



Bayesian Analysis of First Conception through a Probability Model

Ashivani Kumar Yadav

Department of Mathematics and Statistics

D.D.U. Gorakhpur University Gorakhpur, U.P., India, 277123

Abstract: Probability models are widely used in different disciplinary fields. The present paper is an attempt to develop probability model of waiting time of first conception and also analyze under the Bayesian environment of waiting time of first conception under Entropy Loss Function loss function.

Keywords: Probability model, waiting time, conception, Entropy Loss Function

I. Introduction

At the time of marriage a woman is susceptible to conception and the time elapsed before a conception is a random variable determined by fecundability, which is defined as the monthly chance of a conception. It is important here to note the time of first conception after the marriage because the analysis of waiting time of first conception signifies couple's fertility at early stages of married life. This variable is widely used to study fertility characteristic of a woman, since it is independent of effect of amenorrhea period and generally, a woman does not like to use contraceptives to postpone first birth. There is little chance of recall lapse in reporting the time of first birth from the date of first marriage of first birth.

Treating the first conception as the random phenomenon the probabilistic models can be developed. And this variable can be treated as discrete as well as continuous variable depending upon the situation and assumptions made for the study.

For the first time this variable was considered as a discrete variable and Gini (1924) derived the geometric distribution for the waiting time of first conception. He defined the term 'Fecundability' as the monthly chance of conception for women living in the married, fecund and exposed state.

Later the same variable was considered as a continuous variable and hence treating the time elapsed from the marriage or from the beginning of the reproductive process to first conception as continuous makes mathematical treatments more convenient and easy. Singh (1964), Henry (1953) and Vincent (1961) developed some models treating the waiting time of first conception as continuous. The negative exponential distribution plays the role of geometric distribution for studying the waiting time of conceptions after marriage. Thus if X denotes the time of first conception then the density function, say, $f(x)$ is given by

$$f(x) = \delta e^{-\delta t}; \quad x > 0, \delta > 0$$

Where δ is instantaneous fecundability

A number of authors made modifications on the above simple distribution to study realistic situations.

In the present model of waiting time of first conception the time elapsed is defined over the range $(0, \infty)$. But in practical problems the upper limit may be considered as finite; as a woman can conceive up to an age limit. So, there is a need of introducing a new continuous model with finite range. Keeping this in view, an attempt has been made to characterize an existing model derived by Mukherjee-Islam (1983), defined over a finite range for the purpose of life testing analysis but it suits in realistic or real life situations.

II. A FINITE RANGE CONTINUOUS MODEL

A new probability distribution has been considered in the section as a continuous model, introduced by Mukherjee-Islam (1983) for the purpose of studying waiting time

$$f(x, \theta, p) = (p/\theta^p) x^{p-1} \quad p, \theta > 0 \quad x \geq 0$$

The above model is monotonic decreasing and highly skewed to the right. The Graph is J-shaped thereby the unimodal future.

Let us consider the reparametrized finite range distribution whose p.d.f. is given by

$$f(x; \sigma, \theta) = \frac{1}{\theta \sigma} \left(\frac{x}{\sigma} \right)^{\frac{1}{\theta}-1}; \quad \theta > 0, \sigma > 0, \quad 0 \leq x \leq \sigma \quad (2.1.1)$$

where ' θ ' is instantaneous fecundability and σ is considered as age limit beyond which a married woman cannot conceive.

The c.d.f. of the above model is given by

$$\begin{aligned} F(x) &= P(X \leq x) \\ &= \int_0^x f(x) dx \\ &= \left(\frac{x}{\sigma}\right)^{\frac{1}{\theta}} \end{aligned} \quad (2.1.2)$$

The survival function at time x, say $S(x) = P(X > x)$ is given by

$$\begin{aligned} S(x) &= P(X > x) \\ &= 1 - P(X \leq x) \\ &= 1 - \left(\frac{x}{\sigma}\right)^{\frac{1}{\theta}} \end{aligned} \quad (2.1.3)$$

Also, the conception rate function, say, $w(x)$ at time is then given by

$$\begin{aligned} W(x) &= \frac{P(X=x)}{P(X \geq x)} \\ W(x) &= \frac{\frac{1}{\theta x} \left(\frac{x}{\sigma}\right)^{\frac{1}{\theta}-1}}{1 - \left(\frac{x}{\sigma}\right)^{\frac{1}{\theta}}} \\ &= \frac{\frac{1}{\theta} x^{\frac{1}{\theta}-1}}{\frac{1}{\sigma^{\frac{1}{\theta}} - x^{\frac{1}{\theta}}}} \end{aligned} \quad (2.1.4)$$

$$\mu'_r = \frac{\sigma^r}{(1+r\theta)}$$

$$\mu'_1 = E(X) = \frac{\sigma}{1+\theta} \quad (2.1.5)$$

and

$$\mu'_2 = \frac{\sigma^2}{1+2\theta}$$

therefore

$$\mu_2 = V(X) = \frac{(\sigma\theta)^2}{(1+2\theta)(1+\theta)^2} \quad (2.1.6)$$

Maximum Likelihood Estimator:

If $x_1, x_2, \dots, x_n$ having a common pdf as given in (2.1.1), where σ is known, then the joint probability density function is given by

$$f(\underline{x}|\theta) = \left(\frac{1}{\theta}\right)^n \left(\frac{1}{\prod_{i=1}^n x_i}\right) e^{-\frac{z}{\theta}} \quad (2.1.7)$$

Where

$$z = \left[\sum_{i=1}^n \log\left(\frac{\sigma}{x_i}\right) \right]$$

Thus the maximum likelihood estimator (MLE) $\hat{\theta}$ of θ is given as

$$\hat{\theta} = \frac{z}{n} \quad (2.1.8)$$

Bayesian Analysis of The model:

The fundamental problems in Bayesian Analysis are that of the choice of prior distribution $g(\theta)$ and a loss function $L(\hat{\theta}, \theta)$. Let us consider three prior distribution of θ to obtain the Bayes estimators which are as follows:

(i) Quasi-Prior :

For the situation where the experimenter has no prior information about the parameter θ , one may use the quasi density as given by

$$g_1(\theta) = \frac{1}{\theta^d}; \quad \theta > 0, d > 0 \quad (2.1.9)$$

Here $d = 0$ leads to a diffuse prior and $d = 1$, a non informative prior.

(ii) Natural Conjugate Prior of θ :

The most widely used prior distribution of θ is the inverted gamma distribution with parameters α and β (>0) with p.d.f. given by

$$g_2(\theta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{-(\alpha+1)} e^{-\beta/\theta} & ; \theta > 0 \quad (\alpha, \beta) > 0 \\ 0 & ; \text{otherwise} \end{cases}$$

(2.1.10)

The main reason for general acceptability is the mathematical tractability resulting from the fact that inverted gamma distribution is conjugate prior for θ .

(iii) Uniform Prior:

It frequently happens that the life tester knows in advance that the probable values of θ lies over a finite range $[\alpha, \beta]$ but he does not have any strong opinion about any subset of values over this range. In such a case uniform distribution over $[\alpha, \beta]$ may be a good approximation.

$$g_3(\theta) = \begin{cases} \frac{1}{\beta-\alpha}; & 0 < \alpha < \theta \leq \beta \\ 0 & ; \quad \text{otherwise} \end{cases} \quad (2.1.11)$$

Loss Function:

The Bayes estimator $\hat{\theta}$ of θ is of course, optimal relative to the loss function chosen. A commonly used loss function is the squared error loss function (SELF)

$$L(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2, \quad (2.1.12)$$

which is a symmetrical loss function and assigns equal losses to over estimation and underestimation. Canfield (1970) points out that the use of symmetric loss function may be inappropriate in the estimation of reliability function. Over estimation of reliability function or average lifetime is usually much more serious than under estimation of reliability function or mean failure time. Also, an under estimate of the failure rate results in more serious consequence than an overestimation of the failure rate. This leads to statistician to think about asymmetrical loss functions which have been proposed in statistical literature. It is well known that the Bayes estimator under the above loss function, say $\hat{\theta}_s$, is the posterior mean. The squared error loss function (SELF) is often used also because it does not lead to extensive numerical computation but several authors {Ferguson (1967), Varian (1975), Berger (1980), Zellner (1986) and Basu and Ebrahimi (1991)} have recognized the inappropriateness of using symmetric loss function in several estimation problems. These have proposed different asymmetric loss function.

Entropy Loss Function

In many practical situations, it appears to be more realistic to express the loss in terms of the ratio $\frac{\hat{\theta}}{\theta}$. In this case, Calabria and and pulcini (1994) points out that a useful asymmetric loss function is the entropy loss

$$L(\delta) \propto [\delta^p - \log_e(\delta) - 1]$$

Where

$$\delta = \frac{\hat{\theta}}{\theta}$$

And whose minimum occurs at $\hat{\theta} = \theta$ when $p > 0$, a positive error ($\hat{\theta} > \theta$) causes more serious consequences than a negative error, and vice-versa. For small $|p|$ value, the function is almost symmetric when both $\hat{\theta}$ and θ measured in a logarithmic scale, and approximately

$$L(\delta) \propto \frac{p^2}{2} [\log_e(\hat{\theta}) - \log_e \theta]^2$$

Also, the loss function $L(\delta)$ has been used in Dey et al. (1987) and Dey and Liu (1992), in the original form having $p = 1$. Thus $L(\delta)$ can be written as

$$L(\delta) = b[\delta - \log_e(\delta) - 1]; \quad b > 0 \quad (2.1.13)$$

where

$$\delta = \frac{\hat{\theta}}{\theta}$$

The Posterior expectation of loss function in (2.1.13) is

$$E_\pi[L(\delta)] = b \left[E_\pi \left(\frac{\hat{\theta}}{\theta} \right) - E_\pi \left(\log_e \left(\frac{\hat{\theta}}{\theta} \right) \right) - 1 \right] \quad (2.1.14)$$

The value of $\hat{\theta}$ that minimise (2.1.14), denoted by $\hat{\theta}_e$, Bayes estimator of θ under entropy loss function $L(\delta)$, is obtained by solving the following equation

$$\begin{aligned} \frac{d}{d\hat{\theta}} E_\pi[L(\delta)] &= 0 \\ \Rightarrow b \left[E_\pi \left\{ \frac{d}{d\hat{\theta}} \left(\frac{\hat{\theta}}{\theta} \right) \right\} - E_\pi \left\{ \frac{d}{d\hat{\theta}} \log_e \left(\frac{\hat{\theta}}{\theta} \right) \right\} - \frac{d}{d\hat{\theta}} (1) \right] &= 0 \\ \Rightarrow E_\pi \left(\frac{1}{\theta} \right) - E_\pi \left(\frac{1}{\hat{\theta}_e} \right) &= 0 \\ \Rightarrow \hat{\theta}_e &= \left[E_\pi \left(\frac{1}{\theta} \right) \right]^{-1} \end{aligned} \quad (2.1.15)$$

2. Bays Estimator under $g_1(\theta)$:

Under $g_1(\theta)$, the posterior distribution is defined by

$$f(\theta|\underline{x}) = \frac{f(\underline{x}|\theta)g_1(\theta)}{\int_0^\infty f(\underline{x}|\theta)g_1(\theta)d\theta} \quad (2.2.1)$$

Substituting the values of $g_1(\theta)$ and $f(\underline{x}|\theta)$ from equations (2.1.9) and (2.1.7) in (2.2.1) we get, after simplification, as

$$f(\theta|\underline{x}) = \frac{\left(\frac{1}{\theta}\right)^n \left(\frac{1}{\prod_{i=1}^n x_i}\right) e^{-\frac{z}{\theta}} \frac{1}{\theta^d}}{\int_0^\infty \left(\frac{1}{\theta}\right)^n \left(\frac{1}{\prod_{i=1}^n x_i}\right) e^{-\frac{z}{\theta}} \frac{1}{\theta^d} d\theta} \quad (2.2.2)$$

$$= \frac{z^{n+d-1}}{\Gamma(n+d-1)} \theta^{-(n+d)} e^{-z/\theta} \quad ; \theta > 0, n+d > 1.$$

The Bayes estimator under squared error loss function is the posterior mean given by

$$\hat{\theta}_s = \int_0^\infty \theta f(\theta|\underline{x}) d\theta. \quad (2.2.3)$$

Substituting the values of $f(\theta|\underline{x})$ from equation (2.2.2) in equation (2.2.3) and on solving we get

$$\begin{aligned} \hat{\theta}_s &= \int_0^\infty \frac{z^{n+d-1}}{\Gamma(n+d-1)} \theta^{-(n+d)} e^{-z/\theta} d\theta \\ &= \frac{z^{n+d-1}}{\Gamma(n+d-1)} \int_0^\infty \theta^{-(n+d)} e^{-z/\theta} d\theta \\ &= \frac{z^{n+d-1}}{\Gamma(n+d-1)} \frac{\Gamma(n+d-2)}{z^{n+d-2}} \\ \hat{\theta}_s &= \frac{z}{n+d-2} \quad ; \quad n+d > 2. \end{aligned} \quad (2.2.4)$$

The Bayes estimator under entropy loss function, $\hat{\theta}_e$, using the value of $f(\theta|\underline{x})$ from equation (2.2.2) is the solution of equation (2.1.15) given by

$$\begin{aligned} \hat{\theta}_e &= \left[E_\pi \left(\frac{1}{\theta} \right) \right]^{-1} \\ &= \left[\int_0^\infty \frac{1}{\theta} f(\theta|\underline{x}) d\theta \right]^{-1} \\ &= \left[\int_0^\infty \frac{1}{\theta} \frac{z^{n+d-1}}{\Gamma(n+d-1)} \theta^{-(n+d)} e^{-z/\theta} d\theta \right]^{-1} \\ &= \left[\frac{z^{n+d-1}}{\Gamma(n+d-1)} \int_0^\infty \theta^{-(n+d+1)} e^{-z/\theta} d\theta \right]^{-1} \end{aligned}$$

whis, on simplification, leads to

$$\hat{\theta}_e = \frac{z}{(n+d-1)} \quad (2.2.5)$$

3. Bayes Estimator Under $g_2(\theta)$:

Under $g_2(\theta)$, the posterior distribution is defined by

$$f(\theta|\underline{x}) = \frac{f(\underline{x}|\theta)g_2(\theta)}{\int_0^\infty f(\underline{x}|\theta)g_2(\theta)d\theta} \quad (2.3.1)$$

Substituting the values of $g_2(\theta)$ and $f(\underline{x}|\theta)$ from equations (2.1.10) and (2.1.7) in (2.3.1) and simplifying, we get

$$\begin{aligned} f(\theta|\underline{x}) &= \frac{\left(\frac{1}{\theta}\right)^n \left(\frac{1}{\prod_{i=1}^n x_i}\right) e^{-\frac{z}{\theta}} \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{1}{\theta}\right)^{\alpha+1} e^{-\beta/\theta}}{\int_0^\infty \left(\frac{1}{\theta}\right)^n \left(\frac{1}{\prod_{i=1}^n x_i}\right) e^{-\frac{z}{\theta}} \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{1}{\theta}\right)^{\alpha+1} e^{-\beta/\theta} d\theta} \\ &= \frac{(\beta+z)^{n+\alpha}}{\Gamma(n+\alpha)} \theta^{-(n+\alpha+1)} e^{-\frac{1}{\theta}(\beta+z)} \end{aligned} \quad (2.3.2)$$

The Bayes estimator under squared error loss function is the posterior mean given by

$$\hat{\theta}_s = \int_0^\infty \theta f(\theta|\underline{x}) d\theta \quad (2.3.3)$$

Substituting the values of $f(\theta|\underline{x})$ from equation (2.3.2) in equation (2.3.3) and on solving, we get

$$\begin{aligned} \hat{\theta}_s &= \frac{(\beta+z)^{n+\alpha}}{\Gamma(n+\alpha)} \int_0^\infty \theta^{-(n+\alpha)} e^{-\frac{1}{\theta}(\beta+z)} d\theta \\ &= \frac{(\beta+z)^{n+\alpha}}{\Gamma(n+\alpha)} \frac{\Gamma(n+\alpha-1)}{(\beta+z)^{n+\alpha-1}} \\ \hat{\theta}_s &= \frac{\beta+z}{(n+\alpha-1)} \end{aligned} \quad (2.3.4)$$

The Bayes estimator under entropy loss function, $\hat{\theta}_e$, using the value of $f(\theta|\underline{x})$ from equation (2.3.2) is the solution of equation (2.1.15) given by

$$\begin{aligned} &= \left[\int_0^\infty \frac{1}{\theta} f(\theta|\underline{x}) d\theta \right]^{-1} \\ &= \left[\frac{(\beta+z)^{n+\alpha}}{\Gamma(n+\alpha)} \int_0^\infty \theta^{-(n+\alpha-2)} e^{-\frac{1}{\theta}(\beta+z)} d\theta \right]^{-1} \\ &= \left[\frac{n+\alpha}{\beta+z} \right]^{-1} \\ \hat{\theta}_e &= \frac{\beta+z}{(n+\alpha)} \end{aligned} \quad (2.3.5)$$

Bayes Estimator Under $g_3(\theta)$:

Under $g_3(\theta)$, the posterior distribution is defined by

$$f(\theta|\underline{x}) = \frac{f(\underline{x}|\theta)g_3(\theta)}{\int_0^\infty f(\underline{x}|\theta)g_3(\theta)d\theta} \quad (2.4.1.)$$

Substituting the values of $g_3(\theta)$ and $f(\theta|\underline{x})$ from equations (2.1.11) and (2.1.7) in (2.4.1) we get, after simplifying, we get

$$\begin{aligned} f(\theta|\underline{x}) &= \frac{\left(\frac{1}{\theta}\right)^n \left(\frac{1}{\prod_{i=1}^n x_i}\right) e^{-\left(\frac{z}{\theta}\right) \frac{1}{(\beta-\alpha)}}}{\int_\alpha^\beta \left(\frac{1}{\theta}\right)^n \left(\frac{1}{\prod_{i=1}^n x_i}\right) e^{-\left(\frac{z}{\theta}\right) \frac{1}{(\beta-\alpha)}} d\theta}, \\ &= \frac{z^{n-1} \theta^{-n} e^{-z/\theta}}{I_g\left(\frac{z}{\alpha}, n-1\right) - I_g\left(\frac{z}{\beta}, n-1\right)}, \end{aligned} \quad (2.4.2)$$

Where

$I_g(x, n) = \int_0^x e^{-t} t^{n-1} dt$ is the incomplete gamma function

The Bayes estimator under squared error loss function is the posterior mean given by

$$\hat{\theta}_S = \int_\alpha^\beta \theta f(\theta|\underline{x}) d\theta \quad (2.4.3)$$

Substituting the values of $f(\theta|\underline{x})$ from equation (2.4.2) in equation (2.4.3), we get

$$\hat{\theta}_S = \int_\alpha^\beta \theta \frac{z^{n-1} \theta^{-n} e^{-z/\theta}}{I_g\left(\frac{z}{\alpha}, n-1\right) - I_g\left(\frac{z}{\beta}, n-1\right)} d\theta$$

which on simplification leads to

$$\hat{\theta}_S = \left(\frac{I_g\left(\frac{z}{\alpha}, n-2\right) - I_g\left(\frac{z}{\beta}, n-2\right)}{I_g\left(\frac{z}{\alpha}, n-1\right) - I_g\left(\frac{z}{\beta}, n-1\right)} \right) z. \quad (2.4.4)$$

The Bayes estimator under entropy loss function, using the value of $f(\theta|\underline{x})$ from equation (2.4.2) is the solution of equation (2.1.15) given by

$$\begin{aligned} \hat{\theta}_e &= \left[E_\pi \left(\frac{1}{\theta} \right) \right]^{-1} \\ &= \left[\int_\alpha^\beta \frac{1}{\theta} f(\theta|\underline{x}) d\theta \right]^{-1} \\ &= \left[\int_\alpha^\beta \frac{1}{\theta} \frac{z^{n-1} \theta^{-n} e^{-z/\theta}}{I_g\left(\frac{z}{\alpha}, n-1\right) - I_g\left(\frac{z}{\beta}, n-1\right)} d\theta \right]^{-1} \\ \hat{\theta}_e &= \left[\frac{I_g\left(\frac{z}{\alpha}, n-1\right) - I_g\left(\frac{z}{\beta}, n-1\right)}{I_g\left(\frac{z}{\alpha}, n\right) - I_g\left(\frac{z}{\beta}, n\right)} \right] z \end{aligned} \quad (2.4.5)$$

The equation (2.4.4) and (2.4.5) can be solved numerically.

References

- [1] Gini, c. (1924) : "Premiers recherches sur la éconolabite de laemme", Proceedings of the international Mathematic Congress, Toronto, pp 889-892.
- [2] Singh S.N. (1964 a) : "On the time of first birth", Sankhya, vol26(B), pp 95-102.
- [3] Henary, L (1973) : "Fundements theoriques des mesures de la feconolite naturelle," Reveve del' Institute international de statistique, vol 21, pp 135-151.
- [4] Vincent, P. (1961): "Researches sua la Fecodite- Biologique Institute National 'D' Etudes Dempgraphic Paris", Farncce, Press Universitaires De France.
- [5] Mukheerji, S. P. and Islam, R. (1983). A finite range distribution of failures times, Naval Research Logistics Quarterly, Vol. 30, p. 487 – 491.
- [6] Ferguson, T.S.(1967): Mathematical Statistics: A Decision Theoretic Approach. New York: Academic Press.
- [7] Varian Hal R. Two Problems in the Theory of Fairness.
- [8] Berger, J.O.(1998): 'Bayes Factors' . in Encyclopaedia of Statistical Science, Vol.3 (update), Eds S. Kotz et al., New York: Wiley.
- [9] Zellner, A. On assessing prior distributions and Bayesian regression analysis with g-prior distributions. In Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti, (eds. P. K. Goel and A. Zellner), pp. 233-243. North-Holland/Elsevier(1986).
- [10] Basu, A.P. and Ebrahimi, N. (1991): Bayesian approach to life testing and reliability estimation using asymmetric loss function. J. Statist. Plan. Inf. 29, pp 21-31.