Demography India
Vol. 17, No 2 (1988), pp. 204-215

Suddhendu Biswas and Ganga Shrestha

Correlation Between Successive Gonceptive Delays

Introduction

ASSUMING that the distribution of fecundability follows a Beta distribution, Potter and Parker (1964) derived expressions for the mean and variance of conceptive delays and correlation between successive delays in the same woman. They distinguished between conceptive delay and the time required to conceive; a conceptive delay is the exposure months preceding but not including the month of conception whereas the time required to conceive includes that month as well. They have also estimated the parameters of type I distribution from two sets of data and illustrated their results numerically by using the estimated value.

Shops (1964) extended the work of Potter and Parker by assuming the frequency distribution of fecundability to assume virtually any shape. It may be a bimodal too. Sheps assumed a heterogeneous population as an aggregate of homogeneous sub-populations and used geometric distribution for the waiting time for conception for a homogeneous population. For heterogeneous population the geometric distribution is weighted by Beta distribution. Sheps, while deriving the expression for correlation also assumed that the *kth* moment of x (the waiting time for conception) in a heterogeneous population is the expected value of the components of the *kth* moment of x in a homogeneous population. Under these assumptions the mean, variance, the correlation between successive conceptive delays were obtained. The present exercise is carried out under the same assumptions as Sheps (1964). A bivariate exponential for the joint distribution of the waiting time of the first and the second order of conceptions, given that the first has occurred at some time origin, has been developed under the same line as Freund (1961). Freund proposed a bivariate extension of the exponential distribution for certain life-testing proplems in particular to two component systems. la his model among the two components there is a chance for anyone to fail first whereas in our model one has to keep in view that the second conception can occur only after the first has occurred. In case of heterogeneous population fecundability is assumed to vary among couples following Pearson type III distribution and hence this exercise may be taken as an extension to Sheps. The model is developed under the following assumptions in accordance with Potter and Parker and Sheps :

1. Conception is a chance occurrence.
2. The fecundability of each couple in a population remains constant.
3. Fecundability varies between couples in the population.
4. The fecundability of each couple remains constant from month to month until pregnancy of any order.

Development of the Model

Let X_1 and X_2 represent the first two successive conceptive delays from the time of effective marriage. The waiting time distribution for the first conception is given by

$$f(x_1 \mid \lambda) = \lambda e^{-\lambda x_1}. \qquad (2.1)$$

The conditional probability distribution for the second conception given that the first has occurred at X_1 is

$$f(x_2 \mid x_1) = \lambda e^{-\lambda(x_2 - (x_1 + \pi))}, \; x_1 + \pi \leqslant x_2 < \ldots \qquad (2.2)$$

It has been assumed that the fecundability parameter λ remains the same in both order of conceptions and π is the infecundable period including gestation period and period of post-partum amenorrhoea followed by the first live birth.

The joint probability density for the first and second order of conceptive delays is thus

$$f(x_1 x_2) = \lambda^2 e^{-\lambda x_1} e^{-\lambda[x_2 - (x_1 + \pi)]}. \qquad (2.3)$$

The moment generating function is given by

$$M(t_1 t_2) = \int_0^\infty \int_{x_1+\pi}^\infty e^{t_1 x_1 + t_2 x_2} \lambda^2 e^{-\lambda x_1} e^{-\lambda[x_2 - (x_1 + \pi)]} \, dx_1 dx_2$$

$$= \int_0^\infty \int_{x_1+\pi}^\infty \lambda^2 e^{t_1 x_1} e^{(t_2 - \lambda) x_2} e^{\pi\lambda} \, dx_1 \, dx_2$$

$$= e^{\pi t_2} \left(1 - \frac{t_2}{\lambda}\right)^{-1} \left(1 - \frac{t_1 + t_2}{\lambda}\right)^{-1}$$

$$= \left(1 + \pi t_2 + \frac{\pi^2 t^2}{2!} + \dots\right)\left(1 + \frac{t_2}{\lambda} + \frac{t_2^2}{\lambda^2} + \dots\right)$$

$$\left(1 + \frac{t_1 + t_2}{\lambda} + \frac{(t_1 + t_2)^2}{\lambda^2} + \dots\right)$$

$$= 1 + \frac{t_1}{\lambda} + \left(\frac{2}{\lambda} + \pi\right)t_2 + \frac{1}{\lambda^2} t_1^2$$

$$+ \left(\frac{\pi}{\lambda} + \frac{2}{\lambda^2} + \frac{1}{\lambda^2}\right) t_1 t_2 + \left(\frac{\pi^2}{2} + \frac{2\pi}{\lambda} + \frac{3}{\lambda^2}\right) t_2^2. \tag{2.4}$$

Equating the appropriate coefficients of the power series in t_1 and t_2 we have for fixed values of λ

$$E(X_1) = \frac{1}{\lambda} \tag{2.4A}$$

$$E(X_2) + \pi + \frac{2}{\lambda} \tag{2.4B}$$

$$E(X_1^2) = \frac{2}{\lambda^2} \tag{2.4C}$$

$$E(X_2^2) = 2\left(\frac{\pi^2}{2} + \frac{2\pi}{\lambda} + \frac{3}{\lambda^2}\right) \tag{2.4D}$$

$$E(X_1 X_2) = \frac{\pi}{\lambda} + \frac{3}{\lambda^2} \tag{2.4E}$$

$$\text{Var}(X_1) = \frac{1}{\lambda^2} \tag{2.4F}$$

$$\text{Var}(X_2) = \frac{2}{\lambda^2} \tag{2.4G}$$

$$\text{Cov}(X_1, X_2) = \frac{1}{\lambda^2} \tag{2.4H}$$

It can be seen immediately that if λ is fixed the correlation in the waiting time between X_1 and $X_2' = X_2 - X_1$ is zero since we have

$$\gamma X_1, X_2' = \frac{\text{Cov}(x_1, x_2)}{\sqrt{\text{Var}(X_1)}\sqrt{\text{Var}(X_2)}} = 0. \tag{2.5}$$

Similar is the finding of Sheps (1964) for a homogeneous population.

Now, we consider the case when the fecundability parameter λ follows Pearsonian type III distribution that is

$$\phi(\lambda) = \frac{e^{-a\lambda}\lambda^{k-1}a^k}{\Gamma(k)} \tag{2.5}$$

$$\Rightarrow E\left(\frac{1}{\lambda}\right) = \frac{a}{k-1} \tag{2.6}$$

$$E\left(\frac{1}{\lambda^2}\right) = \frac{a^2}{(k-1)(k-2)}\,. \tag{2.7}$$

The variances and covariance of the conceptive delays with the intensity of λ weighted by the above distribution (2.5′) are obtained as follows :

$$\begin{aligned}
\text{Var}(X_1) &= E[\text{Var}(X_1 \mid \lambda)] + \text{Var}\,[E(X_1 \mid \lambda)] \\
&= E\left[\frac{1}{\lambda^2}\right] + \text{Var}\left(\frac{1}{\lambda}\right) \quad \text{from (2.4E) and (2.4A)} \\
&= \frac{2a^2}{(k-1)(k-2)} - \frac{a^2}{(k-1)^2} \\
&= \frac{ka^2}{(k-1)^2(k-2)}.
\end{aligned} \tag{2.8}$$

$$\begin{aligned}
\text{Var}(X_2) &= E[\text{Var}(X_2 \mid \lambda)] + \text{Var}\,[E(X_2 \mid \lambda)] \\
&= E\left[\frac{2}{\lambda^2}\right] + \text{Var}\left(\pi + \frac{2}{\lambda}\right) \\
&= \frac{2a^2(k+1)}{(k-1)^2(k-2)}.
\end{aligned} \tag{2.9}$$

$$\begin{aligned}
\text{Cov}(X_1X_2) &= E[\text{Cov}(X_1X_2 \mid \lambda)] + \text{Cov}\,[E(X_1X_2 \mid \lambda)] \\
&= E\left(\frac{1}{\lambda^2}\right) + E\left(\frac{1}{\lambda}, \pi + \frac{2}{\lambda}\right) - E\left(\frac{1}{\lambda}\right)E\left(\pi + \frac{2}{\lambda}\right) \\
&= \frac{(k+1)a^2}{(k-1)^2(k-2)}.
\end{aligned} \tag{2.10}$$

The correlation between the two overlapping successive conceptive delays for heterogeneous population is thus

$$\gamma X_1X_2 = \frac{\dfrac{08\,(k+1)}{(k-1)^2\,(k-2)}}{\sqrt{\dfrac{ka^2}{(k-1)^2\,(k-2)}}\sqrt{\dfrac{2(k+1)\,a^2}{(k-1)^2\,(k-2)}}}$$

$$= \frac{k+1}{\sqrt{2k\,(k+1)}}.$$

$$\text{Cov}\,(X_1,\,X_2 - X_1) = \text{Cov}\,(X_1X_2) - \text{Var}\,(X_1)$$

$$= \frac{(k+1)\,a^2}{(k-1)^2\,(k-2)} - \frac{ka^2}{(k-1)^2\,(k-2)}$$

$$= \frac{a^2}{(k-1)^2\,(k-2)}. \qquad (2.12)$$

$$\text{Var}\,(X_2 - X_1) = \text{Var}\,(X_2) + \text{Var}\,(X_1) - 2\,\text{Cov}\,(X_1X_2)$$

$$= \frac{2a^2\,(k+1)}{(k-1)^2\,(k-2)} + \frac{ka^2}{(k-1)^2\,(k-2)}$$

$$- \frac{2\,(k+1)\,a^2}{(k-1)^2\,(k-2)} = \frac{ka^2}{(k-1)^2\,(k-2)}. \qquad (2.13)$$

The correlation between the two non-overlapping intervals is obtained as

$$\gamma X_1X_2' = \frac{\dfrac{a^2}{(k-1)^2\,(k-2)}}{\sqrt{\dfrac{ka^2}{(k-1)^2\,(k-2)}}\sqrt{\dfrac{ka^2}{(k-1)^2\,(k-2)}}}$$

$$= \frac{1}{k}. \qquad (2.14)$$

It should be noted that the correlation for k > 2 is less than 0.5 and goes on decreasing as *k* increases, which is in accordance to Sheps. Shops has shown that the correlation between conceptive delays which arises due to heterogeneity of population should never be greater than 0.5. This is exactly again in conformity with our findings, namely for the existence of the variance of the waiting time distributions (from marriage to first conception as well as from first to second conceptions) it is necessary that *k*>2.

In the next place we evolve an alternative mathematical approach following

Sheps which uses a model of geometric distribution representing the waiting months for the conception (a discrete random variable) subject to the condition that the waiting time varies from individual to individual. Following a Beta distribution, we employ a continuous time model given by negative exponential distribution subject to the condition that the intensity of conception varies from individual to individual following a gamma distribution which is the best prior distribution in the Baysian sense. Thus Shep's exercise and the present exercise has an important common approach of considering some suitable prior distribution of the parameter for the variation of fecundability. We have.

$$E(X_1 + X_2')^2 = E(X_1^2) + E(X_2'^2) + 2E(X_1X_2')$$

$$E(X_2'^2) = E(X_2 - X_1)$$
$$= E(X_2^2) + E(X_1^2) - 2E(X_1X_2).$$

Substituting from (2.4C), (2.4B), and (2.4E) we get

$$E(X_2'^2) = \frac{2}{\lambda^2} + 2\left(\frac{\pi^2}{2} + \frac{2\pi}{\lambda} + \frac{3}{\lambda^2}\right) - \frac{2\pi}{\lambda} - \frac{6}{\lambda^2}$$

$$= \pi^2 + \frac{2\pi}{\lambda} + \frac{2}{\lambda^2}. \qquad (2.4D')$$

$$E(X_1X_2') = E(X_1, X_2 - X_1)$$
$$= E(X_1X_2) - E(X_1^2)$$
$$= \frac{\pi}{\lambda} + \frac{1}{\lambda^2} \qquad (2.4E)$$

$$E(X_1 + X_2')^2 = \frac{2}{\lambda^2} + \pi^2 + \frac{2\pi}{\lambda} - \frac{2}{\lambda^2} + \frac{2\pi}{\lambda} + \frac{2}{\lambda^2}$$

$$= \frac{6}{\lambda^2} + \frac{4\pi}{\lambda} + \pi^2. \qquad (2.4I)$$

The waiting time distribution for the first conception given by the exponential weighted by gamma distribution is

$$f(x_1) = \int_0^\infty \lambda\, e^{-\lambda x_1} \frac{a^k e^{-a\lambda} \lambda^{k-1}}{\Gamma(k)}\, d\lambda$$

$$\Rightarrow \frac{ka^k}{(a + x_1)^{k+1}} \tag{2.15}$$

$$\Rightarrow \quad E(X_1) = \frac{a}{k-1}. \tag{2.16}$$

$$E(X_1^2) = \frac{2a^2}{(k-1)(k-2)}. \tag{2.17}$$

$$\text{Var}(X_1) = \frac{ka^2}{(k-1)^2(k-2)}. \tag{2.18}$$

The conditional waiting time distribution for the second conception given that the first has occurred at $X_1 = 0$ by using Palm Probability (Biswas and Pachal 1985) is

$$f(x_2' \mid x_1) = \frac{(k+1)\,a^{k+1}}{(a + x_2' - \pi)^{k+2}}, \quad T_1 \leqslant x_2' < \infty,\ a, k > 0 \tag{2.19}$$

$$\Rightarrow E(X_2' \mid X_1 = 0) = \int_{\pi}^{\infty} \frac{x_2'(k+1)\,a^{k+1}}{(a + x_2' - \pi)^{k+2}}\, dx_2$$

$$= \pi + \frac{a}{k} \tag{2.20}$$

and

$$E(X_2'^2 \mid X_1 = 0) = \frac{k+1}{k-1}a^2 - 2a(a-\pi)\frac{k+1}{k} + (a-\pi)^2. \tag{2.21}$$

$$\text{Var}(X_2' / X_1 = 0) = \frac{k+1}{k-1}a^2 - \frac{k+1}{k}2a^2 + \frac{k+1}{k}2a\pi$$

$$+ a^2 - 2a\pi + \pi^2 - \left(\frac{a}{k} + \pi\right)^2$$

$$= \frac{(k+1)\,a^2}{k^2(k-1)}. \tag{2.22}$$

Now

$$E[E(X_1 + X_2)^2] = E\left[\frac{6}{\lambda^2} + \frac{4\pi}{\lambda} + \pi^2\right] \quad \text{from (2.41))}$$

Substituting from (2.6) and (2.7), we get

$$E[E(X_1 + X_2)^2] = \frac{6a^2}{(k-1)(k-2)} + 4\pi \frac{a}{k-1} + \pi^2. \qquad (2.23)$$

$$E(X_1 X_2') = \frac{1}{2} [E(X_1 + X_2')^2 - E(X_1) - E(X_2'^2)]$$

$$= \frac{1}{2}\Big[\frac{6a^2}{(k-1)(k-2)} + \frac{4\pi a}{k-1}$$

$$+ \pi^2 - \frac{2a^2}{(k-1)(k-2)}$$

$$+ \frac{k+1}{k-1} a^2 + 2a(a-\pi)\frac{k+1}{k} - (a-\pi)^2\Big]$$

$$= \frac{a^2(k^2-2)}{(k(k-1)^2(k-2)} - \frac{\pi a}{k}$$

$$\text{Cov}\,(X_1 X_2) = E(X_1 X_2') - E(X_1)\,E(X_2')$$

$$= \frac{a^2(k^2-2)}{(k(k-1)^2(k-2)} - \frac{\pi a}{k} - \frac{a}{k-1}\left(\pi + \frac{a}{k}\right)$$

$$= \frac{a}{k(k-1)}\left[\frac{(3k-4)a}{(k-1)(k-2)} - \pi(2k-1)\right]$$

$$\gamma X_1 X_2' = \frac{\frac{a}{k(k-1)}\left[\frac{(3k-4)a}{(k-1)(k-2)} - \pi(2k-1)\right]}{\sqrt{\frac{ka^2}{(k-1)^2(k-2)}}\sqrt{\frac{(k+1)a^2}{k^2(k-1)}}}$$

$$= \frac{(3k-4)a - \pi(k-1)(k-2)(2k-1)}{a\sqrt{k(k^2-1)(k-2)}} \qquad (2.24)$$

The correlation obtained directly as given by (2.14) differs from the one given by (2.24), because in the latter case, the technique of Palm Probability has been used to obtain the conditional waiting time distribution of the second conception given the time of the first conception. As shown in Biswas and Pachal (1985), the conditional probability distribution (Palm Probability) is obtained from the corresponding unconditional distribution by replacing k by $k + 1$. As such the mean waiting time under the set-up of Palm Probability is reduced frem $a\backslash k$ — 1 to a/k ($k > 2$). This may be apparently a reason for having the possibility of a negative correlation between the interconceptive delays too. It may further be seen that the correlation is negative if

$$a < \frac{\pi (k-1)(k-2)(2k-1)}{(3k-4)} \quad (k > 2) \quad \text{holds.}$$

Application

The parameters a and k are estimated on the basis of moments based on the primary data collected by Dey (1985). The motivation of her work was to study the demographic trend of the Ao Naga tribe. The field work was carried out in some villages near Dimapur, Nagaland (India). The sampling design was simple random sampling. The questionnaire included social aspects such as living standard including food habits, yearly income, medical and sanitation facilities, marriage (within one's caste or intereaste), type of marriage (arranged or otherwise), literacy level etc. The respondents included in the survey belong to the lower middle class earning about Rs. 12,000 to 24,000 per year. The questionnaire relating to reproduction performance included age of mother at different orders of parity, thus giving the time interval from marriage to first birth, first to second and so on. This particular portion of questions relating to fertility behaviour was asked of 50 females of different households. On the basis of the data giving interlive birth intervals the waiting time for successive conceptions was obtained under the assumption of one-to-one correspondence between conception and live birth. The correlation between the first two successive conceptive delays for the same woman was obtained from the bivariate table constructed from the survey data.

The estimated correlation coefficient is

$$Y = 0.57357$$

Equating the sample estimates of time interval from marriage to first conception and from first to second in years from Dey with the theoretical expressions.

We have,

$$E(X_1) = \frac{a}{k-1} = 2.04166 \qquad \textbf{(A)}$$

$$E(X_2 \mid x_1 = 0) = \pi + \frac{a}{k} = 2.42 \qquad \textbf{(B)}$$

For $\pi = 1$ year and solving (A) and (B)

$$\hat{a} = 4.66356$$

$\hat{k} = 3.2842$

Substituting for a and k in (2.24)

$\hat{l} = .3658$

To test if the observed value of $\hat{\gamma}$ differs significantly from the estimated value $\hat{l}$ Fisher's Z-transformations was used that is,

$$Z = \frac{1}{2} \log_e \frac{1 + \gamma}{1 - \gamma} \quad \text{and}$$

$$\xi = \frac{1}{2} \log_e \frac{1 + \rho}{1 - \rho}$$

under H_0 i,e. there is no significant difference between γ and l

$$Z \approx N\left(\xi, \frac{1}{n-3}\right)$$

$$\Rightarrow \quad \frac{z - \xi}{\sqrt{\frac{1}{n-3}}} \sim N(0, 1)$$

Now, we have

$$Z = \frac{1}{2} \log_e \frac{1.57357}{.42643} = .65282$$

$$\xi = \frac{1}{2} \log_e \frac{1.3658}{.6342} = .38356$$

$$S.E(Z) = \frac{1}{\sqrt{n-3}} = \frac{1}{\sqrt{41}} = .15617$$

$$Z = \frac{Z - \xi}{1 \mid \sqrt{n-3}} = \frac{.26926}{.15617} = 1.72414$$

Since Z is smaller than 1.96 there is no significant difference between $\hat{\rho}$ and $\hat{\gamma}$ at 5 per cent level of significance.

Discussion

As it has been noted that there is no significant difference between the observed correlation coefficient and the estimated one, we may conclude that the model developed for the bivariate distribution of the two successive conceptive delays is more or less justified even though the observed data from Dey (1985) are based on a sample of size 50 out of which only 44 could be used as only the same number of mothers had two or more births whereas the remaining six mothers had only single birth. Another limitation of this study it the assumption of one-to-one correspondence between conception and live birth which may not be always true due to the interruption of the waiting time interval by stillbirth or abortion, so that the actual waiting time for the first conceptions as well as between the consecutive conceptions should be less than the observed value which we have considered. Although the data collected include the total number of foetal wastage during the reproductive period it was not known within which order of conception the pregnancy wastage occurred and hence the exact time interval for conception could not be obtained. Nevertheless, as the purpose of the study is to estimate the correlation between successive conceptive delays overestimation of both the consecutive delays for reasons mentioned above may not subitantially vitiate the estimate of the correlation coefficient.

It has been shown by Biswas and Pachal that the covariance between the waiting time for non-overlapping consecutive conceptions is correlated because of the weighting of the fecundability parameter A by a gamma distribution. In the case of a fixed A the correlation becomes zero which was shown in (2.5') in conformity with the finding of Shops.

Following Sheps, we have maintained that correlation is zero in the case of a homogeneous population. On the other hand, the correlation arising because of heterogeneity in fecundability may often be encountered in practice due to sampling from a mixed population made up of subpopulations consisting of women with varying levels of fecundity with unequal probabilities of success. Within a subpopulation the correlation between the successive conceptive delays may be almost zero. But when these subpopulations are mixed up and considered as one population, the fecundity level does not remain fixed among the couples so that the probability of conceiving varies among the couples following a prior distribution and obviously giving rise to non-zero correlation.

References

Biiwas, S. and Pachal, T. K., 1983, On the Application of Palm Probability for Obtaining Inter-arrival Time Distribution in Weighted Poissen Process, *Calcutta Statistical Association Bulletin,* 32,111-15.

._____ and _____, 1985, On the Application of Palm Probability for Obtaining the Waiting Time Distribution between the First and Higher Order of Conceptions, *Demography India.* 14 (2), 277-312.

D«y, Swarupa, 1985, Genetics and Health Status of the Ao Nagas. Field investigation for Ph.D. work.

Freund J. E., 1961, A Bivariate Extension of the Exponential Distribution, *Journal of American Statistical Association, 56,* 971-77.

Potter, R. G. and M. P. Parker, 1964, Predicting the Time Required to Conceive, *Population Studies,* 18, 99-116.

Sheps, M. C., 1964, On the Time Required for Conception, *Population Studies,* 18, 85-97.