CHAPTER II

MODELS FOR INCUBATION PERIOD

2.0 INTRODUCTION

The HIV incubation period is the random time between the HIV infection and the onset of clinical AIDS symptoms. The probability distribution of this non-negative random variable is known as HIV incubation period distribution. Medley et al (1987) showed that the incubation period of HIV is known to be very long and it is highly variable within and between cohorts.

antiviral drugs and other opportunistic infections. Hence for estimation of the
HIV infection and projection of future HIV prevalence and AIDS, it is very
much important to study the HIV incubation distribution under different
conditions.

In this chapter some of the probability distributions that were used in
the literature for modeling incubation period are presented. Some new models
for incubation period are also proposed. The sensitivity of the backcalculation
estimates to various choices of incubation period distribution are presented.
The sensitivity with respect to specific applications is further discussed in
Chapter IV and V.

2.1  STATISTICAL MODELS FOR INCUBATION PERIOD

The incubation period models are similar to survival models based on
non-negative random variables and can be fitted using either Parametric or
Semi-parametric approach. Here we restrict our attention to only parametric
models for incubation period.

2.1.1  Weibull and Gamma Models

Weibull and gamma models are the most commonly used for many real
data applications and in particular for backcalculation approach. Between the
two, Weibull model is a popular model for HIV incubation period because of
its nice properties viz., proportional hazard as well as accelerated failure time model.

The Weibull distribution function is given by

\[ F(t) = 1 - e^{-\left(\lambda t\right)^{\alpha}} \quad \lambda > 0, \alpha > 0, \ t > 0 \] (2.1.1)

The density function is given by

\[ f(t) = \begin{cases} \alpha \lambda (\lambda t)^{\alpha-1} e^{-\left(\lambda t\right)^{\alpha}} & \lambda > 0, \alpha > 0, \ t > 0 \\ 0 & \text{otherwise} \end{cases} \] (2.1.2)

The hazard function is given by

\[ h(t) = \lambda^\alpha t^{\alpha-1} \] (2.1.3)

The mean and variance of the distribution are

Mean \[ = \frac{\Gamma(1 + 1/\alpha)}{\lambda} \] (2.1.4)

and

\[ \text{Variance} = \frac{\left\{ \Gamma(1 + 2/\alpha) - \left[ \Gamma(1 + 1/\alpha) \right]^2 \right\}}{\lambda^2} \] (2.1.5)

The median of the distribution is

\[ \text{Median} = \left[ \frac{0.69315}{\lambda^\alpha} \right]^{1/\alpha} \] (2.1.6)
The hazard function is increasing with $t$ if $\alpha > 1$ and decreasing if $\alpha < 1$. The Weibull model reduces to negative exponential model if $\alpha = 1$ and has the constant hazard rate for this choice. Naturally Weibull model with increasing hazard ($\alpha > 1$) have been used in many studies for modeling incubation period.

The earliest studies of Weibull incubation period have been attempted by Lui et al (1986) and Medley et al (1987). The study of incubation period distribution for transfusion associated AIDS cases is developed by Lui et al (1986) and is given below:

$$F(t) = 1 - e^{0.0243 t^{2.286}} \quad (2.1.7)$$

These parameter values correspond to a median incubation period of 4.3 years. Medley et al (1987) studied incubation period of patients infected by blood transfusion. The fitted parameter values for the Weibull distribution are as follows:

| Table 2.1.1 |
| Parametric values of Weibull model for patients infected by blood transfusion |

<table>
<thead>
<tr>
<th>Patient group</th>
<th>$\alpha$</th>
<th>$\lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children (0-4 years)</td>
<td>1.9390</td>
<td>0.3843</td>
</tr>
<tr>
<td>Adults (5-59 years)</td>
<td>2.3960</td>
<td>0.1077</td>
</tr>
</tbody>
</table>
Boldson et al (1988) used gamma, Weibull and log-normal models for incubation time of cohort study for San Francisco AIDS cases. The fitted Weibull model for their data is given by

\[ F(t) = 1 - \exp(-0.001296t^{2.5}) \]  \hspace{1cm} (2.1.8)

The Weibull HIV incubation model used by Anderson et al (1986) is given by

\[ F(t) = 1 - \exp(-0.1190t^{1.9974}) \]  \hspace{1cm} (2.1.9)

Brookmeyer and Goedert (1989) used the Weibull incubation period distributions based on the study of haemophiliacs over 20 years of age. The fitted Weibull model for their data is given by

\[ F(t) = 1 - e^{-0.0021t^{2.516}} \]  \hspace{1cm} (2.1.10)

This estimate corresponds to a median incubation of 10 years.

Based on 732 HIV-positive haemophiliacs enrolled in Italian registry, Chiarotti et al (1994) estimated the incubation distributions assuming three different parametric models: uniform \( U_1 \), uniform in three sub intervals \( U_3 \) and truncated Weibull \( W_3 \). Each model has two approaches namely the median (M) and median of three random values (R). So there are six different approaches to estimate the incubation time of individuals. They found that the incubation time obtained using \( U_1 \) and \( U_3 \) is same. Therefore they reported
only four estimates $\mu_1$, $\nu_1$, $\mu_3$ and $\nu_3$. The $\mu_1$ represents the incubation time ascertained by taking median of the interval $(L, R)$ and $\nu_1$ refers to median of the 3 different estimates obtained on the interval $(L, R)$. Similarly $\mu_3$ and $\nu_3$ can be interpreted with reference to Weibull model. The estimates of the four models are given in the following table.

<table>
<thead>
<tr>
<th>Model parameter</th>
<th>Method of Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mu_1$</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>2.9</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.0654</td>
</tr>
<tr>
<td>Median incubation time (years)</td>
<td>13.5</td>
</tr>
</tbody>
</table>

Munoz and Xu (1996), based on a Multicenter AIDS Cohort Study (MACS), obtain the following estimated Weibull model.

$$F(t) = 1 - e^{-0.052087t^{1.285347}}$$

(2.1.11)
The median incubation period corresponding to the above model is 7.5 years. Other important studies which used the Weibull model for HIV incubation period include Mode et al (1988), Isham (1989), Kalbfleisch and Lawless (1989) and Rosenberg and Gail (1990).

The gamma distribution is another important parametric distribution used to model incubation period of HIV/AIDS.

The density function of gamma distribution is

\[
f(t) = \frac{1}{\sigma \Gamma(k)} \left( \frac{t}{\sigma} \right)^{k-1} \exp\left( -\frac{t}{\sigma} \right) \quad t > 0, \, \sigma > 0, \, k > 0 \quad (2.1.12)
\]

The hazard function is

\[
h(t) = \frac{f(t)}{1 - F(t)} = \frac{\frac{1}{\sigma \Gamma(k)} \left( \frac{t}{\sigma} \right)^{k-1} \exp\left( -\frac{t}{\sigma} \right)}{1 - \int_0^t \frac{1}{\sigma \Gamma(k)} \left( \frac{x}{\sigma} \right)^{k-1} \exp\left( -\frac{x}{\sigma} \right) \, dx} \quad t > 0, \, \sigma > 0, \, k > 0 \quad (2.1.13)
\]

The mean and variance of this distribution are

Mean \quad = \quad k\sigma

Variance \quad = \quad k\sigma^2 \quad (2.1.14)

The median \( M \) of this distribution is obtained by solving

\[
\frac{M}{\sigma \Gamma(k)} \left[ \frac{t}{\sigma} \right]^{k-1} \exp\left( -\frac{t}{\sigma} \right) \, dt = 0.5 \quad (2.1.15)
\]
One of the earliest studies that used gamma model for incubation period of HIV is by Medley et al (1987). The parameter estimates for the gamma models for adults and children are as follows:

**Table 2.1.3**

<table>
<thead>
<tr>
<th>Patient Group</th>
<th>$k$</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children (0-4 years)</td>
<td>2.669</td>
<td>0.911</td>
</tr>
<tr>
<td>Adults (5-59 years)</td>
<td>2.473</td>
<td>11.001</td>
</tr>
</tbody>
</table>

The parameter estimates of gamma model obtained by Boldsen et al (1988) based on the San Francisco AIDS data are $k = 3.130$ and $\sigma = 5.715$ years. Freund and Book (1990) fitted gamma model with $k = 3$ (Erlang form) to the San Francisco AIDS data and the estimate obtained for the parameter $\sigma$ is 2.660 years.

### 2.1.2 Log-logistic and Log-normal models

The log-logistic distribution function is

$$F(t) = 1 - [1 + (\lambda t)^v]^{-1} \quad \lambda > 0, \ v > 0, \ t > 0$$  \hspace{1cm} (2.1.16)
The density function of the distribution is

\[ f(t) = \frac{\nu \lambda (t \lambda)^{\nu-1}}{[1 + (t \lambda)^\nu]^2} \quad \lambda > 0, \ \nu > 0, \ t > 0 \quad (2.1.17) \]

The hazard function of the distribution is

\[ h(t) = \frac{\nu \lambda (t \lambda)^{\nu-1}}{[1 + (t \lambda)^\nu]} \quad (2.1.18) \]

The mean and variance of the distribution are

\[
\text{Mean} = \left[ B(1 + \frac{1}{\nu}, 1 - \frac{1}{\nu}) \right] \frac{1}{\lambda} ; \quad \nu > 1
\]

\[
\text{Variance} = \left[ B(1 + \frac{1}{2\nu}, 1 - \frac{1}{2\nu}) - \left\{ B(1 + \frac{1}{\nu}, 1 - \frac{1}{\nu}) \right\}^2 \right] \frac{1}{\lambda^2} ; \quad 2\nu > 1
\]

\[ (2.1.19) \]

The median of the distribution is

\[ \text{Median} = \frac{1}{\lambda} \quad (2.1.20) \]

The earliest application of log-logistic models for incubation period of HIV among homosexual men was adopted by Lui et al (1988). Lawless and Sun (1992) also used the log-logistic model for HIV incubation period. The estimates of the parameters obtained by them are \( \lambda = 0.10 \) and \( \nu = 3.08 \). In addition to Weibull model, Chiarotti et al (1994) used log-logistic model and
generalized exponential model for their data. The methods of estimation MU\(_t\), RU\(_t\), MW\(_t\) and RW\(_t\) are as explained in section 2.1.1. The parameter estimates of the log-logistic model are given below:

Table 2.1.3

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Methods of Estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MU(_t)</td>
</tr>
<tr>
<td>(\lambda)</td>
<td>0.0694</td>
</tr>
<tr>
<td>(\nu)</td>
<td>3.0</td>
</tr>
<tr>
<td>Median incubation period (years)</td>
<td>14.4</td>
</tr>
</tbody>
</table>

The log-normal distribution has been used by Rees (1987) and Boldsen et al (1988) for HIV incubation period. Recent studies by Munoz and Xu (1996) and Munoz et al (1997) have shown that log-normal distribution fits better than Weibull model.

The density function of the log-normal distribution is

\[
f(t; \mu, \sigma^2) = \frac{1}{t\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2} (\log t - \mu)^2\right\} \quad t > 0, \quad -\infty < \mu < \infty, \quad \sigma > 0
\]

(2.1.21)
The distribution function is

$$F(t) = \Phi \left( \frac{\log t - \mu}{\sigma} \right)$$  \hspace{1cm} (2.1.22)

where

$$\Phi (t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2} dx$$  \hspace{1cm} (2.1.23)

denotes the cumulative distribution function of standard normal variate.

The hazard function of the distribution is

$$h(t) = \frac{1}{t \sigma \sqrt{2\pi}} \exp \left\{ -\frac{1}{2\sigma^2} (\log t - \mu)^2 \right\} \frac{1 - \Phi \left( \frac{\log t - \mu}{\sigma} \right)}{1 - \Phi \left( \frac{\log t - \mu}{\sigma} \right)}$$  \hspace{1cm} (2.1.24)

The mean and variance of this distribution are

Mean = $\exp (\mu + \frac{\sigma^2}{2})$

Variance = $\exp (2\mu + \sigma^2) \{ \exp (\sigma^2) - 1 \}$  \hspace{1cm} (2.1.25)

The median of the distribution is

Median = $\exp (\mu)$  \hspace{1cm} (2.1.26)
The parameter estimates obtained by Boldsen et al (1988) are given by $\mu = 1.099$ and $\sigma = 0.322$. This estimate corresponds to a median incubation period of 3 years. The parameter estimates for log-normal model based on the study of Munoz and Xu (1996) is given by $\mu = 2.208$ and $\sigma = 0.683$. This estimate corresponds to a median incubation period of 9.098 years.

2.1.3 Generalized Family (Exponential, Log-logistic and Gamma Distributions)

The probability density function of generalized exponential distribution is

$$f(t) = \nu \lambda \exp \{-\nu \lambda \} [1 - \exp (-\nu \lambda)]^{\nu - 1} \quad t > 0, \quad \lambda > 0, \quad \nu > 0 \quad (2.1.27)$$

The distribution function is

$$F(t) = 1 - [1 - \exp (-\nu \lambda)]^\nu \quad (2.1.28)$$

The hazard function of the distribution is

$$h(t) = \frac{\nu \lambda \exp \{-\nu \lambda \}}{[1 - \exp (-\nu \lambda)]} \quad (2.1.29)$$

The median incubation period is

$$\text{Median} = (-\log_e (1 - (0.5)^{1/\nu}) / \lambda \quad (2.1.30)$$
The parameter estimates for the above model by Chiarotti et al (1994) is given in the following table.

Table 2.1.4

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Methods of Estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MU₁</td>
</tr>
<tr>
<td>λ</td>
<td>0.1266</td>
</tr>
<tr>
<td>μ</td>
<td>4.2</td>
</tr>
<tr>
<td>Median incubation</td>
<td>15.0</td>
</tr>
<tr>
<td>period (years)</td>
<td></td>
</tr>
</tbody>
</table>

The probability density function of the generalized log-logistic distribution is

\[
f(t) = \frac{1}{[\beta(m_1, m_2)\tau]} \left[H(t)\right]^{m_1} \left[1 - H(t)\right]^{m_2} \quad t > 0, \; m_1 > 0, \; m_2 > 0, \; \tau > 0
\]

(2.1.31)

The distribution function is

\[
G(t) = \frac{1}{\beta(m_1, m_2)} \int_{0}^{H(t)} x^{m_1-1} (1 - x)^{m_2-1} dx \quad t > 0, \; m_1 > 0, \; m_2 > 0
\]

(2.1.32)
where \( H(t) = \frac{1}{1 + \exp\left\{-\frac{\log(t) - \mu}{\tau}\right\}} \quad -\infty < \mu < \infty, \quad \tau > 0 \)

and \( \beta(m_1, m_2) \) is beta integral. The generalized log-logistic distribution reduces to log-logistic distribution when \( m_1 = m_2 = 1 \). Singh and George (1987) and Singh et al (1988) have shown that the three parameters generalized log-logistic distribution with \( m_2 = 1 \) fits better than the log-logistic distribution for data on cancer survival analysis. Tan and Byers (1993) have also used the generalized log-logistic distribution as the incubation distribution in their simulation study on stochastic model for HIV epidemic in homosexual population.

Stacy (1962) introduced a generalization of gamma distribution with the three parameters. The density function of the three parameter generalized gamma distribution is given by

\[
f(t) = \frac{k}{\sigma^\alpha} \frac{t^{\alpha-1} e^{-(t/\sigma)^k}}{\Gamma(\alpha/k)} \quad t > 0, \quad \alpha > 0, \quad k > 0 \quad (2.1.33)
\]

This model is a generalization of many survival distributions. For example, the standard gamma density is obtained when \( k = 1 \). The Weibull distribution arises as a particular case when \( \alpha = k \) and also the density reduces to negative exponential when \( \alpha = k = 1 \). Lawless (1980) has shown
that the log-normal distribution can be obtained as a limiting case of the
generalized gamma distribution.

Stacy (1962) has given the convolution of independent generalized
gamma distribution. But the explicit form of the density function is very
complicated and therefore some special cases of the convolution have been
used in the literature as survival models. The convolution of exponential
distribution has been used as incubation model for HIV.

Let \(X_1, X_2, \ldots, X_k\) be the \(k\) independent exponential random variables with
mean \(\sigma_j\). Then the distribution \(T = \sum_{i=1}^{k} X_i\) is a special case of the
convolution of generalized gamma distribution. The probability density
function of \(T\) is given by

\[
f(t) = \sum_{j=1}^{k} B_{1k}(j) \exp(-\lambda_j t) \quad (2.1.34)
\]

where \(\lambda_j = 1/\sigma_j\) and \(B_{1k}(j) = \frac{\prod_{i=1}^{k} \lambda_i}{\prod_{i=1, i \neq j}^{k} (\lambda_i - \lambda_j)}\) \quad (2.1.35)

The mean and variance of the distribution are

\[
\text{Mean} = \sum_{j=1}^{k} \sigma_j
\]

\[
\text{Variance} = \sum_{j=1}^{k} \sigma_j^2.
\quad (2.1.36)
Longini et al (1989) used a staged Markov model to estimate the distribution and mean length of the incubation period from a cohort study of 603 HIV infected individuals who have been followed through various stages of infection. They used the generalized gamma model to describe the transition probabilities of the Markov model. The probability of going from a transient state $i$ to a transient state $k$ at time $t$ is given by

$$
p_{ik}(t) = \frac{(-1)^{k-i} \lambda_i \ldots \lambda_{k-1} \sum_{j=1}^{k} \exp(-\lambda_j t)}{\prod_{l=i}^{k} (\lambda_j - \lambda_l)} \quad i = 1,2,\ldots,n ; \quad k \geq i
$$

(2.1.37)

In this formulation of the model the incubation period covers three stages going from initial infection stage to the third stage of pre-AIDS symptoms. The estimated parameter values for the incubation period distribution are $k = 3$, $\lambda_1 = 0.4571$, $\lambda_2 = 0.019$ and $\lambda_3 = 0.0159$. The mean and median of the incubation period are 9.81 and 8.25 years respectively.

In another study Longini et al (1989) have formulated a six stage Markov model to describe the progression of HIV infection to ultimate death of the individuals. The parameter estimates obtained in their study are $k = 6$, $\lambda_1 = 0.0764$, $\lambda_2 = 0.0665$, $\lambda_3 = 0.0499$, $\lambda_4 = 0.4290$, $\lambda_5 = 0.0408$ and $\lambda_6 = 0.0529$. 
2.2 Mixture Model

One way to accommodate the variation between different groups of the population is by using mixture models. Suppose certain proportions of infected individuals $p$ have an incubation period distribution $F_1(t)$ and the remaining proportions $(1 - p)$ have the incubation period distribution $F_2(t)$. Then the incubation period distribution for the entire population of infected individuals is a mixture of $F_1(t)$ and $F_2(t)$ is given by

$$F(t) = pF_1(t) + (1 - p)F_2(t) \quad 0 < p < 1$$  \hspace{1cm} (2.2.1)

The infection density function is given by

$$f(t) = pf_1(t) + (1 - p)f_2(t) \quad 0 < p < 1$$  \hspace{1cm} (2.2.2)

The above can be generalized to many groups of individuals with the mixture of incubation period density function is given by

$$f(t) = \sum_{i=1}^{k} p_i f_i(t) \quad 0 < p_i < 1 \quad \text{and} \quad \sum_{i=1}^{k} p_i = 1$$  \hspace{1cm} (2.2.3)

Auger et al (1988) have considered a mixture of two Weibull densities for the incubation period of Paediatric AIDS cases. The mixture of two Weibull densities is given by
\[ f(t) = p\alpha_1 \lambda_1 (\lambda_1 t)^{-\alpha_1} \exp\left(-\left(\lambda_1 t\right)^{\alpha_1}\right) + (1-p)\alpha_2 \lambda_2 (\lambda_2 t)^{-\alpha_2} \exp\left(-\left(\lambda_2 t\right)^{\alpha_2}\right) \]

\( 0 < p < 1, \alpha_i > 0, \lambda_i > 0, \ i = 1,2 \ ; \ t > 0 \)

(2.2.4)

The parameter estimates obtained by Auger et al (1988) were

\( p = 0.120, \alpha_1 = 3.540, \lambda_1 = 0.201, \alpha_2 = 1.160 \) and \( \lambda_2 = 0.010 \). The mixture

of two Weibull distributions was also used by Lui et al (1988) in a study of

incubation period distribution of sample individuals drawn from San

Francisco AIDS data.

2.3 STAGING MODEL

Under staging models the incubation period is considered to be

comprised of stages. The progression from the time of infection to AIDS was

assumed to occur in 3 stages by Brookmeyer and Liao (1990). The stage 1

refers to HIV infection without immunology abnormalities, stage 2 is the
development of pre-AIDS disease and stage 3 is the development of clinical

AIDS. The incubation time of an individual by definition is the total time

spent on stage 1 and 2. Therefore different models for these two stages can be

assumed. Let \( h_1(t) \) and \( h_2(t) \) denote the hazard functions of the two stages.

The convolution equation for the incubation period comprising of these two

stages as given by Brookmeyer and Liao (1990) is
\[ F(t) = \int_0^t f_1(u) F_2(t - u) du \]  

(2.3.1)

where

\[ f_1(u) = h_1(u) \exp\left\{ - \int_0^u h_1(s) ds \right\} \]  

(2.3.2)

and

\[ F_2(u) = 1 - \exp\left\{ - \int_0^u h_2(s) ds \right\} \]  

(2.3.3)

Suitable changes should be made in the above formulations to account for calendar time of infection.

Suppose the time spent on the two stages are not independent, then the time spent on the stage 2 can be conditioned on the time spent on stage 1. Under this case Mariotti and Cascioli (1996) have given the survival functions for the second stage as

\[ S_2(u / u) = \exp\left\{ - \int_0^u h_2(s) ds \right\} \]  

(2.3.4)

where \( u \) is the time spent on the first stage. The distribution function \( F_2(.) \) in the convolution equation (2.3.3) should be suitably modified by using the survival function \( S_2(.) \) given in equation (2.3.4).

2.4 CHANGE POINT MODELS

In this section, the author proposes a change point model for incubation period of HIV infection. Suppose the incubation time for an individual is \( t \), it is reasonable to assume that between 0 to \( t \), there is a time
point \( \tau \) at which the hazard of incubation changes. The point \( \tau \) may be the time after infection when the individual realizes the threat of AIDS and seeks some kind of medication. In the following sections some change point hazard models are proposed.

2.4.1 Change Point Model with constant Hazard

Suppose the hazard before and after the change point is constant, then \( h(t) \) is given by

\[
h(t) = \begin{cases} 
\alpha & t \leq \tau \\
\beta & t > \tau 
\end{cases}
\]  

(2.4.1)

The survival function of the change point model is given by

\[
S(t) = \exp \left\{ - \int_0^t h(x) \, dx \right\} = \begin{cases} 
e^{-\alpha t} & t \leq \tau \\
e^{-\alpha \tau} e^{-\beta (t-\tau)} & t > \tau 
\end{cases}
\]  

(2.4.2)

The distribution function is given by

\[
F(t) = \begin{cases} 
1 - e^{-\alpha t} & t \leq \tau \\
1 - e^{-\alpha \tau} e^{-\beta (t-\tau)} & t > \tau 
\end{cases}
\]  

(2.4.3)

The density function of the change point model is given by

\[
f(t) = \begin{cases} 
\alpha e^{-\alpha t} & t \leq \tau \\
e^{-\alpha \tau} \beta e^{-\beta (t-\tau)} & t > \tau 
\end{cases}
\]  

(2.4.4)
The median of the incubation period is given by

\[ M = \begin{cases} \log 2 / \alpha & M \leq \tau \\ (2e^{-\tau \alpha} - 1)/(e^{\tau(\beta-\alpha)}) & M > \tau \end{cases} \quad (2.4.5) \]

### 2.4.2 Change Point Model with Varying Hazard

Let the hazard function before and after the change point be as given below.

\[ h(t) = \begin{cases} \alpha & t \leq \tau \\ \alpha \nu t^{\nu-1} & t > \tau \end{cases} \quad (2.4.6) \]

The survival function is given by

\[ S(t) = \begin{cases} e^{-\alpha t} & t \leq \tau \\ e^{-\alpha (t-\tau)^{\nu}} e^{-\alpha t^{\nu}} & t > \tau \end{cases} \quad (2.4.7) \]

The distribution function is given by

\[ F(t) = \begin{cases} 1 - e^{-\alpha t} & t \leq \tau \\ 1 - e^{-\alpha (t-\tau)^{\nu}} e^{-\alpha t^{\nu}} & t > \tau \end{cases} \quad (2.4.8) \]

The density function is given by

\[ f(t) = \begin{cases} \alpha e^{-\alpha t} & t \leq \tau \\ e^{-\alpha (t-\tau)^{\nu}} e^{-\alpha t^{\nu}} \alpha \nu t^{\nu-1} & t > \tau \end{cases} \quad (2.4.9) \]
The median is given by

\[
M = \begin{cases} \frac{\log(2)}{\alpha} & M \leq \tau \\ \frac{1}{\alpha} \log e \left[ 2e^{-\alpha(t-\tau)} \right]^{1/u} & M > \tau \end{cases}
\]  

(2.4.10)

2.4.3 Change Point Model with Weibull Hazard

Suppose the hazards before and after the change point is that of Weibull distribution then the hazard function of the change point model is given by

\[
h(t) = \begin{cases} \lambda_1 \uparrow_1 t^{\uparrow_1-1} & t \leq \tau \\ \lambda_2 \uparrow_2 t^{\uparrow_2-1} & t > \tau \end{cases}
\]  

(2.4.11)

The survival function is given by

\[
S(t) = \begin{cases} e^{-\lambda_1 t^{\uparrow_1}} & t \leq \tau \\ e^{-\lambda_1 t^{\uparrow_1}} \cdot e^{-\lambda_2 [t^{\uparrow_2} - \tau^{\uparrow_2}]} & t > \tau \end{cases}
\]  

(2.4.12)

The distribution function is given by

\[
F(t) = \begin{cases} 1 - e^{-\lambda_1 t^{\uparrow_1}} & t \leq \tau \\ 1 - e^{-\lambda_1 t^{\uparrow_1}} \cdot e^{-\lambda_2 [t^{\uparrow_2} - \tau^{\uparrow_2}]} & t > \tau \end{cases}
\]  

(2.4.13)

The density function is given by

\[
f(t) = \begin{cases} \uparrow_1 \lambda_1 t^{\uparrow_1-1} e^{-\lambda_1 t^{\uparrow_1}} & t \leq \tau \\ e^{-\lambda_1 t^{\uparrow_1}} e^{-\lambda_2 [t^{\uparrow_2} - \tau^{\uparrow_2}]} \lambda_2 \uparrow_2 t^{\uparrow_2-1} & t > \tau \end{cases}
\]  

(2.4.14)
The median of the above model is given by

\[
M = \begin{cases} 
\left( \frac{\log(2)}{\lambda_1} \right)^{1/v_1} & M \leq \tau \\
\frac{\log(2) - \lambda_1 \tau^{v_1} + \lambda_2 \tau^{v_2}}{\lambda_2}^{1/v_2} & M > \tau 
\end{cases}
\]  (2.4.15)

2.5 IMMUNE INVASION LEVEL MODEL FOR INCUBATION PERIOD

In this section the author proposes a model for incubation period using the concept of invasion to immune system. Suppose at time \( t = 0 \), a member tested for HIV positive for the first time, experiences a random \( N \) number of invasion before the member shows clinical symptoms to AIDS. The number of immune invasions \( N \) experienced by the individual is assumed to follow a Poisson process with parameter \( \lambda (>0) \). Let the probability that the individual who has already experienced \( n \) contacts upto a time \( t \), shows the clinical symptom for AIDS in the interval \((t, t+\Delta t)\) be given by

\[
n\mu \Delta t + o(\Delta t), \quad \mu > 0
\]  (2.5.1)

Then the incubation period of the individual \( T \) can be obtained as follows:

By definition

\[
f(t) = \lim_{\Delta t \to 0} \frac{P[t < T < t + \Delta t]}{\Delta t}
\]  (2.5.2)
Therefore \( f(t) \Delta t \) denotes the probability that the individual becomes an AIDS case in the interval \((t, t+\Delta t)\) after experiencing \(n\) invasions. We assume that the system undergoes at least one invasion before the individual become AIDS in \((t, t+\Delta t)\). Hence

\[
f(t) = e^{-\lambda t} \lambda \odot e^{-(\lambda + \mu)t} \mu + e^{-\lambda t} \lambda \odot (\sum_{n=2}^{\infty} e^{-(\lambda + \mu)t} \lambda \odot \ldots \odot e^{-((\lambda + (n-1)\mu)\mu} n\mu) \right) \tag{2.5.3}
\]

Taking Laplace transform on both sides of (2.5.3), we get

\[
f^*(s) = \frac{\lambda}{s + \lambda} \sum_{n=1}^{\infty} \frac{n\lambda^{n-1} \mu}{(s + \lambda + \mu)(s + \lambda + n\mu)} \tag{2.5.4}
\]

By using partial fraction method, the above equation can be written as

\[
f^*(s) = \lambda \sum_{n=1}^{\infty} \frac{1}{(s + \lambda + (n-1)\mu)} \left( \frac{\lambda}{\mu} \right)^{n-1} \left\{ \sum_{j=0}^{n} \binom{n}{j} (-1)^j \frac{1}{(s + \lambda + j\mu)} \right\} \tag{2.5.5}
\]

Inverting the equation (2.5.5) we obtain the probability density function as

\[
f(t) = \lambda e^{-\lambda t} (1 - e^{-\mu t}) e^{\lambda(1 - e^{-\mu t})/\mu} \tag{2.5.6}
\]

The density function is unimodal and is given by

\[
t_m = \frac{1}{\mu} \log \left( \frac{2\lambda}{2\lambda + \mu - \sqrt{\mu^2 + 4\lambda \mu}} \right) \tag{2.5.7}
\]
The distribution function can be written as

\[ F(t) = \frac{\lambda}{\mu} \int_0^{1-e^{-\mu t}} u(1-u)^{(\lambda/\mu)-1} e^{(\lambda/\mu)u} \, du \]  

(2.5.8)

If \( \lambda = \mu \), then

\[ F(t) = 1 - e^{-\lambda t} e^{1-e^{-\lambda t}} \]  

(2.5.9)

The hazard function of equation (2.5.9) is given by

\[ h(t) = \lambda(1 - e^{-\lambda t}) \]  

(2.5.10)

It can be noted that the hazard rate is increasing function of \( t \). The median of the distribution can be obtained numerically by solving equation (2.5.8). The parameters of \( f(t) \) can be estimated by using the method of maximum likelihood.

2.6 DISCRETE TIME FORMULATION OF INCUBATION PERIOD

At a time point \( t = 0 \), an individual is tested for HIV positive. A discrete formulation of Weibull model for an individual to shows clinical symptom of AIDS after \( t \) units of time has reported by Becker et al (1991) and is given by

\[ F(t) = 1 - \exp\{-\beta(t+1)^\alpha\}, \ t = 0,1,2,\ldots \]  

(2.6.1)

The above distribution function gives the probability that the person become a victim of AIDS after infection in \( t \) time points.
In this section we propose an alternative formulation of the problem mentioned above.

Assume that a time \( t = 0 \), the person is tested for HIV positive. Let the conditional probability that the individual shows first identifiable symptom of AIDS during the \( t^{th} \) year given that the individual has not shown any symptom of AIDS in the previous years be given by

\[
1 - e^{-\eta t} \quad t = 1, 2, \ldots, \quad \mu > 0 \tag{2.6.2}
\]

If \( T \) denotes the incubation period then its probability mass function is given by

\[
P[T = t] = e^{-\mu} e^{-2\mu} \cdots e^{-t\mu} (1 - e^{-\mu})
\]

\[
= e^{-\frac{(t-1)\mu}{2}} - e^{-\frac{t(t+1)}{2}} \quad t = 1, 2, \ldots \tag{2.6.3}
\]

The median of the above distribution is given by

\[
\text{Median} = \frac{\sqrt{\mu^2 + 8 \log 2} - \mu}{2\mu} \tag{2.6.4}
\]

The discrete hazard function of the model is given by

\[
h(t) = P[T = t \mid T \geq t]
\]

\[
= \frac{P[T = t]}{P[T \geq t]} = \frac{P[T = t]}{1 - P[T < t]} = 1 - e^{-\eta t} \tag{2.6.5}
\]
The above equation (2.6.5) is the hazard of AIDS conversion at the $i^{th}$ time point. The mean and variance of the above incubation time is given by

$$\text{Mean} = \sum_{t=0}^{\infty} e^{-\frac{t(t+1)\mu}{2}} \quad (2.6.6)$$

$$\text{Variance} = \left[ \sum_{t=0}^{\infty} (2t+1)e^{-\frac{t(t+1)\mu}{2}} \right] - \left[ \sum_{t=0}^{\infty} e^{-\frac{t(t+1)\mu}{2}} \right]^2 \quad (2.6.7)$$

Method of maximum likelihood can be used to estimate the parameters of the model.

So far, the author has given a descriptive account of the statistical models of the incubation period distribution proposed by many researchers. These models have their own merits and demerits. Keeping this view in mind some of the models described above are considered for projection of HIV/AIDS in India and Tamilnadu, which are presented in Chapters IV and V.