Survival analysis of birth interval of women among rural Ethiopia (Comparison of Shared frailty Models)

Jaleta Abdisa Fufa[1*], Dechasa Bedada Tolessa[1]

Department of Statistics, College of natural science, Jimma University, Jimma, Ethiopia[1]

Email address: jaleabdi100@gmail.com , decheebe@gmail.com

[*] Corresponding Authors: jaleabdi100@gmail.com

ABSTRACT

Background: A birth interval is the length of time between two successive live births.Longer time periods between births allow the next pregnancy and birth to be at full gestation and growth. Births too close together are associated with schizophrenia in offspring and hinder the physiological ability of mothers and, thus, expose them to complications during and after pregnancy (Smits L *et al*, 2003).We know that, birth intervals are believed to affect fertility, mortality and population growth. Thus, childbearing or fertility of population has received serious attention.This study aimed to investigate the potential risk factors affecting the birth interval of women among rural Ethiopia.

Method: Data were taken from CSA (EDHS, 2011) of adult women living rural Ethiopia. Kaplan-Meier estimation method, Cox proportional hazard model and shared frailty model were used for analysis.

Results: The Weibull inverse Gaussian model was preferred over Exponential, and Log-logistic models based on Akaike's Information Criterion evidence. The results implied that contraceptive use, breast feeding and education prolong the time between successive birth intervals.

And also the results indicated that as the age of women increases the time interval between successive children also increases while Muslim and protestant were significantly shorten time interval between successive children.

Conclusion: Finally, the findings of this study implied that age, contraceptive use, breast feeding, marital status, religion and education were major factors affecting birth interval of women among rural Ethiopia

Key words: Survival Data Analysis, Proportional hazard, shared frailty, birth interval

Abbreviations

AIC**:** Akaike Information Criterion

CSA: Central Statistical Agency

EDHS: Ethiopia Demographic and Health Survey

1. Introduction

A birth interval is the length of time between two successive live births (WHO). Longer time periods between births allow the next pregnancy and birth to be at full gestation and growth (Rutstein, 2008). Births too close together are associated with schizophrenia in offspring and hinder the physiological ability of mothers and, thus, expose them to complications during and after pregnancy (Smits *et al.*, 2003). A study from Latin America revealed that short birth intervals adversely affect the health of mothers and their children's chances of survival. Short birth intervals (less than 24 months) increase maternal risks such as third trimester bleeding, toxemia, malnutrition, anemia, and maternal mortality. It can lead to several serious outcomes for neonates as well, such as prematurity, low birth weight, stillbirth,
neonatal mortality and adverse effect on intellectual ability, physical growth and development (Agudelo*et al.*, 2012).

Birth interval is the length of time between two successive live births (WHO). Successive births physiologically deplete the mother of energy and nutrition (Gribble, 1993) which may lead to premature births or pregnancy complications, increasing the risk of infant or maternal

death, or impair the mother's ability to care for her children. Additionally, it has also been argued that women with closely spaced births may still have very young children and, as such, are less likely to attend prenatal care services. Invariably, longer duration of the inter-birth interval has been found to increase profoundly the probability of infant survival (Pedersen, 2000).

We know that, birth intervals are believed to affect fertility, mortality and population growth. Thus, childbearing or fertility of population has received serious attention. There are several reasons for this. Fertility largely determines age structure of a population, high fertility leads to large proportion of child and young population, resulting high dependency of child population and high investment on educations. In fact this was present scenario in most of the less developed countries of Asia, Africa and Latin American countries. When we look back at the background of Ethiopian population growth rate, there has been a steady increase since 1960. As per 1984 census data, population growth was estimated at about 2.3 percent for the 1960-70 period, 2.5 percent for the 1970-80 period, and 2.8 percent for the 1980-85. Population projections compiled in 1988 by the Central Statistical Authority (CSA) projected a 2.83 percent growth rate for 1985-90 and a 2.96 percent growth rate for 1990-95, shows a steady increase in the population in the past period. But 2007 Ethiopia population census shows that, the annual population growth rate within 1994-2007 was estimated as 2.6%, and hence a slow decline in the growth of population. This may be due to decline in the fertility performance of Ethiopian population. This gives an evidence of increase in the practice of birth control measures such as modern contraceptive measures, birth spacing, late marriage etc.

2. Material and Methods

The study was based on the national cross-sectional demographic and health survey (EDHS) conducted in Ethiopia 2011.The 2011 EDHS sample was selected using a stratified, two-stage cluster design and EAs were the sampling units for the first stage. The sample included 624 EAs, 187 in urban areas and 437 in rural areas (EDHS, 2011).

Households comprised the second stage of sampling. A complete listing of households was carried out in each of the 624 selected EAs from September 2010 through January 2011 (EDHS, 2011).But this study takes place on rural parts only. The sample size in this study was 6066 women among rural Ethiopia who gave at least one birth over five year's period.

The dependent (response) variable was birth interval which is defined as the time gap between successive children for each mother measured in months. Birth intervals can be considered as time-to-event data. The intervals are categorized into two. Those are closed interval and open-ended interval.Closed interval means the time between two successive births and Open ended interval means the time between the birth of the youngest child and the interview date. This means when one mother gives only one child over five years period. Then the open-ended intervals are considered as censored while closed intervals are as failure. There are many explanatory variables are used as predictors of birth interval of women. Those are age, level of education of women, wealth index, marital status, religion, work status, breastfeeding and contraceptive use.

2.1 Statistical methods

Survival analysis is a collection of statistical procedures for data analysis for which the outcome variable of interest is time until an event occurs(Allison, 1995). When we say time, we mean years, months, weeks, or days from the beginning of follow-up of an individual until an event occurs; alternatively, time can refer to the age of an individual when an event occurs (Allison, 1995). By event, we mean death, disease incidence, relapse from remission, recovery (e.g., return to work) or any designated experience of interest that may happen to an individual (Allison, 1995)

Survival Functions

Given a random variable T that denotes the survival time, the basic quantity employed to describe time-to-event phenomenon is the Survival Function S(t), and it is defined as:

$S(t) = P[T > t]$ = the probability that an individual survives beyond time t.

Since a unit either fails, or survives, and one of these two mutually exclusive alternatives must occur, the survival function is:

$$S(t) = 1 - F(t) \quad (1)$$

Where, F (t) is CDF. If T is a continuous random variable, then S(t) is a continuous, strictly decreasing function. The survival function is the integral of pdf, that is:

$$S(t) = \int_t^{\infty} f(u)du \quad (2)$$

$$f(t) = -\frac{dS(t)}{dt} (3)$$

2.2 Hazard Function

The hazard function is the probability that an individual will experience an event (like, death) within a small time interval, given that the individual has survived up to the beginning of the interval.

$$\square(t) = -\left(\frac{dlnS(t)}{dt}\right) \quad (4)$$

Where s (t) is survival probability function at time t.

2.3 Kaplan-Meier Methods

The Kaplan-Meier (KM) estimator is the standard non-parametric estimator of the survival functions$s(t)$, proposed by Kaplan and Meier (1958), and is also called the Product-Limit estimator.

The Kaplan-Meier estimator of the survivorship function (or survival probability) at time$t$, $s(t) = P(T > t)$ is defined as:

$$\hat{s}(t) = \prod_{t(j)\leq t}\left(\frac{nj-rj}{nj}\right) = \prod_{t(j)\leq t}\left(1 - \frac{rj}{nj}\right) \quad (5)$$

2.4 Shared Frailty Model

Shared frailty model is a conditional model in which frailty is common to all subjects in a cluster. It is responsible for creating dependence between event times. It is also known as a mixture model because the frailties in each cluster are assumed to be random. Shared frailty model was introduced by Clayton (1978) without using the notion frailty and extensively studied by Hougaard (2000), Therneau and Grambsch (2000), Duchateau*et al.* (2003), and

Duchateau and Janssen (2004). The notation of frailty provides a covenant way to introduce random effect, association and unobserved heterogeneity into models for survival data. In its simplest form, a frailty is an unobserved random proportionality factor that modifies the hazard function of an individual, or related individuals (Wienke*et al.*, 2003).

2.5 The Gamma Frailty Distribution

The gamma distribution is very-well known and has simple densities. It is the most common distribution used for describing frailty. Even though gamma models have closed form expressions for survival and hazard functions, from a computational view, it fits well to frailty data and it is easy to derive the closed form expressions for unconditional survival and hazard functions. There are many applications of the gamma frailty model. Lancaster (1979) suggested this model for the duration of unemployment. Aalen (1987) studied the expulsion of intra-uterine contraceptive devices. Ellermann*et al.* (1992) studied recidivism among criminals using gamma-Weibull model. Andersen *et al*. (1993) used the gamma frailty model to check the proportional hazards assumptions in his study of malignant melanoma. Vaupel*et al*. (1979) used the gamma distribution in their studies on population mortality data from Sweden. From a computational point of view, gamma models fit very well into survival models, because it is easy to derive the formulas for any number of events. This is due to the simplicity of the derivatives of the Laplace transform. The gamma frailty distribution has been widely used in parametric modeling of intra cluster dependency because of its simple interpretation and mathematical tractability (Vaupel*et al*., 1979; Clayton, 1978 and Oakes, 1982).

The density of a gamma-distributed random variable with parameter θ is given by

$$f_z(z) = \frac{z_i \exp(^{-z_i}/_{\theta})}{\theta^{\frac{1}{\theta}} \gamma(\frac{1}{\theta})}, \qquad \theta > 0 \ (6)$$

2.6 Inverse Gaussian Shared Frailty Distribution

The inverse Gaussian (inverse normal) distribution was introduced as a frailty distribution alternative to the gamma distribution by Hougaard (1984). The probability density function of an inverse Gaussian distributed random variable with parameter θ > 0 is given by:

$$f_z(z_i) = \left(\frac{1}{2\pi\theta}\right)^{1/2} z_i^{-3/2} \exp\left(\frac{-(z_i-1)^2}{2\theta z_i}\right) \qquad \theta>0,\ z>0 \qquad (7)$$

In order to investigate effect of the candidate covariates on survival time of women with cervical cancer, we first did univariable parametric frailty models analysis by fitting a separate model for each candidate covariates. Covariates that were found to be significant in the univariable analysis were included in the multivariable analysis. The multivariable parametric frailty models in the study were done by assuming the Exponential, Weibull, and Log logistics distributions for the base line hazard function.

3. Results

Data for the analyses in this study came from the 2011 Ethiopian Demographic and Health Survey with reference to 6066 rural women who gave at least one birth over the five years period. The response variable is birth interval which is defined as time in month between successive children for each mother. Birth intervals can be considered as time-to-event data, when women gave birth it considered as event (failure) and the time between the birth of the youngest child and the interview date is censored. The data consists of 6066 women who gave birth over five years period living among rural Ethiopia. There were about 2413(39.78%) events and 3653(60.22%) censored.

Table 1. Summary results of covariates of birth interval of women among rural Ethiopia (EDHS, 2011).

Covariate/Factor	Category	Event (%)	Censored (%)	Total (%)
	15-19	78(1.29)	287(4.73)	365(6.02)
	20-24	498(8.21)	712(11.74)	1210(19.95)
Mother age	25-29	823(13.57)	894(14.74)	1717(28.31)
	30-34	503(8.29)	677(11.16)	1180(19.45)
	35-39	366(6.03)	603(9.94)	969(15.97)
	40-44	114(1.88)	346(5.70)	460(7.58)
	45-49	31(0.51)	134(2.21)	165(2.72)

	No education	1917(31.62)	2613(43.08)	4530(74.68)
Educational level	Primary	485(7.99)	963(15.87)	1448(23.87)
	Secondary	9(0.15)	49(0.81)	58(0.96)
	Higher	2(0.03)	30(0.49)	30(0.49)
	Poor	1466(24.17)	2006(33.07)	3472(57.23)
Wealth index	Medium	437(7.20)	746(12.30)	1183(19.5)
	rich	510(8.41)	901(14.85)	1411(23.26)
	Single	3(0.05)	34(0.56)	37(0.61)
Marital status	Married	2325(38.33)	3254(53.64)	5579(91.97)
	Widowed	34(0.56)	93(1.53)	127(2.09)
	Divorced	51(0.84)	272(4.48)	323(5.32)
Contraceptive	No	2152(35.48)	2819(46.47)	4971(81.95)
	Yes	261(4.30)	834(13.75)	1095(18.05)
Work status	No	1568(25.85)	2379(39.22)	3947(65.07)
	Yes	845 (13.93)	1274(21.0)	2119(34.93)
Religion	Orthodox	568(9.36)	1322(21.79)	1890(31.16)
	Catholic	16(0.26)	51(0.84)	67(1.1)
	Protestant	504(8.31)	759(12.51)	1263(20.82)
	Muslim	1268(20.9)	1434(23.59)	2702(44.54)
	Traditional	22(0.36)	35(0.58)	57(0.94)
	Others	35(0.58)	52(0.86)	87(1.43)
	No	475(7.83)	1619(26.70)	2094(34.52)
Breast feeding	Yes	1938(31.95)	2034(33.53)	3972(65.48)

According to the table 1, out of women who gave birth over last five years, large number of women 1717(28.31%) were found 25-29 age categories whereas small number of women 165(2.72%) found in age category 45-49. Small number of events 0.51% (31) experienced in the age category 45-49 years,

Whereas large number of events were experienced in the age category of 25-29 years which was 13.57%(823). There were 74.68%, 23.87%, 0.96%, 0.49% of women under study who had no education, primary, secondary and higher respectively, indicating the sample consist of more number of illiterates. The proportion of women who experienced events have no education, primary, secondary and higher level of education were respectively 31.622%, 7.99%, 0.15%, 0.03% respectively. About 57.24% (3472) of the households were classified as poor while 19.5% (1183) had medium income and 23.26% (1411) were rich. Out of these categories poor women experienced large number of events 1466(24.17%), whereas medium wealth index women experienced minimum events 437(7.2%). From the same Table 1 out of women who gave birth over five years period only 37 women were found to be single and among them only 3 women have experienced the event under study remaining were censored, Whereas 5579 women were married and large number of events experienced by them. Out of the entire subjects integrated in this study 4975(82.01%) of women were not using contraceptive method whereas 1095(18.05%) using it. Women who were not using contraceptive method experienced more events which were 2152(43.29%). As indicated on above table among women who gave birth in last five years 2702(44.54%) were Muslim, and they have experienced more events compared to other religious groups. The table 1 results also shows 31.16%, 1.10%, 20.82%, 0.94% and 1.434% were orthodox, catholic, protestant, traditional and others respectively, they experienced less number of events compared to Muslims.

Most of the women under study were jobless (65.07%),remaining 34.93 had job, and the jobless women were have more events compared to jobholders..

The AIC value of weibull Inverse-Gaussian model 22092.56 was the minimum among all the other AIC values of the models indicating that it was the most efficient model to describe birth interval of women among rural Ethiopians' dataset. Analysis based weibull inverse Gaussian model shoes that contraceptive use, marital status, age, breast feeding religion and educational level were statistically significant.

Table 2: The value of AIC for Multivariable Parametric Shared Frailty Models (EDHS, 2011)

Model		**AIC**
Baseline distribution	**Frailty model**	
	Gamma	22101.66
Weibull	Inverse Gaussian	22092.56
Exponential	Gamma	24853.86
	Inverse Gaussian	24853.73
Log logistic	Gamma	22338.68
	Inverse Gaussian	22339.68

The variance of the frailty were significant for all baseline hazard function with an inverse Gaussian shared frailty distribution in the models whereas only exponential gamma shared frailty is significant from baseline hazard function with gamma shared frailty using the same baseline as inverse Gaussian models at 5% level of significance. This indicates that the presence of heterogeneity and necessitates of the frailty models. The estimate for the variance parameter θ in a shared frailty models can be thought as a measure of the degree of correlation and provides information on the variability (the heterogeneity) in the population of clusters. The value of shared frailty distribution (θ) is 0.0116, 0.0456 and 0.0161 for Weibull-inverse Gaussian, exponential-inverse Gaussian, and log-logistic-inverse Gaussian respectively.The corresponding Kendall's tau (τ) values of shared frailty distribution are 0.058, 0.03 and 0.288 respectively.The Kendall's tau (τ) value is used to measure the dependence within the clusters (region).

Model comparisons were presented in Table 2 Accordingly; it suggested that Weibull-inverse Gaussian shared frailty model was selected according to AIC.

Table 3 Results of the multivariable Weibull-Inverse gaussian shared Frailty Model for birth interval dataset (EDHS, 2011)

Variables	Coef	S.E (Coef)	Φ	95%CI LCL UCL	Chi-sq	P-value
(Intercept)	4.603	0.2051	99.783	66.7533,149.1565	0.2022	0.00
Contraceptive						
No	Ref					
Yes	0.231	0.0235	1.26	1.2031, 1.3192	0.0235	0.00
Marital status						
Single	Ref					
Married	-0.465	0.1983	0.628	0.4258, 0.9265	0.1983	1.90e-02
Widowed	-0.322	0.2067	0.725	0.4833, 1.0867	0.2067	1.20e-01
Divorced	-0.149	0.2036	0.861	0.5781, 1.2841	0.2036	4.60e-01
Age						
15-19	Ref					
20-24	-0.027	0.0419	0.973	0.8966, 1.0567	0.0419	5.20e-01
25-29	0.035	0.041	1.036	0.9556, 1.1223	0.041	4.00e-01
30-34	0.067	0.0421	1.069	0.9846, 1.1613	0.0421	1.10e-01
35-39	0.128	0.0431	1.137	1.0445, 1.2367	0.0431	2.90e-03
40-44	0.299	0.0508	1.348	1.2207, 1.4897	0.0507	3.70e-09
45-49	0.395	0.0733	1.484	1.2857, 1.7137	0.0733	7.30e-08
Breast feeding						
No	Ref					
Yes	0.02	0.0046	1.0202	1.0110, 1.0294	18.67	0.00
Religion						
Orthodox	Ref					
Catholic	0.019	0.0904	1.019	0.8537, 1.2167	0.0901	8.40e-01
Protestant	-0.107	0.0303	0.898	0.8467, 0.9535	0.0297	3.90e-04

Muslim	-0.116	0.0263	0.89	0.8457, 0.9376	0.0256	1.00e-05
Traditional	-0.129	0.077	0.879	0.7558, 1.0222	0.0768	9.40e-02
Others	-0.092	0.064	0.912	0.8046, 1.0340	0.0637	1.50e-01
Education						
No education	Ref					
Primary	0.142	0.0549	1.152	1.0393, 1.2782	7.2	7.30e-03
Secondary	0.1268	0.0485	1.135	1.0322, 1.2484	6.84	2.70e-02
Higher	0.1731	0.0604	1.189	1.0562,1.3384	8.26	4.20e-03
Wealth						
Poor	Ref					
Medium	0.034	0.0194	1.034	0.9960, 1.0747	0.0194	8.00e-02
Rich	-0.007	0.0189	0.993	0.9569, 1.0305	0.0189	6.90e-01
frailty					53.29	1.50e-08
$\theta =$ 0.0116		$\lambda = 0.342$		AIC=22092.56		
$\gamma = 2.651$		$\tau = 0.058$				

From Table 3 above the confidence intervals of the acceleration factor for all significant categorical covariates do not include one at 5% level of significance. Hence from the above table the acceleration factors and its 95% confidence interval for using contraceptive is 1.260(1.2031, 1.3192) when we compared to not using contraceptive as reference. In the confidence interval of acceleration factors 1 is not included, implies that contraceptive use was statistically significant for birth interval of women among rural Ethiopia. The corresponding p-value is less than 0.05 which supports that the contraceptive use status of women was significant. Accordingly it prolonged the birth interval by factor of (Φ =1.260) at 5% level of significance.

The acceleration factors and its 95% confidence interval for marital status was 0.628, 0.725, 0.861 ((0.4258, 0.9265), (0.4833, 1.0867), (0.5781, 1.2841)) for the group of married, widowed and divorced respectively using single as reference. The acceleration factor for married does not include one whereas the rest include one.

This implies that out of the group of marital status of women married is statistically significant for birth interval of women among rural Ethiopia. This is also supported by p-value which is less than 0.05. The estimated coefficient of the parameters married was -0.465. The sign of the coefficient is negative which implies that decreasing logged of survival time and hence, shorter expected duration of time to get next successive child.

Some category of age of women is statistically significant; it determines the birth interval of women among rural Ethiopia. The acceleration factors and 95% confidence intervals for statistically significant category of age of women are 1.137, 1.348 and 1.484((1.0445, 1.2367), (1.2207, 1.4897), (1.2857, 1.7137)) for the group of 35-39, 40-44, 45-49 when comparing 15-19 age category as reference.

The 95% confidence interval of the acceleration factor for those women who were breast feeding their children is (1.0110, 1.0294) and its acceleration factors is 1.0202 by using those who had not breast feeding their children as reference category at 5% level of significance. This shows that breast feeding prolong the interval between successive children.

From the variable religion catholic, traditional and others were not statistically significant using orthodox as reference. The acceleration factor and 95% CI for protestant and Muslim were 0.898 and 890 ((0.8467, 0.9535), (0.8457, 0.9376)). It shows that confidence interval of both variables did not include one. This implies that both were statistically significant at 5% level of significance. It also supported by p-value which is less than 0.05. The estimated coefficient of the parameters for Protestant and Muslim were -0.107, -0.116 respectively.

The sign of the coefficient for both were negative which implies that decreasing logged of survival time and hence, shorter expected duration of time to get next successive child.

The 95% confidence interval for acceleration factor of women educational levels was (1.0393, 1.2782), (1.0322, 1.2484) and (1.0562, 1.3384) for the group of primary, secondary and higher education educational level respectively. This confidence interval does not include one in all; indicating that primary, secondary and higher education's were significantly important factors for the timing of birth interval by using not educated women as a reference category.

Accordingly, it prolonged the interval between successive children by a factor of (Φ =1.152, Φ = 1.35 and Φ =1.189) for primary, secondary and higher education respectively at 5% level of significance.

After the model has been fitted, it is desirable to determine whether a fitted parametric model adequately describes the data or not. Therefore, the appropriateness of model with Weibullbaseline can be graphically evaluated by plotting log(-log(s(t)) versus log(time), the Log-logistic baseline by plotting log($\frac{\hat{s}(t)}{1-\hat{s}(t)}$) versus log(time) and for exponential distribution, there is log(s(t))=-λt. If the plot is linear, the given baseline distribution is appropriate for the given dataset.Accordingly, their respective plots are given in figure below and the plot for weibull baseline distribution make straight line better than exponential and Log-logistic baseline distribution. This evidence also strengthens the decision made by AIC value that weibull baseline distribution is appropriate for the given dataset.

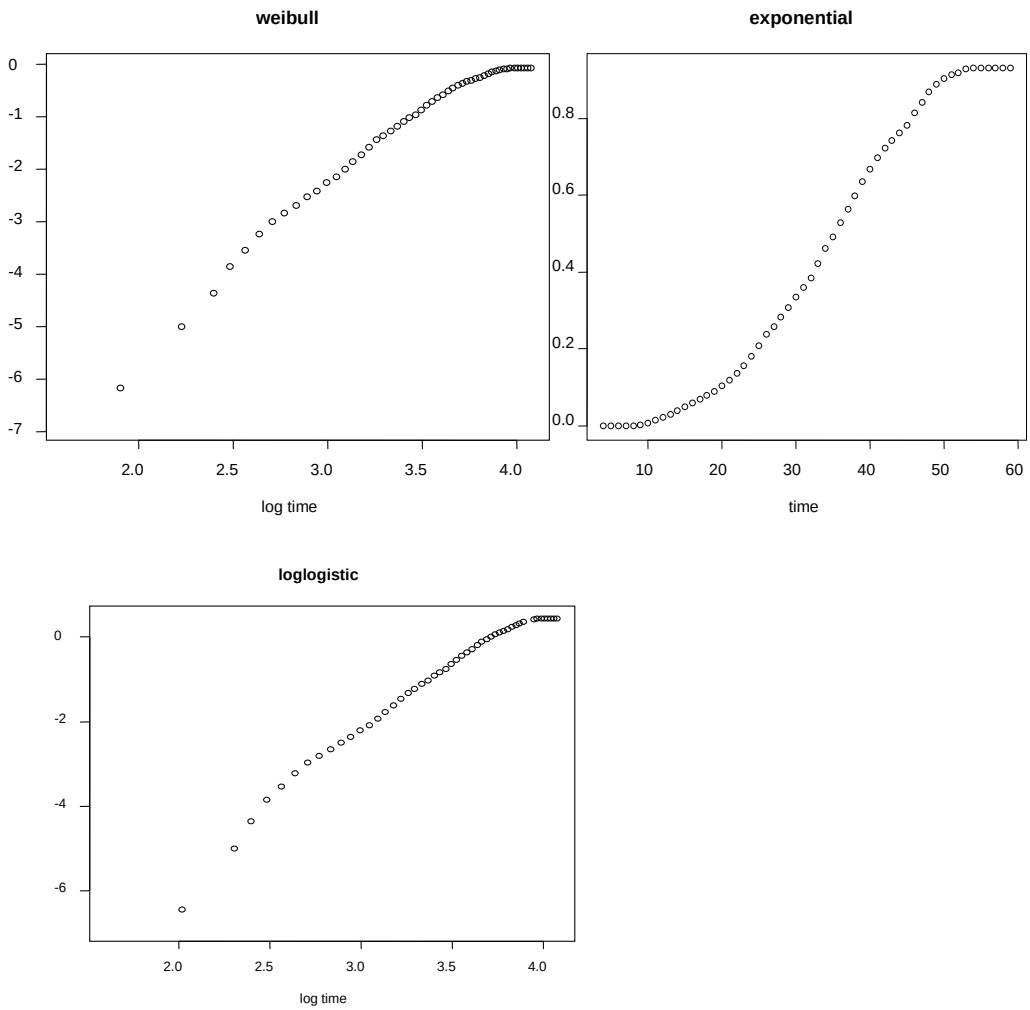

Figure 1.Graphical evaluation of the exponential, weibull and log-logistic assumptions

That the plot of the cumulative hazard function against Cox-Snell residuals (Cox and Snell ,1968)in figure 1 is nearest to the 45 degree straight lines through the origin for weibull Inverse Gaussian model suggesting that it is appropriate for birth interval datasets.

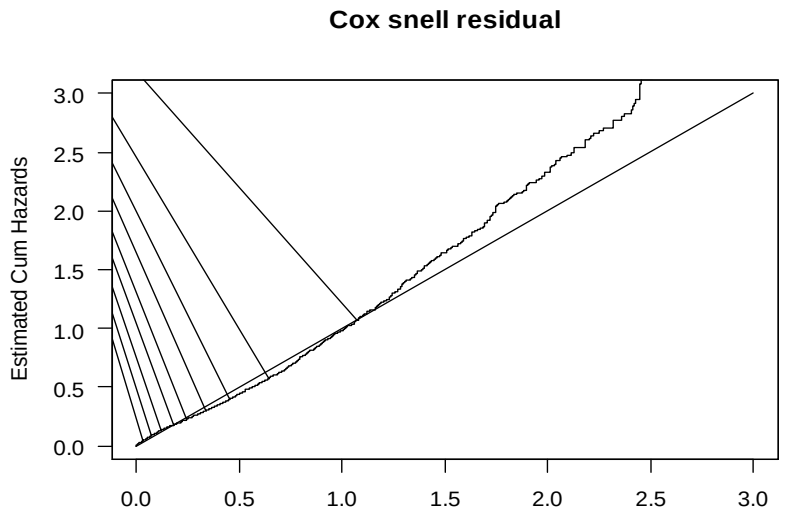

Figure 2**:** Cox- Snell residuals plots of weibull baseline distribution for birth interval dataset

4. Discussion

The main aim of the study was statistical analysis of the determinants of birth interval of women among rural Ethiopia using suitable three baselines parametric and two parametric shared frailty distributions using the 2011 EDHS data. Both univariable and multivariable shared frailty models i.e., the three baseline parametric distribution (exponential, Weibull and log-logistic) with two shared frailty distributions (gamma and inverse Gaussian) were employed to examine the factors that determine birth interval of women. Factors that are concerned for our study were work status of women, contraceptive use, marital status of women, age of women, breast feeding status, religion of women, educational level of women and wealth index. The univariable analysis given in Appendix revealed that except work status of women all of factors were significantly related to birth interval. All significant covariates in univariable analysis were included in multivariable analysis.

The comparisons of different parametric distributions with two shared frailty distribution of the models for different multivariable analysis were done by using the AIC criteria. The model that had minimum AIC criteria is accepted to be the best (Munda, 2012) gives the best fit for parametric shared frailty model. In our study, Weibull-inverse Gaussian shared frailty for the multivariable model had small AIC criteria i.e., 22092.56, which is appropriate model for describing birth interval dataset of women among rural Ethiopia when compared to the rest model.

The findings of this study revealed that contraceptive use of women had a significant effect on the birth interval of women with 5% level of significance and it prolonged time to get next child by the factor of Φ= 1.260 when not using contraceptive was used as reference. The result of the study shows that women who were used contraceptive method had longer time to get next child than those who were not using contraceptive method. Similar effect of contraceptive use has been observed in a study conducted in Southern and Northern Ethiopia where contraceptive users space birth longer than the non-users in each observed births. Our findings also comparable with the findings of Chowdhury*et al*. (2013) and (Gyimah, 2001).

The result of this study suggests that marital status was statistically significant predictive factor for birth interval of women among rural Ethiopia. This shows that women who were married have shorter survival time between successive children than women who were single, widowed and divorced. The findings of this study revealed that age of women had a significant effect on birth interval of women among rural Ethiopia with 5% level of significance. It shows that as age of women increases the time between successive children increases. As women get older and older there isa consistent decrement in the probability of having subsequent births. It is supported by studies donein southern Ethiopia and northern Iran support this finding (Yohannes*eta.l*, 2011 and Singh *et al*.,2010).

As the result of the study shows breast feeding status is statistically significant factor for determining birth interval of women. It shows that the survival time between successive children for women who were breast feeding their children was longer than those who were not breast feeding.

This finding is found to be in the same direction with the findings of Hemochandra*et al.* (2010) in Rural Manipur, India and the study revealed breastfeeding emerged as an important protective covariate that extended the birth interval irrespective of parity.

According to our study Protestant and Muslim were statistically significant factors from the category of religion for determining birth interval of women. It shows that the survival time between successive children for protestant and Muslim was shorter when we used orthodox as reference. This result is similar with the study done byAyanaw A. (2008)studied proximate determinants of birth interval length in Amhara Region: The case of FagitaLekomaWereda, Awi-Zone. Among the various socio- cultural factors influencing fertility, religion has been considered very important and Chowdhury*et al* (2013).

The findings of this study revealed that the educational level of women had a significant effect on the birth interval of women among rural Ethiopia with 5% level of significance and it prolonged interval by the factor of Φ= 1.152, 1.135 and 1.189 for primary, secondary and higher education respectively when not educated women was used as the reference group. The result of the study shows that women with higher educational level had more survived than those uneducated, primary education and secondary. A similar study in Ethiopia, Amhara region by Ayanaw A. (2008) andChowdhury A. *et al*. (2013) in Bangladeshi suggest that education were significantly associatedwith timing of birth interval. This finding is also agreed with studies by Shayan*et al.* (2008).Women with no formal education are more likely to have subsequent birth after the index child as compared to women with some formal education. Such an association between educational status and fertility planning has been observed in previous studies in Ghana and Nigeria where it was found that educated women are less likely to report next birth than uneducated women (AmoakoF, 2008). This could be due to the fact that educated women have better access to family planning information and services than uneducated women. This finding also similar with studies done by Chowdhury A. *et al*. (2013)

In this study, we used the region as a clustering (frailty) effect on modeling the determinants of birth interval of women among rural Ethiopia using 2011 EDHS data. The clustering effect were significant (p-value :< 1.5e-08) in weibull-inverse Gaussian shared frailty model.

This showed that there is heterogeneity between regions by assuming women within the same region share similar risk factors on birth interval i.e., the correlation within regions cannot be ignored and clustering effect was important in modeling the hazard function. In our study the adequacy of baseline distributions are checked by using graphs in figure (2). From the plot of exponential, Weibull and log-logistic distributions; the plot of weibull was more straight line compared with exponential and lo-logistic for birth interval dataset. Also, the goodness of the fit in the baseline models was assessed through residual plots. The results from the plot of estimated cumulative hazard rate of Cox-Snell residual with respect to Cox-Snell residuals also supports the goodness of fit.

5. Conclusions

The results of this study suggested thatWeibull-inverse Gaussian shared frailty model is better fitted to birth interval dataset than other parametric shared frailty models. There is a frailty (clustering) effect on the birth interval dataset that arises due to differences in distribution of time between successive children among region of Ethiopia.

The result of Weibull-inverse Gaussian shared frailty model showed that the factors that determine the timing of interval between successive children are contraceptive use; Age, marital status, Religion, breast feeding status and educational level of women are statistically significant. Goodness of the fit of baseline distribution by means of graphical method and Cox-Snell residuals plots revealed that weibull distribution is better when compared to exponential and log logistic distributions to model birth interval dataset.

Acknowledgement: First, and foremost, I thank almighty God as all is and has been possible through him and I would like to thank CSA.

Ethics approval and Consent of participate

Not applicable because data is taken EDHS, 2011

Consent for publication

Not applicable.

Availability of data Materials

Data is taken from EDHS or .CSA

Competing interests

No competing interests.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' contribution

Our contribution is analysis data in impetrating and identifying the model which best fit the birth interval dataset.

Author details

[1] Department of Statistics, College of Natural Sciences, Jimma University, Jimma, Ethiopia

Acknowledgement

I would like to thank Jimma University for giving me all necessary material when I do this research and all my staff department of statistics for invaluable comments, suggestions during the entire time of the research.

Appendix

Weibull- Inverse-GaussianUnivariable analysis

Variables	Coef	S.E (Coef)	Chi-sq	Φ	95%CI LCL UCL	P-value
Work status						
(Intercept)	3.88946	0.0489	6317.9	48.884	44.4167, 53.8016	0.00
No	Ref					
Yes	0.00551	0.0174	0.1	1.005	0.9718, 1.0404	7.50e-01
Contraceptive						
(Intercept)	3.843	0.0432	7910.4	46.665	42.8767, 50.7886	0.00
No	Ref					
Yes	0.313	0.0271	132.6	1.367	1.2968, 1.4421	0.00
Marital status	4.459	0.238	350.94	86.401	54.1913,137.7554	0.00

(Intercept)	Ref					
Single	-0.593	0.234	6.42	0.553	0.3494, 0.8743	1.10e-02
Married	-0.305	0.244	1.57	0.737	0.4569, 1.1891	2.10e-01
Widowed	-0.16	0.24	0.44	0.852	0.5324, 1.3639	5.10e-01
Divorced						
Age						
(Intercept)	3.8024	0.0658	3338.96	44.809	39.3868, 50.9767	0.00
15-19	Ref					
20-24	-0.0344	0.0483	0.51	0.966	0.8789, 1.0621	4.80e-01
25-29	0.0061	0.0471	0.02	1.006	0.9174, 1.1034	9.00e-01
30-34	0.0817	0.0484	2.85	1.085	0.9869, 1.1931	9.10e-02
35-39	0.1678	0.0496	11.46	1.183	1.0731, 1.3034	7.10e-04
40-44	0.3957	0.0583	46.11	1.485	1.3250, 1.6652	1.10e-11
45-49	0.5801	0.0842	47.46	1.786	1.5145, 2.1067	5.60e-12
Breast feeding						
(Intercept)	4.234	0.0534	6289.5	68.993	62.1366, 76.6051	0.00
No	Ref					
Yes	0.0524	0.0118	19.91	1.0538	1.0297,1.0785	0.00
Religion						
(Intercept)	4.0086	0.0446	8081.81	55.07	50.4601, 60.1004	0.00
Orthodox	Ref					
Catholic	0.0268	0.1065	0.06	1.027	0.8336, 1.2656	8.00e-01
Protestant	-0.1691	0.0349	23.52	0.844	0.7886, 0.9042	1.20e-06
Muslim	-0.1533	0.0298	26.49	0.858	0.8092, 0.9095	2.60e-07
Traditional	-0.124	0.0908	1.87	0.883	0.7394, 1.0554	1.70e-01
Others	-0.1716	0.0753	5.19	0.842	0.7267, 0.9763	2.30e-02
Education						
(Intercept)	3.881	0.0475	6675.82	48.473	44.1636, 53.2022	0.00
No education	Ref					
Primary	0.039	0.0217	3.24	1.04	0.9965, 1.0849	7.20e-02
Secondary	0.241	0.1367	3.1	1.272	0.1367, 1.6635	7.80e-02
Higher	0.722	0.2888	6.26	2.058	1.1688, 3.6257	1.20e-02

Wealth						
(Intercept)	3.8603	0.0499	5992.78	47.48	43.0558, 52.3580	0.00
Poor	Ref					
Medium	0.0648	0.0229	8.01	1.067	1.0201, 1.1159	4.60e-03
Rich	0.08	0.022	13.25	1.083	1.0376, 1.1310	2.70e-04

REFERENCES

Aalen O. (2008).Survival and Event History Analysis. Springer-Verlag, New York.

Agudelo, A. Rosas-Bermúdez, and A. C. Kafury-Goeta, "Birth spacing and risk of adverse perinatal outcomes: a meta-analysis,"

Allison, P. D. Survival analysis using the SAS system. Cary, NC SAS Institute. (1995).

Amoako F, and Madise N, (2008): Examining the geographical Heterogeneity associated with risk of mistimed and unwanted pregnancy in Ghana, Journal of Biosocial Science, pp 1-19.

Andersen PK, Borgan O, Gaill RD, Keiding N (1993) Statistical methods based on counting processes.

Ayanaw A. (2008): Proximate Determinants of Birth Interval Length in Amhara Region: The Case of FagitaLekomaWereda, Awi-Zone. Unpublished M.Sc. Thesis, Population Study Department, Addis Ababa University, Ethiopia.

Central Statistical Agency, (2011). Ethiopian Demographic and Health Survey, Addis Ababa, Ethiopia

Chowdhury A. and Karim A. (2013): Patterns and Differentials of Birth Intervals in Bangladesh, Global Journal Of Science Frontier Research Interdisciplinary, Volume 13,2249-4626.

Clayton, D., &Cuzick, J. (1985). Multivariate generalizations of the proportional hazards model, Journal of the Royal Statistical Society. Series A (General), 82-117.

Cox and Snell. (1968): A general definition of residuals (with discussion).

Duchateau, L. and Janssen, P. (2008): The Frailty Model. Springer: New York.

Duchateau L. and Janssen P. (2004). Penalized partial likelihood for frailties and smoothin splines in time to first insemination models for dairy cows. Biometrics, 60, 608 -614.

Ellermann R., S. Pasquale and J. M. Tien. (1992). An Alternative Approach to Modeling Recidivism Using Quantile Residual Life Functions. Operations Research 40:485-504.

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

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

Hemochandra L, Singh NS, Singh AA.Factors Determining the Closed Birth Interval in Rural Manipur. J Hum Ecol. 2010;29:209–213.

Hougaard P (1984) Life table methods for heterogeneous populations: Distributions describing heterogeneity. Biometrika 71:75–83

Hougaard, P. (2000). Analysis of Multivariate Survival Data,

Kaplan and Meier. (1958*):* Non-parametric estimation from incomplete observations. Journal of American Statistical Association

Lancaster, T. (1979) Econometric Methods for the Duration of Unemployment.Econometrica, 47, 939-956.

Munda M., Rotolo F. and Legrand C. (2012): Parametric Frailty Models in *R*. Journal of American Statistical Association: 55:1-21.

Rutstein SO (2008): Further Evidence of the Effects of Preceding Birth Intervals onNeonatal, Infant, and Under-Five-Years Mortality and Nutritional Status in Developing Countries: Evidence from the Demographic and Health Surveys. In Demography and health research, Volume 41.

Shayan Z., Mohammad S., Zare N. and Moradi F. (2014): Prognostic factors of first birth interval using the parametric survival models. Iranian Journal of Reproductive Medicine Vol.12.No. 2.pp: 125-130

Singh SN, Singh N, Narendra RK.(2010): Demographic and Socio-economic Determinants of Birth Interval Dynamics in Manipur: A Survival Analysis. Online J Health Allied Scs.; 9(4):3.

Smits L, Pedersen C, Mortensen P, van Jim OS (2003): Association between short intervals and schizophrenia in the offspring. Schizophr Res, 70:49–56.

Therneau T. and Grambsch P. (2000).Modeling Survival Data.Springer, New York.

T.M. Therneau and P.M. Grambsch (2000): Modeling Survival Data: Extending the Cox Model, Springer, NewYork

Vaupel J., Manton K. and Stallard E. (1979).The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality.Demography. 16: 439-454

Wienke A. (2011). Frailty Models in Survival Analysis.Chapman & Hall/CRC biostatistics series, London.

WHO, Report of a WHO technical consultation on birth spacing, World Health Organization, Geneva, Switzerland, 2007.

Yohannes S, Wondafrash M., and Abera M, Girma E. (2011): Duration and determinants of birth interval among women of child bearing age in Southern Ethiopia. BMC Pregnancy and Childbirth, 11:38.