ON A MODEL FOR THE MECHANISM OF VARIATION IN LACTATION AMENORRHEA

K.B. Pathak and **R.N. Pandey**
International Institute for Population Studies,
Deonar, Bombay, India

Résumé — A l'aide de quelques simples suppositions, on a développé un modèle de probabilité pour décrire une tendance concernant l'amenorrhée post-partum à la suite d'une naissance vivante. En guise d'éclaircissement, on a appliqué le modèle aux données de Saxena (1966) à l'aide des estimations des paramètres obtenus sur une base de fréquence de première cellule et des deux premiers instants de la distribution observée. On a aussi comparé l'ajustement à ceux obtenus en utilisant des estimations de fréquence aussi bien qu'en minimisant le χ^2.

Abstract — Under some simple assumptions, a probability model has been developed for describing the pattern of postpartum amenorrhea following a live birth. In order to illustrate this, the model has been applied to the observed data of Saxena (1966) with the help of the estimates of the parameters obtained on the basis of first cell frequency and the first two moments of the observed distribution. The fit is also compared to the ones obtained by using frequency estimates as well as by minimising the χ^2 (chi-square).

Key Words — **postpartum amenorrhea, breastfeeding, mixture, moment estimates, modified χ^2 (chi-square)**

Introduction

Postpartum amenorrhea (PPA) is the period of temporary sterility which immediately follows the termination of a pregnancy and during which conception does not take place. The PPA following a live birth is an important segment of the closed birth interval, and the need for the study of the mechanism underlying its variation cannot be over-emphasized for the evaluation of the postpartum contraception pro-gramme. It is very difficult to ascertain the time after the birth when the couple should begin to practice contraception. The intensity of this pro-blem is particularly more acute in the developing countries like India or in many sub-Saharan nations of Africa where majorities of mothers breastfeed their children and do not use contraception in order to space their pregnancies (Page and Lesthaeghe, 1981). It is now well established that lactation is one of the most important factors in delaying the onset of menstruation after birth and in fact lengthens PPA (Jain *et al.*, 1970; Tietze, 1961). Therefore, an assessment of the probable effects of breastfeeding on the period of amenorrhea becomes desirable. For demographers as well as for those involved in providing family planning services, it may be of interest to determine the probability of conception in the absence of breastfeeding. This may be possible if we have a methodology capable of measuring the impact that lactation has on amenorrhea. Since a direct method of ascertaining this impact does not exist, the only way to measure it is to develop a model distribution of the PPA and lactation period.

Several studies describe the distribution of PPA in the past. Srinivasan (1966) assumed that this period has a discrete triangular distribution. Talwar (1965) and Yadava (1966) assumed this period to have asym-metrical triangular and chi-square distributions, respectively. Saxena and Pathak (1977) have fitted a mixture of two truncated chi-square distribu-tions to an observed distribution of PPA. Billewics (1979) devised a func-tion to describe the distribution of PPA depending upon the proportion of breastfeeding women. In 1969, Barrett presented postpartum amenor-rhea as a modified Pascal distribution, or, more precisely, as the sum of a constant term plus two geometric variates. This proposal was subse-quently adopted by several researchers in their models of human reproduction. More recently, Potter and Kobrin (1981) argued some ad-vantages of replacing Barrett's modified Pascal distribution with an analogously modified negative binomial distribution. Ginsberg (1972, 1973) also developed a general stochastic model for studying the effects

of lactation on the resumption of ovulation and postpartum menstruation. His model, however, requires information about the length of partial and full breastfeeding. In practice, determination of partial and full breastfeeding remains highly subjective and depends largely on the age and economic, social and cultural background of the respondent. By using the logit system, Lesthaeghe and Page (1980) recently deduced standard schedules which are meant to describe the essence of the observed distribution of lactation and amenorrhea. Jain *et al.* (1970) used a path model to ascertain the contribution of lactation to the variation in length of amenorrhea for individual women.

None of these attempts were, however, able to extract the patterns of breastfeeding from the simple information on the duration of PPA. The aim of the present paper is to develop a probability model of the PPA period which does not require any information about types of breastfeeding patterns but which enables us to derive the proportion of women using the different lactation patterns. The model is comparatively simple and is amenable to a straightforward and convenient estimation procedure. It is essentially a mixture of two displaced geometric distributions.

The proposed distribution not only enables us to estimate the proportions of women in each breastfeeding pattern, but also their propensity to resume menstruation after the live birth. It is a medical fact that there will be a minimum PPA associated with any live birth irrespective of the breastfeeding pattern of the mother. For illustration, the model has been fitted to the observed data of Saxena (1966) with the help of the estimates obtained by using the first cell frequency and the first two raw moments of the observed distribution. The expected frequencies have also been obtained by using minimised χ^2 estimates as well as the frequency estimates of the parameters.

The Model

Our model has been derived under the following assumptions:

1. The mother breastfeeds the child for at least k months if it survives.

2. The minimum length of amenorrhea after live birth is v months, after which the mother is exposed to the risk of menstruation, where $v \leq k$.

19

3. After k months, the probability that the mother breastfeeds as usual is $1 - \alpha$, and the probability that she decreases the intensity of breastfeeding by supplementation with some outside food or milk is α.

4. The probability that an amenorrhic mother practicing normal breastfeeding resumes menstruation in a unit of time is p. This probability remains constant from one unit of time to another as long as she does not supplement the breastfeeding with outside food.

5. In the case where the amenorrhic mother supplements breastfeeding with outside food, the conditional probability p of resuming menstruation in a unit of time is increased to p_1.

It should be mentioned that in the case of infant death, breastfeeding abruptly stops and menstruation may be resumed more quickly, if it has not already occurred. Such cases may be taken into account by increasing p_1 and α to some extent. For simplicity, it is assumed that infant mortality does not significantly affect the distribution of PPA. However, it is possible to account for such cases by making the distribution more complex.

Let random variable x denote the length of PPA after a live birth. Then, under the above assumptions, the probability distribution of x is determined by

(1)

$$P(X = x) \begin{cases} = pq^{x-v} & \text{for } v \leq x < k, \ 0 < p < 1 \\ \\ = q^{k-v}\left[\alpha p_1 q_1^{x-k} + (1 - \alpha)pq^{x-k}\right] & \text{for } x \geq k \end{cases}$$

$$0 < p_1 < 1$$
$$0 < \alpha < 1$$

For $\alpha = 0$, the above model reduces to a simple geometric distribution, and for $\alpha = 1$, it reduces to a modified geometric distribution. In reality, p and p_1 may vary between women. We can therefore take suitable distributions for p and p_1 to take into account the heterogeneity of women with respect to p and p_1. For the present analysis, we have assum-

ed that the group of women is homogeneous with respect to the risk of exposure to menstruation after live birth *if* their pattern of breastfeeding is the same.

The distribution of PPA may also be treated as continuous. A continuous analogue of Equation 1 is:

$$f(x) \begin{cases} = \lambda e^{-\lambda(x-v)} & \text{for } v \leq x < k, \ \lambda > 0 \\ \\ = e^{-\lambda(k-v)} \left[\alpha\lambda_1 e^{-\lambda_1(x-k)} + (1-\alpha)\lambda e^{-\lambda(x-k)} \right] \end{cases} \tag{2}$$

$$\text{for } x \geq k$$
$$\lambda_1 > 0$$
$$0 < \alpha < 1$$

where λ and λ_1 denote the risks of exposure to menstruation for the mothers who belong in the two patterns of breastfeeding after a live birth.

Application

The distribution given by Equation 1 has five parameters, namely, v, k, α, p and p_1. It may be noted that the maximum likelihood equations become quite intractable for estimating these parameters. Moment estimates in the case of mixture of two or more distributions (especially power series distributions) are however, easier to obtain under certain simplifying assumptions. By assuming the empirical values of v and k, the moment estimates of the parameters α and p_1 can easily be obtained after determining the values of p with the help of first cell frequency of the observed distribution. The first cell frequency implies generally the proportion of mothers resuming menstruation within the minimum period of lactation without food supplements. Our purpose is to provide a simple procedure of estimation so that the estimates so obtained can easily be used as pilot points in obtaining estimates that are more efficient.

In the present case, it is assumed that $v = 1$ because in normal situa-

21

tions no mother will resume menstruation within one month postpartum. The value of k depends upon the lactation practices in different societies. It has been observed that mothers in India normally breastfeed for at least three months, when the child becomes able to eat some solid food. For our purposes, we assume $k = 3$ months. In the case of Western societies, the value of k may be 1.5 or even two months.

Let f_1, μ_1 and μ_2 denote the first cell frequency and first and second raw moments, respectively, of the probability distribution shown by Equation 1. Then

$$f_1 = p + qp \tag{3}$$

$$\mu_1' = \frac{1}{p} - \alpha \left[\frac{1}{p} - p - 2pq \right] + \alpha \frac{q^2}{q_1^2} \left[\frac{1}{p_1} - p - 2p_1 q_1 \right] \tag{4}$$

$$\mu_2' = \frac{2q^2}{p^2} + \frac{q}{p} - \alpha \left[\left(\frac{2q^2}{p^2} + \frac{q}{p} \right) - p - 4pq \right]$$
$$+ \alpha \frac{q^2}{q_1^2} \left[\left(\frac{2q_1^2}{p_1^2} + \frac{q_1}{p_1} \right) - p_1 - 4p_1 q_1 \right] \tag{5}$$

After eliminating α from Equations 4 and 5, we find that

$$\frac{\mu_1' - A}{\mu_2' - B} = \frac{q \left[p_1 - p_1^3 - 2p_1^3 q_1 \right] - Cq_1^2 p_1^2}{q \left[2q_1^2 - q_1^2 p_1^2 - p_1^3 - 4p_1^3 q_1 \right] - Dq_1^2 p_1^2} \tag{6}$$

$$\text{where} \quad A = \frac{1}{p}; \quad B = \frac{2q^2}{p^2} + \frac{q}{p}; \quad C = \frac{1}{p} - p \, 2pq \quad \text{and}$$

$$D = \frac{2q^2}{p^2} + \frac{q}{p} - p - 4pq \tag{7}$$

It may be noted that all the quantities, namely A, B, C and D in Equation 6 are functions of p alone. Now an estimate of p from Equation 3 can be obtained and substituted in Equation 6 to obtain a polynomial in p_1. It can easily be proven that this polynomial will have one consistant solution for p_1 (the one which lies between 0 and 1). Then, after substituting the estimated values of p and p_1 in either of Equations 4 and 5, an estimate of α can be obtained. In application, f_1', μ_1' and μ_2' may be replaced by their observed values f_1', m_1 and m_2, respectively.

The proposed model is applied to Saxena's data (1966) to test its validity in describing the observed distributions. The information collected in the survey relates to the fertility history of the women whose husbands were engaged in the clerical profession in the Banaras Hindu University in India. The information pertaining to PPA was available for 369 women who had given at least one live birth and were 15 to 35 years of age at the time of the survey. The awareness about PPA in this population was sufficient to lend credence to the hope that the resumption of menstruation after childbirth, being an identifiable event, had been correctly reported by the couples (see Saxena and Pathak, 1977).

From the observed distribution of PPA reported in Table 1, we get $f_1' = 0.265583$, $m_1 = 6.2005$ and $m_2 = 56.6504$. Now applying the procedure of estimation discussed earlier, we estimate $\overline{p} = 0.143011$, $\overline{p}_1 = 0.17257$ and $\overline{\alpha} = 0.9000220$. Based on these estimates, the expected number of women corresponding to different durations of PPA are given in column 3 of Table 1. We have also obtained the estimates of these parameters by equating the theoretical and observed relative frequencies in the first three cells of the observed distribution in Table 1. These estimates (which we call the frequency estimates) turn out to be $\hat{p} = 0.14302$, $\hat{p}_1 = 0.18654$ and $\alpha = 0.91502$. The frequencies based on these estimates are given in column 4 of Table 1.

The two fits given in Table 1 are acceptable at a five per cent level of

TABLE 1. OBSERVED AND EXPECTED DISTRIBUTION OF
POSTPARTUM AMENORRHEA FOLLOWING FIRST LIVE BIRTH.

Duration of PPA (in months)	Number of mothers		
	Observed	Expected (Based on moment estimates	Expected (Based on frequency estimates)
1-3	98	98.00	98.00
3-5	90	84.11	90.00
5-7	60	57.95	60.00
7-9	38	39.94	40.03
9-11	27	27.54	26.73
11-13	25	18.99	17.86
13-15	14	13.11	11.95
15-17	8	9.05	8.00
17-19	5)	6.25)	5.36)
19-21	2) 9	4.32) 20.31	3.60) 16.44
22 and over	2)	9.74)	7.48)
Total	369	369.00	369.00
χ^2		8.97	6.39
Modified χ^2		11.50	7.28

significance whether tested by using Pearson's χ^2 or Fraser's modified χ^2 (Fraser, 1968:292). Fraser's modified χ^2 is calculated as

$$\Sigma\, 4\, (\sqrt{O} - \sqrt{E}) \pm$$

where O and E indicate observed and expected frequencies, respectively. The first fit implies that the population of mothers is composed of two groups: one comprising about 90 per cent and the other about 10 per cent. While the first group has an average amenorrhea period of six months, the other has an average amenorrhea of about eight months, imply-

ing the elongation of amenorrhea by two months due to a particular pattern of breastfeeding prevalent in that population. The weighted mean of the PPA of the two groups (as ascertained by the fit in column 3 of Table 1) comes out to be 6.1 months.

It is noteworthy to mention that if we make $\alpha = 0$, the model reduces to a simple geometric distribution, and the value of Pearson's χ^2 based on the frequency estimate $\bar{p} = 0.143$ (obtained by using Equation 3) comes out to be 23.21 (significant value for 7 d.f.). However, the fit obtained on the basis of the moment estimate $\bar{p} = 0.1613$ (by using Equation 4) is satisfactory (Pearson's $\chi^2 = 14.05$ for 7 d.f.). When $\alpha = 1$, the distribution corresponds to a modified geometric distribution, and the estimates of p and p_1 when using Equations 3 and 4 turn out to be 0.143 and 0.2049, respectively. These values give a good fit (Pearson's χ^2 is 8.92 for 6 d.f.).

Calculations suggest that in order to improve the fit towards the tail of the distribution we need to push the value of p upward from 0.143 and that of p_1 from 0.161. Further, p_1 should always be greater than p. The fits so far discussed are based on less efficient estimation procedures. The value of χ^2 may become minimum for some set of the values p, p_1 and α. One procedure to calculate the required values of p, p_1 and α would be to generate the sets of expected frequencies corresponding to all possible values p, p_1 and α in the interval [0, 1]. But this would involve tremendous calculations and would require the use of large high-speed computers. However, the ranges of each of these parameters can be judiciously estimated. From a close look at the different fits obtained on the basis of frequency and frequency *cum* moment estimates of p and p_1, it is evident that p_1 should be always greater than p, and p should not be less than 0.143 because the fit to the tail of the distribution worsens for lesser values. Considering only the value of p, the possible range is [0.143, p_1]. The calculations also show that the value of p beyond 0.250 gives very poor fit to the tail frequencies and produces swelling in the expected frequencies beyond three months. Since the value of p_1 cannot be less than p, we may vary p_1 in the range of [0.144, 0.250]. Thus fixing $\alpha = 0$, the expected frequencies were generated, with the help of a microcomputer, for all the values of $p = 0.143$ (0.001) 0.250, and it was observed that $p = 0.176$ gives the minimum value of χ^2 (Pearson's $\chi^2 = 13.08$ and modified $\chi^2 = 13.74$ on 7 d.f.). This fit is acceptable at a five per cent level of significance and implies that even the geometric distribution can describe the data. However, the fit is tremendously improved when we put $\alpha = 1$ and allow the value of p and p_1 to vary over the

range, as specified earlier. The χ^2 is minimised corresponding to $p = 0.144$ and $p_1 = 0.192$. The value of modified χ^2 is 6.03 (Pearson's $\chi^2 = 6.43$) on 6 d.f. Again, by fixing $\alpha = 0.90022$ (as estimated in Table 1) and repeating the procedure of obtaining a minimum χ^2 value, the estimates are $p = 0.146$ and $p_1 = 0.197$. The value of modified χ^2 turns out to be 6.50 (Pearson's $\chi^2 = 6.86$). This fit is as good as that obtained by putting $\alpha = 1$. The expected frequencies obtained on the basis of the different estimates are shown in Table 2.

The different sets of fits discussed in Tables 1 and 2 lend support to the use of some form of a modified geometric distribution to describe the

TABLE 2. EXPECTED NUMBER OF MOTHERS ACCORDING TO THE PERIOD OF PPA (BASED ON DIFFERENT ESTIMATES OF THE PARAMETERS)

	Expected Number of Mothers				
	$\alpha = 0$		$\alpha = 1$		$\alpha = 0.90022$
Duration of PPA (in months)	Based on Moment Estimates of p	Based on Minimum Estimates of p	Based on Moment Estimates of p and p_1	Based on Minimum Estimates of p and p_1	Based on Minimum Estimates of p and p_1
	$p = .1613$	$p = .176$	$p = .143$ $p_1 = .2049$	$p = .144$ $p_1 = .192$	$p = .146$ $p_1 = .197$
1-3	109.44	118.66	97.98	98.40	99.79
3-5	77.01	80.50	99.74	93.97	93.47
5-7	54.13	54.62	63.03	61.34	60.84
7-9	38.08	37.05	39.82	40.04	39.65
9-11	26.79	25.14	25.16	26.13	25.88
11-13	18.86	17.05	15.87	17.06	16.91
13-15	13.25	11.57	10.07	11.13	11.07
15-17	9.29	7.85	6.31	7.27	7.26
17-19 ⎫ 19-21 ⎬ 21+ ⎭	22.15*	16.56*	11.02*	13.66*	14.14*

* Observed frequency is 9.

variation in the duration of PPA. Our aim in proposing the model has been to take account of a general distribution of PPA under some simple assumptions regarding lactation practices. Only for illustration purposes were the data of Saxena (1966) utilised in examining the validity of the model. We do hope that new sets of data will be able to establish the usefulness of our model. In any case, the pure geometric and the modified geometric distributions are special cases when $\alpha = 0$ and $\alpha = 1$, respectively.

In fact, the reporting of the PPA period generally suffers from the errors of digit preference: the preferred digits are 6, 12 and 18. This may be one reason that cell frequency corresponding to 11-13 is found to be inflated in the observed distribution (when we compare the observed and expected frequencies). The proposed distribution, therefore, may also provide a graduation formula to detect errors in the data on PPA. To make the model more realistic, it would be logical to apply a right-hand truncation on the distribution at a point of, say, two years, rather than allowing the variable to go to infinity.

Acknowledgment

References

Barrett, J.C. 1969. A Monte Carlo simulation of human reproduction. Genus 25:1-22.

Billewics, W.Z. 1979. The timing of postpartum menstruation and breastfeeding. Journal of Biosocial Science 2:141-151.

Fraser, D.A.S. 1968. The Structure of Inference. New York: John Wiley and Sons.

Ginsberg, R.B. 1972. A class of doubly stochastic processes with an application to the effects of lactation on the postpartum anovulatory period. In T.N.E. Greville (ed.), Population Dynamics. New York: Academic Press, pp. 297-332.

_____. 1973. The effect of lactation on the length of postpartum anovulatory period: An application of a bivariate stochastic model. Theoretical Population Biology 4:276-299.

Jain, A.K., T.C. Hsu, R. Freedman and M.C. Chang. 1970. Demographic aspects of lactation and postpartum amenorrhea. Demography 7:255-271.

Lesthaeghe, R. and H.J. Page. 1980. The postpartum nonsusceptible period: Development and application of model schedules. Population Studies 34:143-170.

Page, H.J. and R. Lesthaeghe. 1981. Child Spacing in Tropical Africa: Tradition and Change. New York: Academic Press.

Potter, R.G. and F.E. Kobrin. 1981. Distributions of amenorrhea and anovulation. Population Studies 35:85-99.

Saxena, P.C. 1966. A demographic survey of Banaras Hindu University employees. Unpublished paper. Banaras Hindu University, India.

Saxena, P.C. and K.B. Pathak. 1977. On a distribution of postpartum amenorrhea period following a live birth. Demography - India 6:174-181.

Srinivasan, K. 1966. An application of a probability model to the study of interlive - birth intervals. Sankhya (Series B) 28:175-182.

Talwar, P.P. 1965. A bio-statistical model of the mechanism underlying spacing between births and its application to study the effect of contraception on fertility. Proceedings of the Seminar on New Approaches to the Use of Mathematical Models in Demographic Analysis. Working paper no. 2. Bombay: Demographic Training and Research Centre.

Tietze, C. 1961. The effect of breastfeeding on the rate of conception. Proceedings of the International Conference 2:129-136.

Yadava, R.C. 1966. A probability model for the number of births. In S.N. Singh (ed.), Seminar Volume in Statistics. Varanasi, India: Banaras Hindu University, pp. 85-90.