20.534	17.454	24.158
85	2.547E-06	4.41956	21.313	17.961	25.289
90	2.547E-06	4.41956	22.268	18.566	26.707
95	2.547E-06	4.41956	23.634	19.398	28.795

4.3.2. Age at first birth

The estimates of the scale and shape parameters of the Weibull model were estimated as 1.42×10^{-6} (95% CI: -1.96–1.69×10^{-6}) and 4.42 (95% CI: 4.07–4.73) respectively, with a log likelihood of -32.28 (AIC = 68.56). The confidence interval of the shape parameter does not contain unity and this goes to show that the hazard function for age at first birth cannot be modelled with exponential distribution. Figure 4.3-3 shows the survival function of age at first birth highlighting the estimate of median age at first birth.

Table 4.3-2 shows the estimates of the percentiles of age at first birth. The estimated median age at first birth is 19.60 years (95% CI: 17.02–22.56). Hence, 95% of the time, age at first birth is expected to fall within the range of about 17 years to 23 years. The interval is less precise compared to the estimate median from Kaplan-Meier estimates but the results go to show that age at first birth could be estimated using the Weibull distribution. It is observed that width of the 95% CI increased with the percentiles.

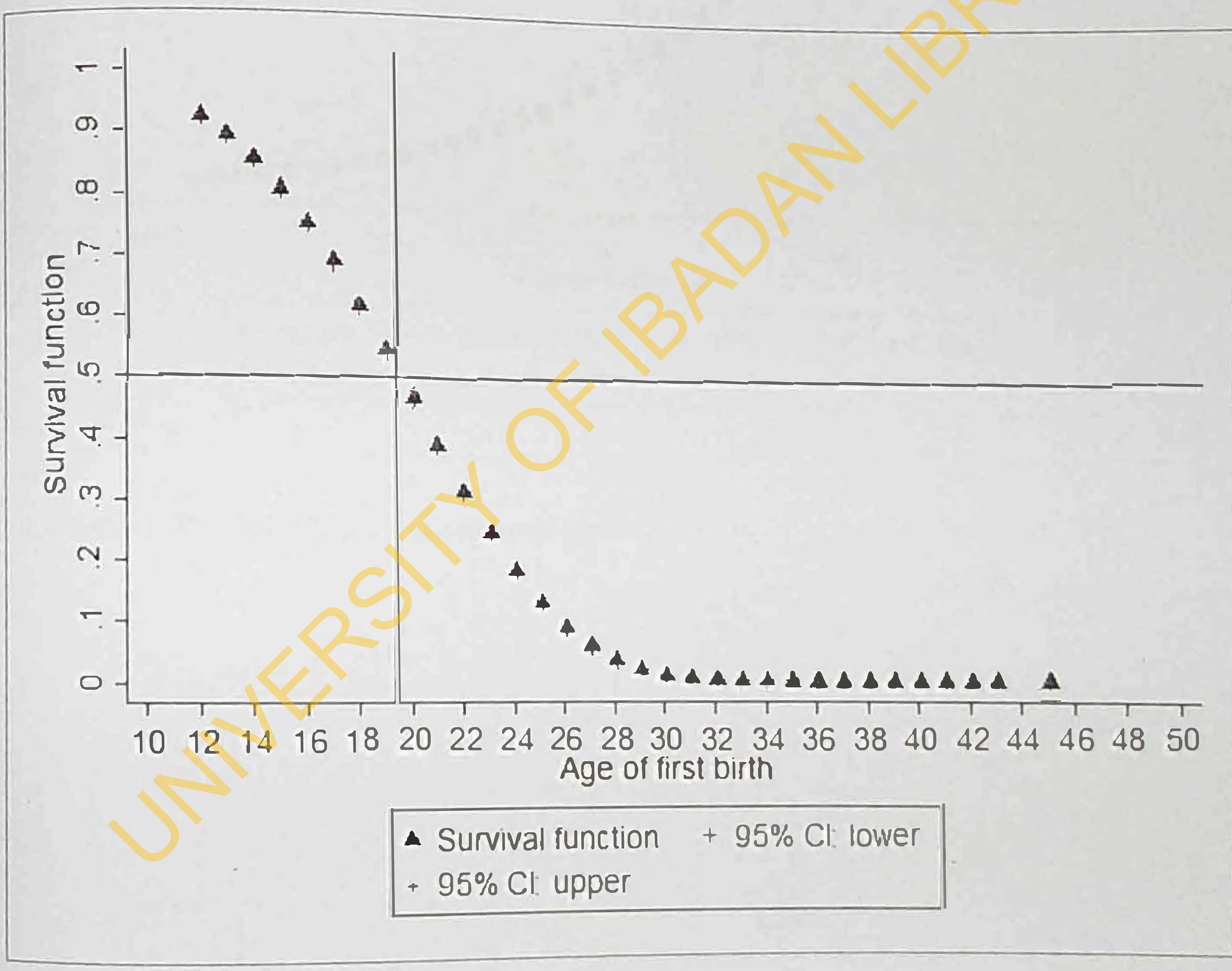


Figure 4.3-3: Survival function of age at first birth highlighting the median age at first birth using Weibull estimates

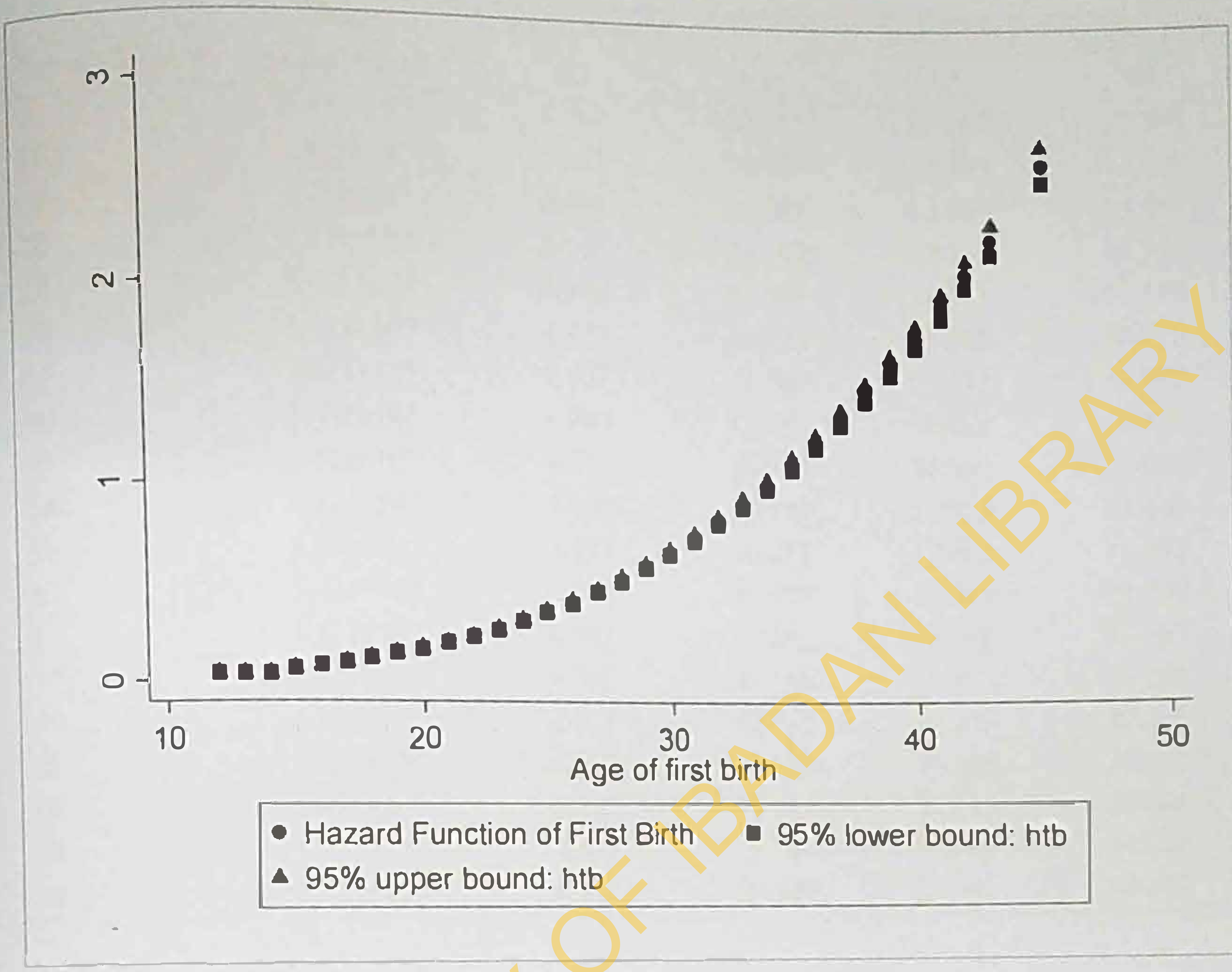


Figure 4.3-4: Predicted hazard function of age at first birth

Table 4.3-2: Maximum Likelihood estimates of Weibull parameters for age at first birth

Percentile (p)	Shape parameter (a)	Scale parameter (b)	t(p)	Lower 95% CI	Upper 95% CI
5	1.421×10^{-6}	4.402	10.847	10.165	11.574
10	1.421×10^{-6}	4.402	12.773	11.789	13.840
15	1.421×10^{-6}	4.402	14.095	12.867	15.441
20	1.421×10^{-6}	4.402	15.147	13.703	16.744
25	1.421×10^{-6}	4.402	16.047	14.403	17.880
30	1.421×10^{-6}	4.402	16.850	15.015	18.910
35	1.421×10^{-6}	4.402	17.589	15.568	19.872
40	1.421×10^{-6}	4.402	18.283	16.079	20.790
45	1.421×10^{-6}	4.402	18.948	16.560	21.681
50	***1.421×10^{-6}***	4.402	***19.596***	***17.020***	***22.561***
55	1.421×10^{-6}	4.402	20.236	17.468	23.442
60	1.421×10^{-6}	4.402	20.878	17.909	24.339
65	1.421×10^{-6}	4.402	21.534	18.352	25.267
70	1.421×10^{-6}	4.402	22.214	18.803	26.244
75	1.421×10^{-6}	4.402	22.937	19.273	27.299
80	1.421×10^{-6}	4.402	23.728	19.775	28.471
85	1.421×10^{-6}	4.402	24.632	20.335	29.836
90	1.421×10^{-6}	4.402	25.740	21.000	31.549
95	1.421×10^{-6}	4.402	27.325	21.911	34.077

4.3.3. Interval between age at first sex and age at first birth

The estimates of the scale and shape parameters of the Weibull model were estimated as 0.22 (95% CI: -1.74--0.23) and 1.26 (95% CI: 1.25–1.27) respectively. The confidence interval of the shape parameter does not contain unity and this goes to show that the hazard function for age at first birth cannot be modelled with exponential distribution. Figure 4.3-5 shows the distribution of interval between sex debut and first birth. Table 4.3-3 shows the estimates of the percentiles. The estimated median first sex to first birth interval was 2.47 years (95% CI: 2.66–2.86). Hence, 95% of the time, the women have their first child after 3 years. The interval is more precise compared to the estimate median from Kaplan-Meier estimates and the results go to show that age at first birth could be estimated using the Weibull distribution.

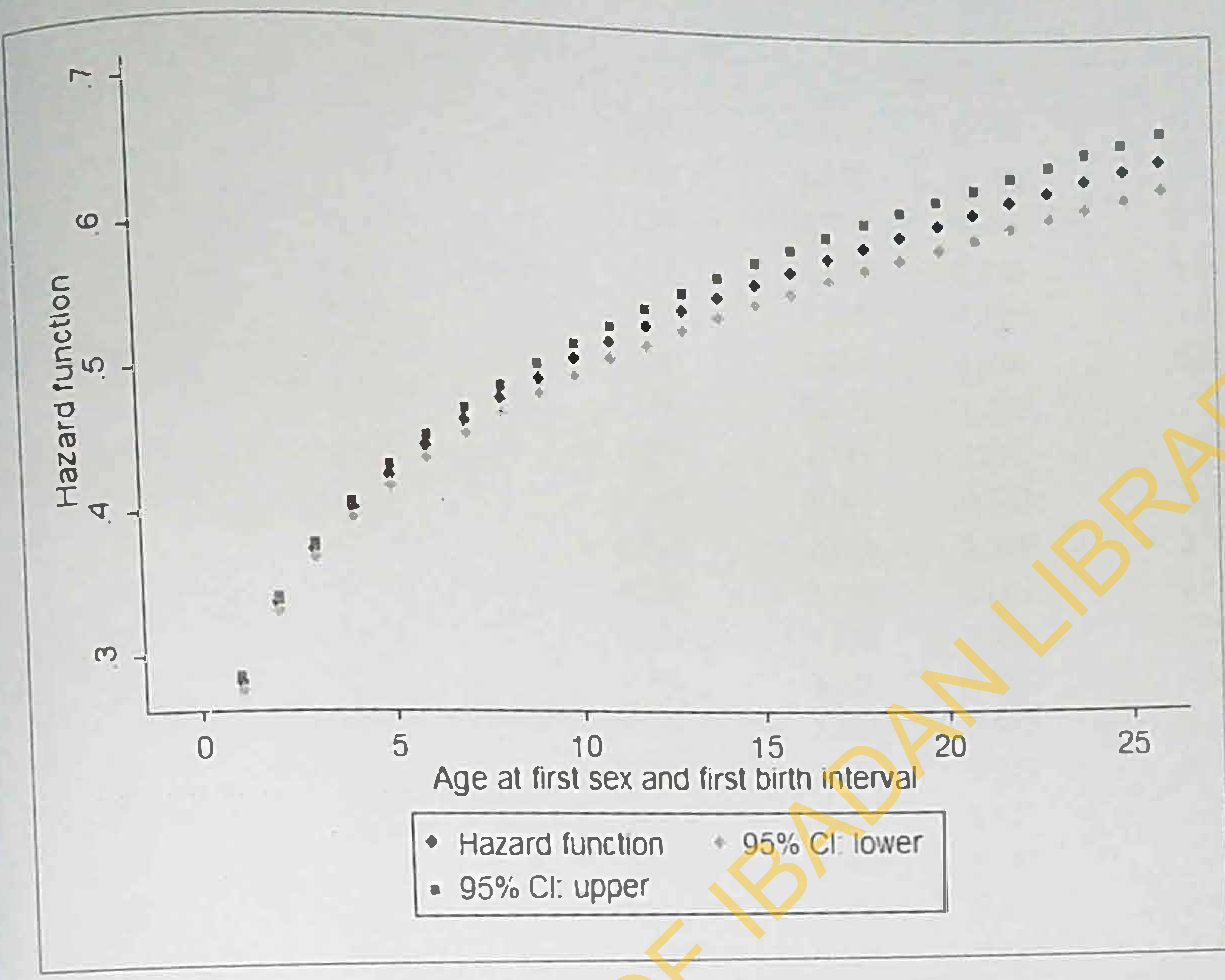


Figure 4.3-5: Hazard function of interval between first sex and first birth

Table 4.3-3: Maximum Likelihood Estimates of Weibull parameters for interval between age at first and age at first birth

Percentile (p)	Scale parameter (a)	Shape parameter (b)	t(p)	Lower 95% CI	Upper 95% CI
5	0.223	1.259	0.312	0.310	0.313
10	0.223	1.259	0.552	0.548	0.556
15	0.223	1.259	0.779	0.771	0.787
20	0.223	1.259	1.002	0.989	1.015
25	0.223	1.259	1.226	1.207	1.245
30	0.223	1.259	1.454	1.428	1.482
35	0.223	1.259	1.690	1.654	1.726
40	0.223	1.259	1.935	1.887	1.983
45	0.223	1.259	2.192	2.132	2.254
50	**0.223**	**1.259**	**2.465**	**2.389**	**2.544**
55	0.223	1.259	2.759	2.663	2.858
60	0.223	1.259	3.077	2.959	3.201
65	0.223	1.259	3.428	3.282	3.582
70	0.223	1.259	3.823	3.641	4.014
75	0.223	1.259	4.276	4.049	4.516
80	0.223	1.259	4.814	4.527	5.119
85	0.223	1.259	5.486	5.115	5.884
90	0.223	1.259	6.398	5.896	6.943
95	0.223	1.259	7.886	7.131	8.722

4.4. Modeling of age at interval of first sex to first birth

4.4.1. Single-level modeling of the interval between age at first sex and age at first birth

Interval between first sex and first birth were fitted into a Level 1 Weibull regression model using the parametric proportional hazard of the model, as presented on Table 4.4-1, with LR $\chi^2(20)$ = 3134.85, $p < 0.001$, AIC = 56115.94. The hazard ratio compares women with wide intervals with women with narrow intervals. As expected, age at first sex was a significant predictor of the interval (HR = 1.08, p<0.001), which implies that the interval between first sex and first birth widens by 8% as the ages of the women increased i.e. women that had their first sex at early ages have narrower interval between first sex and first birth. This is also similar to what is observed with the birth cohort. Older women tend to have wider interval compared to the younger ones. Place of residence was not significant in the model (HR = 0.999, p = 0.996). Compared to women that were Catholic, all the other religions had smaller intervals but the probability of narrower interval between first sex and first birth was not significantly different between Catholics and other Christians (HR = 1.03, p = 0.19). With regards to the educational attainment, the interval increased with increasing education although the interval is lower for women with less than secondary education compared to those with no education.

Having controlled marital status, ever used contraceptives, and ever terminated pregnancy, the result shows that the hazard of narrower first sex to first birth is 15% higher for women that had never been in a union (HR = 1.15, p = 0.001), providing an indication that married women controlled their first birth better than the unmarried women . Women that used contraceptives gave birth significantly later than women that do not use contraceptives after their first sex (HR = 0.94 p = 0.003). Women that ever terminated a pregnancy had small interval of sex debut to first birth compared to women that never terminated a pregnancy (HR = 1.07, p < 0.001).

4.4. Modeling of age at interval of first sex to first birth

4.4.1. Single-level modeling of the interval between age at first sex and age at first birth

Interval between first sex and first birth were fitted into a Level 1 Weibull regression model using the parametric proportional hazard of the model, as presented on Table 4.4-1, with LR $\chi^2(20)$ = 3134.85, $p < 0.001$, AIC = 56115.94. The hazard ratio compares women with wide intervals with women with narrow intervals. As expected, age at first sex was a significant predictor of the interval (HR = 1.08, p<0.001), which implies that the interval between first sex and first birth widens by 8% as the ages of the women increased i.e. women that had their first sex at early ages have narrower interval between first sex and first birth. This is also similar to what is observed with the birth cohort. Older women tend to have wider interval compared to the younger ones. Place of residence was not significant in the model (HR = 0.999, p = 0.996). Compared to women that were Catholic, all the other religions had smaller intervals but the probability of narrower interval between first sex and first birth was not significantly different between Catholics and other Christians (HR = 1.03, p = 0.19). With regards to the educational attainment, the interval increased with increasing education although the interval is lower for women with less than secondary education compared to those with no education.

Having controlled marital status, ever used contraceptives, and ever terminated pregnancy, the result shows that the hazard of narrower first sex to first birth is 15% higher for women that had never been in a union (HR = 1.15, p = 0.001), providing an indication that married women controlled their first birth better than the unmarried women . Women that used contraceptives gave birth significantly later than women that do not use contraceptives after their first sex (HR = 0.94 p = 0.003). Women that ever terminated a pregnancy had small interval of sex debut to first birth compared to women that never terminated a pregnancy (HR = 1.07, p < 0.001).

The estimated values of the scale and shape parameters were a = 0.314 (SE = 0.005) and b = 1.369 (SE=0.006) respectively. With this model, the median of the interval was estimated as 1.78 (95% CI: 1.75–1.82), after controlling for demographic and reproductive health factors, as shown on Table 4.5-2.

Table 4.4-1: Weibull Model fitting for interval between first sex and first birth

Factor	HR	SE	p	95% CI Lower	95% CI Upper
Age at first sex	1.079	0.002	0.000	1.074	1.083
Birth cohort					
15-19					
20-24	0.730	0.025	0.000	0.682	0.782
25-29	0.498	0.017	0.000	0.466	0.532
30-34	0.399	0.014	0.000	0.372	0.427
35-39	0.345	0.012	0.000	0.322	0.371
40-44	0.332	0.012	0.000	0.308	0.357
45-49	0.342	0.013	0.000	0.317	0.367
Type of residence					
Rural					
Urban	0.999	0.016	0.966	0.969	1.030
Religion					
Catholic					
Other Christian	1.033	0.026	0.191	0.984	1.085
Islam	1.226	0.033	0.000	1.163	1.293
Traditionalist	1.156	0.081	0.037	1.008	1.326
Other	3.417	1.212	0.001	1.705	6.846
Educational attainment					
No education					
Incomplete primary	1.151	0.035	0.000	1.085	1.221
Complete primary	1.065	0.024	0.005	1.019	1.112
Incomplete secondary	0.910	0.024	0.000	0.864	0.959
Complete secondary	0.738	0.019	0.000	0.702	0.776
Higher	0.537	0.017	0.000	0.504	0.572
Ever been in a union					
No					
Yes	1.151	0.049	0.001	1.058	1.252
Ever used contraceptives					
No					
Yes	0.944	0.018	0.003	0.909	0.981
Ever terminated pregnancy					
No					
Yes	1.071	0.018	0.000	1.035	1.107
Constant	0.106	0.007	0.000	0.093	0.120
/ln_p	0.314	0.005	0.000	0.304	0.323
p	1.369	0.007		1.356	1.382
1/p	0.731	0.003		0.724	0.738

The model was used to estimate the interval between age at first sex and first birth (Table 4.4-2). Median of the interval was estimated at 1.78 years (95% CI: 1.75 – 1.82). Figure 4.4-1 shows the trend in the interval between sex debut and first birth with a growing 95% CI.

Table 4.4-2: Estimates of median interval between sex debut and first sex using fixed effect Weibull model

Percentile (p)	Shape parameter (a)	Scale parameter (b)	t(p)	Lower 95% CI	Upper 95% CI
5	0.314	1.369	0.266	0.265	0.267
10	0.314	1.369	0.450	0.448	0.453
15	0.314	1.369	0.618	0.614	0.622
20	0.314	1.369	0.779	0.772	0.786
25	0.314	1.369	0.938	0.928	0.948
30	0.314	1.369	1.098	1.084	1.111
35	0.314	1.369	1.260	1.242	1.278
40	0.314	1.369	1.427	1.404	1.450
45	0.314	1.369	1.601	1.572	1.629
50	***0.314***	***1.369***	***1.783***	***1.748***	***1.819***
55	0.314	1.369	1.977	1.934	2.021
60	0.314	1.369	2.186	2.134	2.240
65	0.314	1.369	2.415	2.351	2.481
70	0.314	1.369	2.669	2.591	2.750
75	0.314	1.369	2.959	2.863	3.058
80	0.314	1.369	3.299	3.180	3.423
85	0.314	1.369	3.721	3.569	3.878
90	0.314	1.369	4.286	4.086	4.496
95	0.314	1.369	5.194	4.902	5.504

Figure 4.4-1: Distribution of percentile estimates of the interval between first sex and first birth flanked by the 95% confidence intervals

4.4.2. Comparing Weibull PH and Cox PH

Table 4.4-3 shows comparison between the Weibull PH and the Cox PH. In general, the effect of each of the explanatory variables on the outcome variable is quite comparable except for religion where the *HR* of Traditionalist to Catholic was no longer significant ($HR = 1.096$, $p = 0.187$). Also, the Cox PH provided higher p-values generally compared with the Weibull; and it is inconclusive that the estimates of the *HR* for the Cox PH model was more precise than those of the Weibull PH model, as some of the 95% CI for the Weibull, e.g. *HR* for the cohorts, education level, and contraceptive use from the Weibull PH model are more precise than those of the Cox PH model. Comparing the information criteria, though being used to compare nested models, the AIC for the Cox PH model is about 8 times higher than that of the Weibull PH model. Hence, the results show that the Weibull is amenable modeling the interval between first sex and first birth.

Table 4.4-3: Comparing hazard ratios for the Weibull and Cox PH models

Factor	Weibull PH					Cox PH				
	HR	SE	*p*	95% CI Lower	95% CI Upper	HR	SE	*p*	95% CI Lower	95% CI Upper
Age at first sex	1.079	0.002	0.000	1.074	1.083	1.053	0.002	0.000	1.048	1.057
Birth cohort										
15-19										
20-24	0.73	0.025	0.000	0.682	0.782	0.824	0.029	0.000	0.769	0.882
25-29	0.498	0.017	0.000	0.466	0.532	0.643	0.022	0.000	0.601	0.687
30-34	0.399	0.014	0.000	0.372	0.427	0.566	0.020	0.000	0.528	0.606
35-39	0.345	0.012	0.000	0.322	0.371	0.516	0.018	0.000	0.481	0.553
40-44	0.332	0.012	0.000	0.308	0.357	0.505	0.019	0.000	0.470	0.543
45-49	0.342	0.013	0.000	0.317	0.367	0.523	0.019	0.000	0.486	0.562
Type of residence										
Rural										
Urban	0.999	0.016	0.966	0.969	1.03	1.004	0.016	0.793	0.974	1.035
Religion										
Catholic										
Other Christian	1.033	0.026	0.191	0.984	1.085	1.015	0.025	0.543	0.967	1.066
Islam	1.226	0.033	0.000	1.163	1.293	1.133	0.030	0.000	1.075	1.194
Traditionalist	1.156	0.081	0.037	1.008	1.326	1.096	0.076	0.187	0.956	1.257
Other	3.417	1.212	0.001	1.705	6.846	2.186	0.775	0.027	1.091	4.381
Educational attainment										
No education										
Incomplete primary	1.151	0.035	0.000	1.085	1.221	1.110	0.033	0.001	1.047	1.177
Complete primary	1.065	0.024	0.005	1.019	1.112	1.055	0.023	0.016	1.010	1.101
Incomplete secondary	0.91	0.024	0.000	0.864	0.959	0.944	0.025	0.031	0.897	0.995
Complete secondary	0.738	0.019	0.000	0.702	0.776	0.816	0.021	0.000	0.777	0.857
Higher	0.537	0.017	0.000	0.504	0.572	0.650	0.021	0.000	0.610	0.692
Ever been in a union										
No										
Yes	1.151	0.049	0.001	1.058	1.252	1.093	0.047	0.039	1.005	1.189
Ever used contraceptives										
No										
Yes	0.944	0.018	0.003	0.909	0.981	0.956	0.019	0.019	0.920	0.993
Ever terminated pregnancy										
No										
Yes	1.071	0.018	0.000	1.035	1.107	1.045	0.018	0.011	1.010	1.080
Constant	0.106	0.007	0.000	0.093	0.12					
/ln_p	0.314	0.005	0.000	0.304	0.323					
p	1.369	0.007		1.356	1.382					
1/p	0.731	0.003		0.724	0.738					
LR $\chi^2(20)$	3134.58		0.000			1357.37		0.000		
AIC	56115.94					432078.3				

LR = Likelihood Ratio; AIC = Akaike Information Criteria; HR = Hazard Ratio

4.4.3. Multilevel mixed-effect modeling of interval between age at sex debut and first birth

4.4.3.1. Adjusting for age at first cohabitation

The primary sampling unit (PSU) was included in the parametric model as a mixed effect to control of level to variation the observed data. Table 4.4-4 shows results of the mixed effect model controlling of the primary sampling unit as second-level variable. Unlike what obtained with the fixed effect model, hazard ratio of age at first birth is less than unity ($HR = 0.96$, $p < 0.001$); there is a significant hazard ratio between the urban and rural women ($HR = 1.08$, $p = 0.02$); with regards to religion, only the Muslim women significant HR ($HR = 0.89$, $p < 0.001$) with respect to the Catholic women and the other religions are not significantly different from the Catholic women. Also, the hazard ratio between women with no education and those with incomplete secondary education is not significant ($HR = 0.96$, $p = 0.12$); and hazard of narrower interval was not significant between women that ever terminated pregnancy and those that never did ($HR = 0.985$, $p = 0.43$). Estimated values of the Weibull scale and shape parameters were 0.15 and 6.53 respectively. However, including the age at first cohabitation in the model forced marital status out of the model because of colinearity. Hence another model was run to account for marital status.

4.4.3. Multilevel mixed-effect modeling of interval between age at sex debut and first birth

4.4.3.1. Adjusting for age at first cohabitation

The primary sampling unit (PSU) was included in the parametric model as a mixed effect to control of level to variation the observed data. Table 4.4-4 shows results of the mixed effect model controlling of the primary sampling unit as second-level variable. Unlike what obtained with the fixed effect model, hazard ratio of age at first birth is less than unity ($HR = 0.96$, $p < 0.001$); there is a significant hazard ratio between the urban and rural women ($HR = 1.08$, $p = 0.02$); with regards to religion, only the Muslim women significant HR ($HR = 0.89$, $p < 0.001$) with respect to the Catholic women and the other religions are not significantly different from the Catholic women. Also, the hazard ratio between women with no education and those with incomplete secondary education is not significant ($HR = 0.96$, $p = 0.12$); and hazard of narrower interval was not significant between women that ever terminated pregnancy and those that never did ($HR = 0.985$, $p = 0.43$). Estimated values of the Weibull scale and shape parameters were 0.15 and 6.53 respectively. However, including the age at first cohabitation in the model forced marital status out of the model because of colinearity. Hence another model was run to account for marital status.

Table 4.4-4: Weibull baseline hazard estimates with mixed effects for interval between first sex and first birth

Factor*	HR	SE	p	95% CI Lower	95% CI Upper
Age at first birth	0.955	0.003	0.000	0.949	0.961
Birth cohort					
15-19					
20-24	0.694	0.025	0.000	0.648	0.745
25-29	0.501	0.017	0.000	0.468	0.537
30-34	0.403	0.014	0.000	0.376	0.432
35-39	0.343	0.012	0.000	0.320	0.369
40-44	0.276	0.010	0.000	0.256	0.297
45-49	0.272	0.010	0.000	0.252	0.293
Type of residence					
Rural					
Urban	1.082	0.016	0.000	1.050	1.115
Religion					
Catholic					
Other Christian	1.043	0.026	0.087	0.994	1.094
Islam	0.886	0.023	0.000	0.841	0.933
Traditionalist	0.899	0.061	0.115	0.788	1.026
Other	1.799	0.638	0.098	0.898	3.603
Educational attainment					
No education					
Incomplete primary	1.250	0.037	0.000	1.180	1.323
Complete primary	1.135	0.025	0.000	1.088	1.184
Incomplete secondary	0.960	0.025	0.118	0.912	1.010
Complete secondary	0.881	0.022	0.000	0.839	0.925
Higher	0.794	0.025	0.000	0.746	0.845
Age at first cohabitation	0.841	0.002	0.000	0.837	0.846
Ever used contraceptives					
No					
Yes	1.232	0.021	0.000	1.192	1.274
Ever terminated pregnancy					
No					
Yes	0.985	0.019	0.429	0.949	1.022
Constant	0.000	0.000	0.000	0.000	0.000
Random effect Factor					
Primary Sampling Unit (Community) variance	0.047				
/ln_p	1.876	0.004	0.000	1.868	1.884
p	6.526				
1/p	0.153				
AIC	-26504.65				

*Ever been in a union was excluded because of colinearity

4.4.3.2. *Adjusting of marital status*

In order to account for differences between women that were ever in a union and those that never, age at cohabitation was removed in lieu of ever had a union or not. The result is present in Table 4.4-5. Effect of age at first sex, birth cohorts, educational attainment, and contraceptive use are similar to the previous mixed effect model but generally with lower estimates of hazard ratios, more precise confidence intervals, but larger AIC. However, the region is now significant in the model, which implies that marital status strongly affects the interval between first birth and birth sex as a result of the fact that more than four-fifth of the women were had ever been in a union, although not all the births were within marriage. Hazard of narrower interval between ages at first sex and first birth was 12% ($HR = 1.12$, $p = 0.04$) higher for women that had ever been in a union compared to women that had never been in a union.

Table 4.4-5: Weibull baseline hazard estimates with mixed effects for interval between first sex and first birth controlled for women ever in a union

Factor	Weibull PH Mixed Model				
	HR	SE	*p*	95% CI Lower	95% CI Upper
Age at first birth	0.834	0.002	0.000	0.831	0.838
Birth cohort					
15-19					
20-24	0.707	0.024	0.000	0.662	0.755
25-29	0.496	0.016	0.000	0.465	0.529
30-34	0.368	0.013	0.000	0.344	0.394
35-39	0.303	0.011	0.000	0.283	0.324
40-44	0.273	0.010	0.000	0.254	0.293
45-49	0.269	0.010	0.000	0.251	0.289
Type of residence					
Rural					
Urban	0.999	0.015	0.947	0.970	1.029
Religion					
Catholic					
Other Christian	1.128	0.027	0.000	1.076	1.182
Islam	1.324	0.034	0.000	1.258	1.393
Traditionalist	1.117	0.075	0.100	0.979	1.274
Other	2.591	0.866	0.004	1.345	4.989
Educational attainment					
No education					
Incomplete primary	1.096	0.032	0.002	1.035	1.160
Complete primary	1.048	0.022	0.029	1.005	1.092
Incomplete secondary	0.850	0.022	0.000	0.808	0.893
Complete secondary	0.708	0.017	0.000	0.676	0.743
Higher	0.531	0.016	0.000	0.499	0.564
Ever been in a union					
No					
Yes	1.120	0.044	0.004	1.036	1.211
Ever used contraceptives					
No					
Yes	1.113	0.018	0.000	1.078	1.149
Ever terminated pregnancy					
No					
Yes	0.986	0.019	0.470	0.951	1.024

Factor	Weibull PH Mixed Model				
				95% CI	
	HR	SE	*p*	Lower	Upper
Constant	0.000	0.000	0.000	0.000	0.000
Random Effect					
PSU (sd)	0.000	0.000	0.000	0.000	0.000
/ln_p	1.792	0.004	0.000	1.784	1.800
P	6.000				
1/p	0.167				
AIC	-100199				

Chapter Five

Discussions

5.0. Introduction

This section provides insight into the result that were obtained from the analyses and also discusses some of the limitations to the study. First, the description of the data is discussed then the model fitting for the baseline hazard and lastly, the parametric multilevel modelling. The objectives of the study were finally discussed.

The study had looked at fitting the baseline hazard of age at first sex, age at firth birth, and their interval to a parametric survival model – the Weibull model, using both graphical and analytical approaches, using data obtained from the Nigeria DHS 2013. Description of the individual-level and community-level variables was provided. Study variables were also described and factor-described by individual and community level variables. Cumulative log hazard plot was used to investigate the suitability of the data to a Weibull distribution and a general parametric survival model that uses Guassian quadrature to estimate the cumulative hazard function was used to fit the baseline hazard to Weibull distributions. Parameter estimates were used to predict quintiles of the survival variables, which were compared with the estimates from the Kaplan-Meier analysis. Afterwards, fixed effect and random effect parametric models were fitted and compared with the semi-parametric Cox proportional hazards model. All parameter estimates were done using maximum likelihood estimation.

5.1. Results from exploratory data analyses

Although the data was obtained from a secondary source (Nigeria DHS 2013) among women of reproductive age that ever had sex, the results obtained from the data provide additional insight in

Chapter Five

Discussions

5.0. Introduction

This section provides insight into the result that were obtained from the analyses and also discusses some of the limitations to the study. First, the description of the data is discussed then the model fitting for the baseline hazard and lastly, the parametric multilevel modelling. The objectives of the study were finally discussed.

The study had looked at fitting the baseline hazard of age at first sex, age at firth birth, and their interval to a parametric survival model – the Weibull model, using both graphical and analytical approaches, using data obtained from the Nigeria DHS 2013. Description of the individual-level and community-level variables was provided. Study variables were also described and factor-described by individual and community level variables. Cumulative log hazard plot was used to investigate the suitability of the data to a Weibull distribution and a general parametric survival model that uses Guassian quadrature to estimate the cumulative hazard function was used to fit the baseline hazard to Weibull distributions. Parameter estimates were used to predict quintiles of the survival variables, which were compared with the estimates from the Kaplan-Meier analysis. Afterwards, fixed effect and random effect parametric models were fitted and compared with the semi-parametric Cox proportional hazards model. All parameter estimates were done using maximum likelihood estimation.

5.1. Results from exploratory data analyses

Although the data was obtained from a secondary source (Nigeria DHS 2013) among women of reproductive age that ever had sex, the results obtained from the data provide additional insight in to the reports of the 2013 Nigeria DHS. Age at sex debut, age at first birth and interval between the two have been described as reproductive health milestones. Literature showed that these two events could be used in lieu of age at first marriage to describe the fertility of women especially given the increasing incidences of premarital births.

The analyses covered women in the reproductive age range 15-49 years but the results also showed that there is considerable variation in the reproductive health milestones. Despite the efforts of the government to increase literacy rates among women and with 63% literacy rate among women (NPC [Nigeria] and ICF International, 2014) one would expect that the age at first birth and sex debut would rise but the reverse is the case. Need for educational attainment and career fulfilment have been described to be protective of age at first sex and first marriage but the current study shows that this is nor largely the case. However, the observation is due to the low educational status of women in Northeastern and Northwestern geopolitical zones of Nigeria. Early exposure to sexual intercourse and low contraceptive use was quite higher among the Muslims, majority of whom are from the Northeastern and Northwester regions of Nigeria, compared to other religions. Women that had never been in a union and women that had never had a birth were also found to have been using contraceptives more than women that were married.

5.2. Description of study RH milestones

Although the essence of this study was not to model the RH events in themselves but their baseline hazard of occurrence, it was necessary to provide insight into the nature and characteristics of the raw data. The median age at these RH events as well as their interval has declined over time. With respect to the birth cohorts, median age at first sex declined from 17 years among the 40-44 years birth cohort to 15 years among the 15-19 years birth cohort and the median age at first birth also decreased from 20 years to 16 years among the birth cohorts aforementioned respectively. This is despite the fact that the expected median age at first marriage in Nigeria is 18 years among women. The results corroborate the assertion in the NDHS that men and women in Nigeria tend to initiate sex before marriage (NPC [Nigeria] and ICF International, 2014). Current trend in non-marital cohabitation would continue to encourage early sexual debut, long-term exposure to multiple sexual partners, and may not curtail the low contraceptive prevalence. The interval between first sex and first birth has been shrinking pointing to the fact that contraceptive use is low but does not decry the fact that a number of births now take place within the ambient of marital union. The shrinking interval corroborates the work of Feng and Quanhe (1996) where they showed a reducing trend in interval between marriage and first birth. On the other hand, unlike the works of Feng & Quanhe (1996) and those of Finer & Philbin (2014) where they showed a growing trend in age at first birth and age at first marriage, this study shows the reverse. Women, over time, are giving experience sexual debut earlier, marry earlier, and consequently give birth earlier.

5.3. Graphical model fitting

The first objectives of this study were to investigate the probability distribution of baseline hazard of age at first birth, age at sex debut and the interval between them. Attempt was made to fit the baseline hazard to Weibull distribution. The log cumulative hazard plot showed that Weibull distribution could fit the baseline hazard of the RH events. The survival function was transformed into a log cumulative hazard equation that could be evaluated as a linear equation and the log cumulative hazard plot obtained was near straight line. Estimates of the shape and scale parameters were used to predict the percentiles of RH events. Compared to the median values predicted by the Kaplan-Meier estimates the estimates of the parametric distribution were less precise. For instance the values of the KM estimates for median age at sex debut was 16 years (95% CI: 15.94-16.05) while that of Weibull model was 16.6 years (95% CI: 16.43–16.80). On the other hand, the

KM estimates of the median interval between sex debut and first birth was 2 years (95% CI: 1.97–2.03) and less precise compared to the parametric estimate of 1.98 (95% CI: 1.96–2.01) years. Although the interval between sex debut and first birth could have been assumed to follow a normal distribution, fitting a Weibull distribution showed more likelihood of a Weibull given the results.

5.4. Analytical model fitting

Similar to the graphical model fitting, the baseline hazard of the RH events was estimated using the Weibull distribution. The estimation of the hazard function using maximum likelihood estimation of the Weibull parameters was assessed. The Weibull distribution showed a good capacity to estimate the median values of the age at first sex, age at first birth and the interval between them. The estimates of the interval were more precise than the other variables. The stgenreg estimation proposed by Crowther and Labert (2013) did not particularly produce better estimates than the streg module of Stata however, stgenreg is very amenable to most parametric survival time models.

The findings from the comparison of Weibull PH model and Cox PH model is very sunilar to the finding of Stewart (2010, pp. 226-228) who also found that the result of the parametric method is similar to Cox PH with major disparities when comparing birth cohort years, although she compared the single-level Cox with multi-level logistic regression model and her estimations were done by penalized least squares estimation on MLwiN. Her work also attests to the fact that inferences from the parametric methods are similar to the Cox PH, which produces lower estimates.

The Weibull distribution has been described as being very robust to accommodate several survival data. First it is a generalization of the exponential model and these two are the only ones than can be estimated using the accelerated failure time forms.

5.5. Multilevel modeling

Fixed factor modelling showed that age at first sex is a significant predictor of the interval between first sex and first birth. Women that had sex early tend to have lower interval. Furthermore, the interval increased with birth cohort. Older women had larger interval compared to the young cohorts but the interval was not different between rural and urban women. Catholics and other Christians had higher sex debut to first birth interval, which could be attributed to the fact that most Muslims lived in northern parts of Nigeria where there is very low contraceptive use and under-age marriages are common. Women in this zones of the country usually experience sexual debut during their first marriage because of strict religious demands about sexuality that could have severe consequences if violated. Increased education has been linked with later age at sex debut and later birth. This fixed factor model shows that that the interval between first sex and first birth could also be increased by education attainment of the women with more education implying increasing interval. Authors have showed that educational status has been linked with contraceptive use and the univariate analyses show that women with higher education use contraceptives twice as much as women with primary education and 12 times more than women with no education.

Ever being in a union and contraceptive use increased the interval while pregnancy termination decreased the interval, which also is similar to the findings of Subaiya and Johnson (2008), who found that having a birth prior to first union, initiating sexual activity while still in school, number of lifetime partners, terminating a pregnancy, and practicing contraception as significantly associated with longer intervals between first sex and first marriage, which eventually influences first birth as birth is usually planned with marriage.

The estimated median interval was 1.78 years which is less than 2 years from the Kaplan Meier estimates. However the fixed factor estimate has been controlled for demographic and RH characteristics. Given early marriages in some parts of the country, change in age at sex debut and first birth over time then it is expected that the interval be less than that predicted by the KM estimates.

Controlling of community level variation did not particularly alter the results obtained in the fixed model. The estimates of the hazard ratios for the multi-level model were more precise and hence more reliable. Furthermore, the multilevel modeling showed that accounting for the hierarchical nature of the data provided better insight into the nature of the variables. For instance, age at first sex was predictive in the fixed effect model but became protective in the random effect model, the hazard ratio of urban-rural became significant in the model, only the Muslims because significantly different from the Catholics unlike what we had in the fixed effect model. Exclusion of marital status greatly impacted on the estimates of other variables in the model as well as the information criteria for the model. Hence the variable appeared to be of utmost influence. This is so because the experience of child birth between women that ever married or were in a union is not generally comparable to those that had never married. This is also supported by the contraceptive use that is different between the groups.

Comparing model with age at first cohabitation and model with marital status provided insight into how much difference the two variables can introduce into a model. Including age at first cohabitation showed that even in the mixed effect model, the interval between age at first sex and first birth may vary by the age of the women which is different between the urban-rural divide, but the interval is not significantly different across the urban-rural divide when controlling for marital status (i.e. ever in a union). Hirschman & Rindfuss (1980), perhaps one of the earliest to study the interval between age at first marriage and first birth at national level across three cultures, mentioned the high correlation between age at first marriage and first birth and if age at first marriage is included in a regression of age at first birth on other covariates, age at first marriage mediates almost all the other covariates. Hence, if age at first birth too had been included in the model, it would increase the hazard ratio of age at first sex by 400% and mediate the other factors. Hence, it was excluded. In most culture, marriage at child birth a usually planned together (Hirschman & Rindfuss, 1980).

Chapter Six

Conclusions and Recommendations

6.1. Conclusion

The Weibull distribution fitted age at first sex, age at first birth and their interval. Age at first sex, education, religion, women's birth cohort, contraceptive use, marital status, and ever terminated a pregnancy were significant predictors of the interval between first sex and first birth.

Although the Weibull proportional hazard model produced less precise estimates compared to the Cox proportional hazard model, the conclusions are quite comparable. Also the estimates of the Weibull model are also not different from Kaplan-Meier estimates. Furthermore, the estimates of the Akaike Information Criteria (AIC) showed that parametric proportional hazard modelling of survival data in reproductive health, in this case the Weibull proportional hazard model, is preferable to the Cox PH model.

6.2. Recommendations

This study has taken a bold step to investigate the parametric distribution of baseline hazard of sexual reproductive events and their intervals. I would love to recommend that

1. Further studies be engaged to compare other parametric models, although the Weibull has been described to be very robust.
2. the study be investigated with statistical software that would accommodate more than two-level parametric modeling of survival data
3. the multilevel mixed effects parametric survival modeling utilizing MLE through adaptive and non-adaptive Gauss-Hermite quadrature is still relatively new with few studies having

Chapter Six

Conclusions and Recommendations

6.1. Conclusion

The Weibull distribution fitted age at first sex, age at first birth and their interval. Age at first sex, education, religion, women's birth cohort, contraceptive use, marital status, and ever terminated a pregnancy were significant predictors of the interval between first sex and first birth.

Although the Weibull proportional hazard model produced less precise estimates compared to the Cox proportional hazard model, the conclusions are quite comparable. Also the estimates of the Weibull model are also not different from Kaplan-Meier estimates. Furthermore, the estimates of the Akaike Information Criteria (AIC) showed that parametric proportional hazard modelling of survival data in reproductive health, in this case the Weibull proportional hazard model, is preferable to the Cox PH model.

6.2. Recommendations

This study has taken a bold step to investigate the parametric distribution of baseline hazard of sexual reproductive events and their intervals. I would love to recommend that

1. Further studies be engaged to compare other parametric models, although the Weibull has been described to be very robust.
2. the study be investigated with statistical software that would accommodate more than two-level parametric modeling of survival data
3. the multilevel mixed effects parametric survival modeling utilizing MLE through adaptive and non-adaptive Gauss-Hermite quadrature is still relatively new with few studies having

tried it out. I suggest that studies on SRH milestones be studied more effectively with this new development in Stata.

4. Recall information introduces a lot of bias to large-scale surveys such as the Nigeria DHS. I would suggest that calendars and response cards be employed as tools to enhance quality of response for past events (McDonald, et al., 2003)

References

Abdelkader, Y., & Al-Marzouq. (2010, February). Probability distribution relationships. *International Journal of Basic & Applied Sciences, 10*(1), 48-56. doi:1001-91310-3434

Adebowale, S., Fagbamigbe, F., Okareh, T., & Ganiyu, O. (2012). Survival analysis of timing of first marriage among women of reproductive age in Nigeria: Regional Difference. *African Journal of Reproductive Health, 16*(4), 95-107.

Agbim, K., & Ikyemum, J. (2013). Fertility and fertility determinants: A nexus for understanding the need to manage the fertility of women In Awka Town, Anambra State, Nigeria. *The International Journal Of Engineering And Science (IJES), 2*(8), 48-56.

Akpa, O., & Ikpotokin, O. (2012). Modeling the determinants of fertility among women of childbearing age in Nigeria: Analysis using generalized linear modeling approach. *International Journal of Humanities and Social Science, 2*(18), 167-176. Retrieved January 2015, from www.ijhssnet.com

Albright, J., & Marinova, D. (2010). *Estimating multilevel models using SPSS, Stata, SAS and R.* Retrieved Januar 2015, from http://ww.indiana.edu/~statmath/stat/all/hlm/hlm.pdf

Al-Fawzan, M. (2000). *Methods of estimating the parameters of the Weibull distribution.* (A. Godbole, Ed.) Retrieved January 2015, from InterStat: http://interstat.statjournals.net/YEAR/2000/articles/0010001.pdf

Amin, S., & Bajracharya, A. (2011). Marriage and first birth intervals in early and late marrying societies: An exploration of determinants. Washington DC: Population Council.

Amusan, A., & Mohd, K. (2014, April). Survival modeling of first birth interval after marriage. *Life Science Journal, 11*(7), 299-307. Retrieved January 2015, from http://www.lifesciencesite.com/lsj/life1107/035_23903life110714_299_307.pdf

Angeles, G., & Mroz, T. (2001). A guide to using multilevel models for the evaluation of program impacts. *Measure Evaluation Working Paper Series No. 33, 01*(33).

Aremu, O. (2013, December 3). The influence of socioeconomic status on women's preference for modern contraceptive providers in Nigeria: a multilevel choice modeling. *Patient Preference and Adherence, 7*, 1213-1220.

Atuhaire, L. (2011, May). Application of current-status survival analysis methodology to estimation of age at first sex: Uganda. *The African Statistical Journal, 12*, 141-150.

Bartkute, V., & Sakalauskas, L. (2008). The method of three-parameter Weibull distribution estimation. *Acta et Commentationes Universitatis Tartuensis de Mathematica, 12*, 65-78.

Basia, P., Ties, B., Elizabet, P., & Nahum, B. (2002). Estimation of levels and trends in age at first sex from surveys using survival analysis. *MEASURE Evaluation.*

Chandrasekhar, S. (2010, July). Factors affecting age at first marriage and age at first birth in India. *Journal of Quantitative Economics, 8*(2), 81-97.

Collet, D. (1994). The Weibull model for survival data. In D. Collet, C. Chatfield, & J. Zidek (Eds.), *Modelling Survival Data in Medical Research* (pp. 107-153). London: Chapman and Hall.

Cook, J. (2014). *Diagram of distribution relationships*. Retrieved January 20 2015, from http://www.johndcook/blog/distribution_chart

Copen, C., K., D., Vespa, J., & Mosher, W. (2012). *First Marriages in the United States: Data From the 2006-2010 National Survey of Family Growth. National Health Statistics Report.* Center for Disease Control and Prevention (CDC). Hyattsville, MD: National Center for Health Statistics.

Cremin, I., Mushati, P., Hallet, T., Mupambireyi, Z., Nyamukapa, C., & Garnett, G. (2009). Measuring trends in age at first sex and age at marriage in Manicaland, Zimbabwe. *Sex Transm Infect, 85*(1), 34-40.

Crowther, M., & Lambert, P. (2013). stgenreg: A Stata package for general parametric survival analysis. *Journal of Statistical Software, 53*(12), Online. Retrieved 2015, from http://www.jstatsoft.org/v53/i12/paper

Crowther, M., Look, M., & Riley, R. (2014, May 1). Multilevel mixed effects parametric survival models using adaptive Gauss-Hermite quadrature: with application to recurrent events and IPD meta-analysis. *Statistics in Medicine*. doi:10.1002/sim.6191

Dätwyler, C., & Stucki, T. (2011, May 9). *Parametric Survival Models*. Retrieved November 20, 2014, from Seminar for Statistics: https://stat.ethz.ch/education/semesters/ss2011/seminar/contents/handout_9.pdf

Eric, Y., & Eleanor, M. (2008, December). Factors influencing the timing of first sexual intercourse among young people in Nyanza, Kenya. *International Family Planning Perspectives, 34*(4), 177-188.

Fan, J., & Jiang, J. (2008). Non- and semi-parametric modeling in survival analysis. In J. Fan, X. Lin, & J. Liu, *New Developments in Biostatistics and Bioinformatics* (pp. 3-33). New Jersy: World Scientific.

Feng, W., & Quanhe, Y. (1996). Age at marriage and the first birth interval: The emerging change in sexual behavior in young couples in China. *Population and Development Review, 22*(2), 29-320.

Feyisetan, B., & Bankole, A. (2002). Fertility transition in Nigeria: trends and prospects. *Completeing The Fertility Transition* (pp. 506-528). New York: United Nations Department of Economic and Social Affairs. Retrieved from www.un.org/esa/population/publications/completingfertility/RevisedBANKOLEpaper.pdf

Finer, L., & Philbin, J. (2014, February 3). Trends in ages at key reproductive transitions in the United States, 1951–2010. *Women's Health Issues, xxx*(xx), e1-e9.

Garenne, M. (2004). Age at marriage and modernisation in Sub-Saharan Africa. *South African Journal of Demography,* 9(2), 56-79. Retrieved 2015, from http://www.commerce.uct.ac.za/Organisations/Demography/SA_Journal_of_Demography/sajd/volume%209%202/sajdem_2004_9_2_garenne.pdf

Gayawan, E., & Adebayo, S. (2013, June 26). A Bayesian semiparametric multilevel survival modelling of age at birth in Nigeria. *Demographic Research, 28*(45), 1339-1372. doi:10.4054/DemRes.2013.28. 45

Gayawan, E., & Adebayo, S. (2014). Spatial pattern and determinants of age at marriage in Nigeria using a geo-additive survival model. *Mathematical Population Studies: An International Journal of Mathematical Demography, 21*(2), 112124. doi:10.1080/08898480.2014.892336

Gharoro, E., & Igbafe, A. (2002). Maternal age at first birth and obstetric outcome. *Nigerian Journal of Clinical Practice, 5*(1), 20-24. Retrieved from www.ajol.info/index.php/njcp/article/view/11079

Hallett, T., Gregson, S., & Lewis, J. (2007). Behaviour change in generalised HIV epidemics, impact of reducing cross-generational sex and delaying age at sexual debut. *Sex Transm Infect, 83*(1), 50-54.

Hirschman, C., & Rindfuss, R. (1980). Social, cultural and economic determinants of age at birth of first child in Peninsular Malaysia. *Population Studies, 34*(3), 507-518. Retrieved 2015

Hogan, D., Sun, R., & Cornwell, G. (2000, September). Sexual and fertility behaviors of American females aged 15-19 years: 1985, 1990, and 1995. *American Journal of Public Health, 90*(9), 1421-1425. Retrieved 2015, from http://ajph.aphapublications.org/doi/pdf/10.2105/AJPH.90.9.1421

IBBSS 2010. (2010). *HIV Intergated Biological and Behavioural Surveillance Survey.* Abuja: Federal Ministry of Health [Nigeria].

IBM Corporation. (2013). IBM SPSS Statistics Version 22.

Jensen, R., & Thornton, R. (2003, July). Early female marriage in the developing world. *Gender and Development, 11*(2), 9-19. Retrieved 2015, from http://www.jstor.org/stable/4030636

Johnson, L., Dorrington, R., Bradshaw, D., Wyk, V., & Rehle, T. (2009, September 10). Sexual behaviour patterns in South Africa and their association with the spread of HIV: Insights from mathematical model. *Demographic Research, 21*(11), 289-340.

Kaestle, C., Halpern, C., Miller, W., & Ford, C. (2005). Young age at first sexual intercourse and sexually transmitted infections in adolescents and young adults. *American Journal of Epidemiology, 161*(8), 774-780. doi:10.1093/aje/kwi095

Khan, J., Rosenthal, S., Succop, P., Ho, G., & Burk, R. (2002). Mediators of the association between age of first sexual intercourse and subsequent human papillomavirus infection. *PEDIATRICS, 109*(1), 1-8. Retrieved November 30, 2014, from http://www.pediatrics.org/cgi/content/full/109/1/ e5

Krishnamoorthy, K., Mathew, T., & Xu, Z. (2014). Standardized likelihood inference for the mean and percentiles of a lognormal distribution-based on samples with multiple detection limits. *Journal of Environmental Statistics, 6*(5), 1-18. Retrieved 2015, from http://www.jenvstat.org/v06/i05/paper

Kumchulesi, G., Palamuleni, M., & Kalule-Sabiti, I. (2011). Factors affecting age at first marriage in Malawi. *Sixth African Population Conference.* Ouagadougou - Burkina Faso: Pampa.

Leemis, L., & McQuestion, J. (2008, February). Univariate distribution relationships. *The American Statistician - Teacher's Corner, 62*(1), 45-53. doi:10.1198/000313008X27040448

Lei, Y. (2008). Evaluation of three methods for estimating the Weibull distribution parameters of Chinese pine (Pinus tabulaeformis). *Journal of Forest Science, 54*(12), 566-571.

Lisonkova, S., Jansses, P., Sheps, S., Lee, K., & Dahlgren, L. (2010). The effect of maternal age on adverse birth outcomes: does parity matter? *Journal of Obstetrics and Gyneacology Canada, 32*(6), 541-548. Retrieved from www.jogc.com/abstracts/full/201006_Obstetrics_1.pdf

Lofstedt, P., Ghilagaber, G., Shusheng, L., & Johansson, A. (2005, September). Changes in marriage age and first birth interval in Huaning County, Yunnan Province, PR China.

Southeast Asian J. Trop Med Public Health, 36(5), 1329-1338. Retrieved December 10, 2014

Luke, N. (2003). Age and economic asymmetries in t he sexual relationships of adolescent girls in Sub-Saharan Africa. *Studies in Family Planning, 34*(2), 67-86.

MacQuarrie, K. (2014). *Marriage timing, marriage regime and timing the first birth in India.* Retrieved November 2014, from International Union for the Study of Population: http://iussp.org/sites/default/files/event_call_for_papers/MacQuarrie%20Marriage%20IUSSP%20Abstract.pdf

McDonald, J., Burnett, N., Coronado, V., & Johnson, R. (2003). Questionnaire Desgin. In U. D. Services, *Reproductive Health Epidemiology Seires- Module 4.* Atlanta, Georgia, USA: Department of Health and Human Services, CDC.

Mensch, B., & Singh, S. C. (2005). Trends in the timing of first marriage amond men and women in the developing world. *Policy Research Division Working Papers.* Retrieved from www.popcouncil.org/publications/wp/prd/rdwplist.html

Microsoft Corporation. (2012). Microsoft Office Excel 2013. Santa Rosa, California.

NARHS Plus. (2013). *National HIV & AIDS and Reproductive Health Survey 2012 (NARHS Plus).* Abuja: Federal Ministry of Health, Nigeria.

Ngalinda, I. (1998). *Age at first birth, fertility and contraception in Tanzania.* Berlin: Humbolt University of Berlin.

Nigeria GARPR. (2014). *Nigeria Global AIDS Response Country Progress Report.* Abuja: National AIDS Control Agency.

NIST/SEMATECH e-Handbook of Statistical Methods. (2013). Lognormal Distribution. In NIST/SEMATECH, *Engineering Statistics Handbook.* NIST. Retrieved 2015, from http://www.itl.nist.gov/div898/handbook/eda/section3/eda3669.htm

NPC [Nigeria] and ICF International. (2014). *Nigeria Demographic and Health Survey 2013.* Abuja, Nigeria and Rockville, Maryland, USA: Nigeria Population Commission (NPC) [Nigeria] and ICF International.

NPC [Nigeria] and ICF Macro. (2009). *NIgerian Demograhic and Health Survey 2008.* Abuja, Nigeria: Nigerian Population Commission and ICF Macro.

NPC and ICF International. (2014). *Nigeria Demographic and Health Survey 2013.* Abuja, Nigeria and Rockville, Maryland, USA: Nigeria Population Commission (NPC) [Nigeria] and ICF International.

Okonufua, F. (2000). Adolescent reproductive health in Africa: future challenges. *African Journal or Reproductive Health, 4*(1), 7-12.

Oladokun, A., Morhason-Bello, I., Enakpene, C., Owonikoko, K., Akinyemi, J., & Obisesan, J. (2007). Sexual behaviour and contraceptive usage of secondary school adolescents in Ibadan, Nigeria. *Journal of Reproduction and Contraception, 18*(4), 279-288. doi:10.1016/S1001-7844(07)60034-0

Olalekan, A. (2008, May 30). Geographical variations and contextual effects on age of initiation of sexual intercourse among women in Nigeria: a multilevel and spatial analysis. *International Journal of Health Geographics, 7*(27).

Oyedokun, A., & Odumegwu, C. (2013). Factors influencing first sexual intercourse for South African youth . *International Union for Scientific Study of Population.* Busan, Republic of Korea.

Oyefara, J. (2012). Women's age at first birth and knowledge of family planning methods in Yoruba society, Nigeria. *Journal of Sociological Research, 3*(2), 249-271. doi:10.5296/jsr.v3i2.2394

Rodríguez, G. (2010). *Parametric Survival Models.* Retrieved Novmber 2014, from Germán Rodríguez - Pop509: http://data.princeton.edu/pop509/ParametricSurvival.pdf

Shayan, Z., Ayatollahi, S., Zare, N., & Moradi, F. (2014, February). Prognostic factors of first birth interval using the parametric survival models. *Iranian Journal of Reproductive Medicine, 12*(2), 125-130. Retrieved 2015, from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4009565/pdf/ijrm-12-125.pdf

Shretha, G., & Shretha, S. (2008). Nonparametric distribution fitting of age at first marriage of Nepalese women through scatter plot smoother. *Nepal Journal of Science and Technology, 9*, 179-186.

Simeon, A., & Khalid, Z. (2014). Survival modelling of first birth interval after marriage. *Life Science Journal, 11*(7), 299-307. Retrieved from http://www.lifesciencesite.com

Singh, S., & Samara, R. (1996). Early marraige among women in developing countries. *International Family Planning Perspectives, 22*(4), 148-157. doi:http://www.alanguttmacherinstitute.com/pubs/journals/2214896.pdf

Song, W. (2005). Relationships among some univariate distributions. *IIE Transactions, 37*, 651-656. doi:10.1080/07408170590948512

StataCorp LP. (2012, February). Stata/IC 12.1 for Windows. College Station, Texas, USA.

Stewart, C. (2010). *Multilevel modelling of event history data: comparing methods appropriate for large datasets* (PhD Thesis ed.). Glasgow Thesis Service, University of Glasgow. Retrieved from http://theses.gla.ac.uk/

Stover, J. (1998). Revising the proximate determinants of fertility framework: what have we learned in the past 20 years? *Stud Fam Plann, 29*, 255-67.

Subaiya, L., & Johnson, K. (2008). What's in the Gap? Factors associated with the interval between age at first sex and age at first marriage in Cameroon and their implications for reproductive health and women's empowernment. *Working Paper- DHS Working Paper Series, 43*. Retrieved 2014, from http://www.measuredhs.com/pubs/pdf/WP43/WP43.pdf

Trading Economics. (2014). *Nigerian Population 1960-2014*. Retrieved November 2014, from Trading Economics: http://www.tradingeconomics.com/nigeria/population

Ueyama, M., & Yamauchi, F. (2009, April). Marriage behavior response to prime-age adult mortality: evidence from Malawi. *Demography, 46*(1), 43-63.

Weinken, A. (2007). Survival Analysis. In A. Weinken, J. Haerting, U. Mansmann, & L. Duchateau (Eds.), *Frailty Models in Survival Analysis* (pp. 5-20). Halle-Wittenberg: Habilitation. Retrieved 2015, from http://sundoc.bibliothek.uni-halle.de/habil-online/07/07H056/index.htm

Zaba, B., Boerma, T., Pisani, E., & Baptiste, N. (2002). Estimation of levels and trends in age at first sex from surveys using survival analysis. *Measure Evaluation Working Paper Series, 02*(51).

Zaba, B., Pisani, E., Slaym aker, E., & Ties Boerma, J. (2004, September 6). Age at first sex: Understanding recent trends in African demographic surveys. *Sex Transm Infect, 80*(II), ii28-ii35. doi:10.1136/sti.2004.012674

Appendices – Data Analysis Program Syntax Codes

1.0. SPSS syntax commands for exploratory data analysis and Kaplan-Meier survivorship estimation

```
FREQUENCIES VARIABLES=V531 V511 V212 FXFB
 /FORMAT=NOTABLE
 /NTILES=4
 /STATISTICS=STDDEV MINIMUM MAXIMUM MEAN MEDIAN MODE SKEWNESS SESKEW KURTOSIS SEKURT
 /HISTOGRAM
 /ORDER=ANALYSIS.

COMPUTE NHX = (V531 > 0). /*Used the inputed values of ever had sex. identify missing cases*/.
RECODE NHX(SYSMIS=0).
EXECUTE.
FREQUENCIES NHX.
SORT CASES BY NHX(D).

DATASET ACTIVATE DataSet1.
FREQUENCIES VARIABLES=V013
  /ORDER=ANALYSIS.

FREQUENCIES VARIABLES=V012
 /FORMAT=NOTABLE
 /NTILES=4
 /STATISTICS=STDDEV MINIMUM MAXIMUM MEAN MEDIAN MODE SKEWNESS SESKEW KURTOSIS SEKURT
 /ORDER=ANALYSIS.

FREQUENCIES VARIABLES=V013
  /ORDER=ANALYSIS.

WEIGHT OFF.

* Descriptive statistics.
CTABLES
  /VLABELS VARIABLES=V101 V102 V130 V149 DISPLAY=LABEL
  /TABLE V101 [C][COUNT F40.0, COLPCT.COUNT PCT40.1] + V102 [C][COUNT F40.0, COLPCT.COUNT PCT40.1]
    + V130 [C][COUNT F40.0, COLPCT.COUNT PCT40.1] + V149 [C][COUNT F40.0, COLPCT.COUNT PCT40.1]
  /CATEGORIES VARIABLES=V101 V102 V130 V149 ORDER=A KEY=VALUE EMPTY=EXCLUDE.

FREQUENCIES v133 v149.

FILTER OFF.
USE ALL.
SELECT IF (V133 ~= 0).
EXECUTE.

ONEWAY V133 BY V149
 /STATISTICS DESCRIPTIVES
 /MISSING ANALYSIS.
```

```
* Custom Tables.
CTABLES
 /VLABELS VARIABLES=V155 V190 DISPLAY=LABEL
 /TABLE V155 [COUNT F40.0, COLPCT.COUNT PCT40.1] + V190 [COUNT F40.0, COLPCT.COUNT PCT40.1]
 /CATEGORIES VARIABLES=V155 V190 ORDER=A KEY=VALUE EMPTY=EXCLUDE.

CTABLES
 /VLABELS VARIABLES=V535 V501 V228 V502 V503 DISPLAY=LABEL
 /TABLE V535 [C][COUNT F40.0, COLPCT.COUNT PCT40.1] + V501 [C][COUNT F40.0, COLPCT.COUNT PCT40.1]
   + V228 [C][COUNT F40.0, COLPCT.COUNT PCT40.1] + V502 [C][COUNT F40.0, COLPCT.COUNT PCT40.1] +
V503
   [C][COUNT F40.0, COLPCT.COUNT PCT40.1]
 /CATEGORIES VARIABLES=V535 V501 V228 V502 V503 ORDER=A KEY=VALUE EMPTY=EXCLUDE.

* Custom Tables.
CTABLES
 /VLABELS VARIABLES=V535 V501 V228 V502 V503 V013 DISPLAY=LABEL
 /TABLE V535 [C][COLPCT.COUNT PCT40.1] + V501 [C][COLPCT.COUNT PCT40.1] + V228 [C][COLPCT.COUNT
   PCT40.1] + V502 [C][COLPCT.COUNT PCT40.1] + V503 [C][COLPCT.COUNT PCT40.1] BY V013 [C]
 /CATEGORIES VARIABLES=V535 V501 V228 V502 V503 ORDER=A KEY=VALUE EMPTY=EXCLUDE
 /CATEGORIES VARIABLES=V013 ORDER=A KEY=VALUE EMPTY=EXCLUDE TOTAL=YES POSITION=AFTER.


* Custom Tables.
CTABLES
 /VLABELS VARIABLES=V013 V149 V501 V503 V502 V228 DISPLAY=LABEL
 /TABLE V013 + V149 [C] + V501 [C] + V503 [C] + V502 [C] BY V228 [C][COUNT F40.0, ROWPCT.COUNT
   PCT40.1]
 /CATEGORIES VARIABLES=V013 ORDER=A KEY=VALUE EMPTY=INCLUDE TOTAL=YES POSITION=AFTER
 /CATEGORIES VARIABLES=V149 V501 V503 V502 ORDER=A KEY=VALUE EMPTY=EXCLUDE TOTAL=YES
POSITION=AFTER
 /CATEGORIES VARIABLES=V228 ORDER=A KEY=VALUE EMPTY=EXCLUDE TOTAL=YES .

* Custom Tables.
CTABLES
 /VLABELS VARIABLES=V013 V149 V502 V714 V714A DISPLAY=LABEL
 /TABLE V013 [C][COUNT F40.0, ROWPCT.COUNT PCT40.1] + V149 [C][COUNT F40.0, ROWPCT.COUNT
PCT40.1]
   + V502 [C][COUNT F40.0, ROWPCT.COUNT PCT40.1] BY V714 [C] + V714A [C]
 /CATEGORIES VARIABLES=V013 ORDER=A KEY=VALUE EMPTY=INCLUDE TOTAL=YES POSITION=AFTER
 /CATEGORIES VARIABLES=V149 V502 V714 V714A ORDER=A KEY=VALUE EMPTY=EXCLUDE TOTAL=YES
   POSITION=AFTER.

* Custom Tables.
CTABLES
 /VLABELS VARIABLES=V013 V149 V502 V101 V102 V212 V511 V525
   DISPLAY=LABEL
 /TABLE V013 [C] + V149 [C] + V502 [C] + V101 [C] + V102 [C] BY V212 [S][PTILE 25, MEDIAN, PTILE
   75] + V511 [S][PTILE 25, MEDIAN, PTILE 75] + V525 [S][PTILE 25, MEDIAN, PTILE 75]
 /CATEGORIES VARIABLES=V013 ORDER=A KEY=VALUE EMPTY=INCLUDE TOTAL=YES POSITION=AFTER
 /CATEGORIES VARIABLES=V149 V502 V101 V102 ORDER=A KEY=VALUE EMPTY=EXCLUDE TOTAL=YES
   POSITION=AFTER.
```

```
*First sex and first birth interval.
COMPUTE FXFB = V212 - V531.
COMPUTE FXFM = V511 - V531.
EXECUTE.

FREQUENCIES FXFB FXFM.
MISSING VALUES   FXFB (-1) FXFM (LO TO -1).

*Identify all missing cases.
COMPUTE DUMMY = (FXFB +FXFM + FMFB).
RECODE DUMMY (SYSMIS=0) (ELSE=1).
EXECUTE.
FREQUENCIES   DUMMY.

* Distribution of missingness.
CTABLES
 /VLABELS VARIABLES=V013 V101 V149 V502 V714 DUMMY DISPLAY=LABEL
 /TABLE V013 [C] + V101 [C] + V149 [C] + V502 [C] + V714 [C] BY DUMMY [C][COUNT F40.0,
   ROWPCT.COUNT PCT40.1]
 /CATEGORIES VARIABLES=V013 V101 V149 V502 V714 ORDER=A KEY=VALUE EMPTY=INCLUDE TOTAL=YES
   POSITION=AFTER
 /CATEGORIES VARIABLES=DUMMY ORDER=A KEY=VALUE EMPTY=EXCLUDE TOTAL=YES POSITION=AFTER.

*Missing on first sex and first birth.
RECODE FXFB (SYSMIS=0) (ELSE=1) INTO fxbie.
EXECUTE.

*Never had a birth; INDICATOR VARIABLE.
RECODE FMFB (SYSMIS=0) (ELSE =1) INTO FMBIE.
VARIABLE LABELS FMBIE 'Never had a birth'.
FORMATS FMBIE(F5.0).
VARIABLE LEVEL FMBIE(NOMINAL).
VALUE LABELS FMBIE 0 'Never had a birth' 1 'Had a birth'.
EXECUTE.
FREQUENCIES FMBIE.

*Checking for consistency in records on birth history.
CROSSTABS
 /TABLES=V201 BY V501
 /FORMAT=AVALUE TABLES
 /CELLS=COUNT COLUMN
 /COUNT ROUND CELL.

* EXPLORATORY DATA ANALYSIS.
CTABLES
 /VLABELS VARIABLES=V013 V101 V102 V149 V502 V212 V511 V531 FXFB
   DISPLAY=LABEL
 /TABLE V013 [C] + V101 [C] + V102 [C] + V149 [C] + V502 [C] BY V212 [S][MEDIAN] + V511
   [S][MEDIAN] + V531 [MEDIAN] + FXFB [S][MEAN]
 /CATEGORIES VARIABLES=V013 V101 V102 V149 V502 ORDER=A KEY=VALUE EMPTY=INCLUDE TOTAL=YES
   POSITION=AFTER.

EXAMINE VARIABLES=V531 V511 V212 FXFB
```

```
 /PLOT NONE
 /PERCENTILES(5,10,25,50,75,90,95) HAVERAGE
 /STATISTICS DESCRIPTIVES EXTREME
 /CINTERVAL 95
 /MISSING PAIRWISE
 /NOTOTAL.

*MODEL FITTING.
*Estimates of the cumulative hazard function.
KM V531 /*Age at 1st sex*./
 /STATUS=NHX(1)
 /PRINT TABLE MEAN
 /PLOT SURVIVAL HAZARD
 /SAVE HAZARD. /*Save the cumulative hazard function and use its values in teh next step to plot a straight line
with time.

KM V212 /*Age at birth*./
 /STATUS=NHX(1)
 /PRINT TABLE MEAN
 /PLOT SURVIVAL HAZARD
 /SAVE HAZARD. /*Save the cumulative hazard function and use its values in the next step to plot a straight line
with time.

KM FXFB /*Interval*./
 /STATUS=NHX(1)
 /PRINT TABLE MEAN
 /PLOT SURVIVAL HAZARD
 /SAVE HAZARD. /*Save the cumulative hazard function and use its values in the next step to plot a straight line
with time.


*log cumulatuve hazard.
COMPUTE LOGHZ = LN(HAZ_2).
COMPUTE LOGSX = LN(FXFB).
EXECUTE.
*Since the log-cum Hazard and log of survival time are related bt a straight line equation, we estimate the
parameters of Weibull using OLS linear regression.
REGRESSION
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS BCOV R ANOVA
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT LOGHZ
 /METHOD=ENTER LOGSX.

* Chart Builder.
GGRAPH
 /GRAPHDATASET NAME="graphdataset" VARIABLES=LOGSX LOGHZ MISSING=LISTWISE
REPORTMISSING=NO
 /GRAPHSPEC SOURCE=INLINE.
BEGIN GPL
 SOURCE: s=userSource(id("graphdataset"))
```

```
DATA: LOGSX=col(source(s), name("LOGSX"))
DATA: LOGHZ=col(source(s), name("LOGHZ"))
GUIDE: axis(dim(1), label("Log of Age at first Sex"))
GUIDE: axis(dim(2), label("Log-Cumulative Hazard"))
ELEMENT: point(Position(LOGSX*LOGHZ))
END GPL.
```

```
REGRESSION
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS CI(95) BCOV R ANOVA
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT LOGHZ
 /METHOD=ENTER LOGSX.
```

2.0. Stata commands for estimating cumulative hazard function curve parameter estimates

```
///generate log cum haz FIRST BIRTH
stset V212 ///V212 is the age at first birth
sts generate cumhaz_bir = na
gen ln_cumhaz_bir = ln(cumhaz_bir)
label variable ln_cumhaz_bir "log cumulative hazard function"
drop cumhaz_bir
gen ln_bir = ln(V212)
label variable ln_bir "log Age at first birth"
twoway (line ln_cumhaz_bir ln_bir, sort)
regress ln_cumhaz_bir ln_bir
estat ic, n(31074) ///AIC
estat vce  ///Variance-covariance matrix

stset V531, scale(1) ///declare V531 as survival data
stdescribe /// describe survival data
streg, dist(loglogistic) nolog
streg, dist(lognormal) nolog
streg, dist(Weibull) time nolog
streg, dist(exponential) time nolog
streg, dist(gamma)
nlcom (theta:exp([ln_p]_b[_cons])) (beta:exp(-_b[_cons])) ///beta and theta
mat list r(V) ///variance-covariance matrix
```

MODEL FITTING - stgenreg

```
stset FXFB  ////Interval between first birth and first sex
stgenreg,  loghazard([ln_lambda]  :+  [ln_gamma]  :+  (exp([ln_gamma])-1):*log(#t))  ln_lambda(NHX)  ///NHX indicates censored and observed data
nlcom (theta:exp([ln_gamma]_b[_cons])) (beta:exp(-[ln_lambda]_b[_cons])) ///beta and theta
estat vce  ///variance-covariance matrix
```

Multiple Regression

```
/*Fixed effect model*/
stset FXFB, failure(EHAB=1)
streg V531 i.V013 i.V102 i.V130 i.V149 i.V502 i.V228 i.V302A, dist(weibull) nolog
```

```
estat ic ///AIC

/*Cox Rrgression*/
stcox V531 i.V013 i.V102 i.V130 i.V149 i.V502 i.V228 i.V302A, nolog
estat ic ///AIC

/*Random Effect Model*/
stset FXFB, failure(EHAB=1) ///set FXFB as survival time, Ever Had A Birth (EHAB) as censored observation
xi: stmixed V531 V102 i.V130 i.V149 V511 V228 V302A V502 || V021: , dist(weibull) nolog
```

```
estat ic ///AIC

/*Cox Rrgression*/
stcox V531 i.V013 i.V102 i.V130 i.V149 i.V502 i.V228 i.V302A, nolog
estat ic ///AIC

/*Random Effect Model*/
stset FXFB, failure(EHAB=1) ///set FXFB as survival time, Ever Had A Birth (EHAB) as censored observation
xi: stmixed V531 V102 i.V130 i.V149 V511 V228 V302A V502 || V021: , dist(weibull) nolog
```