Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v19n1p62

Beta-Binomial-Poisson Mixture Model and Its Parameter Estimation: An Application to Number of Female Childbirths

By Tripathi, Misra, Kumar

March 15, 2026

Electronic Journal of Applied Statistical Analysis
Vol. 19, Issue 01, March 2026, 62-78
DOI: 10.1285/i20705948v19n1p62

Beta-Binomial-Poisson Mixture Model and Its Parameter Estimation: An Application to Number of Female Childbirths

Dhairya Tripathi[a], Amit Kumar Misra[b], and Anup Kumar* [c]

[a,b] *Department of Statistics, Babasaheb Bhimrao Ambedkar University, Lucknow, India, 226025*
[c] *Department of Biostatistics and Health Informatics, S.G.P.G.I.M.S., Lucknow, India, 226025*

March 15, 2026

This study investigates parameter estimation for the beta-binomial-Poisson mixture model, which accounts for overdispersion and heterogeneity in count data, making it relevant for demographic stuties such as number of female births. The model was applied to NFHS-V data from five Indian states, where son preference influences reproductive behavior. Traditional methods, including Maximum Likelihood Estimation (MLE) and the Method of Moments (MoM), face challenges due to latent variables and model complexity. We explore the Expectation-Maximization (EM) algorithm as a robust alternative. Results show that EM yields more stable and realistic parameter estimates, while MoM often produces poor fits or unrealistic values. Chi-square tests indicate that EM achieves a substantially better fit than MoM, although discrepancies remain, suggesting the need for more advanced modelling.

Keywords: EM algorithm, Method of Moments, Mixture model, Son preference

1 Introduction

Mixture models are a powerful tool in statistical analysis for modeling heterogeneous data. The beta-binomial-Poisson mixture model is particularly effective for modeling

*Corresponding authors: anup.stats@gmail.com

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

count data, allowing for both overdispersion and heterogeneity (see, McLachlan and Peel (2000)). This model combines the Poisson distribution, which describes the occurrence of events like births, with the beta-binomial distribution, which is employed when the probability of success is random, such as when the probability of female births varies across different families. This adaptability makes the beta-binomial-Poisson mixture model especially relevant in demographic research, where fertility patterns often exhibit variability due to both cultural and biological influences (see, for example, Singh et al. (2012, 2015); Kumar (2020); Roy et al. (2023), and references cited therein).

There are many researchers who have given significant contribution in stochastic modelling of female birth data and their estimation. Earlier contributions by Dandekar (1955) and Pathak (1966) laid the groundwork for probabilistic fertility models by modifying Binomial and Poisson distributions to better align with real-world birth patterns. However, these models often assumed independent birth events or uniform probabilities across populations, limiting their applicability in more heterogeneous contexts. Later models developed by Singh et al. (2015); Yadava et al. (2013) and Yadava (2016) introduced elements of son preference and stochastic variation into fertility models. More recent work by Rai et al. (2014); Kumar (2020) and Roy et al. (2023) applied the beta-binomial-Poisson mixture model to estimate the parameters representing the number of female births among couples, providing a more detailed understanding of fertility patterns.

Despite these advancements, efforts to estimate the parameters of this mixture model have primarily utilized Maximum Likelihood Estimation (MLE) and the Method of Moments (MoM). However, these approaches frequently encounter difficulties, particularly due to the model's complexity, which involves latent variables. Studies indicate that while MLE is theoretically sound, it can struggle to converge to true parameter values in the presence of overdispersion, small sample sizes, or high computational complexity (see, for example, Zhu et al. (2003); McLachlan and Peel (2008); Redner and Walker (1984)). Similarly, the MoM, despite being straightforward, can produce biased parameter estimates when the assumed model differs from the actual underlying distribution (see, Fisher (1922); Casella and Berger (2002)). Therefore, there remains a need for a more comprehensive evaluation of parameter estimation methods for this model, particularly in the context of son preference-driven fertility behavior.

To address these estimation challenges, this paper investigates the Expectation-Maximization (EM) algorithm applications for parameter estimation in the beta-binomial-Poisson mixture model. Dempster et al. (1977) has first proposed the EM algorithm which has shown great efficacy in handling models with latent variables and incomplete data. Through iterative parameter updates, the EM algorithm provides a more robust alternative to MLE and MoM, especially in complex mixture models (Zhu et al., 2003; Sammaknejad et al., 2019; Rahim et al., 2021). Its relevance in demographic research is underscored by Mamun et al. (2016), who applied the EM algorithm to handle missing data in the Bangladesh Demographic and Health Survey, demonstrating its utility in survey-based studies. In this study, both the EM algorithm and MoM are applied to estimate parameters for the beta-binomial-Poisson model using fertility data, focusing on female births in India, where son preference significantly impacts reproductive behavior.

This research aims to address the gaps in the models presented by Kumar (2020) and Roy et al. (2023), introducing refinements through empirical evaluations of female birth incidences in India. By comparing the effectiveness of the EM algorithm and MoM in modeling complex fertility patterns, this study seeks to enhance methodological approaches for demographic research related to son preference. Using data from the National Family Health Survey-V (NFHS-V) across five Indian states, model parameters are estimated, and model adequacy is evaluated through chi-square goodness-of-fit tests. The findings from this research are expected to provide valuable insights for policymakers and researchers on best practices for analyzing fertility data, ultimately guiding more informed policy decisions.

In summary, this paper aims to compare two methodologies for obtaining more accurate estimates of the parameters involved in modeling female births. For this purpose, Section 2 describes the probability model used to represent the number of female births per couple. Section 3 outlines the estimation methods employed for estimating the parameters of the beta-binomial-Poisson mixture model, specifically the MoM and EM algorithm. Section 4 presents the chi-square goodness of fit test used to assess the adequacy of the model. Section 5 provides the results and discussion, while Section 6 summarizes the conclusions drawn from the study. Finally, the tables and graphs supporting the estimation procedure and results are presented.

2 The Probability Model

This study employs the probability model presented by Rai et al. (2014); Singh et al. (2015); Kumar (2020) for parameter estimation. For this, consider a woman with n children, where the occurrence of a female birth is defined as a success, while a male birth is a failure. Let X represent the occurence of female births, and let p denote the probability of having a female child at a specific parity. The probability distribution of X when the n and p are given, is expressed as:

$$P(X = x|n, p) = \binom{n}{x} p^x (1-p)^{n-x}, \quad 0 \leq p \leq 1,\ n > 0,\ x = 0, 1, 2, \ldots, n.$$

The probability p of giving birth to a female child can vary among individuals and follows a beta distribution characterized by shape parameters a and b. The probability density function (pdf) for p is represented as:

$$f(p) = \frac{1}{B(a,b)} p^{(a-1)} (1-p)^{(b-1)}, \quad 0 \leq p \leq 1,$$

where

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

Additionally, the total number of children ever born, denoted as n, is modeled as a random variable and follows a Poisson distribution, described by the pdf:

$$P[n = k] = \frac{e^{-\lambda}\lambda^k}{k!}, \quad k = 0, 1, \ldots, n \quad \text{and} \quad \lambda > 0.$$

The joint distribution of X and p for a specified n can be formulated as:

$$\begin{aligned} P[X = x, P = p \mid n] &= P[X = x \mid n, p] f(p) \\ &= \binom{n}{x} p^x (1-p)^{n-x} \frac{1}{B(a,b)} p^{(a-1)} (1-p)^{(b-1)} \\ &= \binom{n}{x} p^{x+a-1} (1-p)^{n-x+b-1}. \end{aligned}$$

To obtain the marginal distribution of X for a fixed n, we integrate over p:

$$\begin{aligned} P[X = x \mid n] &= \int_0^1 \binom{n}{x} \frac{1}{B(a,b)} p^{x+a-1} (1-p)^{n-x+b-1} \, dp \\ &= \frac{\binom{n}{x} B(x+a, n-x+b)}{B(a,b)}. \end{aligned}$$

Consequently, the marginal distribution of X can be expressed as:

$$\begin{aligned} P[X = x] &= \sum_{n=x}^{\infty} P[X = x, N = n] \\ &= \int_0^1 \binom{n}{x} \frac{1}{B(a,b)} p^{x+a-1} (1-p)^{n-x+b-1} \frac{e^{-\lambda}\lambda^n}{n!} \, dp. \end{aligned}$$

Upon solving this integral, we obtain

$$P[X = x] = \frac{\lambda^x}{B(a,b) \cdot x!} \int_0^1 e^{-\lambda p} p^{x+a-1} (1-p)^{b-1} \, dp. \tag{1}$$

This equation provides the probability mass function (pmf) for the number of female births to a couple.

3 Method of Estimation

There are two estimation methods for determining the parameters given in equation (1). The following subsections provide a comprehensive description of both methods, including their theoretical foundations, and the steps involved in their implementation.

3.1 Method of Moments

The method of moments is utilized to estimate the parameters (λ, a, and b) described in equation (1), which defines the distribution of female births across different parity levels. The first three raw moments of the probability model in equation (1) are computed as

follows

$$E(X) = \frac{\lambda a}{a+b}, \tag{2}$$

$$E(X^2) = \frac{\lambda^2(a+1)a}{(a+b+1)(a+b)} + \frac{\lambda a}{a+b}, \tag{3}$$

$$E(X^3) = \frac{\lambda^3(a+2)(a+1)a}{(a+b+2)(a+b+1)(a+b)} + \frac{3\lambda^2(a+1)a}{(a+b+1)(a+b)} + \frac{\lambda a}{a+b}. \tag{4}$$

To estimate the parameters a, b, and λ, we solve the system of equations using the method of moments by matching the theoretical moments with the sample moments. For this, let $a+b=y$, then denoting the first sample moment by m_1, we have

$$m_1 = \frac{\lambda a}{y}. \tag{5}$$

Substituting $a+b=y$ and simplifying Equation(3), the second sample moment m_2 becomes

$$m_2 = \frac{\lambda + m_1 y}{y+1} \cdot m_1 + m_1. \tag{6}$$

Again, simplifying equation(4) we have

$$m_3 = \frac{(m_1 y + 2\lambda)}{y+2} \cdot (m_2 - m_1) + 3(m_2 - m_1) + m_1. \tag{7}$$

By solving these three equations from Equation(5) to Equation (7), the estimate of a, b, and λ are as follows:

$$\hat{\lambda} = \frac{(m_3 - m_1 - m_1^2)y - m_1 + m_2}{m_1},$$
$$\hat{a} = \frac{m_1 \cdot y}{\hat{\lambda}},$$
$$\hat{b} = y - \hat{a},$$

where

$$y = \frac{(2m_1^2 - 2m_2^2 - 2m_1 m_2 + 2m_1 m_3)}{(m_1^3 + 2m_2^2 - m_1 m_2 - m_1^2 m_2 - m_1 m_3)}.$$

However, when applied to real data from the NFHS-V for female births, this method encountered significant challenges. In some states, the model produced unrealistic negative parameter estimates. This will occur only when the sample mean is smaller than the sample variance, reflecting the high variability in the data (see, (Casella and Berger, 2002, p. 314)). This issue highlights the limitations of the method in accurately modeling real-world demographic data.

Thus consequently, for estimating the parameters of the beta-binomial-Poisson model, this paper adopts the methodology presented by Kumar (2020). The analysis is re-

stricted to women of all parity levels who have been married for at least seven years, ensuring a homogenous sample and reliable probability estimates for childbirth. Additionally, women who have not yet given birth are included to maintain alignment with the model's assumptions. With these modifications, we estimated λ, representing the mean number of children ever born to females with at least seven years of marriage exposure, calculated as

$$\lambda = \frac{T}{n}, \tag{8}$$

where T is the total number of children ever born, and n is the number of females of all parity levels exposed to at least seven years of marriage. Solving Equations (2), (3), (4), and (11) yielded the estimates for a, b and λ. We also employed the EM algorithm as an alternative estimation method to address the limitations observed with the method of moments in fitting the empirical data accurately.

3.2 Expectation Maximization Algorithm

The Expectation-Maximization (EM) algorithm is a powerful tool for estimating maximum likelihood parameters in statistical models, especially when dealing with incomplete or missing data. It was first introduced by Dempster et al. (1977). It is particularly effective for latent variable models, where certain influencing factors are unobservable (Dempster et al., 1977). This algorithm has found widespread application across various domains, including machine learning, data mining, Bayesian statistics etc.

Let the parameter set be $\boldsymbol{\theta} = (a, b, \lambda)$, and consider the observations $(x_1, x_2, \ldots, x_m)$ and $(n_1, n_2, \ldots, n_m)$. The joint probability distribution for the variables $\{\mathbf{x}, \mathbf{n}\}$ is given by

$$\begin{aligned} f(\mathbf{x}, \mathbf{n}; \boldsymbol{\theta}) &= \prod_{i=1}^{m} f(x_i, n_i; \boldsymbol{\theta}) \\ &= \int_0^1 \binom{n_i}{x_i} \frac{1}{B(a,b)} p_i^{x_i+a-1} (1-p_i)^{n_i-x_i+b-1} \frac{e^{-\lambda}\lambda^{n_i}}{n_i\,!}\, dp_i \end{aligned}$$

and the log-likelihood function is

$$\begin{aligned} l(\boldsymbol{\theta}; \mathbf{x}, \mathbf{n}) \propto\ & \log(\Gamma a + b) - \log(\Gamma a) - \log(\Gamma b) - \lambda + n_i \log(\lambda) \\ & + \log\left(\int_0^1 p_i^{x_i+a-1} (1-p_i)^{n_i-x_i+b-1}\, dp_i \right) \end{aligned} \tag{9}$$

Maximizing this log-likelihood function directly is challenging due to the complexity of the integral term. Consequently, the EM algorithm is employed by treating the unobservable success probabilities p_i as missing data, which simplifies the estimation process (Zhu et al., 2003).

3.2.1 Theoretical Framework of the EM Algorithm

Before applying the EM algorithm, it is crucial to outline its theoretical framework. Given the observed data $\mathbf{x} = (x_1, x_2, \ldots, x_m)$ and $\mathbf{n} = (n_1, n_2, \ldots, n_m)$, with missing observations $\mathbf{p} = (p_1, p_2, \ldots, p_m)$, the complete probability distribution is

$$\begin{aligned} f(\mathbf{x}, \mathbf{n}, \mathbf{p}; \boldsymbol{\theta}) &= \prod_{i=1}^{m} f(x_i, n_i, p_i \mid \boldsymbol{\theta}) \\ &= \prod_{i=1}^{m} P[X = x_i \mid n_i, p_i].f(p_i).P[N = n_i] \\ &\propto \prod_{i=1}^{m} \frac{p_i^a \ (1-p_i)^b \ e^{-\lambda}\lambda^{n_i}}{B(a,b)}. \end{aligned}$$

Thus, the complete log-likelihood function becomes

$$l \propto a\sum_{i=1}^{m}\log(p_i)+b\sum_{i=1}^{m}\log(1-p_i)+m\log(\Gamma a+b)-m\log(\Gamma a)-m\log(\Gamma b)-m\lambda-\sum_{i=1}^{m}n_i\log(\lambda).$$

The estimation process proceeds with the iterative E-step and M-step, crucial for parameter estimation

E Step:

In the E-step, we compute the expected value of the complete log-likelihood given the observed data $\mathbf{x}$, $\mathbf{n}$, and the initial parameter estimates $\hat{\boldsymbol{\theta}} = (a, b, \lambda)$:

$$\begin{aligned} E\left[l(\boldsymbol{\theta}; \mathbf{x}, \mathbf{n}, \mathbf{p}) \mid \mathbf{x}, \mathbf{n}; \hat{\boldsymbol{\theta}}\right] \ &\propto a\sum_{i=1}^{m} E\left[\log(p_i) \mid x_i, n_i; \hat{\boldsymbol{\theta}}\right] + b\sum_{i=1}^{m} E\left[\log(1-p_i) \mid x_i, n_i; \hat{\boldsymbol{\theta}}\right] \\ &\quad + m\log(\Gamma(a+b)) - m\log(\Gamma(a)) - m\log(\Gamma(b)) - m\lambda \\ &\quad - \sum_{i=1}^{m} n_i\log(\lambda) \propto a\sum_{i=1}^{m}\hat{P}_i + b\sum_{i=1}^{m}\hat{Q}_i + m\log(\Gamma(a+b)) \\ &\quad - m\log(\Gamma(a)) - m\log(\Gamma(b)) \quad - m\lambda - \sum_{i=1}^{m} n_i\log(\lambda), \end{aligned}$$

where

$\hat{P}_i = E\left[\log(p_i) \mid x_i, n_i; \hat{\boldsymbol{\theta}}\right]$ and $\hat{Q}_i = E\left[\log(1-p_i) \mid x_i, n_i; \hat{\boldsymbol{\theta}}\right]$.

We can define

$$E_1 = a\sum_{i=1}^{m}\hat{P}_i + b\sum_{i=1}^{m}\hat{Q}_i + m\log(\Gamma(a+b)) - m\log(\Gamma(a)) - m\log(\Gamma(b)),$$
$$E_2 = m\lambda - \sum_{i=1}^{m} n_i \log(\lambda).$$

M-Step:

The updated parameters a and b are obtained by maximizing $E_1(a,b)$, while λ is maximized through $E_2(\lambda)$. To compute $\hat{P}_i$ and $\hat{Q}_i$, we utilize the formula

$$\begin{aligned} E[\phi(p_i) \mid x_i, n_i; \hat{\theta}] &= \int_0^1 \phi(p_i) f(p_i \mid x_i, n_i, \hat{\theta})\, dp_i \\ &= \frac{\int_0^1 \phi(p_i) f(x_i, n_i, p_i; \hat{\theta})\, dp_i}{\int_0^1 f(x_i, n_i, p_i; \hat{\theta})\, dp_i}, \end{aligned}$$

where $\phi(.)$ is a real-valued function. We set $\phi(p_i)$ as $\log(p_i)$ for $\hat{P}_i$ and $\log(1-p_i)$ for $\hat{Q}_i$. Now, the Monte-Carlo method will be used to approximate $E[\phi(p_i) \mid x_i, n_i; \hat{\theta}]$ by

$$\begin{aligned} &\frac{\sum_{m=1}^{M} \phi(p_i^{(m)}) f(x_i \mid n_i, p_i^{(m)}) f(n_i \mid p_i^{(m)})}{\sum_{m=1}^{M} f(x_i \mid n_i, p_i^{(m)}) f(n_i \mid p_i^{(m)})} \\ &= \frac{\sum_{m=1}^{M} \phi(p_i^{(m)}) P[X = x_i \mid n_i, p_i^{(m)}] P[N = n_i]}{\sum_{m=1}^{M} P[X = x_i \mid n_i, p_i^{(m)}] P[N = n_i]} \\ &= \frac{\sum_{m=1}^{M} \phi(p_i^{(m)})\ p_i^{x_i - 1}\ (1-p_i)^{n_i - x_i}}{p_i^{x_i-1}\ (1-p_i)^{n_i-x_i}}, \end{aligned}$$

where $p_i^{(1)}, p_i^{(2)}, \ldots, p_i^{(M)}$ are random samples drawn from $Beta(\hat{a}, \hat{b})$. A large value of M is necessary for convergence. Finally, the functions $E_1(a,b)$ and $E_2(\lambda)$ are maximized using the Newton-Raphson algorithm.

3.2.2 Starting Values

The EM algorithm is known for its enhanced numerical stability in comparison to direct maximization approaches (see, Zhu et al. (2003)). However, the effectiveness of the EM algorithm is influenced by the choice of initial parameter values. To establish starting points for the parameters a and b, we utilize a standard beta-binomial model, represented by the following log-likelihood function

$$\ell(a,b;x,n) \propto m\log\Gamma(a+b) + \sum_{i=1}^{m}\log\Gamma(a+x_i) + \sum_{i=1}^{m}\log\Gamma(b+n_i-x_i)$$
$$- m\log\Gamma(a) - m\log\Gamma(b) - \sum_{i=1}^{m}\log\Gamma(a+b+n_i).$$

We generated 1000 values ranging from 0.1 to 100 and created a 3D graph in R software, illustrating the relationship between parameters a and b, and the log-likelihood (z) of the beta-binomial model. The point at which z was maximum, served as our initial values for a and b.

4 Results and Discussion

This study conducted a comprehensive analysis, utilizing data from the National Family Health Survey (2019-21) across five Indian states—Bihar, Kerala, Mizoram, Uttarakhand, and West Bengal. The adequacy of the proposed beta-binomial-Poisson mixture model is evaluated through the use of chi-square goodness-of-fit tests given by Pearson (1900). The tables 1,2,3,4,5 reveal significant variation in the parameters λ, a, b across different states, reflecting regional differences in fertility and female birth patterns. The mean number of children ever born (λ) varies from 1.818 in West Bengal to 2.868 in Bihar, highlighting differences in mean number of children ever born to a women who have had at least one child. . The variation in the parameters a and b across different states, under both the EM and MoM estimation methods, highlights important distinctions in their ability to capture the underlying distribution of female births. The EM method demonstrates significant regional variation in a and b, indicating differences in the dispersion of female birth probabilities. For instance, in Bihar, the values of $a = 5.603$ and $b = 5.974$ reflect moderate variability, while in West Bengal, these parameters increase dramatically to $a = 101.51$ and $b = 86.92$, underscoring much greater heterogeneity in fertility behavior. In contrast, the MoM consistently underestimates both a and b, yielding values such as $a = 0.69$ and $b = 0.83$ for Bihar, and $a = 0.30$ and $b = 0.34$ for West Bengal. These lower estimates from MoM result in substantially higher chi-square values, indicating poor model fit and suggesting that MoM fails to adequately capture the complexity of birth patterns. The EM method, by comparison, provides more robust and accurate estimates, demonstrating its superiority in modeling the variability in female birth distributions across diverse regions. Figures 6,7,8,9,10 presents the curves illustrating the comparison between the EM algorithm and the MoM for various states.

However, the findings reveal notable disparities between the expected and observed frequencies of female childbirths, underscored by high chi-square values generated from both EM algorithm and the MoM. These results indicate that neither method achieved a satisfactory fit to the data across the various states examined. While the EM algorithm consistently outperformed the MoM in estimating parameters, the chi-square values remained relatively high. For example, in Bihar, the chi-square statistic was 40.68 when

using the EM method (as shown in Table 1). This indicates that, while the EM algorithm is better at capturing the overall pattern of female births, it still does not fully explain the complexities of the data. Similar patterns were observed in the other states as well (see, Tables 2, 3, 4, 5), showing that significant gaps persist between the observed and expected frequencies.

A significant reason for these elevated chi-square values may be attributed to the fact that, particularly for substantial total frequencies, even slight deviations from the observed values can result in a considerable chi-square value Lin et al. (2013); Ross (2014). This highlights the sensitivity of the chi-square statistic to deviations between observed and expected frequencies. Moreover, these discrepancies could be influenced by stopping rules, which introduce dependencies between the number of children and the sex composition of previous births, adding further complexity to the data. Additionally, sex-selective practices, including abortion, further complicate the estimation, requiring more advanced statistical methods to capture such dynamics. Such dependencies further complicate the relationship between expected and observed frequencies, suggesting that the dynamics of human fertility behavior are not adequately captured by the beta-binomial-Poisson mixture model employed in this analysis.

In addition, a notable observation in Bihar (Table 1) is the very high frequency of zero female births, pointing to the presence of excess zeros in the data. Addressing this pattern through zero-inflated or hurdle extensions of the model would be a valuable direction for future research (see, Kassahun et al. (2014)), allowing a more comprehensive representation of fertility behavior.

5 Conclusion

In conclusion, while the EM algorithm provides a relatively better fit than the MoM, both methods exhibit significant shortcomings, as evidenced by the high chi-square values. This underscores the necessity for employing more sophisticated models that can better encapsulate the complexities of female birth distributions in these contexts. The findings underscore the need to consider additional factors, such as the impact of stopping rules, to improve the model's fit and better capture the complexities of human fertility behavior in these regions. Future research endeavors could focus on refining the model to achieve a better fit with the observed data. Exploring potential adjustments or extensions to the model may enhance its accuracy and applicability in capturing the nuances of fertility patterns among Indian females. Moreover, it is imperative to explore how stopping rules, along with socio-economic and biological factors, influence son preferences within society. Regional heterogeneity could be more effectively captured through hierarchical or multilevel frameworks, while Bayesian methods based on MCMC would allow full posterior inference and provide richer uncertainty quantification. Moreover, the limitations of the Method of Moments could be alleviated by adopting the Generalized Method of Moments (GMM), which offers greater flexibility in handling moment conditions and reduces the risk of negative or unstable estimates. Complementary evaluation tools, including likelihood-based selection criteria such as AIC and BIC, as well

as alternative goodness-of-fit checks based on bootstrap or residual analysis, may further strengthen assessment of model adequacy. In addition, while this paper concentrated on point estimation through EM and MoM, future research should incorporate interval estimation procedures. Confidence intervals and standard errors derived from the observed information matrix or bootstrap resampling would provide more rigorous insights into parameter uncertainty, thereby enhancing the robustness of inference. Integrating such interval-based approaches would not only improve interpretability but also strengthen the reliability of model comparisons in applied demographic contexts.

By understanding these dimensions, researchers can develop more comprehensive models that consider the multifaceted nature of fertility decisions, ultimately contributing to more effective policy formulations aimed at addressing gender imbalances in birth rates.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Table 1: **Bihar: Distribution of observed and expected frequency of female childbirths**

Number of female births	Observed frequency	Expected frequency (EM)	Expected frequency (MoM)
0	2272	2157	3707
1	3731	3193	2410
2	2309	2375	1618
3	894	1184	959
4	263	445	491
5	55	134	219
6	6	34	85
7	2	7	30
Total	9532	9532	9532
Estimates		λ=2.868 a=5.603 b=5.974 χ^2=40.68	λ=2.868 a=0.69 b=0.83 χ^2=1907.334

Table 2: **Kerala: Distribution of observed and expected frequency of female childbirths**

Number of female births	Observed frequency	Expected frequency (EM)	Expected frequency (MoM)
0	884	1056	1196
1	1272	1008	898
2	537	487	435
3	61	159	163
4	3	39	50
6	1	1	3
Total	2758	2758	2758
Estimates		λ=1.88	λ= 1.88
		a=40.08	a=1.94
		b=37.95	b=2.02
		χ^2=195.92	χ^2=370.414

Table 3: **Mizoram: Distribution of observed and expected frequency of female childbirths**

Number of female births	Observed frequency	Expected frequency (EM)	Expected frequency (MoM)
0	417	420	611
1	697	593	519
2	434	420	325
3	143	199	166
4	32	71	71
5	6	20	26
6	1	5	9
Total	1730	1730	1730
Estimates		λ=2.53	λ= 2.53
		a=110.42	a=1.34
		b=86.80	b=1.38
		χ^2=68.90	χ^2=206.3075

Table 4: **Uttarakhand: Distribution of observed and expected frequency of female childbirths**

Number of female births	Observed frequency	Expected frequency (EM)	Expected frequency (MoM)
0	890	898	1199
1	1257	1062	913
2	610	631	496
3	149	251	217
4	31	75	80
5	2	18	7
Total	2939	2939	2939
Estimates		λ=2.218	λ=2.218
		a=100.46	a=1.44
		b=86.94	b=1.63
		χ^2=118.061	χ^2=290.34

Table 5: **West Bengal: Distribution of observed and expected frequency of female childbirths**

Number of female births	Observed frequency	Expected frequency (EM)	Expected frequency (MoM)
0	1852	1880	2668
1	2197	1833	1144
2	785	898	674
3	135	295	325
4	23	73	128
5	4	14	42
Total	4094	4094	4094
Estimates		λ=1.818 a=101.51 b=86.92 χ^2=215.0892	λ=1.818 a=0.30 b=0.34 χ^2=1468.681

Starting Values for Different States

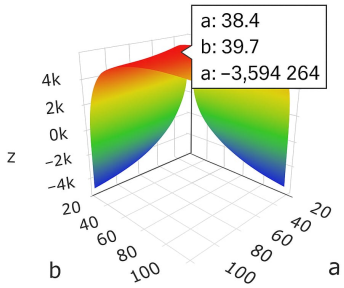

Figure 1: Starting Values for Kerala

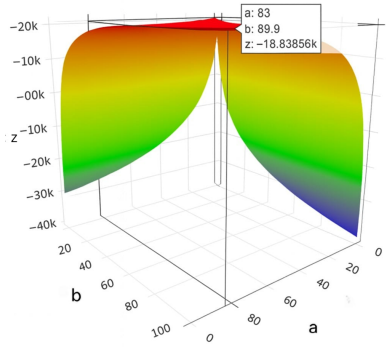

Figure 2: Starting Values for Bihar

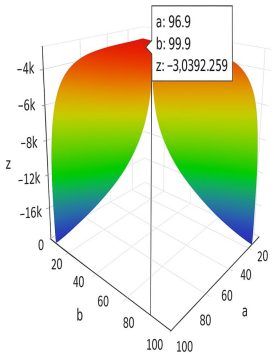

Figure 3: Starting Values for Mizoram

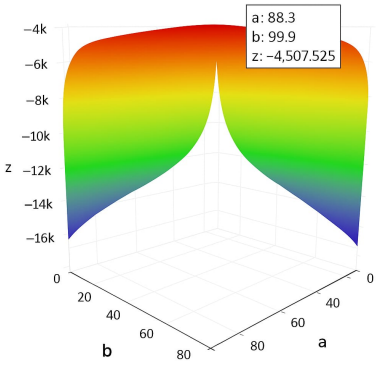

Figure 4: Starting Values for Uttarakhand

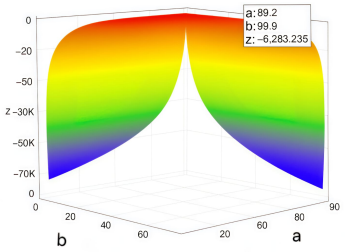

Figure 5: Starting Values for West Bengal

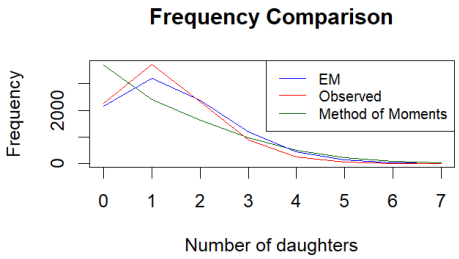

Figure 6: Bihar: Comparison between Methods of Estimation and EM

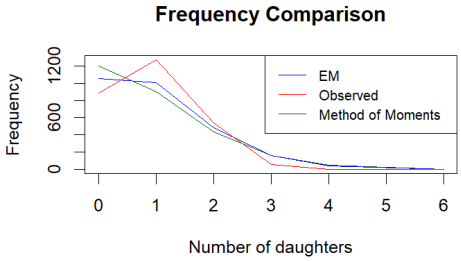

Figure 7: Kerala: Comparison between Methods of Estimation and EM

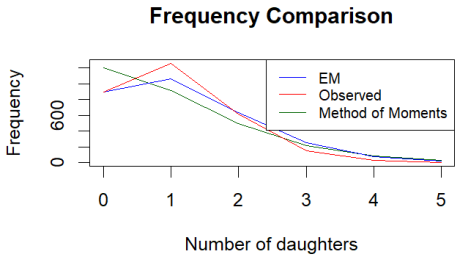

Figure 8: Uttarakhand: Comparison between Methods of Estimation and EM

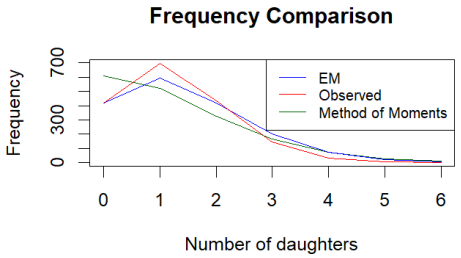

Figure 9: Mizoram: Comparison between Methods of Estimation and EM

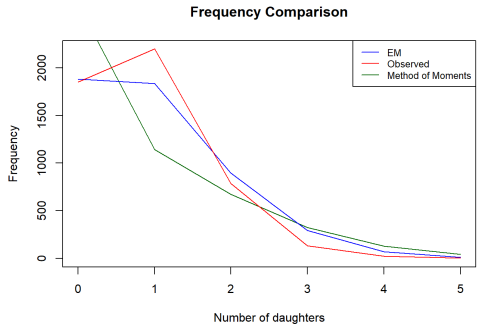

Figure 10: West Bengal: Comparison between Methods of Estimation and EM

Comparison between Methods of Estimation and EM for Various States

References

Casella, G. and Berger, R. L. (2002). *Statistical Inference*. Thomson Learning, Australia; Pacific Grove, CA, 2nd edition.

Dandekar, V. . M. (1955). Certain modified forms of binomial and poisson distributions. *Sankhyā: The Indian Journal of Statistics (1933-1960)*, 15(3):237–250.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1):1–22.

Fisher, R. A. (1922). On the mathematical foundations of theoretical statistics. *Philosophical transactions of the Royal Society of London. Series A, containing papers of a mathematical or physical character*, 222(594-604):309–368.

Kassahun, W., Neyens, T., Molenberghs, G., Faes, C., and Verbeke, G. (2014). Marginalized multilevel hurdle and zero-inflated models for overdispersed and correlated count data with excess zeros. *Statistics in medicine*, 33(25):4402–4419.

Kumar, A. (2020). A probability model for the number of female child births. *Journal of Statistics Applications & Probability*, 9(3):525–534.

Lin, M., Lucas Jr, H. C., and Shmueli, G. (2013). Research commentary—too big to fail: large samples and the p-value problem. *Information Systems Research*, 24(4):906–917.

Mamun, A., Zubairi, Y., Hussin, A., and Rana, S. (2016). A comparison of missing data handling methods in linear structural relationship model: evidence from bdhs2007 data. *Electronic Journal of Applied Statistical Analysis*, 9(1):122–133.

McLachlan, G. J. and Peel, D. (2000). *Finite Mixture Models*. Wiley-Interscience.

McLachlan, G. J. and Peel, D. (2008). *Finite Mixture Models*. John Wiley & Sons.

Pathak, K. (1966). A probability distribution for the number of conceptions. *Sankhyā: The Indian Journal of Statistics, Series B*, 28:213–218.

Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302):157–175.

Rahim, S. A., Manson, G., and Aziz, M. (2021). Data clustering based on gaussian mixture model and expectation-maximization algorithm for data-driven structural health monitoring system. *International Journal of Integrated Engineering*, 13(7):167–175.

Rai, P. K., Pareek, S., and Joshi, H. (2014). On the estimation of probability model for the number of female child births among females. *Journal of Data Science*, 12(3):137–156.

Redner, R. A. and Walker, H. F. (1984). Mixture densities, maximum likelihood and the em algorithm. *SIAM Review*, 26(2):195–239.

Ross, S. M. (2014). *Introduction to probability models*. Academic Press.

Roy, S., Sharma, P., Singh, K., and Srivastava, R. (2023). On a statistical model useful for demographics: Estimating the mean number of children ever born through the

distribution of male births with an application to data from india. *Journal of Reliability and Statistical Studies*, 16(1):57–80.

Sammaknejad, N., Zhao, Y., and Huang, B. (2019). A review of the expectation maximization algorithm in data-driven process identification. *Journal of Process Control*, 73:123–136.

Singh, B. P., Maheshwari, S., and Gupta, P. K. (2015). A probability model for sex composition of children in the presence of son preference. *Demography India*, 44(1 & 2):50–57.

Singh, K., Singh, B. P., and Singh, N. (2012). A probabilistic study of variation in number of child deaths. *Journal of Rajasthan Statistical Association*, 1(1):54–67.

Yadava, R. C. (2016). Stochastic models for human fertility. *Demography India*, 45(1 & 2):1–16.

Yadava, R. C., Kumar, A., and Srivastava, U. (2013). Sex ratio at birth: A model based approach. *Mathematical Social Sciences*, 65(1):36–39.

Zhu, J., Eickhoff, J. C., and Kaiser, M. S. (2003). Modeling the dependence between number of trials and success probability in beta-binomial-poisson mixture distributions. *Biometrics*, 59(4):955–961.