Pak. J. Statist.
2025 Vol. 41(3), 269-288

ESTIMATING FECUNDABILITY IN BANGLADESH: A MODELING APPROACH

Mohammad Manjur Alam[1§], Md. Abdul Karim[2]
and **Mohammad Salim Zahangir[2]**
[1] Department of Computer Science and Engineering
International Islamic University, Chittagong, Bangladesh
[2] Department of Statistics, University of Chittagong, Bangladesh
[§] Corresponding author Email: manjufse@iiuc.ac.bd

ABSTRACT

Fecundability is an important concept in both demography and reproductive health, as it is a key determinant of fertility. In this study, an attempt has been made to estimate the current fecundability level in Bangladesh. The study is based on data from the nationally representative Bangladesh Demographic and Health Survey (BDHS) 2017-18. A total of 6,140 currently married women out of 20,170 ever-married women were extracted for this study. Both geometric and beta-geometric models were applied in this study. The findings of this study reveal that the overall fecundability level in Bangladesh is 0.053. The study further shows that the estimated fecundability level is 0.056 when reproductive behavior of women's is treated as homogeneous. However, when the cohort is assumed to be heterogeneous, the estimated fecundability levels are 0.056 and 0.057 for the methods of moments (MOM) and maximum likelihood (MLE), respectively. A more robust estimate of fecundability is found by the Bayesian approach and is 0.053. This study will help policy makers, healthcare providers, population scientists and epidemiologists to adopt effective policies and programs for the individuals and couples achieving their desired family size.

KEYWORDS

Fecundability, Estimates, Homogeneous, Heterogeneous, Bayesian estimation, Bangladesh.

1. INTRODUCTION

One of the most important and challenging areas of study in demography and reproductive health is fecundability. The term "fecundability" referred to the time to conception, is defined as the chance of conception per menstrual cycle (Konishi et al., 2021). In biology, fecundity is defined as the reproductive potential of an individual. Epidemiological studies measure fecundity by using the waiting time to pregnancy as fecundity within a 12-month period or cycles. Generally, a woman's fecundability is estimated based on the first conception delays. It has been argued that the first birth is less affected by memory lapses and free from errors normally connected with the amenorrhea period common to the closed birth intervals.

© 2025 Pakistan Journal of Statistics

Earlier studies have examined several methods of estimating fecundability. These methods include: (i) estimation of fecundability by coital frequency and the duration of viability of sperm and ovum (Tietze, 1960; Lachenbruch, 1967); (ii) by observing the proportion of individual who conceive during a one-month exposure (Gini, 1924; Henry, 1953; Sheps, 1965; Potter, 1961); (iii) by fitting models to the distribution of waiting times to first conception (Jain, 1969; Majumdar and Sheps, 1970; Sheps, 1973; Mode, 2012; Maheshwari et al., 2015; Islam et al., 2015); (iv) fitting models to birth interval distributions (Singh, 1969; Mazharul and Yadava, 1997); (v) fitting models to the distribution of parities attained in a given time interval by a group of women (Singh, 1969; Balakrishnan, 1979); (vi) estimate the mean fecundability using various symmetric and asymmetric loss functions (Podder, 2019).

Among these methods, the first two have the constraint that their outcomes are highly dependent on assumptions about the length of the fertile period and the amount of the fertile period that coincides with subsequent recognizable pregnancy, but there are no precise estimates for either of these factors. In addition, the estimates for the second method are likely to be too high, as some of the respondents who were already pregnant at the time of wedding may state that they became pregnant immediately afterwards. Models of the 3rd type produced fairly consistent estimates of fecundability and are based on relatively few assumptions, but they often rely on the woman's memory to determine the number and timing of conceptions, and also have the restraint that they only apply to the commencement of conjugal life. Methods four and five are associated to a considerable number of assumptions and data input, which may not be capable, for example, the length of the postpartum anovulation period. Finally, the most significant drawback of method six is that it was accompanied without incorporating any data. Following the discussion above, it is evident that the third approach is the most suitable option due to its limited constraint and compatibility with the data employed in this study. The only assumption of the third method is postpartum sterility after foetal loss, and is less susceptible to inaccurate data from respondents who may have become pregnant before marriage.

There have been few attempts to estimate fecundability, especially in Bangladesh, because measuring fecundability is very complex as it is an indirect phenomenon. Using a geometric distribution and Pearson's type 1 distribution, this study estimates fecundability owing to the reciprocal of mean conception delays (Sheps, 1973).In addition, considering the 2nd method, a simple procedure was used to estimate fecundability based on observations of the proportion of women who become pregnant during the first month in which they are exposed to the risk of conception. Both classical and non-classical methods are used to estimate the parameters of the distributions. In classical approach, we use the MOM and the method of MLE, while the non-classical approach uses a Bayesian estimation procedure in which the results obtained by the MOM and MLE method are considered priors. In case of geometric distribution, the women reproductive characteristics are considered as homogeneous, while in Pearson's type 1 distribution the same is considered as heterogeneous.

The rationale for using fecundability in this study lies in its role as an important biological measure of a couple's probability of conception within a menstrual cycle. In the context of Bangladesh, where fertility patterns are undergoing significant demographic and socio-economic transitions, estimating fecundability compromises crucial insight into the

underlying biological and behavioral factors that influence reproductive outcomes. By applying a robust modeling approach, this study aims to produce accurate estimates of fecundability, which can help policy makers, public health officials, reproductive health programmers, demographers and researchers in addressing fertility issues and improving maternal and child health.

2. MATERIALS AND METHODS

In order to effectively conduct the proposed research, the current BDHS 2017-18 dataset was used. Out of a total of 20,170 ever-married women, a sample of 6,140 currently married women was extracted from this dataset. The BDHS is a two-stage, nationally representative survey across eight administrative divisions of Bangladesh. The data set is adequate and trustworthy because it is conducted by the National Institute for Population Research and Training (NIPORT) and sponsored by the United States Agency for International Development (USAID). The survey gathered up-to-date information on a number of key factors related to reproductive and maternal health.

The outcome variable 'Fecundability' is estimated based on the data on first conception delays, which is usually calculated by subtracting 9 months from marriage to first birth interval, as it is assumed that women take substantial period of 9 months on an average for each birth (Bongaarts, 1975). Thus,

i. Conception interval (CI) for live birth (closed first CI) = Date of first live birth (in CMC) –Date of first cohabitation (in CMC) – 9 months (gestational period for live birth).

ii. CI (non-live birth) = Date of first pregnancy terminated– Gestational age at the time of termination – Date of first cohabitation (in CMC).

iii. CI for currently pregnant women (Open first birth interval) = Date of interview — Gestational age at the time of survey– Date of first cohabitation (in CMC).

The details description of the sample selection procedure is set out below:

We began with 20,170 ever-married women from the BDHS dataset. After applying the appropriate sampling weight, we retained 18,894 currently married women. Among them, 1,882 women had no live birth, while 17,102 had at least one live birth.

The 1,882 women without a live birth were categorized into two groups:

- Those who were pregnant at the time of the interview (314 respondents) and
- Those whose first pregnancies had terminated (spontaneous and nonspontaneous abortion, stillbirth and miscarriage) before the interview (155 respondents).

From the 155 terminated pregnancy cases, we calculated the conception interval using the formula:

CI (non-live birth) = Date of first pregnancy termination – Gestational age at the time of termination – Date of first cohabitation (in CMC).

From the remaining 1,727 respondents (i.e. 1,882 minus 155), 314 were pregnant at the time of the survey. For these respondents, the conception interval was calculated as follows:

CI for currently pregnant women = Date of interview – Gestational age at the time of survey – Date of first cohabitation (in CMC).

Among the 17,102 women with at least one live birth, we calculated the conception interval for the first birth using the formula:

CI for live birth (closed first CI) = Date of first live birth (CMC) – 9 months (gestational period for live birth) – Date of first cohabitation (CMC).

We then excluded women who gave birth more than five years prior to the survey and those who did not give birth within the first 10 years of marriage. After excluding 11,431 women, a total of 5,671 valid sample units remained in this group.

In total, our final analytical sample comprised:

- 5,671 women with at least one live birth,
- 155 women with a terminated first pregnancy, and
- 314 currently pregnant women,

This resulted in a final sample size of 6,140 respondents for the study.

Due to the potential for misreporting of pregnancy outcomes and dates, women who gave birth more than five years prior to the survey and those who did not give birth within the first 10 years of marriage (1.1% of the sample) were excluded from the analysis considering them as primarily sterile. Therefore, we have considered the women having at least one recognizable conception and whose outcome consists of live birth, non-live births, and also who are currently pregnant. Since we have considered the first recognizable conception, the closed birth interval has been considered in this analysis for determining the conception wait for live birth and non-live birth of the first conception, and open birth interval for those who are currently pregnant for the first time. Finally, estimation of the parameters a, and b of the Pearson's type 1 model (Beta-geometric) was carried out using the C++ language.

2.1 A Model of Fecundability for Homogeneous Women

Consider a homogeneous population, in which fecundability is constant not only among women but also from month to month. Let the parameter $P; 0 < P < 1$, which represents fertility in a given month, and assume that each menstrual cycle represents an independent trial. We also assume that (i) all these women marry on the same time; (ii) they have no premarital conceptions; (iii) they have no desire to delay their first birth and (iv) marriage takes place at the beginning of the cycle.

The month in which conception occurs is indicated by the random variable X, then, $Prob\ (X = 1) = P$ which is the probability that the conception will occur in the first month. If $X = x$, so that $x > 1$, then in the preceding $(x - 1)$ months conception has not occurred with probability $(1 - P)^{x-1}$ and first conception occurs in x month with probability P. Therefore,

$$\boldsymbol{P}[X = x] = pq^{x-1}\,; x = 1,2,3 \ldots\ldots \quad (1)$$

where, $q = 1 - p$, formula (1) is the desired formula for random variable X which is known as geometric distribution with fecundability parameter P. The probability that conception does not occur in the first x months is

$$\boldsymbol{P}[X > x] = q^x \,; x = 1,2,3 \ldots\ldots$$

The survival function S(x) of this distribution is

$$\boldsymbol{S(x)} = (1-p)^x = q^x$$

By definition the mean of the distribution is,

$$\boldsymbol{E[X]} = \boldsymbol{m} = \frac{1}{p}$$

where, m specifies the average number of ovulations before conception and the variance of the distribution is $\boldsymbol{V}[\mathbf{T}] = \frac{q}{p^2}$

Thus we can find $\bar{X} = \frac{\mathbf{1}}{\boldsymbol{n}} \sum_{\boldsymbol{i=1}}^{\boldsymbol{N}} \boldsymbol{X_i}$ then moment estimator of P is $\frac{1}{\bar{X}}$ which is also same by the method of maximum likelihood estimator.

2.2 A Model of Fecundability for Heterogeneous Women

The type 1 geometric model ensuring the following assumptions that fecundability 'p' of all couples remains stable from month to month until pregnancy. Among couples, fecundability is distributed according to Pearson type 1 model. Conception is a chance event, depends on fecundability. The number of couples is sufficient. In reality, however, fecundability is not the same for all women, as they have intercourse at different frequencies and have different biological characteristics.

Thus, the fecundability, P of a particular woman may be considered as a random variable. Hence, the distribution of X is

$$\boldsymbol{P}[X = x/p] = P(1-P)^{x-1}, 0 \leq P \leq 1 \,; x = 1,2,3 \ldots\ldots$$

This is known as the conditional distribution of X. Now if p varies among woman to woman according to beta distribution then p has the following probability density function

$$f(p) = \frac{1}{\beta(a,b)} p^{a-1}(1-p)^{b-1}; 0 < p < 1, a, b > 0 \qquad (2)$$

where, $\beta(a,b)$ in equation (2) is the well-known Beta type 1 function defined by the integral

$$\beta(a,b) = \int_0^1 p^{a-1}(1-p)^{b-1} dp$$

The fundamental equation of the relationship between gamma and beta function is

$$\beta(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

It is known that the beta function can be expressed by the gamma function, which is defined as follows

$$\Gamma(a) = \int_0^{\alpha} x^{a-1} e^{-x} dx$$

The unconditional probability of the conception wait X is

$$g(x) = P[X = x] = \int_0^1 P[X = x / p] f(P) dp$$

$$= \int_0^1 f(P) f(x / p) dp \tag{3}$$

$$= \int_0^1 pq^{x-1} f(x / p) dp$$

$$= \int_0^1 pq^{x-1} \frac{1}{\beta(a,b)} P^{a-1}(1-P)^{b-1} dp$$

$$= \frac{\beta(a+1, x+b+1)}{\beta(a,b)}$$

$$g(x,a,b) = \begin{cases} \frac{a}{a+b}; & if \quad x = 1 \\ \frac{ab(b+1)(b+2).....(b+x-2)}{(a+b)(a+b+1)(a+b+2).....(a+b+x-1)}; & \text{if } x \geq 2 \end{cases}$$

And the expected number of women conceiving during the first month of exposure is given by

$$n(1) = n.P(x = 1)$$

The expected number of women conceiving during the xth month is

$$n(x) = n.P(X = x);\ x \geq 2$$

Therefore, the resulting mixture of distribution defined by the integral (3) is known as beta geometric distribution. This distribution is also known as type 1 geometric distribution.

Thus in the heterogeneous case the mean waiting time for first conception is

$$E[T] = \frac{1}{Harmonic\ mean\ of\ P}$$

and in the homogeneous case,

$$E[T] = \frac{1}{Arithmetic\ mean\ of\ P}.$$

3. ESTIMATION OF PARAMETERS OF THE MODEL

The appropriate method for estimating fecundability parameters will be determined by the available data as well as the models used. The following outlines the procedure for parameter estimation using three distinct methods:

3.1 The Method of Moments (MOM)

Suppose, m_1 is the first and m_2 is the second observed sample raw moments of the month required for women to become pregnant for the first time after their cohabitation. Therefore,

$$E(x) = \sum_{x=1}^{\infty} xp(1-p)^{x-1} = \frac{1}{P},$$

and

$$E(x^2) = \sum_{x=1}^{\infty} x^2 p(1-p)^{x-1} = \frac{2}{p^2} - \frac{1}{p} = \frac{2-p}{p^2}$$

To get the unconditional moments of X we have to put

$$\mu_1' = E(x) = \int E\left(\frac{1}{p}\right) f(p)dp$$

$$= \frac{\Gamma a + b}{\Gamma a \Gamma b} \int_0^1 p^{a-2}(1-p)^{b-1} dp$$

$$\mu_1' = \frac{a+b-1}{a-1} \text{ and, } \mu_2' = \frac{2(a+b-1)(a+b-2)}{(a-1)(a-2)} - \frac{a+b-1}{a-1}$$

By equating m_1 with μ_1' & m_2 with μ_2' we get

$$\hat{a} = \frac{2s^2}{s^2 - m_1^2 + m_1} = \frac{2s^2}{s^2 - m_1(m_1 - 1)} = f_1(m_1, m_2) = f_1$$

and

$$\hat{b} = (\hat{a} - 1)(m_1 - 1) = f_2(m_1, m_2) = f_2$$

where, $s^2 = m_2 - m_1^2$ be the observed sample variance of the months required to conceive for the first time after their cohabitation. It is important to note $s^2 > m_1^2 - m_1$. Otherwise the value of $\hat{a}$ will be smaller than 2 for that the theoretical second raw moment and variance become undefined and subsequently impossible to derive a moment estimate. Thus, the technique is imperfect where the estimated value of a is greater than 2. Substituting the estimated values of 'a' and 'b' in the formulae of mean, mode and variance of fecundability we get, $\bar{P}_a$, P_m and v^2.

Thus the estimating $\bar{P}_a$ is

$$\text{Estimated } \bar{p}_a = \frac{2s^2}{\left[s^2(m_1+1)+m_1(m_1+1)^2\right]}$$

For large sample, the asymptotic variance $\hat{a}$, $\hat{b}$ and covariance of $\hat{a}$ and $\hat{b}$ can be found using the technique explained by (Rao, 1952):

$$\text{Var}(\hat{a}) = \left\{\left(\frac{\partial f_1}{\partial m_1}\right)^2\Big|_{m=\mu'}\right\} V(m_1) + \left\{\left(\frac{\partial f_1}{\partial m_2}\right)^2\Big|_{m=\mu'}\right\} V(m_2) + \left\{2\left(\frac{\partial f_1}{\partial m_1}\right)\right\}\left\{\left(\frac{\partial f_1}{\partial m_1}\right)\Big|_{m=\mu'}\right\} Cov(m_1, m_2)$$

Thus, Var ($\hat{a}$) = 14.27. Where m and μ' are vectors of the respective moments and

$$\left(\frac{\partial f_1}{\partial m_1}\right)_{m=\mu'} = \left[\frac{2(2m_1m_2 - m_1^2 - m_2)}{(m_2 - 2m_1^2 + m_1)^2}\right]_{m=\mu'} = \frac{2(2\mu_1'\mu_2' - \mu_1^2 - \mu_2')}{(\mu_2' - 2\mu_1'^2 + \mu_1')^2}$$

$$= \frac{2(2\times 18.84\times 737.05 - 18.84^2 - 737.05)}{(737.05 - 2\times 18.84^2 + 18.84)^2} = 25.21$$

$$\left(\frac{\partial f_1}{\partial m_2}\right)_{m=\mu'} = \frac{-2\mu_1'(\mu_1'-1)}{(\mu_2' - 2\mu_1'^2 + \mu_1')^2} = \frac{-2\times 18.84(18.84-1)}{(737.05 - 2\times 18.4^2 + 18.84)^2} = -0.312$$

Similarly the asymptotic variance of $\hat{b}$ is

$$\text{Var}(\hat{b}) = \left\{\left(\frac{\partial f_2}{\partial m_1}\right)^2\Big|_{m=\mu'}\right\} V(m_1) + \left\{\left(\frac{\partial f_2}{\partial m_2}\right)^2\Big|_{m=\mu'}\right\} V(m_2) + \left\{2\left(\frac{\partial f_2}{\partial m_1}\right)\right\}\left\{\left(\frac{\partial f_2}{\partial m_1}\right)\Big|_{m=\mu'}\right\} Cov(m_1, m_2)$$

Thus, Var ($\hat{b}$) = 4689.85. The asymptotic covariance of $\hat{a}$ and $\hat{b}$ is

$$\text{Cov}(\hat{a}, \hat{b}) = \left\{\left(\frac{\partial f_1}{\partial m_1}\right)\left(\frac{\partial f_2}{\partial m_1}\right)\Big|_{m=\mu'}\right\} V(m_1) + \left\{\left(\frac{\partial f_1}{\partial m_2}\right)\left(\frac{\partial f_2}{\partial m_2}\right)\Big|_{m=\mu'}\right\} V(m_2) + \left\{\left(\frac{\partial f_1}{\partial m_1}\right)\left(\frac{\partial f_2}{\partial m_2}\right) + \left(\frac{\partial f_1}{\partial m_2}\right)\left(\frac{\partial f_2}{\partial m_1}\right)\Big|_{m=\mu'}\right\} Cov(m_1, m_2)$$

and cov($\hat{a}$ $\hat{b}$) = 258.62. where,

$$\left(\frac{\partial f_2}{\partial m_1}\right)_{m=\mu'} = \frac{\mu_2'^2 + 2\mu_2'\mu_1'^2 - 6\mu_1'\mu_2' + 2\mu_2' + \mu_1'^2}{(\mu_2' - 2\mu_1'^2 + \mu_1')^2}$$

$$= \frac{737.05^2 + 2\times 737.05\times 18.84^2 - 6\times 18.84\times 737.05 + 2\times 737.05 + 18.84^2}{(737.05 - 2\times 18.84^2 + 18.84)^2} = 465.57$$

$$\left(\frac{\partial f_2}{\partial m_2}\right)_{m=\mu'} = \frac{-2\mu_1'(\mu_1' - 1)^2}{(\mu_2' - 2\mu_1'^2 + \mu_1')^2} = \frac{-2\times 18.84(18.84-1)^2}{(737.05 - 2\times 18.4^2 + 18.84)^2} = -5.668$$

If m_r and μ_r' are respectively, the rth sample and population moments about zero then by Kandall and Stuart (1958)

$$E(m_r) = \mu_r' \text{ and, } V(m_r) = \frac{1}{n}\left(\mu_{2r}' - \mu_r'^2\right)$$

Therefore,

$$V(m_1) = \frac{1}{n}\left(\mu_2' - \mu_1'^2\right) = 0.062 \text{ and, } V(m_2) = \frac{1}{n}\left(\mu_4' - \mu_2'^2\right) = 596.71$$

and $Cov(m_1, m_2) = \frac{1}{n}\left(\mu_3' - \mu_1'\mu_2'\right) = 5.29$

Here,

$$\mu_1' = \frac{\hat{a}+\hat{b}-1}{\hat{a}-1} = 18.84, \ \mu_2' = \frac{2(\hat{a}+\hat{b}-1)(\hat{a}+\hat{b}-2)}{(\hat{a}-1)(\hat{a}-2)} - \frac{\hat{a}+\hat{b}-1}{\hat{a}-1} = 737.05$$

$$\mu_3' = \frac{6(\hat{a}+\hat{b}-1)(\hat{a}+\hat{b}-2)(\hat{a}+\hat{b}-3)}{(\hat{a}-1)(\hat{a}-2)(\hat{a}-3)} - \frac{6(\hat{a}+\hat{b}-1)(\hat{a}+\hat{b}-2)}{(\hat{a}-1)(\hat{a}-2)} + \frac{\hat{a}+\hat{b}-1}{\hat{a}-1}$$
$$= 46418.55$$

$$\mu_4' = \frac{24(\hat{a}+\hat{b}-1)(\hat{a}+\hat{b}-2)(\hat{a}+\hat{b}-3)(\hat{a}+\hat{b}-4)}{(\hat{a}-1)(\hat{a}-2)(\hat{a}-3)(\hat{a}-4)}$$
$$-36\frac{(\hat{a}+\hat{b}-1)(\hat{a}+\hat{b}-2)(\hat{a}+\hat{b}-3)}{(\hat{a}-1)(\hat{a}-2)(\hat{a}-3)}$$
$$+14\frac{(\hat{a}+\hat{b}-1)(\hat{a}+\hat{b}-2)}{(\hat{a}-1)(\hat{a}-2)} - \frac{\hat{a}+\hat{b}-1}{\hat{a}-1} = 4207030.90$$

Hence the variance-covariance matrix (V_1) of the estimates of the parameters a, and b obtained by the MOM is

$$V_1 = \begin{pmatrix} \text{var}(\hat{a}) & \text{cov}(\hat{a},\hat{b}) \\ \text{cov}(\hat{a},\hat{b}) & \text{var}(\hat{b}) \end{pmatrix} = \begin{pmatrix} 14.27 & 258.62 \\ 258.62 & 4689.85 \end{pmatrix} = 39.85$$

The correlation co-efficient between the estimators of the two parameters given by

$$r = \frac{cor(\hat{a},\hat{b})}{\sqrt{\text{var}(\hat{a})\,\text{var}(\hat{b})}} = 0.9997$$

Arithmetic mean of fecundability

$$(\hat{p}_a) = \frac{\hat{a}}{\hat{a}+\hat{b}} = 0.056$$

Harmonic mean of fecundability

$$(\hat{p}_h) = \frac{\hat{a}-1}{\hat{a}+\hat{b}-1} = 0.053$$

and the estimate of the variance of fecundability is

$$v = \{\text{var}(\hat{p})\} = \frac{\hat{a}\hat{b}}{(\hat{a}+\hat{b})^2(\hat{a}+\hat{b}+1)} = 0.0002$$

3.2 Method of Maximum Likelihood Estimators (MLE)

Estimation of parameters using the MOM have been observed that the MOM of $\hat{a}$ and $\hat{b}$ are moderately reliable only within a certain range of values of a. Within the upper limit of a, the estimates are either very inefficient or their variances are not clearly defined. The method of MLE does not suffer the complications found in the MOM and is well-defined for all admissible values of a, and b.

Let us consider a probability density function $g(x;\ a,b)$ for conception in month $x(x=1,2,3,.....)$ conditional on a, and b. Thus the likelihood function (L) is

$$L = \prod_{i=1}^{n} g(x_i;a,b)$$

Now $f(x)$ represents the observed frequency of conception in month x and w denotes the highest recorded value of x, such that $\sum_{x=1}^{w} f(x) = n$; then we have

$$\ln L = \sum_{x=1}^{w} f(x)\ln g(x_i;a,b)$$

$$= \sum_{x=1}^{w} f(x)\begin{cases} \ln a - \ln(a+b); x=1 \\ \left[n-f(1)\right]\ln a + \sum_{x=2}^{w} f(x)\sum_{j=0}^{x-2}\ln(b+j) - \sum_{x=2}^{w} f(x)\sum_{j=o}^{x-1}\ln(a+b+j); x\geq 2 \end{cases}$$

$$= \begin{cases} f(1)\left[\ln a - \ln(a+b)\right] \; ; x = 1 \\ \ln a + \sum_{j=0}^{x-2} \ln(b+j) - \sum_{j=0}^{x-1} \ln(a+b+j); x \geq 2 \end{cases}$$

The first order derivatives of $\ln L$ with respect to a, and b represented by s_1 and s_2 are given by

$$s_1 = \frac{\delta \ln L}{\delta a} = \sum_{x=1}^{w} f(x) \frac{\delta \ln g(x;a,b)}{\delta a}$$

$$s_2 = \frac{\delta \ln L}{\delta b} = \sum_{x=1}^{w} f(x) \frac{\delta \ln g(x;a,b)}{\delta b}$$

Solving the equations s_1 and s_2 by using commonly employed procedure is very complicated and it becomes clumsy without the assistance of a computer. There are various existing methods for deriving the MLE by iterative process to solve the simultaneous equations $s_1 = 0$, and $s_2 = 0$. In the current problem, we apply the C++ programming language to determine the optimal estimates of parameters a, and b that maximize the likelihood function lnL. The method of calculating the largest value of lnL consists of entering the initial values of a_0 and b_0 as well as a small quantity (δ), which is gradually reduced in the course of the iteration. The value of $\ln L$ is computed by using a_0, b_0 and $a_0(1 \pm \delta)$, $b_0(1 \pm \delta)$. The values of lnL are compared and the iteration is continued until the maximum value of lnL for the specific value of a and b is found. This program also provides numerical representations of second order derivatives.

The variance – covariance matrix of the MLE is obtained by using the Cramer-Raw lower bound. The likelihood estimates of a, and b as follows:

$$\ln L = \sum_{x=1}^{w} f(x) \ln g(x_i;a,b) = \begin{cases} f(1)\left[\ln a - \ln(a+b)\right] \; ; x = 1 \\ \ln a + \sum_{j=0}^{x-2} \ln(b+j) - \sum_{j=0}^{x-1} \ln(a+b+j); x \geq 2 \end{cases}$$

Taking partial derivatives

$$\frac{\delta \ln L}{\delta a} = \sum_{x=1}^{w} f(x) \frac{\delta}{\delta a} \{\ln g(x;a,b)\} = \sum_{x=1}^{w} f(x) \left(\frac{1}{a} - \sum_{j=0}^{x-1} \frac{1}{a+b+j} \right)$$

$$= \frac{n}{a} - \sum_{x=1}^{w} f(x) \sum_{j=0}^{x-1} \frac{1}{a+b+j} \; ; \; [\text{Since}, \sum_{x=1}^{w} f(x) = n]$$

$$\frac{\delta^2 \ln L}{\delta a^2} = -\frac{n}{a^2} + \sum_{x=1}^{w} f(x) \sum_{j=0}^{x-1} \frac{1}{(a+b+j)^2}; x \geq 1 \qquad (4)$$

$$\frac{\delta^2 \ln L}{\delta a \delta b} = \sum_{x=1}^{w} f(x) \sum_{j=0}^{x-1} \frac{1}{(a+b+j)^2}; x \geq 1 \tag{5}$$

$$\frac{\delta \ln L}{\delta b} = \sum_{x=1}^{w} f(x) \frac{\delta}{\delta b} \{\ln g(x;a,b)\}$$

$$= \sum_{x=1}^{w} f(x) \begin{cases} -\frac{1}{a+b}; x=1 \\ \sum_{j=0}^{x-2} \frac{1}{b+j} - \sum_{j=0}^{n} \frac{1}{a+b+j}; x \geq 2 \end{cases}$$

$$\frac{\delta^2 \ln L}{\delta b^2} = \begin{cases} \frac{1}{(a+b)^2}; x=1 \\ -\sum_{x=2}^{w} f(x) \sum_{j=0}^{x-2} \frac{1}{(b+j)^2} + \sum_{x=2}^{w} f(x) \sum_{j=0}^{x-1} \frac{1}{(a+b+j)^2}; x \geq 2. \end{cases} \tag{6}$$

The asymptotic covariance matrix of the MLE of a, and b is obtained as follows. Let for $p, q = 1, 2$

$$I_{p,q} = -E\left(\frac{\delta^2 \ln L}{\delta \lambda_p \delta \lambda_q}\right) = -\sum\left(\frac{\delta^2 \ln L}{\delta \lambda_p \delta \lambda_q}\right)$$

where $\lambda_1 = a$ and $\lambda_2 = b$. The matrix $I = \left(I_{p,q}\right)$ is called the information matrix. The asymptotic variance-covariance matrix of the MLE, $\hat{a}$ and $\hat{b}$ is the inverse of $I = \left(I_{p,q}\right)$, with elements:

$$Var(\hat{a}) = I^{11} = \frac{1}{I_{11}\left(1 - I_{12}^{\ 2} / I_{11} I_{22}\right)},\ Var(\hat{b}) = I^{22} = \frac{1}{I_{22}\left(1 - I_{12}^{\ 2} / I_{11} I_{22}\right)}$$

and

$$Cov(\hat{a}, \hat{b}) = I^{12} = I^{21} = -\frac{I_{12}}{\left(I_{11} I_{22} - I_{12}^{\ 2}\right)}$$

Taking into expectation over the equations (4), (5) and (6), we get the information matrix that is involved in many complex calculations. In order to avoid complexity and uncomfortably to express the information matrix, we have to define it in an easy way:

$$\hat{I}_{p,q} = \frac{\delta^2 \ln L}{\delta \hat{\lambda}_p \delta \hat{\lambda}_q}$$

The numerical value of (4), (5) and (6) is obtained by using C++ language and then the information matrix $I=\left(I_{p,q}\right)$, as well as the variance- covariance matrix $\left(I^{-1}\right)$. Hence, the information matrix is

$$I=\begin{pmatrix} 28.25 & -1.675 \\ -1.675 & 0.0896 \end{pmatrix}, \text{ and } I^{-1}=\begin{pmatrix} 0.33 & 5.80 \\ 5.80 & 102.94 \end{pmatrix}$$
$$=\begin{pmatrix} \text{var}(\hat{a}) & \text{cov}(\hat{a},\hat{b}) \\ \text{cov}(\hat{a},\hat{b}) & \text{var}(\hat{b}) \end{pmatrix}$$

Thus, the correlation coefficient between the parameters a, and b is

$$r=\frac{\text{cov}(\hat{a},\hat{b})}{\sqrt{v(\hat{a})v(\hat{b})}} = 0.9951$$

Arithmetic mean of fecundability

$$(\hat{p}_a)=\frac{\hat{a}}{\hat{a}+\hat{b}} =0.057$$

Harmonic mean of fecundability

$$(\hat{p}_h)=\frac{\hat{a}-1}{\hat{a}+\hat{b}-1} = 0.053$$

and the estimates of the variance of fecundability is

$$V=\left\{Var(\hat{p})=\frac{\hat{a}\hat{b}}{(\hat{a}+\hat{b})^2(\hat{a}+\hat{b}+1)}\right\} = 0.00021$$

3.3 Bayesian Method of Estimation

Once the parameters have been estimated using both MOM and MLE, we will obtain two different a priori Beta densities. These density functions will then be used in the Bayesian estimation of fecundability, assuming the two aforementioned estimates as priors for the fecundability parameter. Moreover, prior distributions are usually gathered based on the subjective beliefs of the investigators themselves (Susilo, 2021). Therefore, it is reasonable that the distribution of P follows the succeeding two-parameter beta distribution, and it is natural the conjugate prior distribution in the Bayesian sense.

$$f(P)=\frac{1}{B(a,b)}P^{a-1}(1-P)^{b-1} \quad ;0<P<1; \quad a,b>0 \tag{7}$$

The joint density function of the observed sample $x_1, x_2 \ldots\ldots x_n$ is

$$L(x,p)=P^n(1-P)^{\sum_i x_i-n}\frac{1}{B(a,b)}P^{a-1}(1-P)^{b-1} \tag{8}$$

Thus the posterior density of P is

$$g(p|x) = \frac{L(x|p)\, f(p|a,b)}{\int_0^1 L(x|p)\, f(p|a,b) d\theta} \tag{9}$$

Substituting the equation (7) and (8) in equation (9), we obtained the posterior density of P is

$$f(P|x) = \frac{1}{\beta\left(n+a, \sum_i x_i + b - n\right)} P^{n+a-1}(1-P)^{\sum_i x_i + b - n - 1}$$

Thus for the squared error loss, the Bayes estimator is the posterior mean and the mean fecundability is

$$\hat{P} = \frac{n+a}{\sum_i x_i + a + b}$$

where it is assume that the value of a, and b obtained by MOM and MLE.

(i) Using the MOM result, a reasonable prior density of P is

$$f(P) = \frac{1}{\beta\left(16.61,\ \ 278.48\right)} P^{16.61-1}(1-P)^{278.48-1}$$

The posterior density of p is

$$g(p|x) = \frac{1}{\beta(n+16.61,\ \sum x + 278.48 - n)} P^{(n+16.61)-1}(1-P)^{(\sum x + 278.48 - n)-1}$$

Thus mean fecundability is

$$\hat{P} = \frac{6140 + 16.61}{115686 + 16.61 + 278.48} = 0.053$$

ii) Using the MLE results, a reasonable prior density of P is

$$f(P) = \frac{1}{\beta\left(14.32,\ \ 238.48\right)} P^{14.32-1}(1-P)^{238.48-1}$$

The posterior density of P is

$$g(p|x) = \frac{1}{\beta(n+14.32,\ \ \sum x + 238.48 - n)}$$

$$P^{(n+14.32)-1}(1-P)^{(\sum x + 238.48 - n)-1}$$

Thus mean fecundability is

$$\hat{P} = \frac{6140 + 14.32}{115686 + 14.32 + 238.48} = 0.053.$$

4. RESULTS

4.1 Fitting of the Homogenous Model to Data

Let $U_1, U_2, \ldots\ldots, U_k$ be the random variable of the number of women who become pregnant in the first month. Then an estimator of P based on U_k is denoted by the $P_k{}^*$, with relative frequency $P_k{}^* = \frac{U_k}{k}$, since not all data is used one can reasonably be sure that that the estimator $P_k{}^*$ may not be as good as $P_k{}^*$ The following analysis will provide evidence that these suspects are well established.

Table 1
Estimates of Fecundability Based on Different Approaches, 2017-18

Estimates	Mean	Variance	Fecundability
Sample Mean ($\bar{X}$)	18.84	382.12	0.053
Exposure to the risk of conception in the first month ($P_k{}^*$)	11.63	123.58	0.086
Geometric distribution $(\hat{\boldsymbol{P}}_{\boldsymbol{K}})$	17.85	301.02	0.056

From the Table 1 we have seen that, the sample mean and variance are estimated as 18.84 and 382.12 respectively. Using the value of sample mean $\bar{X} = 18.84$, the estimates of fecundability is 0.053. When considering the observed frequency of women of conceiving in the first month (526) in the formula of $P_k{}^*$ we got another estimate, $P_k{}^* = 0.086$. Now we need to justify whether the sample can be considered as coming from a geometric distribution by comparing the estimates of the theoretical mean and variance with the observed values. The estimate of the mean and variance based on $\hat{P}_K$ is 17.85 and 301.02 respectively. Comparing with the observed mean and variance we have seen that fecundability estimates based on $P_k{}^* = 0.086$ differ significantly. The estimate of the mean based on $\hat{P}_K = 0.056$almost corresponds with the observed mean, but the estimate of the variance based on $\hat{p}_k$ again varies considerably from the observed variance. The results indicate that the geometric distribution may not be a suitable model for the data set analyzed in this study.

4.2 Fitting of the Heterogeneous Model to Data

The analysis of the data relating to beta geometric distribution has been described extensively in Section 2.2. Both MOM and MLE were used to estimate the parameters of the model in Section 2.2. At first estimates 'a' and 'b' were obtained by conventional method of moments. These estimates were used as initial values in the iterative procedure to derive estimates $\hat{a}$ and $\hat{b}$ by the method of MLE. Based on the estimates obtained by

the methods of MOM and MLE, levels of natural fecundability along with coefficient of variation were calculated. To examine the relationship between estimates, coefficient of correlation between $\hat{a}$ and $\hat{b}$, and also the standard errors of estimates are also calculated as per procedures and the findings are presents in Table 2.

Table 2
Estimates and Their Standard Errors, Coefficient of Correlation, Levels of Fecundability and Coefficient of Variation

Estimates	Method of Moments	Method of Maximum Likelihood
$\hat{a}$	16.61 ± 3.78	14.32 ± 0.57
$\hat{b}$	278.48 ± 68.48	238.48 ± 0.14
r	0.9997	0.9951
$\hat{p}_a$	0.056 ± 0.001	0.057 ± 0.003
$\hat{p}_h$	0.053 ± 0.0043	0.053 ± 0.0002
$Var(p)$	0.0002 ± 0.00005	0.00021 ± 0.0000093
CV	0.057	0.065

Note: $\hat{a}$ and $\hat{b}$ are the parameters of the beta function $f(p)$; r is correlation coefficient; cv is coefficient of variation.

The findings reflect that all estimation methods are close to the observed value of fecundability (0.053). In classical approach, the MLE give lower estimates of both the parameters a and b, higher estimates of the arithmetic mean and variance of P, and equal estimates of the harmonic mean of p. Since, estimated fecundability for Bayesian estimate (Section 2.3.3) is exactly coincide to the observed value (Table 1) of the parameters, and therefore, we conclude that among all methods, Bayesian estimation technique considerably perform better in terms of observed value. Furthermore, the values of correlation coefficient between a, and $b(r)$, through both MOM and MLE, provide validation for the model employed in this analysis. The coefficient of variation in both methods also indicates the existence of variability among the respondents. The findings suggest that estimating fecundability using the conventional method, such as the reciprocal of mean waiting time to first conception (homogeneous), would be an erroneous approach.

5. DISCUSSION

The results of the study reveal that the observed mean and variance of the sample are 18.84 and 382.12, respectively, and that the estimated fecundability is 0.053. The estimated

fecundability considering a first month of exposure to the risk of conception is 0.086 and differs considerably from the observed mean. The result further reveals that the estimated fecundability obtained from the geometric distribution is 0.056 considering that women are homogeneous in terms of their reproductive features. The estimated fecundability is 0.056 for the method of moments and 0.057 for the MLE and 0.053 for the Bayesian method considering a heterogeneous cohort of women. The fecundability obtained by the classical and non-classical estimation methods is close to the observed value of fecundability and ranges between 0.053 and 0.057. Thus, based on the different estimates, the overall level of fecundability of Bangladeshi women is determined to be 0.053. Moreover, the study found that the coefficient of variation of fecundability is extremely high, indicating significant heterogeneity among the respondents in terms of their reproductive health.

Compared to other countries in the world, where fecundability levels are comparatively high, our estimate is very low. Some Indian studies have also reported very low fecundability, ranging from 0.04 to 0.05 (Singh, 1969; Maheshwari et al., 2015; Singh et al., 2014). However, low fecundability is also evident from the results of several Bangladesh National Representative Surveys over the years. Using BDHS data 2011, (Islam et al., 2015) found the mean fecundability of Bangladeshi women is 0.051. Chowdhury and Umbach (2012) as well as Islam et al. (2005) estimated the level of fecundability in Bangladesh using data from BDHS, 1999-2000. Both studies obtained a similar result of 0.046 for the level of fecundability. The fecundability levels of Bangladesh were 0.06 in 1975 and 0.05 in 1989, based on data from nationally representative surveys in Bangladesh (Mazharul and Yadava, 1997). The lower fecundability of women in Bangladesh compared to their counterparts in Western societies may be attributed to a combination of social, biological and cultural factors as well as various sexual constraints. In Bangladesh, the female age at first marriage is much lower, especially in the rural areas and most of the women, mainly in the rural areas belonging to low socio-economic stratum, which is pronounced compared to other countries. Temporary and prolonged migration may be another important factor in low level of fecundability among Bangladeshi women. For instance, a large number of married men also leave their families and go abroad for livelihood. Many rural Bangladeshi husbands migrate to their main destination due to economic hubs or in search of better employment, leaving their families behind in their origins. This period of separation may naturally affect the coital frequency, which may in turn reduce a woman's likelihood of conception. In addition, memory lapses, misreporting of age at first cohabitation, termination of first pregnancy, and early spontaneous abortion may affect fecundability.

6. CONCLUSION

Fecundability, which is an important fertility parameter, is a complex process that requires the use of different estimation techniques. We used a large, nationally representative BDHS data set to estimate fecundability, which increases the robustness and generalizability of the results. The empirical distributions of fecundability, the proportion of women who conceive during the first month of exposure to the risk of conception, as well as geometric and beta-geometric distribution by the methods of moments, MLE and

Bayesian are fitted to estimate fecundability. By fitting empirical data to both geometric and beta-geometric distributions, we apprehended a more accurate picture of reproductive behavior and heterogeneity in the population. A comparative evaluation of these methods revealed that while the MOM is very simple and easier to calculate and provides initial estimates for obtaining the MLE, the moment estimates are not defined across all permissible values of a, and b. On the other hand, MLE have optimum properties and is preferable than MOM because of its desirable properties of consistency, efficiency, sufficiency and biasedness and also it can be defined for all permissible values of a, and b. Furthermore, when the sample size is large, as in this study's sample size of 6140, Bayesian inference often provides results for parametric models that are superior to the results produced by the other commonly used methods. Not only that, Bayesian model can easily accommodate unobserved variables such as "Fecundability" and make inferences without relying on conditional and precise, asymptomatic approximations on the data. All these evidence suggests that fecundability estimates by using Bayesian method are better than the moment and maximum likelihood methods.

Limitations:

The variable ‘fecundability’ is based on the estimation of first conception delays, which is highly dependent on the exact reporting of the dates of the first pregnancy outcome and first cohabitation. However, there is a tendency of misreporting the date of termination of the first pregnancy, owing to memory lapses of the respondents and digit preference, albeit a lack of a comprehensive vital registration system in Bangladesh.

Future Work:

Therefore, future research should cogent to record the exact date of relevant vital events, the date of contraceptive use, and the discontinuation of contraceptive use, including the respondents, their partners and their offspring, to compute an actual estimate of fecundability level in Bangladesh. Although our findings indicate that the Bayesian method performs better than the method of moments and MLE in terms of flexibility and precision, further studies should examine performance under different data conditions, including smaller sample sizes, the presence of censoring, and different prior distributions.

In addition, future research could benefit from analyzing longitudinal data to track attempts to conceive over time and account for background heterogeneities (e.g., region, education, contraceptive use, and socioeconomic status) to generate more nuanced estimates of fecundability across different population subgroups.

Policy Implications:

Our findings provide updated estimates of fecundability levels in Bangladesh based on a large nationally representative dataset. These findings can help policymakers, public health professionals, and reproductive health program managers to comprehensively understand the timing of first conception. This, in turn, can inform strategies to support individuals and couples in achieving their desired family size, improve fertility counseling,

and develop targeted interventions, particularly for populations with delayed or impaired fertility.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the authority of the National Institute of Population Research and Training (NIPORT), Medical Education and Family Welfare Division, Ministry of Health and Family Welfare for granting public access to the BDHS data. No funding was provided for the research, authorship or publication of this article.

CONFLICTS OF INTEREST

The authors affirm that they have no conflicts of interest to disclose.

REFERENCES

1. Balakrishnan, T.R. (1979). Probability of conception, conception delay, and estimates of fecundability in rural and semiurban areas of certain Latin American countries. *Social Biology*, 26(3), 226-231.
2. Bongaarts, J. (1975). A method for the estimation of fecundability. *Demography*, 12(4), 645-660.
3. Chowdhury, M.R. and Umbach, D. (2012). Some Bayesian Analyses of Fecundability. *Pak. J. Statist*, 28(3), 293-302.
4. Gini, C. (1924). Premières recherches sur la fécondabilité de la femme. In *Proceedings of the International Mathematical Congress* (Vol. 2, pp. 889-892). North-Holland: Toronto.
5. Henry, L. (1953). Fécondité des Mariages: Nouvelle Méthode de Mesure. Institut National d'Etudes Demographiques, Travaux et Documents. Cahier, 16
6. Islam, S.M.S., Karim, M.A., Chowdhury, M.S. and Hossain, M.S. (2005). Fitting and estimation of type-I geometric distribution by using C language. *Asian Journal of Information Technology*, 4(1), 112-126.
7. Islam, R., Islam, N., and Sheikh, D.N. (2015). Estimation of conception waits and fecundability: levels and trainds in Bangladesh. *International Journal of Advanced Research in Management and Social Sciences*, 4(10), 105-118.
8. Jain, A.K. (1969). Fecundability and its relation to age in a sample of Taiwanese women. *Population Studies*, 23(1), 69-85.
9. Konishi, S., Kariya, F., Hamasaki, K., Takayasu, L. and Ohtsuki, H. (2021). Fecundability and sterility by age: estimates using time to pregnancy data of Japanese couples trying to conceive their first child with and without fertility treatment. *International Journal of Environmental Research and Public Health*, 18(10), 5486.
10. Lachenbruch, P.A. (1967). Frequency and timing of intercourse: its relation to the probability of conception. *Population Studies*, 21(1), 23-31
11. Maheshwari, S., Singh, B.P. and Pudir, P.S. (2015). Analysis of Retrospective Time to First Pregnancy Data in Uttar Pradesh under Various Dimensions. *Journal of Institute of Science and Technology*, 20(1), 51-58.
12. Majumdar, H. and Sheps, M.C. (1970). Estimators of a type I geometric distribution from observations on conception times. *Demography*, 7, 349-360.

13. Mazharul Islam, M. and Yadava, R.C. (1997). Estimation of fecundability: Levels and trends in Bangladesh. *Journal of Family Welfare*, 43, 13-21.
14. Mode, C.J. (2012). *Stochastic processes in demography and their computer implementation* (Vol. 14). Springer Science & Business Media.
15. Podder, C.K. (2019). Estimation of fecundability in Bayesian approach using different loss functions. *Pak. J. Statist.* 35(1), 1-9
16. Potter, R.G. (1961). Length of the fertile period. *The Milbank Memorial Fund Quarterly*, 39(1), 132-162.
17. Radhakrishna Rao, C. (1952). Advanced statistical methods in biometric research. Willey, Newyork
18. Sheps, M.C. (1965). An analysis of reproductive patterns in an American isolate. *Population studies*, 19(1), 65-80.
19. Sheps, M.C., Menken, J.A., Menken, J. and Radick, A. P. (1973). *Mathematical models of conception and birth*. University of Chicago Press.
20. Singh, B.P., Pudir, P.S. and Maheshwari, S. (2014). Parameter estimation of beta-geometric model with application to human fecundability data. *arXiv preprint arXiv:1405.6392*.
21. Singh, S.N. (1969). Low estimates of fecundability. In *Proceedings of General Conference, London. IUSSP, Liege*.
22. Susilo, T., Kurniasari, D. and Aziz, D. (2021). Characteristics of Bayes Estimator in the Geometric Distribution with Prior Beta. In *Journal of Physics: Conference Series* (Vol. 1751, No. 1, p. 012020). IOP Publishing.
23. Tietze, C. (1960). Probability of pregnancy resulting from single unprotected coitus. *Fertility and Sterility*, 11(5), 485-488.