5.1 Introduction:

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. 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. Majmudar and Sheps (1970) pointed out the limitations of such estimates and gave a method for obtaining maximum likelihood 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 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.

Singh proposed the following model (1972).

\[(5.1.1) \quad g(t,a,h) = \frac{ah^a}{(h + t)^{a+1}} \quad a > 0, \ h > 0\]

\[= \frac{a}{h} \left[1 + \frac{t}{h}\right]^{-(a+1)}.
\]

This distribution is a particular case of beta distribution.

\[(5.1.2) \quad B(t,k,q,b) = \frac{1}{\beta(k,q)} \cdot \frac{b^k t^{k-1}}{(1 + bt)^{k+q}}.
\]

where \(b = \frac{1}{h}\), \(k = 1\) and \(a = q\).

In this chapter, the gamma distribution is suggested for the distribution of time of the first conception. Further, it is shown how gamma distribution is obtained from (5.1.1) as a limiting distribution. The gamma distribution is fitted
to the observed data collected from the case card records of leading maternity hospitals of Anand and Anand Municipal Hospital, Anand. For sake of comparison, the gamma model is also fitted to Hutterite (Sheps (1967) pp. 129 - 132) data.

5.2 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 cotonions during any time interval $(o, t)$ of length $t$ is a random variable and follows the Poisson distribution with parameter $t\lambda_1$ where $\lambda_1$ is a positive constant.

(b) Cottons are mutually independent and $p_1$ the probability of a cotton 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 $\lambda$ 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. Then $g(t, a, h)$ is given by

$$g(t, a, h) = \frac{ah^a}{(h + t)^{a+1}} \quad a > 0, \ h > 0$$

[Singh's Model, 1972]

$$= \frac{a}{h} \left[ 1 + \frac{t}{h} \right]^{-a+t}$$

We write the above distribution (5.2.1) as a particular case of beta distribution as follows.
(5.2.2) \[ B(t,k,q,b) = \frac{1}{\beta(k,q)} \frac{b^k t^{k-1}}{(1 + bt)^k + q}, \]

where \( b = \frac{1}{h}, \ k = 1 \) and \( q = a \). Now we shall consider the limiting case of equation (5.2.2) when \( b \to 0, q \to \infty \) such that \( bq = c \), a constant.

(5.2.3) \[ g(t,c,k) = \lim_{b \to 0, q \to \infty} B(t,k,q,b) \]

\[ = \frac{t^{k-1}}{\Gamma(k)} \lim_{b \to 0, q \to \infty} \frac{\Gamma(k+q) - (k+q-1)}{\Gamma(q)} \frac{b^k}{(1+bt)^k + q} \]

\[ = \frac{t^{k-1}}{\Gamma(k)} \lim_{q \to \infty} \frac{\Gamma(2r+q-k-1/2)(p-k)}{\Gamma(2r+q-1/2)} e^{-(q-1)} \]

\[ = \frac{c^k}{k q (1 + t/c)^{k+q}} \]

\[ = \frac{c^k}{t^k} \cdot t^{k-1} e^{-ct} \quad k > 0, c > 0, \]

\[ 0 < t < \infty \]

which is the required gamma distribution. For \( k = 1 \), we get exponential distribution.
5.3 **Fitting of the model:**

For fitting the model (5.2.3), the parameters of (5.2.3) are estimated by the method of moments and by the method of maximum likelihood.

### 5.3.1 The Method of Moments:

The $r$th moment $\mu_r'$ of (5.2.3) is given by

\[(5.3.1) \quad \mu_r' = \frac{k(k+1)\ldots(k+r-1)}{c^r} \]

from which one obtains

\[(5.3.2) \quad k = \frac{\mu_2'}{\mu_2' - \mu_1'}, \quad c = \frac{\mu_1'}{\mu_2' - \mu_1'} \]

Replacing $\mu_1'$ and $\mu_2'$ by sample moment $m_1$ and $m_2$ respectively, the estimates of the parameters $k$ and $c$ are obtained as

\[(5.3.3) \quad \hat{k} = \frac{m_2^2}{m_2 - m_1}, \quad \hat{c} = \frac{m_1}{m_2 - m_1} \]

In the usual notation, the asymptotic covariance matrix $V$ of $(\hat{k}, \hat{c})$ is
5.3.2 The Method of Maximum likelihood:

Let \( T_1, T_2, T_3, \ldots, T_n \) be the random sample of size \( n \) from the population with the probability density function (5.2.3). Then the likelihood function given by

\[
L = \left( \frac{c^k}{k!} \right)^n \prod_{i=1}^{n} t_i^{k-1} e^{-ct_i}
\]

Equating the partial derivatives of \( \log L \) w.r.t. \( c \) and \( k \) to zero, we obtain the following equations for estimating \( c \) and \( k \).

\[
\frac{\partial \log L}{\partial c} = \frac{nk}{c} - c \sum t_i = 0
\]

\[
\frac{\partial \log L}{\partial k} = n \log c - n \frac{1}{k} \frac{\partial k}{\partial k} + c \log t_i = 0.
\]
From (5.3.9) we obtain

\[(5.3.11) \hat{c} = \frac{k}{\lambda}\]

where \( \lambda = \frac{\sum t_i}{n} \) = Arithmetic mean of the observations.

From (5.3.10) we obtain

\[(5.3.12) \eta = y\]

where \( \eta = \log k - \frac{1}{\sqrt[k]{k}} \quad \frac{\partial \log k}{\partial k} \)

and \( y = \log A - \log G \)

\[G = \text{Geometric mean}\]
\[= \left( \prod t_i \right)^{1/n}\]

Greenwood (1960) has tabulated the values of \( \eta(k) \)
against \( \eta \). The value of \( k \) satisfying the equation
\[(5.3.12) \]
is obtained from the Tables of Greenwood (1960)
by dividing the value of \( \eta(k) \) corresponding to the
observed value \( y \) by the observed value \( y \). Using this
value of \( k \), the M.L. estimate of \( c \) is obtained from
equation (5.3.11).

In the usual notation the asymptotic covariance
matrix, \( \Gamma^{-1} \), of the maximum likelihood estimates is
given by (5.3.14) where
Using the large sample theory, estimates of $V(\hat{k})$, $V(\hat{c})$ and $\text{cov}(\hat{k}, \hat{c})$ are obtained by replacing $k$ and $c$ respectively by $\hat{k}$ and $\hat{c}$ in equations (5.3.15),

where

(5.3.15) $V(\hat{k}) = \frac{1}{np}$

(5.3.16) $V(\hat{c}) = \frac{c^2}{n} \left( \frac{1}{k} + \frac{1}{kp} \right)$

(5.3.17) $\text{cov}(\hat{k}, \hat{c}) = \frac{c}{kn p}$

$$p = \frac{d^2 \log |F|}{dk^2} - \frac{1}{k}$$

$$\eta = \log k - \frac{d}{dk} \log |F|$$

$$\frac{d\eta}{dk} = \eta' = \frac{1}{k} - \frac{d^2}{dk^2} \log |F| = -p.$$
The value of \[ \frac{1}{p} \] i.e., \[- \frac{1}{\eta} \] at \( k = \hat{k} \) is obtained by equating \[ \frac{1}{\eta} \] by the first difference approximation to \[ \frac{1}{y} \] where

\[
\frac{1}{y} = y^{-1} \left[ \frac{y_{1} - y_{0}^{k}}{y_{1} - y_{0}} - \hat{k} \right]
\]

where \( y_{0} < y < y_{1} \) and \( y_{1} - y_{0} = 0.01 \) and \( y \) is the observed value of \( \eta \) at \( k = \hat{k} \).

5.4 Application:

The Data:

The continuous model has been fitted to two distributions relating to first conception taken from Anand data and Hutterite data (Shepa (1967) pp. 129 - 132).

The Results:

For illustration the following calculations are given for the data in Table 5.1. The first two raw moments are \( m_{1} = 5.1692 \) and \( m_{2} = 47.3384 \). The moment estimates of \( k \) and \( \hat{c} \) obtained with the help of the equation (5.3.3) are \( \hat{k} = 1.2959 \), \( \hat{c} = 0.2507 \). The
expected frequencies based on these values of \( \hat{k} \) and \( \hat{c} \) are calculated and are given in column 3 of Table 5.1. Substituting estimates of \( \hat{k} \) and \( \hat{c} \) in the expressions (5.3.5), (5.3.6) and (5.3.7) we obtain the estimates

\[
V(\hat{k}) = 0.0458, \quad V(\hat{c}) = 0.0021, \quad \text{cov}(\hat{k}, \hat{c}) = 0.0056,
\]

and \( r = 0.8979 \), where \( r \) is the coefficient of correlation.

The value of maximum likelihood estimates \( \hat{k} \) and \( \hat{c} \) are obtained with the help of Tables of

\[
\mathcal{F} \left[ \log \mathcal{F} - \frac{\mathcal{P}^t}{\mathcal{P}} \right]
\]

given by Greenwood (1960). The values are given below. \( \hat{k} = 1.4962 \) and \( \hat{c} = \frac{k}{x} = 0.2894 \).

The expected frequencies in column 4 of Table 5.1 are obtained with the help of these values of \( \hat{k} \) and \( \hat{c} \). From equations (5.3.10), (5.3.11) and (5.3.12) we obtain the estimates

\[
V(\hat{k}) = 0.0285, \quad V(\hat{c}) = 0.0014, \quad \text{cov}(\hat{k}, \hat{c}) = 0.0055, \quad \text{and} \quad r = 0.8594
\]

The same procedure has been followed in the other case (Table 5.2). The estimates by the both methods, for the two sets of data, with the corresponding standard errors are given in Table 5.3.
<table>
<thead>
<tr>
<th>Time from marriage to first conception (in months)</th>
<th>Observed frequency</th>
<th>Expected frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Method of moment</td>
<td>Method of maximum</td>
</tr>
<tr>
<td></td>
<td>$k = 1.2959$</td>
<td>$k = 1.4962$</td>
</tr>
<tr>
<td></td>
<td>$c = .2507$</td>
<td>$c = .2894$</td>
</tr>
<tr>
<td>0 - 2</td>
<td>32</td>
<td>34.50</td>
</tr>
<tr>
<td>2 - 4</td>
<td>40</td>
<td>31.40</td>
</tr>
<tr>
<td>4 - 6</td>
<td>18</td>
<td>22.30</td>
</tr>
<tr>
<td>6 - 8</td>
<td>12</td>
<td>15.40</td>
</tr>
<tr>
<td>8 - 10</td>
<td>11</td>
<td>9.80</td>
</tr>
<tr>
<td>10 - 12</td>
<td>5</td>
<td>6.30</td>
</tr>
<tr>
<td>12 - 14</td>
<td>5</td>
<td>4.20</td>
</tr>
<tr>
<td>14 - 16</td>
<td>3</td>
<td>2.40</td>
</tr>
<tr>
<td>18</td>
<td>2</td>
<td>1.70</td>
</tr>
<tr>
<td>22</td>
<td>1</td>
<td>1.70</td>
</tr>
<tr>
<td>24</td>
<td>1</td>
<td>0.60</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>130</strong></td>
<td><strong>130.00</strong></td>
</tr>
</tbody>
</table>

$\chi^2_{(4 \text{ d.f})}$

4.6675

4.3686
<table>
<thead>
<tr>
<th>Time from marriage to first conception (in months)</th>
<th>Observed frequency</th>
<th>Method of Moments</th>
<th>Expected frequency</th>
<th>Method of maximum likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>based on</td>
<td>based on</td>
<td>based on</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{a} = 4.81$</td>
<td>$(5.1.1)$</td>
<td>$\hat{a} = 4.81$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{b} = 14.30$</td>
<td></td>
<td>$\hat{b} = 16.51$</td>
</tr>
<tr>
<td>$1$</td>
<td>$2$</td>
<td>$3$</td>
<td>$4$</td>
<td>$5$</td>
</tr>
<tr>
<td>$0 - 1$</td>
<td>103</td>
<td>86.1</td>
<td>84.32</td>
<td>114.26</td>
</tr>
<tr>
<td>$1 - 2$</td>
<td>53</td>
<td>61.2</td>
<td>60.45</td>
<td>47.23</td>
</tr>
<tr>
<td>$2 - 3$</td>
<td>43</td>
<td>44.4</td>
<td>44.13</td>
<td>35.12</td>
</tr>
<tr>
<td>$3 - 4$</td>
<td>27</td>
<td>32.7</td>
<td>32.74</td>
<td>26.33</td>
</tr>
<tr>
<td>$4 - 5$</td>
<td>30</td>
<td>24.5</td>
<td>24.64</td>
<td>19.65</td>
</tr>
<tr>
<td>$5 - 6$</td>
<td>09</td>
<td>18.6</td>
<td>18.81</td>
<td>12.41</td>
</tr>
<tr>
<td>$6 - 7$</td>
<td>12</td>
<td>14.3</td>
<td>14.52</td>
<td>16.89</td>
</tr>
<tr>
<td>$7 - 8$</td>
<td>09</td>
<td>11.1</td>
<td>11.34</td>
<td>11.25</td>
</tr>
<tr>
<td>$8 - 9$</td>
<td>06</td>
<td>8.7</td>
<td>8.94</td>
<td>9.16</td>
</tr>
<tr>
<td>$9 - 10$</td>
<td>08</td>
<td>6.9</td>
<td>7.12</td>
<td>7.83</td>
</tr>
<tr>
<td>$10 - 11$</td>
<td>10</td>
<td>5.5</td>
<td>5.71</td>
<td>5.95</td>
</tr>
</tbody>
</table>

contd...
<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>11 - 12</td>
<td>05</td>
<td>4.5</td>
<td>4.62</td>
<td>5.47</td>
<td>4.2</td>
<td>4.29</td>
<td>5.40</td>
</tr>
<tr>
<td>12 - 15</td>
<td>09</td>
<td>9.1</td>
<td>9.42</td>
<td>11.49</td>
<td>8.7</td>
<td>8.92</td>
<td>11.00</td>
</tr>
<tr>
<td>15 - 18</td>
<td>07</td>
<td>5.2</td>
<td>5.40</td>
<td>6.84</td>
<td>5.1</td>
<td>5.32</td>
<td>6.24</td>
</tr>
<tr>
<td>18 - 24</td>
<td>07</td>
<td>5.0</td>
<td>5.28</td>
<td>7.37</td>
<td>5.3</td>
<td>5.59</td>
<td>6.26</td>
</tr>
<tr>
<td>24 - 48</td>
<td>04</td>
<td>4.2</td>
<td>4.56</td>
<td>4.75</td>
<td>5.8</td>
<td>6.33</td>
<td>2.93</td>
</tr>
</tbody>
</table>

| Total | 342 | 342.0 | 342.00 | 342.00 | 342.0 | 342.00 | 342.00 |

\( \chi^2 \) (13 d.f.) 19.5 (15 d.f.) (12 d.f.) (13 d.f.) (13 d.f.) (12 d.f.)

18.74 16.3913 17.2 15.92 15.90

---

\( a \) - Taken from Majmudar and Shepa (1970) for Type I Geometric distribution
### TABLE 5.3
ESTIMATES AND THEIR STANDARD ERRORS FOR THE CONTINUOUS MODEL

<table>
<thead>
<tr>
<th>Estimate</th>
<th>Method of moments</th>
<th>Method of maximum likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N = 130</td>
<td>N = 130</td>
</tr>
<tr>
<td>Anand</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Town</td>
<td>c: 0.2507 ± 0.0458</td>
<td>c: 0.2894 ± 0.0374</td>
</tr>
<tr>
<td></td>
<td>k: 1.2959 ± 0.2140</td>
<td>k: 1.4982 ± 0.1688</td>
</tr>
<tr>
<td></td>
<td>r: 0.8979</td>
<td>r: 0.8594</td>
</tr>
<tr>
<td>Hutterite</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data</td>
<td>N = 342</td>
<td>N = 342</td>
</tr>
<tr>
<td></td>
<td>c: 0.1350 ± 0.0195</td>
<td>c: 0.2005 ± 0.0171</td>
</tr>
<tr>
<td></td>
<td>k: 0.5845 ± 0.0735</td>
<td>k: 0.8678 ± 0.0571</td>
</tr>
<tr>
<td></td>
<td>r: 0.9037</td>
<td>r: 0.8888</td>
</tr>
</tbody>
</table>

From Tables 5.1 and 5.2 it is seen that the model derived in (5.2.3) gives a better fit.