DEMOGRAPHY Volume 9, Number 2 May 1972

MAXIMUM LIKELIHOOD ESTIMATES FOR THE PARAMETERS OF A CONTINUOUS TIME MODEL FOR FIRST CONCEPTION

S. N. Singh
T. Bhaduri
Demographic Research Centre, Banaras Hindu University, Varanasi, India

Abstract—The duration of time between two successive births or between marriage and first birth is an indicator of the level of fertility of a couple. Potter and Parker (1964) and Singh (1961, 1967) have suggested the Type I Geometric as a distribution appropriate for representing the length of interval to first conception leading to a live birth. Potter and Parker estimated the parameters of this distribution with the help of the first two moments. Majumdar and Sheps (1970) pointed out the limitations of these moment estimates and gave a method to obtain maximum likelihood estimates, based on formulas which are too involved for solution without the help of a computer.

Singh proposed a continuous probability distribution based on another set of assumptions for the above situation. He outlined a method to obtain best asymptotically normal estimates of the parameters. These estimates are obtained after several iterations starting from any set of consistent estimates.

The objective of this paper is to show that it is relatively easier to obtain maximum likelihood estimates of the parameters of the continuous model, which describes the data on duration to first conception as well as does the discrete model. Simple expressions for the moment and maximum likelihood estimates with the corresponding covariance matrices are obtained. Application is made to three sets of data.

INTRODUCTION

For analyzing data on the first conception leading to a live birth Potter and Parker (1964) and Singh (1961, 1967) suggested the Type I Geometric distribution as a useful model. Potter, Parker, and Singh estimated the parameters using the first two moments of the conception months. Majumdar and Sheps (1970) pointed out the limitations of such estimates and gave a method for obtaining maximum likelihood (M.L.) estimates, which possess certain optimal properties. But it is difficult to compute these estimates without the help of a computer.

Singh (1961, 1964) proposed a continuous model for the duration of time from marriage to first conception, and obtained best asymptotically normal (B.A.N.) estimates of the parametres. The B.A.N. estimates are obtained after several iterations starting from a set of consistent estimates. The procedure is very lengthy, and the initial estimates are obtained on the basis of the first two cell frequencies, which introduces some arbitrariness.

The objective of this paper is to show

that M.L. estimates of the parameters and expected frequencies relating to the continuous model can easily be obtained. Therefore, whenever data on the time of first conception are available, it seems advisable to represent them by a continuous distribution. For illustration, this approach has been applied to the Princeton Fertility Study (see Westoff et al., 1961, pp. 47–69), Hutterite data (Sheps, 1967, pp. 129–132), and the Varanasi Survey (Singh, 1964).

The Model

Let T be the period from marriage to the first conception leading to a live birth, when the female is exposed to the risk of conception. The distribution of T as given in Singh (1961, 1964) is derived under the following assumptions:

(a) The number of coitions during any time interval $(0, t)$ of length t is a random variable and follows the Poisson distribution with parameter $t\lambda_1$, where λ_1 is a positive constant.
(b) Coitions are mutually independent and p_1, the probability of a coition resulting in conception, is constant.
(c) Conceptions are mutually independent and p_2, the probability of a conception resulting in a live birth, is constant.

Under the assumptions (a), (b), and (c) the number of live births follows a Poisson distribution with parameter $\lambda = \lambda_1 p_1 p_2$, if conceptions are assumed to be instantaneous, i.e., the related periods of temporary sterility (gestation and postpartum amenorrhea) are zero. In the case of a first birth it is enough to assume that a conception not resulting in a live birth is instantaneous.

In Singh (1961) the parameters are assumed to be constant for the simplicity of the derivation of the model. Singh also assumed that

(d) λ follows a Pearson Type III distribution with parameters a and h.

Let $g(t, a, h)$ be the probability density function of T. Under the above assumptions,

$$g(t, a, h) = \frac{ah^a}{(h + t)^{a+1}} \qquad \frac{a > 0}{h > 0}. \tag{2.1}$$

It should be noted that λ differs from couple to couple, but for each couple it is constant over time.

Let P_i be the probability of the event that T lies between t_i and t_{i+1} for $i = 1, 2, \cdots, n - 1$, with t_1 equal to zero, and let P_n be the probability that T is greater than t_n.

From (2.1),

$$P_i = \frac{h^a}{(h + t_i)^a} - \frac{h^a}{(h + t_{i+1})^a} \qquad i = 1, 2, \cdots, (n - 1) \tag{2.2}$$

$$P_n = \frac{h^a}{(h + t_n)^a}. \tag{2.3}$$

Estimation

Maximum Likelihood Estimates

Let $T_1, T_2, \cdots, T_N$ be a random sample of size N from the population with density (2.1). The logarithm of the joint density of $T_1, T_2, \cdots, T_N$ is given by

$$\log L = N(\log a + a \log h) - (a + 1)[\log (h + t_1) + \log (h + t_2) + \cdots + \log (h + t_N)]. \tag{3.1}$$

The M.L. estimates $\hat{a}$ and $\hat{h}$ are obtained from the equations

$$S_1 = \frac{\partial \log L}{\partial a} = 0$$

$$S_2 = \frac{\partial \log L}{\partial h} = 0. \tag{3.2}$$

Solving the above equations, we obtain

$$\hat{a} = \frac{N}{\sum_{i=1}^{N} \log \left(1 + \frac{t_i}{\hat{h}}\right)} \tag{3.3}$$

and

$$\frac{N}{\sum_{i=1}^{N} \log\left(1 + \frac{t_i}{\hat{h}}\right)} = \frac{\sum_{i=1}^{N} \frac{1}{\hat{h} + t_i}}{\frac{N}{\hat{h}} - \sum_{i=1}^{N} \frac{1}{\hat{h} + t_i}}. \tag{3.4}$$

The asymptotic covariance matrix, I^{-1}, of the M.L. estimates is given by (3.6), where

$$I = \begin{bmatrix} E\left(-\frac{\partial^2 \log L}{\partial a^2}\right) & E\left(-\frac{\partial^2 \log L}{\partial a\, \partial h}\right) \\ E\left(-\frac{\partial^2 \log L}{\partial a\, \partial h}\right) & E\left(-\frac{\partial^2 \log L}{\partial h^2}\right) \end{bmatrix} \tag{3.5}$$

$$I^{-1} = \begin{bmatrix} V(\hat{a}) & \operatorname{cov}(\hat{a}, \hat{h}) \\ \operatorname{cov}(\hat{a}, \hat{h}) & V(\hat{h}) \end{bmatrix}. \tag{3.6}$$

One easily obtains

$$E\left(-\frac{\partial^2 \log L}{\partial a^2}\right) = \frac{N}{a^2} \tag{3.7}$$

$$E\left(-\frac{\partial^2 \log L}{\partial h^2}\right) = \frac{aN}{(a+2)h^2} \tag{3.8}$$

$$E\left(-\frac{\partial^2 \log L}{\partial a\, \partial h}\right) = -\frac{N}{h}\frac{1}{(a+1)} \tag{3.9}$$

$$V(\hat{a}) = \frac{a^2}{N}(a+1)^2 \tag{3.10}$$

$$V(\hat{h}) = \frac{h^2(a+1)^2(a+2)}{aN} \tag{3.11}$$

$$\operatorname{cov}(\hat{a}, \hat{h}) = \frac{ah(a+1)(a+2)}{N}. \tag{3.12}$$

The Method of Moments

From equation (2.1),

$$E(T) = \frac{h}{a-1},$$

$$E(T^2) = \frac{2h^2}{(a-1)(a-2)},$$

$$V(T) = \frac{ah^2}{(a-1)^2(a-2)}.$$

Equating the first two sample raw moments m_1 and m_2 to the corresponding population moments $E(T)$ and $E(T^2)$, we obtain the moment estimates, $\bar{a}$, $\bar{h}$, as follows:

$$\bar{a} = \frac{2(m_2 - m_1^2)}{m_2 - 2m_1^2}, \quad \bar{h} = \frac{m_1 m_2}{m_2 - 2m_1^2}. \tag{3.13}$$

In the usual notation the asymptotic covariance matrix of $(\bar{a}, \bar{h})$ is

$$V = \begin{bmatrix} V(\bar{a}) & \operatorname{cov}(\bar{a}, \bar{h}) \\ \operatorname{cov}(\bar{a}, \bar{h}) & V(\bar{h}) \end{bmatrix}, \tag{3.14}$$

where

$$V(\bar{a}) = \frac{4v_1^2}{N(v_2 - 2v_1^2)^4} \cdot \{4v_2^3 - v_1^2 v_2^2 - 4v_1 v_2 v_3 + v_1^2 v_4\} = \frac{a(a-1)^2(a-2)(a^2 - a + 6)}{N(a-3)(a-4)} \tag{3.15}$$

$$V(\bar{h}) = \frac{1}{N(v_2 - 2v_1^2)^4} \cdot \{v_2^5 + 3v_2^4 v_1^2 + 4v_1^4 v_2^3 - 4v_1^3 v_2^2 v_3 + 4v_1^6 v_4 - 8v_1^5 v_2 v_3\} = \frac{h^2 a(a-1)^2(a^2 - 3a + 4)}{N(a-2)(a-3)(a-4)} \tag{3.16}$$

$$\operatorname{cov}(\bar{a}, \bar{h}) = \frac{2v_1}{N(v_2 - 2v_1^2)^4}\{2v_2^4 + 3v_1^2 v_2^3 - 6v_1^3 v_2 v_3 - v_1 v_2^2 v_3 + 2v_1^4 v_4\} = \frac{ha(a-1)^2(a^2 - 2a + 4)}{N(a-3)(a-4)} \tag{3.17}$$

and v_r is the rth population moment about zero.

Application

The Data

The continuous model has been fitted to three distributions relating to first conception, taken from the Princeton

Fertility Study, Hutterite data, and the Varanasi Survey. Details of the data are given in the respective papers.

The Results

For illustration the following calculations are given for the data in Table 1. The first two raw moments are $m_1 = 3.302$ and $m_2 = 27.936$. The moment estimates of a and h, obtained with the help of equations (3.13), are $\bar{a} = 5.56$ and $\bar{h} = 15.06$. The expected frequencies based on these values of a and h are calculated using equations (2.2) and (2.3), and are given in column 5 of Table 1. Substituting estimates of a and h in the expressions (3.15), (3.16), and (3.17) we obtain the estimates

$$V(\bar{a}) = 17.561, \qquad V(\bar{h}) = 182.766,$$

$$\text{cov}\,(\bar{a}, \bar{h}) = 56.378, \quad \text{and} \quad r = 0.995,$$

where r is the coefficient of correlation.

The value of M.L. estimate $\hat{h}$ is 11.98, which has been obtained with the help of equation (3.4) after some iterations. Substituting the above value of $\hat{h}$ in equation (3.3), we obtain $\hat{a} = 4.60$. The expected frequencies in column 4 of Table 1 are obtained with the help of these values of $\hat{a}$ and $\hat{h}$. From equations (3.10), (3.11), and (3.12) we obtain the estimates

$$V(\hat{a}) = 3.624, \qquad V(\hat{h}) = 35.160,$$

$$\text{cov}\,(\hat{a}, \hat{h}) = 11.108, \quad \text{and} \quad r = 0.985.$$

The same procedure has been followed in the other two cases (Tables 2 and 3). The estimates by both methods, for the three sets of data, with the corresponding standard errors (wherever possible) are given in Table 4.

TABLE 1.—Observed and Expected Frequencies of Conception by Year, Varanasi Survey

Time from marriage to first conception	Observed frequency	Expected frequency		
		B.A.N.[a] $\hat{a}$=2.3356 $\hat{h}$=5.0236	M.L. $\hat{a}$=4.61 $\hat{h}$=11.98	Method of moment $\bar{a}$=5.56 $\bar{h}$=15.06
1	2	3	4	5
0.0 - 0.8	49	53.71	47.38	45.98
0.8 - 2.8	73	64.91	66.69	66.74
2.8 - 4.8	27	26.96	30.96	31.77
4.8 - 6.8	13	13.51	15.77	16.34
6.8 - 8.8	7	7.62	8.64	8.93
8.8 - 10.8	2)	4.67)	5.02)	5.14)
10.8 - 12.8	4)	3.07)	3.06)	3.09)
12.8 - 14.8	4)	2.08)	1.93)	1.93)
16.8	2)	)	)	)
20.8	1)	7.47)	4.55)	4.08)
22.8	1)			
24.8	1)			
Total	184	184.00	184.00	184.00
χ^2 (4 d.f.)		1.86	3.47	4.69

Note: For χ^2 the frequencies have been grouped as indicated in the table.
a-Taken from Singh (1964).

TABLE 2.—Observed and Expected Frequencies of Conception by Month, Hutterite Data

Month of conception	Observed frequency	Expected frequency			
		M.L.[a] $\hat{a}=3.40$ $\hat{b}=9.19$	M.L. $\hat{a}=3.26$ $\hat{h}=10.0$	Method of moment[a] $\bar{a}=4.81$ $\bar{b}=14.30$	Method of moment $\bar{a}=4.81$ $\bar{h}=16.51$
0 - 1	103	92.3	91.34	86.1	84.32
1 - 2	53	62.4	61.91	61.2	60.45
2 - 3	43	43.6	43.37	44.4	44.13
3 - 4	27	31.3	31.19	32.7	32.74
4 - 5	30	23.0	23.04	24.5	24.64
5 - 6	9	17.3	17.26	18.6	18.81
6 - 7	12	13.2	13.25	14.3	14.52
7 - 8	9	10.2	10.31	11.1	11.34
8 - 9	6	8.0	8.13	8.7	8.94
9 - 10	8	6.4	6.50	6.9	7.12
10 - 11	10	5.2	5.25	5.5	5.71
11 - 12	5	4.2	4.29	4.5	4.62
12 - 15	9	8.7	8.92	9.1	9.42
15 - 18	7	5.1	5.32	5.2	5.40
18 - 24	7	5.3	5.59	5.0	5.28
24 - 48	4	5.8	6.33	4.2	4.56
Total	342	342.0	342.00	342.0	342.00
χ^2 (13 d.f.)		17.2	15.92	19.5	18.74

a-Taken from Majumdar and Sheps (1970) for Type I Geometric distribution.

A Modification

For the data in Table 3 the expected frequencies have also been calculated on the assumption that the first cell frequency is slightly inflated, for the reasons pointed out by Majumdar and Sheps (1970). They are of the opinion that, even if all dates given are accurate, there is a complicated relationship between the month of marriage (or of observation) and the order of the ovarian cycle to which it corresponds. Variations in the duration of cycles and in the timing of marriage within a month introduce errors into these data even if all the reported dates are accurate. This factor is, however, expected to affect primarily the results for month one. To allow for this fact the first interval has been extended up to 1.25 months and the corresponding corrections have been introduced in the limits of other intervals. The expected frequencies resulting from M.L. and moment estimates under this assumption are given in columns 8 and 9 respectively. In both cases the value of χ^2 is considerably decreased, which shows that the modification is useful.

CONCLUSION

Two models, one based on the Type I Geometric distribution and one based on a continuous distribution, have been proposed for representing time of first conception leading to a live birth after marriage. If an abortion or stillbirth intervenes, the model may still be applied after substracting the related period of temporary sterility from the affected duration of time. The models have been found useful in explaining several sets of observed data. On the basis of their

TABLE 3.—Observed and Expected Frequencies of Conception by Month, Princeton Fertility Study

Month of conception	Month of conception	Observed frequency	Expected frequency					
			M.L.[a] $\hat{a}$=1.29 $\hat{b}$=2.17	M.L. $\hat{a}$=1.54 $\hat{h}$=3.19	Method of moment[a] $\bar{a}$=2.75 $\bar{b}$=7.82	Method of moment $\bar{a}$=2.97 $\bar{h}$=9.37	M.L.[b] $\hat{a}$=1.82 $\hat{h}$=4.50	Method of moment[b] $\bar{a}$=3.08 $\bar{h}$=10.35
1	2	3	4	5	6	7	8	9
0 - 1	0 - 1.25	380	357.0	328.79	249.2	249.13	344.78	284.03
1 - 2	1.25 - 2.25	153	173.8	176.87	168.5	169.57	155.20	151.72
2 - 3	2.25 - 3.25	94	100.9	107.61	118.2	119.45	101.84	109.60
3 - 4	3.25 - 4.25	45	65.2	71.07	85.5	86.55	70.61	81.10
4 - 6	4.25 - 6.25	93	78.1	86.20	111.7	113.00	89.29	108.40
6 - 9	6.25 - 9.25	51	60.1	66.17	90.0	90.57	70.85	89.46
9 - 12	9.25 - 12.25	46	32.2	34.87	47.2	46.94	37.86	47.53
12 - 24	12.25 - 24.25	68	49.4	51.15	62.7	60.74	54.84	52.99
24 - 48	24.25 - 48.25	20	23.5	21.92	19.4	17.63	21.91	18.60
48 - 72	48.25 - 72.25	8	17.8	13.35	5.6	4.42	10.82	4.57
Total		958	958.0	958.00	958.0	958.00	958.00	958.00
χ^2 (7 d.f.)			34	36	115	119	24	72

a–Taken from Majumdar and Sheps (1970) for Type I Geometric distribution.
b–Computed for the intervals in column (2).

TABLE 4.—Estimates and Their Standard Errors (Where Available) for the Continuous Model

Estimate	Method of moments	Method of maximum likelihood
Varanasi survey		
N = 184		
a . .	5.56 ± 4.19	4.61 ± 1.90
h . .	15.06 ± 13.52	11.98 ± 5.92
r . .	0.995	0.985
Hutterite data		
N = 342		
a . .	4.81 ± 1.51	3.26 ± 0.75
h . .	16.51 ± 6.17	10.00 ± 2.92
r . .	0.996	0.972
P.F.S.		
N = 958		
a . .		1.54 ± 0.13
h . .		3.19 ± 0.40
r . .		0.920

critical analysis of various properties of the moment and M.L. estimates of parameters of the Type I Geometric distribution, Majumdar and Sheps (1970) recommended M.L. estimates over moment estimates despite greater difficulties of calculation. Here we have examined the relative merits of the moment and M.L. estimates for a continuous distribution. The moment estimates in the present case have disadvantages similar to those of corresponding estimates for the Type I Geometric distribution. (The moment estimates do not exist for $a \leq 2$ and the variances of such estimates are undefined for $a \leq 4$). But the situation is different for M.L. estimates. It is much easier to compute M.L. estimates, the asymptotic covariance matrix, and expected frequencies for the continuous distribution than for the discrete Type I Geometric distribution. Moreover, the present model provides as good fit as the Type I geometric. It is, therefore, preferable to use the continuous model with M.L. estimates when analyzing data on time of first conception leading to a birth.

ACKNOWLEDGMENTS

This investigation is supported by grant No. 67-99 of the Population Council, New York. The authors are indebted to Mrs. J. Majumdar for her valuable help in the preparation of the paper.

REFERENCES

Majumdar, H., and M. C. Sheps. 1970. Estimators of Type I Geometric distribution from observations on conception times. Demography 7:349–360.

Potter, Robert G., and M. P. Parker. 1964. Predicting the time required to conceive. Population Studies 18:99–116.

Sheps, Mindel C. 1967. Uses of stochastic models in the evaluation of population policies. I. Theory and approaches to data analysis. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 4. Berkeley: University of California Press.

Singh, S. N. 1961. Dissertation submitted for the degree of Ph.D. at the University of California, Berkeley.

———. 1964. On the time of first birth. Sankhya 26B:95–102.

———. 1967. On the first success. The Journal of Scientific Research, B.H.U. 18:248–252.

Westoff, C. F., R. G. Potter, P. C. Sagi, and E. G. Mishler. 1961. Family Growth in Metropolitan America. Princeton: Princeton University Press.