Shelf-Life: Designing and Analysing Stability Trials

A thesis submitted for the degree of Doctor of Philosophy by

ANDREAS KIERMEIER

DECEMBER 2003
Abstract

All pharmaceutical products are required by law to display an expiry date on the packaging. The period between the date of manufacture and expiry date is known as the label shelf-life. The label shelf-life indicates the period of time during which the consumer can expect the product to be safe and effective.

Methods for determining the label shelf-life from stability data are discussed in the guidelines on the evaluation of stability data issued by the International Conference for Harmonization. These methods are limited to data that can be analysed using linear model methods. Furthermore, in the situation where a number of batches are used to determine a label shelf-life, the current regulatory method (unintentionally) penalizes good statistical design. In addition, the label shelf-life obtained this way may not be a reliable guide to the properties of future batches produced under similar conditions.

In this thesis it is shown that the current definition of the label shelf-life may not provide the consumer with the desired level of confidence that the product is safe and effective. This is especially the case when the manufacturer has performed a well designed stability study with many assays. Consequently, a new definition for the label shelf-life is proposed, such that the consumer can be confident that a certain percentage of the product will meet the specification by the expiry date. Several methods for obtaining such a label shelf-life under linear model and generalized linear model assumptions are proposed and evaluated using simulation studies.

The new definition of label shelf-life is extended to allow a label shelf-life to be obtained from stability studies that make use of many batches, such that a proportion of product
over all batches can be assured to meet specifications by the expiry date. Several methods for estimating the label shelf-life in the multi-batch case are proposed and evaluated with the help of simulation studies.
Declaration

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

Signed: ___________________________  Date: ___________
Acknowledgements

I wish to thank my supervisors Assoc. Prof. Arūnas P. Verbyla and Dr. Richard G. Jarrett for their support, guidance and friendship during this research.
# Contents

Abstract \hspace{1cm} i
Declaration \hspace{1cm} iii
Acknowledgements \hspace{1cm} iv

1 Introduction \hspace{1cm} 1
1.1 Label Shelf-life Estimation for a Single Batch \hspace{1cm} 2
1.2 Testing for Batch-to-Batch Variability \hspace{1cm} 9
1.3 Label Shelf-life Estimation for Multiple Batches \hspace{1cm} 11
  1.3.1 The Fixed Effects Model: Current Regulatory Guidelines \hspace{1cm} 12
  1.3.2 The Random Effects Model \hspace{1cm} 15
1.4 Addressing the Issues \hspace{1cm} 18

2 Estimation of Shelf-Life: The Single Batch Case \hspace{1cm} 20
2.1 A New Definition \hspace{1cm} 20
2.2 Estimation of the Label Shelf-Life \hspace{1cm} 23
  2.2.1 Normal Approximation \hspace{1cm} 24
  2.2.2 Various Profile Likelihood Based Approaches \hspace{1cm} 26
  2.2.3 Constrained Profile Likelihood (CPL) Based Approaches \hspace{1cm} 37
  2.2.4 Non-central t-distribution (NCT) Approach \hspace{1cm} 45
2.3 Simulation Results \hspace{1cm} 46
2.4 Discussion \hspace{1cm} 70
3 Estimation of Shelf-Life: The Multi-Batch Case 74

3.1 A New Definition .................................................. 74
3.2 Estimation of the Label Shelf-Life ................................. 78
  3.2.1 Normal Approximation ........................................ 79
  3.2.2 Various Profile Likelihood Based Approaches .............. 82
  3.2.3 Constrained Profile Likelihood Based Approaches .......... 88
  3.2.4 Non-central t-distribution (NCT) Approach ................ 91
3.3 Simulation Results .................................................. 94
3.4 Discussion .......................................................... 115

4 Estimation of Shelf-Life: The GLM Single Batch Case 118

4.1 Introduction .......................................................... 118
4.2 Likelihood based approaches ....................................... 121
  4.2.1 Constrained Profile Likelihood ............................... 122
  4.2.2 Example: The Normal Distribution ......................... 125
  4.2.3 Example: The Gamma Distribution ........................... 128
4.3 Simulation Results .................................................. 132
  4.3.1 The Normal Distribution .................................... 132
  4.3.2 The Gamma Distribution .................................... 133
4.4 Discussion .......................................................... 142

5 Conclusions .......................................................... 144

5.1 Other Applications ................................................ 145
5.2 Future Work ....................................................... 147
Pharmaceutical companies are required by law to indicate the expiry date of each pharmaceutical product on the immediate container. This expiry date indicates the end of the period of time, known as shelf-life, during which the product can be expected to meet specifications. The guidelines which govern how the label shelf-life, and hence the expiry date, is determined are set out by the International Conference on Harmonisation (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH, 2003; 2002). These guidelines are adopted by the regulatory authorities of the European Union, Japan and the USA. Other countries also adopt these guidelines with little or no modification.

Before a product can be released to the market, manufacturers need to apply for a label shelf-life to the appropriate authority — the Therapeutic Goods Administration (TGA) in Australia and the Food and Drug Administration (FDA) in the United States. The label shelf-life is only granted if the manufacturer can demonstrate that the product will remain within specifications during that period. After approval, the degradation of the product must also be monitored to ensure that the shelf-life remains appropriate. These requirements are usually achieved by storing the product under suitable conditions and monitoring the degradation. This type of study is referred to as a stability study.

General overviews of the analysis of stability data are given in Kohberger (1988), Lin
(1990), Buncher and Tsay (1993), and Chow and Liu (1995b), while Carstensen and Rhodes (2000) give an overview which is more focused on the development of drug products. The ICH guidelines attempt to give guidance on analysing stability data, and while several areas covered in these guidelines have received further attention in the literature, there are still problems with these recommendations. These include the estimation of a label shelf-life for a single batch, testing for variability between batches, and the estimation of a label shelf-life which is applicable to all future batches. These three areas are reviewed below.

1.1 Label Shelf-life Estimation for a Single Batch

Long term stability studies are typically carried out after approval for a new drug product has been given, but they can also be run as part of a new drug application. In long term stability studies the expectation is that the drug product remains within specifications for at least 12 months after manufacture. The ICH (2003) suggests that testing is performed every three months during the first year, every six months during the second year, and on an annual basis thereafter. Denote the testing times by \( x_1, \ldots, x_N \) and the corresponding assay results by \( y_1, \ldots, y_N \), where \( N \) is the total number of assays. These assay results may be expressed as a percentage of the claimed label amount and log transformed if necessary (ICH, 2003). While sampling times are not all required to be distinct, assay results are assumed to be independent.

An appropriate statistical model for the degradation relationship may take the form

\[
y_i = x_i^T \beta + e_i,
\]

or in matrix notation

\[
y = X\beta + e,
\]

where

- \( y_i \) is the response from the \( i \)-th assay for \( i = 1, \ldots, N \), on the arithmetic or logarithmic scale, and \( y \) is the column vector of \( y_i \)'s,
The responses $y_i$ will hereinafter be referred to as potency, but they could equally refer to any other measurement.

The current definition of *shelf-life* (ICH, 2003) is

> “the time interval that a drug product is expected to remain within the approved shelf-life specification provided that it is stored under the conditions defined on the label in the proposed containers and closure”.

This definition is equivalent to the mathematical definition for the *true shelf-life* of a single batch (Chow and Shao, 1990; Shao and Chow, 1994), namely the time $\tau$ such that

$$
\tau = \inf \{ t : x_t^T \beta \leq \kappa \},
$$

(1.2)

where $\inf$ denotes the infimum or greatest lower bound, $\kappa$ is the specification limit, or acceptance criterion, and $x_t$ is the vector of regressors at time $t$. Note that a similar definition applies to responses that increase with time.

Different ways of estimating the shelf-life have been proposed. Tatke (1980) used an approach based on survival analysis which relied on a drug product having failed, that is, having a potency less than the specification limit. Such an approach is of little practical value since pharmaceutical companies usually require an estimate of the shelf-life long before any of the products have actually failed. A more useful approach made use of degradation measurements over time and fitted an $L_1$ regression model to this degradation data (Tapon, 1993). This method was chosen as being more robust than least squares regression, which tends to be the method in the literature and regulatory guidelines.
Using least squares regression, the estimated shelf-life is given by the intersection of the mean response and the specification limit, i.e. \( \hat{\tau} \) satisfying
\[
\hat{\tau} = \inf \left\{ t : \mathbf{x}^T \hat{\beta} \leq \kappa \right\},
\]
where \( \hat{\beta} \) denotes the least squares estimate of \( \beta \). The guidelines state that the approach for determining the label shelf-life is

“to determine the time at which the 95% one-sided confidence limit for the mean curve intersects the acceptance criterion”.

This approach attempts to give assurance that the label shelf-life is smaller than the true shelf-life with probability \( 1 - \alpha \). Hence, the label shelf-life, \( \hat{\tau}_{L(\alpha)} \), is a \( 100(1 - \alpha)\% \) lower confidence bound for the true shelf-life and is given by the time \( t \) at which the lower confidence bound for \( \mathbf{x}^T \hat{\beta} \) intersects the specification limit (Shao and Chow, 1994)
\[
\hat{\tau}_{L(\alpha)} = \inf \left\{ t : \mathbf{x}^T \hat{\beta} + t_{\alpha,N-p} s \sqrt{\mathbf{x}^T (X^T X)^{-1} \mathbf{x}} \leq \kappa \right\},
\]
where \( t_{\alpha,N-p} \) is the \( \alpha \) quantile of the \( t \)-distribution with \( N - p \) degrees of freedom, and \( s^2 \) is the least squares estimate of \( \sigma^2 \). This approach is depicted in Figure 1.1.

The current regulatory definition and much of the literature implicitly assume that the variability in the measurements is solely due to the assaying process. It completely ignores the fact that each unit of drug product, a tablet, say, is not exactly identical, but varies about some true potency. In particular, the assay variability is likely to be very small under controlled conditions (Shao and Chow, 2001a). Consequently, variability observed in stability studies is more likely to be due to unit-to-unit variability than due to assay variability. Consequently, the current regulatory approach can be shown to be potentially seriously flawed.

The problem arises in large sample circumstances, where the mean response at a given time, and hence the estimated shelf-life, are determined more accurately. In the limiting case as \( N \to \infty \) the label shelf-life approaches the true shelf-life. This is clearly a desirable large sample property. However, under the assumption of normality, this also implies that in the limit only 50\% of the product remains within specification at time \( \hat{\tau} \).
Figure 1.1: Current regulatory approach to estimating the shelf-life and the label shelf-life using a linear regression model.

This point is illustrated in Figures 1.2 and 1.3; in the latter graph \( r = 25 \) samples are assayed at each time. The distribution of the response \( Y_t \) at time \( t \) is given by \( Y_t \sim N(\mathbf{x}_t^T \beta, \sigma^2) \), and for the moment assume that the mean and variance are known. Consequently, the sample mean at time \( t \) is distributed as \( \bar{Y}_t \sim N(\mathbf{x}_t^T \beta, \sigma^2/r) \). Now, let \( L(\alpha) \) be the \( 100(1 - \alpha)\% \) lower confidence bound for the true mean based on the distribution of \( \bar{Y}_t \), that is, \( L(\alpha) = \mathbf{x}_t^T \beta + z_{\alpha} \sigma/\sqrt{r} \). The proportion of individual responses \( Y_t \) less than \( L(\alpha) \) is then given by

\[
P \left( Z \leq \frac{L(\alpha) - \mathbf{x}_t^T \beta}{\sigma} \right) = P \left( Z \leq \frac{z_{\alpha}}{\sqrt{r}} \right).
\]

This proportion depends on \( \alpha \) and on the sample size \( r \). When \( \alpha = 0.05 \), it increases from 0.05 when \( r = 1 \) to 0.371 when \( r = 25 \), and approaches 0.5 as \( r \to \infty \). Consequently, if the true shelf-life is based on the mean response and many samples are used in the estimation, the percentage of drug product that does not meet the specification at the label shelf-life approaches 50%.

A large proportion of drug product falling below the specification at the label shelf-life may not be a serious problem if the unit-to-unit variability, \( \sigma \), is small, since the specification limit will not have been breached by a large amount. However, if the variability is large,
Figure 1.2: The 95% lower confidence bound for the mean and the 95% lower prediction bound when no replication is used at each time point.

Figure 1.3: The 95% lower confidence bound for the mean and the 95% lower prediction bound when 25 replicates are used at each time point.
then some products may fall far short of the specification. For products that increase in toxicity, say, this relates to a large proportion of product exceeding a safe specification limit. This may have serious consequences — even if the product is used within its shelf-life.

An alternative approach, which overcomes these shortcomings, is based on the prediction interval of a future observation (Carstensen and Nelson, 1976). Carstensen and Nelson define what will be referred to as a modified label shelf-life, \( \hat{\tau}_{L(\alpha)}^* \). This modified label shelf-life is the time at which the \( 100(1 - \alpha)\% \) lower prediction bound intersects the specification limit, that is,

\[
\hat{\tau}_{L(\alpha)}^* = \inf \{ t : \mathbf{x}_t^T \hat{\beta} + t_{\alpha, N-p} s \sqrt{1 + \mathbf{x}_t^T (X^T X)^{-1} \mathbf{x}_t} \leq \kappa \}.
\]

The prediction bound considers the distribution of a single future observation. Hence this approach attempts to make a statement about the ability of an individual item (a tablet, say) to meet the specifications at time \( t \), rather than the ability of the mean of all items of a batch to meet the specifications. This leads to a more conservative, that is, shorter shelf-life.

Comparing the label shelf-life and the modified label shelf-life, in Figure 1.2, with \( r = 1 \), the label shelf-life would be set at about 20 months, yet the same product in Figure 1.3, with \( r = 25 \), would give a label shelf-life of about 24 months. However, using (1.4) would lead to a modified label shelf-life of about 14 months in each case and would guarantee that 95% of the product exceed the specifications at that time. Which label shelf-life is used is clearly of importance to the consumer.

Another undesirable property of the current regulatory approach is the lack of influence the product variability has on the definition of shelf-life and consequently on the estimation process. In particular, since primary interest lies in the mean response, increasing the number of samples can compensate for a manufacturing process with large unit-to-unit variability \( \sigma^2 \). This is often cheaper than attempting to improve the manufacturing process, but it does not put enough emphasis on good manufacturing practices and commitment to quality.
The modified label shelf-life, on the other hand, is not consistent with the current definition of the true shelf-life (Chow and Shao, 1990) in large samples. Instead it has the large sample property which ensures that a proportion of \(1 - \alpha\) of the drug product falls above the specification limit at the time the modified label shelf-life has expired. This follows from the fact that as \(N \to \infty\), equation (1.4) reduces to

\[
\hat{\tau}_\alpha^* = \inf \{ t : \mathbf{x}_t^T \hat{\beta} + z_\alpha \sigma \leq \kappa \}, \tag{1.5}
\]

and the left hand side of the condition which needs to be satisfied is the \(\alpha\) quantile of the distribution of \(Y\) given \(\mathbf{x}_t\). Consequently, this quantile may be a more appropriate basis for the true shelf-life.

Finally, the regulatory guidelines recommend that factors which may influence the stability of the product be included in a stability study and that the product be assayed every three months during the first year, every six months during the second year and yearly thereafter. Bracketing designs and fractional factorial or matrix designs are recommended in the guidelines. The latter have also been endorsed by the literature due to their ability to reduce the total number of assays that need to be tested (Nordbrock, 1992; Stead et al., 1994; Nordbrock, 1994; Barron, 1994; Ju and Chow, 1995; Chow and Liu, 1995b). The choice of sampling times, however, has not received much attention. Only Chow and Pong (1995a) comment that the sampling times are not adequate when the degradation relationship is not linear.

The sampling times recommended by the ICH have most likely been chosen to enable estimation of a label shelf-life to be carried out after the first year. However, for long term stability studies, observations collected late in the study have a large influence on the estimates of the regression line. They are frequently termed influential points. Hence, errors in the measurement process can greatly change the label shelf-life, as can misspecification of the model, e.g. using a linear model when an exponential decay may be more appropriate.

In addition, the sampling effort should be directed near the time when the product is expected to expire. This will result in a more accurate estimate of the shelf-life and thereby give a longer label shelf-life. To illustrate this point, consider two stability studies. Both
1.2 Testing for Batch-to-Batch Variability

Pharmaceutical products are generally manufactured in batches. Ideally, these batches are identical, but in practice, variation from batch to batch will be unavoidable. Different
raw materials for example may be an important source of this variation.

Due to the time delay that is involved in the determination of a label shelf-life it is not practical to wait until a label shelf-life is obtained for a batch before it is released onto the market. In fact, as the ICH (2002) states:

> “Where applicable, an appropriate statistical method should be employed to analyze the long-term primary stability data in an original application. The purpose of this analysis [of stability data] is to establish, with a high degree of confidence, a retest period or shelf life during which a quantitative attribute will remain within acceptance criteria for all future batches manufactured, packaged, and stored under similar circumstances. This same method could also be applied to commitment batches to verify or extend the originally approved retest period or shelf life.”

The guidelines indicate that a minimum of three primary batches is required for stability studies. The guidelines (ICH, 2003; , 2002) indicate that if batch-to-batch variability is small then three batches can be analysed as one. This is established by testing for equality of slopes and intercepts, using a significance level of 0.25. However, if significance tests show that batches cannot be combined, then the label shelf-life should be based on the minimum of the three label shelf-lives obtained from analysing each batch separately.

The choice of significance level of 0.25 is unusual, and has likely been chosen in order to increase the power of the test (Ruberg and Stegeman, 1991). However, Ruberg and Hsu (1990) correctly point out that “a penalty is paid for doing a good study, whereas a poor study allows one to pool more batches into the shelf-life calculation”. Consequently, they propose an approach that is analogous to one used for establishing bioequivalence (Ruberg and Hsu, 1990; , 1992). This proposed test uses a “multiple comparison procedure with the worst” as a decision rule for combining batches. An alternative procedure, which is based on fixing the power of the test in advance, was proposed by Ruberg and Stegeman (1991). Another approach is described by Yoshioka et al. (2002), who calculate a shelf-life for each batch and then use the range of shelf-lives to decide
1.3 Label Shelf-life Estimation for Multiple Batches

In the case that batch variability is shown to be significant, various approaches have been suggested for the estimation of a label shelf-life. These approaches are based on either a fixed or random effects model. The current regulatory guidelines only consider the fixed effects model (ICH, 2002), which cannot be used to make statements about future batches. The random effects model, which is a more realistic model, assumes that batches are random, with some underlying distribution. Based on the random effects model, probability statements about all future batches can then be made.
Both these models are now considered in turn.

### 1.3.1 The Fixed Effects Model: Current Regulatory Guidelines

In the situation when there is significant batch-to-batch variability, Shao and Chow (1994) pointed out that the “minimum of three” approach lacks statistical justification. Since batches are treated as fixed effects, conclusions drawn from the batches used in the stability trial cannot be generalised to all future batches. This can only be achieved by treating batches as random effects. At the same time, using more batches in the analysis can result in the smallest estimate of shelf-life (obtained separately for each batch) becoming smaller, possibly to some limit.

The statement about the lack of statistical justification is examined here with the help of a simulation study. The simulations were structured as those in Section 3.3 of Chapter 3. A summary of the structure is as follows.

- The number of batches $k$ takes values 3, 4, 5, 6, 7, 8, 9, 10, 20 and 50.

- A random sample of $k$ batch effects $b_j$ were drawn from $N(\beta, D)$, where $\beta = (100, -0.3)^T$, $D$ is a $2 \times 2$ matrix with diagonal elements equal to 1 and 0.003, and off-diagonal elements equal to $d_2 = -0.03$, 0 or 0.03. This allows for negative, no and positive correlation between the intercepts and slopes.

- A random sample of size $r = 1$ was drawn for each batch at times $x_i = 0, 3, 6, 9, 12, 18, 24$ and 36 from the distribution $N(x_i^T b_j, \sigma^2)$, where $x_i^T = (1, x_i)$ and $\sigma^2 = 0.5^2$.

- This was repeated 1000 times.

The restrictions on $r$ and $\sigma^2$ are made to better suit the approach taken in the regulatory guidelines. Note that under the current definition, the true shelf-life is determined by the time at which the mean degradation curve intersects the specification limit. In these simulations the true shelf-life equals $\tau_{0.5} = (90 - 100)/(-0.3) = 33.3$ months.
A summary of the simulations is given in Tables 1.1–1.3, which present the coverage probability, the bias and the mean squared error (MSE) of the regulatory approach under the simulation conditions. The coverage probability is calculated as the proportion of label shelf-lives that are less than the true shelf-life. The bias is obtained by taking the difference between the average label shelf-life and the true shelf-life, while the MSE is given by

\[ \text{MSE} = \text{Var [Label Shelf-life]} + \text{Bias}^2. \]

It should be noted that the bias is not a “true bias” in the strict sense as it is calculated using a lower confidence bound instead of an estimator. However, it does indicate how on average a label shelf-life falls short of the true shelf-life.

Table 1.1 shows that the coverage probability quickly approaches 1 as the number of batches increases. This is expected as the label shelf-life decreases as the number of batches increases. This is confirmed by the estimates of bias in Table 1.2. From this table it can be seen that the bias increases as the number of batches increases. This clearly suggests that a longer label shelf-life can be obtained by running a stability study with as few batches as possible — a minimum of three. Table 1.2 also shows that the bias is worse when the batch intercept and slope are positively correlated than when they are negatively correlated.

This same observation also applies to the Mean Squared Error (MSE). From Table 1.3 it can be seen that the MSE also increases as the number of batches increases. While this is partially due to the increase in the bias, the variance of the label shelf-life also increases. Consequently, it can be concluded that the minimum shelf-life approach, as advocated by the regulatory guidelines, does not have desirable large sample properties, and should therefore not be used.
1.3. LABEL SHELF-LIFE ESTIMATION FOR MULTIPLE BATCHES

Table 1.1: Coverage probabilities of the current regulatory approach as evaluated using a simulation study with $\beta_1 = -0.3$ and $\sigma = 0.5$.

<table>
<thead>
<tr>
<th>$d_2$</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>20</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.03</td>
<td>0.964</td>
<td>0.982</td>
<td>0.997</td>
<td>0.998</td>
<td>0.999</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0.959</td>
<td>0.978</td>
<td>0.992</td>
<td>0.997</td>
<td>0.997</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>0.03</td>
<td>0.936</td>
<td>0.979</td>
<td>0.992</td>
<td>0.997</td>
<td>1.000</td>
<td>0.998</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td></td>
</tr>
</tbody>
</table>

Table 1.2: Bias of the current regulatory approach as evaluated using a simulation study with $\beta_1 = -0.3$ and $\sigma = 0.5$.

<table>
<thead>
<tr>
<th>$d_2$</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>20</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.03</td>
<td>-5.10</td>
<td>-6.13</td>
<td>-6.77</td>
<td>-6.88</td>
<td>-7.19</td>
<td>-7.41</td>
<td>-7.65</td>
<td>-7.79</td>
<td>-8.89</td>
<td>-9.91</td>
</tr>
<tr>
<td>0.03</td>
<td>-7.12</td>
<td>-8.23</td>
<td>-8.97</td>
<td>-10.00</td>
<td>-10.14</td>
<td>-10.58</td>
<td>-11.04</td>
<td>-11.08</td>
<td>-12.91</td>
<td>-14.61</td>
</tr>
</tbody>
</table>

Table 1.3: Mean Squared Error of the current regulatory approach as evaluated using a simulation study with $\beta_1 = -0.3$ and $\sigma = 0.5$.

<table>
<thead>
<tr>
<th>$d_2$</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>20</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.03</td>
<td>34.55</td>
<td>43.58</td>
<td>51.11</td>
<td>51.97</td>
<td>55.97</td>
<td>58.84</td>
<td>62.40</td>
<td>64.09</td>
<td>81.67</td>
<td>99.99</td>
</tr>
<tr>
<td>0</td>
<td>56.34</td>
<td>66.95</td>
<td>74.58</td>
<td>82.00</td>
<td>85.81</td>
<td>92.15</td>
<td>97.38</td>
<td>99.84</td>
<td>129.92</td>
<td>163.88</td>
</tr>
<tr>
<td>0.03</td>
<td>71.18</td>
<td>84.89</td>
<td>94.47</td>
<td>111.73</td>
<td>112.69</td>
<td>122.00</td>
<td>130.79</td>
<td>130.66</td>
<td>172.28</td>
<td>217.25</td>
</tr>
</tbody>
</table>
1.3.2 The Random Effects Model

Random effects models have been considered by several authors in one form or another. A general form which is frequently used is

\[ Y_{ij} | b_j = x_{ij}^T b_j + e_{ij}, \]  

where

- \( Y_{ij} \) denotes the \( i \)-th assay result from the \( j \)-th batch for \( i = 1, \ldots, N_j \) and \( j = 1, \ldots, k \), on the arithmetic or logarithmic scale, such that the total number of observations is \( N = \sum_{j=1}^k N_j \),
- \( x_{ij} \) is a \( p \)-dimensional predictor vector corresponding to the \( i \)-th observation from the \( j \)-th batch,
- \( b_j \) is a \( p \)-dimensional vector of random effects for the \( j \)-th batch, distributed as \( b_j \sim N(\beta, D) \), and
- \( e_{ij} \) are i.i.d. errors independent of \( b_j \) such that \( e_j = (e_{1j}, \ldots, e_{N_j})^T \sim N(0, \sigma^2 I_j) \), where \( I_j \) is the identity of size \( N_j \).

The distribution of \( Y_j = (Y_{1j}, \ldots, Y_{N_j})^T \), conditional on \( b_j \), is then given by

\[ Y_j | b_j \sim N(X_j b_j, \sigma^2 I_j) \]

where \( X_j \) is the \( N_j \times p \) model matrix of full rank for batch \( j \), with \( x_{ij} \) in the \( i \)-th row. Unconditionally,

\[ Y_j \sim N(X_j \beta, \Sigma_j) , \]

where \( \Sigma_j = \sigma^2 I_j + X_j D X_j^T \). The joint unconditional distribution of all observations over all batches is now given by

\[ Y \sim N(X\beta, \Sigma) , \]  

where \( Y = (Y_1^T \ldots Y_k^T)^T \), \( X = [X_1^T \ldots X_k^T]^T \) and \( \Sigma = \text{diag}(\Sigma_1, \ldots, \Sigma_k) \). As usual, the realization of \( Y_{ij} \) is denoted by \( y_{ij} \), such that \( Y_j \) and \( Y \) become \( y_j \) and \( y \), respectively.
It will be assumed hereafter that batches degrade linearly over time, in which case \( \mathbf{\beta} = (\beta_0, \beta_1)^T \), and

\[
\mathbf{D} = \begin{bmatrix}
    d_1 & d_2 \\
    d_2 & d_3
\end{bmatrix}.
\]

This variance matrix for the \( \mathbf{b}_i \) describes the form of the underlying response from one batch to another. The value of \( d_1 \), the variance of \( b_{0j} \), describes how the average activity varies from batch to batch at time \( t = 0 \), the time of manufacture; the value of \( d_3 \), the variance of \( b_{1j} \), describes how the rate of degradation varies from one batch to another; and the value of the covariance \( d_2 \) indicates how \( b_{0j} \) affects \( b_{1j} \). That is, batches that start with more active ingredient degrade faster if \( d_2 < 0 \) and degrade slower when \( d_2 > 0 \). Consequently, positive values of \( d_2 \) have the effect of increasing the variability of the response at time \( t \), where the variance of the response due to batch variation is \( d_1 + 2d_2t + d_3t^2 \). This is likely to lead to shorter label shelf-lives due to the greater uncertainty.

Norwood (1986) considered a model with a random intercept, that is, \( \mathbf{b}_j = b_{0j} \) and \( b_{0j} \sim N(\beta_0, d_1) \), because the modified shelf-life (Carstensen and Nelson, 1976) does not hold for models which included batch effects. Hence the unconditional distribution of \( \mathbf{Y}_j \) is \( N(\mathbf{X}_j\mathbf{\beta}, \Sigma_j) \), where \( \Sigma_j = \sigma^2\mathbf{I} + d_1\mathbf{1}\mathbf{1}^T \). He used weighted least squares to find estimates for the fixed effects and calculated estimates for the residual and random effect variances from the residual sum of squares. These estimates were used to calculate a confidence bound for the intercept which in turn was used to obtain the label shelf-life.

A more general model with independent random effects \( (d_2 = 0) \) is proposed by Chow and Shao (1991), along with a way to estimate the random effects variances \( d_1 \) and \( d_3 \). A way for obtaining a lower confidence bound for the average shelf-life based on the coefficient of variation being smaller than some constant \( \rho \), say, is also presented by them. It should however be noted that a model with \( d_2 = 0 \) is not invariant under transformation.

Model (1.7) is used by Shao and Chow (1994) to define the true shelf-life of the \( j \)-th batch as

\[
\tau_j = \inf \left\{ t : \mathbf{x}_t^T\mathbf{b}_j \leq \kappa \right\},
\]

(1.8)
which reduces to the current regulatory definition (1.2) of shelf-life for a single batch when $D = 0$. Since batches are random, the true shelf-life of the $j$-th batch will also be random with some (non-symmetric) distribution. Shao and Chow suggest that the true shelf-life for all batches should be the median of all shelf-lives for individual batches. This equals the time at which the average degradation curve over all batches intersects the specification limit.

Shao and Chow then based the label shelf-life on the time at which no more than a proportion $\epsilon$ of the mean degradation curves for individual batches fall below the specification limit. This is done by finding a $100(1 - \alpha)\%$ lower confidence bound for

$$
\tau_\epsilon = \inf \left\{ t : \textbf{x}_T^T \beta + z_\epsilon \sigma(t) \leq \kappa \right\},
$$

(1.9)

where $\sigma(t) = \sqrt{\textbf{x}_T^T D \textbf{x}_t}$. In the balanced case, the label shelf-life can then be shown to equal

$$
\tilde{\tau}_\epsilon = \inf \left\{ t : \textbf{x}_T^T \bar{\textbf{b}} - Q_{1-\alpha} \sqrt{v(t) / k} \leq \eta \right\},
$$

(1.10)

where $Q_{1-\alpha} = t'_{1-\alpha,k-1}(-z_\epsilon \sqrt{k})$, the $1 - \alpha$ quantile of the non-central t-distribution with non-centrality parameter $\delta = -z_\epsilon \sqrt{k}$; $v(t) = \textbf{x}_T^T \hat{\Omega} \textbf{x}_t$ is the variance of $\textbf{x}_T^T \bar{\textbf{b}}$; and $\bar{\textbf{b}}$ and $\hat{\Omega}$ are the estimates of the mean vector $\beta$ and variance matrix $\Omega = D + \sigma^2 (X^T X)^{-1}$. These are obtained by treating the $\textbf{b}_j$ as fixed effects and calculating the mean vector and variance matrix of their least squares estimates (c.f. Chapter 3). A similar definition applies to the unbalanced case.

Note that this definition is based on the distribution of $\textbf{x}_T^T \hat{\textbf{b}}_j$ and as such concentrates on whether the average of a batch can meet the specifications. Consequently, the issues related to the product variability within a batch, which have been identified in Section 1.1, still apply here.

Ho et al. (1992) compared this method with the method suggested by the regulatory guidelines and the method suggested by Ruberg and Hsu (1990). They found that no one method performs equally well in all circumstances. Sun et al. (1999) show that the distribution of the label shelf-life is approximately normal as the number of batches tends to infinity. This is hardly surprising when considering the central limit theorem. An
alternative way of estimating the distribution of the mean batch shelf-life has also been suggested (Lu and Meeker, 1993). This approach makes fewer distributional assumptions and uses bootstrapping to obtain the distribution function of the mean batch shelf-life as well as the relevant confidence bound.

Regardless of the method used to obtain a label shelf-life, the covariance matrix for the random effects needs to be estimated. Several methods for estimating $D$ have been suggested. Chow and Wang (1994) used an unrealistic random effects model where $D = \sigma^2 \beta I_2$. An alternative approach for estimating $D$ was using the EM algorithm (Chen et al., 1995).

To summarise, in the single batch case, all approaches, except for the modified label shelf-life approach, implicitly assumed that the residual variability is purely due to assay variability. This has already been shown to be deficient in Section 1.1. In addition, in the batch case, all approaches, except for those by Norwood (1986) and Lu and Meeker (1993), rely on a test for establishing significant batch variability, even though Shao and Chow’s approach and definition reduces to their definitions in the single batch case as $D \to 0$.

## 1.4 Addressing the Issues

As discussed in the previous sections, there are areas in the analysis of pharmaceutical stability trials which require attention. In particular, there is a need for definitions of shelf-life and label shelf-life which take into account unit-to-unit as well as batch-to-batch variability. Approaches for estimating these shelf-lives are then needed, and these should not be reliant on tests of significance. A definition of shelf-life for single batch data and methods for estimating a label shelf-life are proposed in Chapter 2. Chapter 3 generalises the definition and methods to the case when a label shelf-life needs to be obtained from several batches.

Another question that has not been answered in the literature is what to do when the errors $e_i$ of the degradation model are not normally distributed or do not have constant variance. These questions can be investigated with the help of generalised linear model
assumptions, which is done in Chapter 4.

While the framework of stability studies will be used to motivate the theoretical developments, it should be noted that this work is not limited to pharmaceutical stability studies. In fact, these methods can be applied to many cases where degradation, or accumulation, to some limit is of interest; examples of situations like these are given in Lu and Meeker (1993).
CHAPTER 2

Estimation of Shelf-Life:
The Single Batch Case

The analysis of stability data from a single batch of drug product has been outlined in the regulatory guidelines (ICH, 2002; , 2003) and by various authors. However, as was discussed in Chapter 1, there are some potential problems associated with these regulatory guidelines. A possible solution, in terms of a new definition of shelf-life, is given in this Chapter. Several methods for the estimation of a label shelf-life are proposed and evaluated using a simulation study.

2.1 A New Definition

In Chapter 1 it was shown that when many samples are taken in a stability study the current definition of shelf-life, as defined by (1.2), can result in up to 50% of drug product falling below the specification limit by the time the expiry date has been reached. This is because the current definition is based on the mean response over time and the assumption that assay variability is the only important source of variation.

However, variation in the measured potency of a tablet can arise from at least two sources
in the single batch case. The first is due to measurement error in the assaying process. The second is the unit-to-unit variability, as exemplified by the variation in the level of active ingredient from one tablet to another, say. This second source is due to the manufacturing process. Clearly, this source of variability should be taken into account in the definition of shelf-life as it is inherent in the product. Assay variability on the other hand should not feature in the definition as it is a property of the measurement process.

Following the recommendations of the current guidelines will result in only a single estimate of variability, which is a combination of product and assay variation. The latter is a nuisance source of variability and needs to be eliminated from the data. This can be done, for example, by choosing a more accurate assaying technique or apparatus. Alternatively, a potentially less costly approach is to assay all tablets in duplicates or triplicates. The mean assay result can then be used as an estimate of the true tablet potency, leaving only product variability in the data. Consequently, it will be assumed from now on that the only source of variability is this unit-to-unit variability, \( \sigma^2 \).

As mentioned above, the problem with the current regulatory approach is that it ignores the product variability and consequently the definition of the label shelf-life is based on the 95% confidence bound for the mean degradation curve. In Chapter 1, an approach presented by Carstensen and Nelson (1976) was discussed. This approach incorporates product variability and leads to the modified label shelf-life. This approach is based on the 95% prediction bound for a future observation. However, while the estimation of the modified label shelf-life takes into account product variability, it does not reduce to the regulatory definition of shelf-life in large sample situations. Consequently, a more consistent approach is to base the definition of shelf-life on (1.5), the asymptotic equivalent of (1.4), and the label shelf-life on the confidence bound for (1.5).

The following new definitions of true shelf-life, estimated shelf-life and label shelf-life are proposed in order to overcome the problems of the current regulatory label shelf-life and the modified label shelf-life.

**Definition 1** Under model (1.1), the true shelf-life, denoted by \( \tau_e \), is the minimum time
at which 100\% of the drug product from the current batch has an activity level that is less than or equal to the pre-determined specification limit, \( \kappa \), that is,

\[
\tau_\epsilon = \inf \{ t : x_t^T \beta + z_\epsilon \sigma \leq \kappa \},
\]

where \( x_t^T \beta \) is the mean activity level at time \( t \) and \( z_\epsilon \) is the \( \epsilon \) quantile of the standard normal distribution.

Note that \( \tau_{0.5} \) represents the true shelf-life currently used by regulatory bodies; this can also be obtained when there is no variability in the drug product, that is, when \( \sigma = 0 \).

The estimated shelf-life, \( \hat{\tau}_\epsilon \), is then defined as the time at which the estimated \( \epsilon \) quantile of \( Y \) intersects the specification limit.

**Definition 2** Based on Definition 1, the label shelf-life, denoted by \( \hat{\tau}_{\epsilon, L(\alpha)} \), is the \( 100(1 - \alpha)\% \) lower confidence bound for the true shelf-life, \( \tau_\epsilon \).

The new definitions are presented graphically in Figure 2.1.

![Figure 2.1](image_url)  

**Figure 2.1:** Estimating the shelf-life using the proposed definition with \( \epsilon = 0.05 \).
Irrespective of the method of estimation that is used, several general points can be made about these definitions.

1. With $100(1 - \alpha)\%$ confidence, at least $100(1 - \epsilon)\%$ of drug product will meet the specification when the label shelf-life has been reached. Hence, the label shelf-life is a $100(1 - \alpha)\%$ lower tolerance bound for the $100\epsilon$ percentile of the distribution of shelf-lives.

2. The true shelf-life can be estimated more accurately by increasing sample size.

3. The true and label shelf-life can be extended by reducing the variability in the manufacturing process.

Several possibilities for obtaining a label shelf-life are available. These are detailed in the next section for linear degradation — some can easily be extended to cover multiple regression models. Unless otherwise indicated, the terms true shelf-life, estimated shelf-life and label shelf-life will henceforth refer to the new definitions.

### 2.2 Estimation of the Label Shelf-Life

Suppose the mean response is $\mu_t = x_t^T \beta = \beta_0 + \beta_1 t$, where $t$ denotes time. Interest lies in the $\epsilon$ quantile of the response, given $t$, which under (1.1) is given by

$$q_{t,\epsilon} = x_t^T \beta + z_\epsilon \sigma,$$

assuming that the variability does not depend on $t$. If there is a mean-variance relationship, the methods in Chapter 4 may be more appropriate.

There are at least three possible approaches that can be used to estimate and find confidence bounds for the true shelf-life as defined in Definition 1. These are based on

1. several *Normal Approximation* approaches,
2.2. ESTIMATION OF THE LABEL SHELF-LIFE

2. a variety of Profile Likelihood approaches, and

3. a Non-central t-distribution approach.

These approaches are discussed in the following sections. This work is presented in Kiermeier et al. (2004).

2.2.1 Normal Approximation

There are various ways of estimating (2.1). Easterling (1969) suggested that the general form of a point estimate for the shelf-life can be based on

$$\hat{q}_{t, \epsilon} = \hat{\beta}_0 + \hat{\beta}_1 t + \delta s,$$

where $\hat{\beta}_0$, $\hat{\beta}_1$ and $s^2$ denote the least squares estimates of $\beta_0$, $\beta_1$ and $\sigma^2$, and various choices for $\delta$ are available. Easterling’s choice for an exact interval estimate is presented in Section 2.2.4.

It is well known that

$$\frac{\nu s^2}{\sigma^2} \sim \chi^2_{\nu}$$

where $\chi^2_{\nu}$ denotes a chi-squared distribution with $\nu$ degrees of freedom (here $\nu = N - 2$), and thus

$$\frac{\sqrt{\nu} s}{\sigma} \sim \chi_{\nu},$$

where $\chi_{\nu}$ is a chi distribution with $\nu$ degrees of freedom. The moments of a $\chi_{\nu}$ variate are given on page 421 in Johnson et al. (1994a). It can be shown that the mean and variance of $s$ are given by

$$\text{E} [s] = \sigma C_{\nu} \quad \text{and} \quad \text{Var} [s] = \sigma^2 \left(1 - C_{\nu}^2\right),$$

(2.2)

where $\nu = N - 2$ in the linear case, and

$$C_{\nu} = \sqrt{\frac{2}{\nu}} \frac{\Gamma \left(\frac{\nu+1}{2}\right)}{\Gamma \left(\frac{\nu}{2}\right)} \approx 1 - \frac{1}{4\nu}.$$
Since the sample mean and the sample standard deviation are independent, it follows that

\[
\begin{align*}
E[\hat{q}_{t,\epsilon}] &= x_t^T \beta + \delta \sigma C_\nu \\
\text{Var}[\hat{q}_{t,\epsilon}] &= \sigma^2 \left[ x_t^T (X^T X)^{-1} x_t + \delta^2 \left( 1 - C_\nu^2 \right) \right],
\end{align*}
\]

and for large sample sizes \( \hat{q}_\tau \) will be close to normally distributed with this mean and standard deviation.

The performance of the normal approximation for \( \hat{q}_{t,\epsilon} \) now relies on (a) the choice of \( \delta \), and (b) the closeness to normality of \( s \). By the Central Limit Theorem it is reasonable to expect the distribution of \( \hat{q}_{t,\epsilon} \) to be approximately normal for a given value of \( t \) and choice of \( \delta \) as the sample size \( N \to \infty \).

Possible choices for \( \delta \), and their resulting value, are

\[
\begin{align*}
\delta &= z_\epsilon \\
\delta s &= \text{the maximum likelihood estimate for } z_\epsilon \sigma : \delta = z_\epsilon \sqrt{\frac{\nu}{N}}, \\
\tilde{\delta} s &= \text{the unbiased estimate for } z_\epsilon \sigma : \delta = \frac{z_\epsilon}{C_\nu}.
\end{align*}
\]

An estimate of the variance of \( \hat{q}_{t,\epsilon} \) can be obtained by either substituting the least squares estimate for \( \sigma \), if using (2.3) or (2.5), or the maximum likelihood estimate for \( \sigma \), if using (2.4). The estimated shelf-life is given by

\[
\hat{\tau}_\epsilon = \inf \{ t : \hat{q}_{t,\epsilon} \leq \kappa \},
\]

and the label shelf-life is the 100(1 - \( \alpha \))% lower confidence bound for the time at which \( q_{t,\epsilon} \) intersects the specification limit, that is,

\[
\hat{\tau}_{\epsilon,L(\alpha)} = \inf \left\{ t : \hat{q}_{t,\epsilon} + Q_\alpha \hat{\sigma} \sqrt{\frac{x_t^T (X^T X)^{-1} x_t + \delta^2 \left( 1 - C_\nu^2 \right)}{N}} \leq \kappa \right\},
\]

where \( \hat{\sigma} = s \) for (2.3) and (2.5), and \( \hat{\sigma} = s \sqrt{\frac{\nu}{N}} \) for (2.4). Subsequent simulations will use both \( z_\alpha \) and \( t_{\alpha,\nu} \) as possible choices for \( Q_\alpha \).

In the linear regression case it can be shown that \( \hat{\tau}_{\epsilon,L(\alpha)} \) is the smaller positive root of \( t \).
for which
\[
\left\{ \hat{\beta}_1^2 - \frac{Q_2^2 \hat{\sigma}_s^2}{s_{xx}} \right\} t^2 + 2 \left\{ \hat{\beta}_1 \left( \hat{\beta}_0 - \kappa + \delta s \right) + \frac{Q_2^2 \hat{\sigma}_s^2}{s_{xx}} \right\} t \\
+ \left\{ \left( \hat{\beta}_0 - \kappa + \delta s \right)^2 - Q^2_2 \hat{\sigma}_s^2 \left[ \frac{1}{N} + \delta^2 (1 - C^2_{N-2}) + \frac{\hat{\sigma}_s^2}{s_{xx}} \right] \right\} = 0 , \quad (2.6)
\]
where \( s_{xx} = \sum_{i=1}^{N} (x_i - \bar{x})^2 \).

### 2.2.2 Various Profile Likelihood Based Approaches

The label shelf-life \( \hat{\tau}_{\epsilon,L(\alpha)} \) is the smallest time for which the null hypothesis
\[
H_0 : \beta_0 + \beta_1 \tau_\epsilon + z_\epsilon \sigma = \kappa \quad (2.7)
\]
is not rejected on a one-sided 100\(\alpha\)% level test. Since \( \kappa \) and \( z_\epsilon \) are known, the likelihood can be reformulated as a function of the parameters \( \{ \beta_1, \sigma^2, \tau_\epsilon \} \), by substituting \( \beta_0 = \kappa - \beta_1 \tau_\epsilon - z_\epsilon \sigma \). Hypotheses about \( \tau_\epsilon \) can then be tested using the profile likelihood for \( \tau_\epsilon \), denoted by
\[
\tilde{L}(\tau_\epsilon) = \tilde{L}(\tau_\epsilon, \hat{\beta}_{1,\tau_\epsilon}, \hat{\sigma}^2_{\tau_\epsilon}; y) 
\]
where \( \hat{\beta}_{1,\tau_\epsilon} \) and \( \hat{\sigma}^2_{\tau_\epsilon} \) are the maximum likelihood estimates of \( \beta_{1,\tau_\epsilon} \) and \( \sigma^2_{\tau_\epsilon} \) for given values of \( \tau_\epsilon \). This likelihood can be used like a genuine likelihood, and thus, the maximum likelihood estimate of \( \tau_\epsilon \) is equal to the overall maximum likelihood estimate \( \hat{\tau}_\epsilon \), and confidence intervals for \( \tau_\epsilon \) can be constructed via the likelihood ratio test (Hinkley et al., 1991).

### The Basic Profile Likelihood (PL) Approach

The log-likelihood based on (1.1) is given by
\[
L(\beta, \sigma^2; y) = -\frac{N}{2} \log 2\pi - \frac{N}{2} \log \sigma^2 - \frac{1}{2 \sigma^2} (y - X\beta)^T (y - X\beta) ,
\]
which in the linear case expands to
\[
L(\beta_0, \beta_1, \sigma^2; y) = -\frac{N}{2} \log 2\pi - \frac{N}{2} \log \sigma^2 - \frac{1}{2 \sigma^2} \sum_{i=1}^{N} (y_i - \beta_0 - \beta_1 x_i)^2 . \quad (2.8)
\]
Eliminating $\beta_0$ and ignoring the constant term gives the log-likelihood involving $\tau_\epsilon$

$$L(\tau_\epsilon, \beta_1, \sigma^2; y) = -\frac{N}{2} \log \sigma^2_{\tau_\epsilon} - \frac{1}{2\sigma^2_{\tau_\epsilon}} \sum_{i=1}^{N} \left( y_i - \beta_{1,\tau}(x_i - \tau_\epsilon) - \kappa + z_i \sigma_{\tau_\epsilon} \right)^2.$$ 

For notational simplicity, let $y_i^* = y_i - \kappa$ and $x_i^* = x_i - \tau_\epsilon$, giving

$$L(\tau_\epsilon, \beta_1, \sigma^2; y) = -\frac{N}{2} \log \sigma^2_{\tau_\epsilon} - \frac{1}{2\sigma^2_{\tau_\epsilon}} \sum_{i=1}^{N} (y_i^* - \beta_{1,\tau} x_i^* + z_i \sigma_{\tau_\epsilon})^2. \quad (2.9)$$

Differentiating with respect to $\beta_{1,\tau}$ and equating to zero gives

$$\frac{\partial L}{\partial \beta_{1,\tau}} = \frac{1}{\sigma^2_{\tau_\epsilon}} \sum_{i=1}^{N} (y_i^* - \beta_{1,\tau} x_i^* + z_i \sigma_{\tau_\epsilon}) x_i^* = 0,$$ 

which has the solution

$$\hat{\beta}_{1,\tau} = \frac{N z_i \sigma_{\tau_\epsilon} \bar{x}^* + \sum_{i=1}^{N} x_i^* y_i^*}{s_{xx} + N(\bar{x}^*)^2}. \quad (2.11)$$

Similarly, differentiating with respect to $\sigma^2_{\tau_\epsilon}$ and setting equal to zero yields

$$\frac{\partial L}{\partial \sigma^2_{\tau_\epsilon}} = -\frac{N}{2\sigma^2_{\tau_\epsilon}} + \frac{1}{2\sigma^4_{\tau_\epsilon}} \sum_{i=1}^{N} (y_i^* - \beta_{1,\tau} x_i^* + z_i \sigma_{\tau_\epsilon})^2 - \frac{z_i}{2\sigma^3_{\tau_\epsilon}} \sum_{i=1}^{N} (y_i^* - \beta_{1,\tau} x_i^* + z_i \sigma_{\tau_\epsilon}) = 0. \quad (2.12)$$

Multiplying by $-2\sigma^4_{\tau_\epsilon}$ yields

$$N \sigma^2_{\tau_\epsilon} - \sum_{i=1}^{N} (y_i^* - \beta_{1,\tau} x_i^* + z_i \sigma_{\tau_\epsilon})^2 + z_i \sigma_{\tau_\epsilon} \sum_{i=1}^{N} (y_i^* - \beta_{1,\tau} x_i^* + z_i \sigma_{\tau_\epsilon}) = 0,$$ 

which reduces to

$$N \sigma^2_{\tau_\epsilon} - N z_i (\bar{y}^* - \beta_{1,\tau} \bar{x}^*) \sigma_{\tau_\epsilon} - \sum_{i=1}^{N} (y_i^* - \beta_{1,\tau} x_i^*)^2 = 0, \quad (2.13)$$

where $\bar{y}^* = \frac{1}{N} \sum_{i=1}^{N} y_i^* = \bar{y} - \kappa$ and $\bar{x}^* = \frac{1}{N} \sum_{i=1}^{N} x_i^* = \bar{x} - \tau_\epsilon$.

Notice that the maximum likelihood estimates for $\beta_{1,\tau}$ and $\sigma_{\tau_\epsilon}$ rely on each other, which is different from the unconstrained linear regression case, where $\beta_1$ can generally be estimated independently of $\sigma$. Consequently, a solution could be found by employing some iterative scheme like Fisher scoring.

In this situation it is however possible to find an analytical solution. Write (2.11) as

$$\hat{\beta}_{1,\tau} = \frac{z_i \bar{x}^*}{s_{xx} V} \sigma_{\tau_\epsilon} + \left\{ \frac{\hat{\beta}_1}{NV} + \frac{\bar{x}^* \bar{y}^*}{s_{xx} V} \right\},$$ 

$$= A \sigma_{\tau_\epsilon} + B,$$ 

27
where $\hat{\beta}_1$ is the usual maximum likelihood estimate of $\beta_1$ that does not depend on $\tau$, obtained from (2.8), and

$$V = \frac{1}{N} + \frac{(\bar{x}^\star)^2}{s_{xx}}.$$ 

Note that (2.14) can also be written as

$$\hat{\beta}_{1,\tau} = \hat{\beta}_1 + \frac{\bar{x}^\star}{s_{xx}V} \left( \bar{y}^\star - \hat{\beta}_1 \bar{x}^\star + z_i \hat{\sigma}_\tau \right),$$

which better illustrates the modification of $\hat{\beta}_1$. Substituting (2.14) into (2.13) gives

$$N\sigma^2_{\tau} - Nz_{\epsilon}(\bar{y}^\star - [A\sigma_{\tau} + B]\bar{x}^\star)\sigma_{\tau} - \sum_{i=1}^N (y_i^\star - [A\sigma_{\tau} + B]x_i^\star)^2 = 0.$$ 

Expanding the square and collecting terms results in

$$N\sigma^2_{\tau} - Nz_{\epsilon}(\bar{y}^\star - B\bar{x}^\star)\sigma_{\tau} - \sum_{i=1}^N (y_i^\star - Bx_i^\star)^2 = 0. \quad (2.15)$$

The coefficient of $\sigma_{\tau}$ can be rewritten as

$$-Nz_{\epsilon}(\bar{y}^\star - B\bar{x}^\star) = -\frac{z_{\epsilon}}{V}(\bar{y}^\star - \hat{\beta}_1 \bar{x}^\star). \quad (2.16)$$

Similarly, the constant term can be expanded into

$$-\sum_{i=1}^N (y_i^\star - Bx_i^\star)^2 = -\sum_{i=1}^N \left\{ (y_i - \bar{y}) - B(x_i - \bar{x}) \right\}^2 - \frac{1}{NV^2} \left\{ \bar{y}^\star - \hat{\beta}_1 \bar{x}^\star \right\}^2. \quad (2.17)$$

The first term on the right hand side can again be expanded as follows

$$-\sum_{i=1}^N \left\{ (y_i - \bar{y}) - B(x_i - \bar{x}) \right\}^2$$

$$= -\sum_{i=1}^N \left\{ (y_i - \bar{y}) - \frac{\hat{\beta}_1}{NV}(x_i - \bar{x}) - \frac{\bar{x}^\star \bar{y}^\star}{s_{xx}V}(x_i - \bar{x}) \right\}^2$$

$$= -\sum_{i=1}^N \left\{ (y_i - \bar{y}) - \frac{\hat{\beta}_1}{NV}(x_i - \bar{x}) \right\}^2 - \frac{(\bar{x}^\star \bar{y}^\star)^2}{s_{xx}V^2} + \frac{2 \bar{x}^\star \bar{y}^\star}{s_{xx}V} \left\{ \frac{s_{xy}}{s_{xx}} - \frac{\hat{\beta}_1}{NV} \right\}. $$

28
where \(s_{xy} = \sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})\) and \(s_{yy} = \sum_{i=1}^{N} (y_i - \bar{y})^2\). Consequently,

\[
- \sum_{i=1}^{N} \{(y_i - \bar{y}) - B(x_i - \bar{x})\}^2
\]

\[
= - \frac{1}{N^2 \sigma^2} \sum_{i=1}^{N} \left\{N^2 (y_i - \bar{y}) - \hat{\beta}_1 (x_i - \bar{x})\right\}^2 + \frac{2 (\bar{x}^*)^2 \hat{y}^* \hat{\beta}_1}{s_{xx} \sigma^2} - \frac{(\bar{x}^* \hat{y}^*)^2}{s_{xx} \sigma^2}
\]

\[
+ \frac{2 (\bar{x}^*)^2 \hat{y}^* \hat{\beta}_1}{s_{xx} \sigma^2} - \frac{(\bar{x}^* \hat{y}^*)^2}{s_{xx} \sigma^2}.
\]

Note that \(N \sigma^2 = \sum_{i=1}^{N} \{(y_i - \bar{y}) - \hat{\beta}_1 (x_i - \bar{x})\}^2 = s_{yy} - \hat{\beta}_1 s_{xx}\), where \(\sigma^2\) is the usual maximum likelihood estimate of \(\sigma^2\) that does not depend on \(\tau\), so

\[
- \sum_{i=1}^{N} \{(y_i - \bar{y}) - B(x_i - \bar{x})\}^2
\]

\[
= - \frac{1}{N^2 \sigma^2} \left\{N^2 (\bar{x}^*)^2 \sigma^2 \left\{\frac{1}{N} + \frac{(\bar{x}^*)^2}{s_{xx} \sigma^2}\right\} - \frac{(\bar{x}^*)^4 \hat{\beta}_1}{s_{xx} \sigma^2} + \frac{2 (\bar{x}^*)^3 \hat{y}^* \hat{\beta}_1}{s_{xx} \sigma^2} - \frac{(\bar{x}^* \hat{y}^*)^2}{s_{xx} \sigma^2}\right\}
\]

\[
+ \frac{2 (\bar{x}^*)^2 \hat{y}^* \hat{\beta}_1}{s_{xx} \sigma^2} - \frac{(\bar{x}^* \hat{y}^*)^2}{s_{xx} \sigma^2},
\]

which leads to

\[
- \sum_{i=1}^{N} \{(y_i - \bar{y}) - B(x_i - \bar{x})\}^2
\]

\[
= - \frac{N \sigma^2}{\sigma^2} \left\{\frac{1}{N} + \frac{(\bar{x}^*)^2}{s_{xx} \sigma^2}\right\} - \frac{(\bar{x}^*)^4 \hat{\beta}_1}{s_{xx} \sigma^2} + \frac{2 (\bar{x}^*)^3 \hat{y}^* \hat{\beta}_1}{s_{xx} \sigma^2} - \frac{(\bar{x}^* \hat{y}^*)^2}{s_{xx} \sigma^2}
\]

\[
= - N \sigma^2 - \frac{(\bar{x}^*)^2}{s_{xx} \sigma^2} \left\{\bar{y}^* - \hat{\beta}_1 \bar{x}^*\right\}.
\]

Substituting back into (2.17) gives

\[
- \sum_{i=1}^{N} (y_i^* - Bx_i^*)^2 = - N \sigma^2 - \frac{1}{V} \left\{\bar{y}^* - \hat{\beta}_1 \bar{x}^*\right\}^2,
\]

and after using (2.16), the quadratic in \(\sigma\), given in (2.15), becomes

\[
N \sigma^2 \frac{z}{V} (\bar{y}^* - \hat{\beta}_1 \bar{x}^*) \sigma_{\tau} - N \sigma^2 - \frac{1}{V} \left\{\bar{y}^* - \hat{\beta}_1 \bar{x}^*\right\}^2 = 0.
\]
The estimate of $\sigma_{\tau_e}$ is the positive root to this equation, given by

$$
\hat{\sigma}_{\tau_e} = \sqrt{\hat{\sigma}^2 + \frac{(\bar{y}^* - \hat{\beta}_1 \bar{x}^*)^2}{4N^2V^2} (z^2 + 4N\nu) + \frac{z_c(\bar{y}^* - \hat{\beta}_1 \bar{x}^*)}{2N\nu}},
$$

(2.18)

which only depends on the maximum likelihood estimates of $\beta_1$ and $\sigma^2$. Consequently, the estimate of $\beta_{1,\tau_e}$ is found by substituting (2.18) into (2.14). Note that since $z_c < 0$, the last term in (2.18) shifts the estimate of $\sigma_{\tau_e}$ back toward $\hat{\sigma}$.

The maximum likelihood estimate of the shelf-life under this approach is given by

$$
\hat{\tau}_e = \inf \left\{ t : x_i^T \hat{\beta} + z_c \hat{\sigma} \leq \kappa \right\},
$$

where $\hat{\beta}$ and $\hat{\sigma}$ are the maximum likelihood estimates of $\beta$ and $\sigma$, respectively.

Hence, the profile likelihood for $\tau_e$ is

$$
\tilde{L}(\tau_e) = -\frac{N}{2} \log \hat{\sigma}_{\tau_e}^2 - \frac{1}{2\hat{\sigma}_{\tau_e}^2} \sum_{i=1}^{N} \left( y_i - \hat{\beta}_{1,\tau_e} (x_i - \tau_e) - \kappa + z_c \hat{\sigma}_{\tau_e} \right)^2.
$$

(2.19)

To find the label shelf-life, let $\tilde{L}(\hat{\tau}_e)$ be the likelihood evaluated at the maximum likelihood estimates and let $\tilde{L}(\tau_e)$ be the likelihood evaluated, for a given value of $\tau_e$. Then, asymptotically, the generalised likelihood ratio statistic is given by

$$
w(\tau_e) = 2 \left( \tilde{L}(\hat{\tau}_e) - \tilde{L}(\tau_e) \right),
$$

(2.20)

which is distributed approximately $\chi^2_1$. Consequently, the label shelf-life, based on the $100(1 - \alpha)\%$ lower confidence bound, is given by

$$
\hat{\tau}_{e,L(\alpha)} = \inf \left\{ t : (t \leq \hat{\tau}_e) \& (w(t) \leq \chi^2_{1,1-2\alpha}) \right\}.
$$

(2.21)

The results on likelihood ratio tests are based on large sample properties, and therefore confidence bounds based on the likelihood ratio test may not be very accurate for small samples. To overcome this problem, a modification to the profile likelihood has been proposed (Barndorff-Nielsen, 1983). It will be discussed in the following section.
The Modified Profile Likelihood (MPL) Approach

Consider the general likelihood problem in which scalar parameter $\tau_e$ needs to be estimated in the presence of nuisance parameter vector $(\beta_1, \sigma^2)$. A modified profile likelihood (Barndorff-Nielsen, 1983) appropriate for estimation of $\tau_e$ is given by

$$\tilde{L}_m(\tau_e) = \tilde{L}(\tau_e) - \frac{1}{2} \log \left| I(\hat{\beta}_1, \hat{\sigma}_e^2) \right| - \log \left| \frac{\partial(\hat{\beta}_1, \hat{\sigma}_e^2)}{\partial(\hat{\beta}_1, \hat{\sigma}_e^2)} \right|,$$  

(2.22)

where $I(\cdot)$ denotes the observed information matrix based on $\tilde{L}(\tau_e)$ — the use of the expected information instead of the observed is also an option (Barndorff-Nielsen, 1983). The final term in (2.22) is the log-determinant of a matrix of derivatives, known as a Jacobian. This Jacobian is found by differentiating (2.14) and (2.18) with respect to the maximum likelihood estimates $\hat{\beta}_1$ and $\hat{\sigma}_e^2$, that is

$$\frac{\partial(\hat{\beta}_1, \hat{\sigma}_e^2)}{\partial(\hat{\beta}_1, \hat{\sigma}_e^2)} = \begin{bmatrix} \frac{\partial \hat{\beta}_1}{\partial \beta_1} & \frac{\partial \hat{\beta}_1}{\partial \sigma_e^2} \\ \frac{\partial \hat{\sigma}_e}{\partial \beta_1} & \frac{\partial \hat{\sigma}_e}{\partial \sigma_e^2} \end{bmatrix}.$$

From (2.10) and (2.12) it follows that the elements of the observed information matrix are given by

$$I_o(\beta_1, \sigma_e^2, \beta_1, \sigma_e^2) = s_{xx} + N(\bar{x}^*)^2$$

$$I_o(\beta_1, \sigma_e^2) = \frac{1}{\sigma_e^2} \left( \frac{\partial L}{\partial \beta_1} \right) - \frac{N z_e \bar{x}^*}{2 \sigma_e^2}$$

and

$$I_o(\sigma_e^2, \sigma_e^2) = -\frac{N(2 + z_e^2)}{4 \sigma_e^4} + \frac{1}{\sigma_e^4} \sum_{i=1}^N (y_i^* - \beta_1, \tau_e, x_i^* + z_e \sigma_e)^2 - \frac{5N z_e (y^* - \beta_1, \tau_e, \bar{x}^* + z_e \sigma_e)}{4 \sigma_e^5}.$$

Taking expectations yields the elements of the expected information matrix

$$I_e(\beta_1, \sigma_e^2, \beta_1, \sigma_e^2) = \frac{s_{xx} + N(\bar{x}^*)^2}{\sigma_e^2}$$

$$I_e(\beta_1, \sigma_e^2) = -\frac{N z_e \bar{x}^*}{2 \sigma_e^2}$$

$$I_e(\sigma_e^2, \sigma_e^2) = \frac{N(2 + z_e^2)}{4 \sigma_e^4}.$$
The elements of the Jacobian, the last term in (2.22), are generally very difficult to obtain. In this case however exact results are possible by differentiating (2.14) and (2.18) with respect to the maximum likelihood estimates \( \hat{\beta}_1 \) and \( \hat{\sigma}^2 \). They are

\[
\frac{\partial \hat{\sigma}^2_{\tau, \epsilon}}{\partial \hat{\beta}_1} = \frac{\frac{z_{\epsilon} \bar{x}^*}{2N} + \frac{D(z_{\epsilon}^2 + 4NV) \bar{x}^*}{4N^2V^2}}{\sqrt{\hat{\sigma}^2 + \frac{D^2(z_{\epsilon}^2 + 4NV)}{4N^2V^2}}}
\]

\[
\frac{\partial \hat{\beta}_1_{\tau, \epsilon}}{\partial \hat{\sigma}^2} = \frac{z_{\epsilon} \bar{x}^*}{2\hat{\sigma}_{\tau, \epsilon} s_{xx}V} \left( \frac{\partial \hat{\sigma}^2_{\tau, \epsilon}}{\partial \hat{\beta}_1} \right) + \frac{1}{NV}
\]

where \( D = \bar{y}^* - \hat{\beta}_1 \bar{x}^* \).

The modified version of (2.20) is given by

\[
w_m(\tau_\epsilon) = 2 \left( \bar{L}_m(\hat{\tau}_\epsilon) - \bar{L}_m(\tau_\epsilon) \right),
\]

which is distributed approximately \( \chi^2_1 \), where \( \hat{\tau}_\epsilon \) is obtained by maximizing (2.22). Consequently, the label shelf-life, based on the 100(1 − \( \alpha \))% lower confidence bound, is given by

\[
\hat{\tau}_{\epsilon, L(\alpha)} = \inf \left\{ t : (t \leq \hat{\tau}_\epsilon) & (w_m(t) \leq \chi^2_1, 1 - 2\alpha) \right\}.
\] (2.23)

In more general circumstances explicit solutions of the conditional parameters \( (\beta_{1, \tau, \epsilon}, \sigma^2_{\tau, \epsilon}) \) in terms of the maximum likelihood parameters \( (\beta_1, \sigma^2) \) may not exist. Consequently, it may not be feasible to calculate the Jacobian in those cases and the only possible solution is to ignore this term in such instances. This will be referred to as the truncated modified profile likelihood (TMPL).

An alternative modification to the PL, which does not require the calculation of the Jacobian term, was presented by Cox and Reid (1987). The application of their approach to stability testing is presented in the next section.
The Approximate Conditional Profile Likelihood (ACPL) Approach

The approximate conditional profile likelihood (Cox and Reid, 1987) is

$$\tilde{L}_c(\tau_\epsilon) = \tilde{L}(\tau_\epsilon, \lambda) - \frac{1}{2} \log |I(\hat{\lambda})|,$$

where $\tau_\epsilon$ and $\lambda$ are orthogonal parameters, and $I(\hat{\lambda})$ denotes the observed information matrix of $\lambda$, obtained from $\tilde{L}(\tau_\epsilon, \lambda)$, evaluated at $\hat{\lambda}$ — again a possible alternative is to use of the expected information instead of the observed information. This conditional PL is very similar to the MPL, except that the term involving the Jacobian is ignored, since it is of order $1/N$ when the parameters are orthogonal. The ACPL is not invariant under transformations of $\lambda$, but Cox & Reid have argued that the orthogonal parameterization reduces this lack of invariance.

A set of parameters $\lambda^T = (\lambda_1, \lambda_2)$, which are orthogonal to $\tau_\epsilon$, can be found from the original parameters $(\beta_{1,\tau_\epsilon}, \sigma^2_{\tau_\epsilon})$ by solving the differential equation (Cox and Reid, 1987, equation (4) on page 3)

$$I_e(\beta_{1,\tau_\epsilon}, \sigma^2_{\tau_\epsilon}) \begin{bmatrix} \frac{\partial \beta_{1,\tau_\epsilon}}{\partial \tau_\epsilon} \\ \frac{\partial \sigma^2_{\tau_\epsilon}}{\partial \tau_\epsilon} \end{bmatrix} = \mathcal{E} \begin{bmatrix} \frac{\partial^2 \tilde{L}}{\partial \beta_{1,\tau_\epsilon} \partial \tau_\epsilon} \\ \frac{\partial^2 \tilde{L}}{\partial \sigma^2_{\tau_\epsilon} \partial \tau_\epsilon} \end{bmatrix},$$

where $I_e(\beta_{1,\tau_\epsilon}, \sigma^2_{\tau_\epsilon})$ denotes the expected information matrix of $\beta_{1,\tau_\epsilon}$ and $\sigma^2_{\tau_\epsilon}$, obtained from (2.9). This is equivalent to solving

$$\begin{bmatrix} s_{xx} + N\bar{x}^* - \frac{Nz_2\bar{x}^*}{2\sigma^2_{\tau_\epsilon}} \\ -\frac{Nz_2\bar{x}^*}{2\sigma^2_{\tau_\epsilon}} + \frac{N(2+z^2_{\tau_\epsilon})}{4\sigma^2_{\tau_\epsilon}} \end{bmatrix} \begin{bmatrix} \frac{\partial \beta_{1,\tau_\epsilon}}{\partial \tau_\epsilon} \\ \frac{\partial \sigma^2_{\tau_\epsilon}}{\partial \tau_\epsilon} \end{bmatrix} = \begin{bmatrix} \frac{N\beta_{1,\tau_\epsilon}\bar{x}^*}{\sigma^2_{\tau_\epsilon}} \\ -\frac{Nz_2\beta_{1,\tau_\epsilon}\bar{x}^*}{2\sigma^2_{\tau_\epsilon}} \end{bmatrix}.$$

Consequently, the first equation is

$$\frac{\partial \beta_{1,\tau_\epsilon}}{\partial \tau_\epsilon} = \frac{2N\beta_{1,\tau_\epsilon}\bar{x}^*}{\sqrt{g(\tau_\epsilon)}},$$

where $g(\tau_\epsilon) = s_{xx}(2 + z^2_{\tau_\epsilon}) + 2N(\bar{x}^*)^2$. Hence

$$\int \frac{1}{\beta_{1,\tau_\epsilon}} d\beta_{1,\tau_\epsilon} = \int \frac{2N\bar{x}^*}{\sqrt{g(\tau_\epsilon)}} d\tau_\epsilon.$$
Let \( u = 2N(\bar{x}^*)^2 \), which means that \( du = -4N\bar{x}^* d\tau_\epsilon \) and hence

\[
\log \beta_{1,\tau_\epsilon} = -\frac{1}{2} \int \frac{1}{s_{xx}(2 + z^2_\epsilon) + u} \, du \\
= -\frac{1}{2} \log g(\tau_\epsilon) + a_1(\lambda).
\]

Now \( a_1(\lambda) \) is an arbitrary function of \( \lambda \). Choosing \( a_1(\lambda) = \log \lambda_1 \) gives

\[
\lambda_1 = \beta_{1,\tau_\epsilon} \sqrt{g(\tau_\epsilon)}.
\]

The second equation is

\[
\frac{\partial \sigma^2}{\partial \tau_\epsilon} = -2\sigma_{\tau_\epsilon} \frac{z_\epsilon s_{xx} \beta_{1,\tau_\epsilon}}{g(\tau_\epsilon)},
\]

and substituting for \( \beta_{1,\tau_\epsilon} \) and integrating yields

\[
\int \frac{1}{2\sigma_{\tau_\epsilon}} d\sigma^2_{\tau_\epsilon} = \int -\frac{z_\epsilon s_{xx} \lambda_1}{g(\tau_\epsilon)^{3/2}} \, d\tau_\epsilon.
\]

This reduces to

\[
\sigma_{\tau_\epsilon} = \frac{z_\epsilon s_{xx} \lambda_1}{(2N)^{3/2}} \int \frac{1}{[c^2 + (\bar{x}^*)^2]^{3/2}} \, d\tau_\epsilon,
\]

where \( c^2 = \frac{s_{xx}(2 + z^2_\epsilon)}{2N} \). Now, let \( \bar{x}^* = c \tan \theta \), such that \( -d\tau = \frac{c}{\cos^2 \theta} \, d\theta \), giving

\[
\sigma_{\tau_\epsilon} = \frac{z_\epsilon s_{xx} \lambda_1}{(2N)^{3/2}} \int \frac{1}{c^2 \cos \theta} \, d\theta \\
= \frac{z_\epsilon s_{xx} \lambda_1}{(2N)^{3/2}} \int \frac{1}{c^2 \cos \theta} \, d\theta \\
= \frac{z_\epsilon s_{xx} \lambda_1}{c^2(2N)^{3/2}} [\sin \theta + a_2(\lambda)] \\
= \frac{z_\epsilon \lambda_1}{(2 + z^2_\epsilon)^{1/2}} [\sin \theta + a_2(\lambda)],
\]

where again \( a_2(\lambda) \) is an arbitrary function of \( \lambda \). Since, \( \bar{x}^* = c \tan \theta \) and \( 1/\cos \theta = \sqrt{\sec^2 \theta} = \sqrt{1 + \tan^2 \theta} \) it follows that

\[
\sin \theta = \frac{\bar{x}^*/c}{\sqrt{1 + (\bar{x}^*)^2/c^2}} = \frac{\bar{x}^*}{\sqrt{c^2 + (\bar{x}^*)^2}}.
\]

Letting \( a_2(\lambda) = \lambda_2 \) gives

\[
\sigma_{\tau_\epsilon} = \frac{z_\epsilon \bar{x}^* \lambda_1}{(2 + z^2_\epsilon)^{1/2} \sqrt{g(\tau_\epsilon)}} + \frac{z_\epsilon \lambda_1 \lambda_2}{(2 + z^2_\epsilon)^{1/2} \sqrt{2N}},
\]

34
and re-arranging for \(\lambda_2\) gives

\[
\lambda_2 = \sqrt{2N} \frac{\sigma_{\tau_e}(2 + z_2^2) - z_\epsilon \beta_{1,\tau_e} \bar{x}^*}{z_\epsilon \beta_{1,\tau_e} \sqrt{g(\tau_e)}}.
\]

The value of \(L(\tau_e, \hat{\lambda})\) in (2.24) will be the same, irrespective of whether the parameterization is in terms of \((\beta_{1,\tau_e}, \sigma_{\tau_e}^2)\) or \((\lambda_1, \lambda_2)\). The derivatives of \(\tilde{L}\) with respect to \(\lambda\) are found by using the chain rule, that is,

\[
\frac{\partial \tilde{L}}{\partial \lambda_1} = \left( \frac{\partial \tilde{L}}{\partial \beta_{1,\tau_e}} \right) \left( \frac{\partial \beta_{1,\tau_e}}{\partial \lambda_1} \right) + \left( \frac{\partial \tilde{L}}{\partial \sigma_{\tau_e}^2} \right) \left( \frac{\partial \sigma_{\tau_e}^2}{\partial \lambda_1} \right)
\]

\[
\frac{\partial \tilde{L}}{\partial \lambda_2} = \left( \frac{\partial \tilde{L}}{\partial \sigma_{\tau_e}^2} \right) \left( \frac{\partial \sigma_{\tau_e}^2}{\partial \lambda_2} \right).
\]

Similarly, the second derivatives are given by

\[
\frac{\partial^2 \tilde{L}}{\partial \lambda_1 \partial \lambda_2} = \left( \frac{\partial^2 \tilde{L}}{\partial (\sigma_{\tau_e}^2)^2} \right) \left( \frac{\partial \sigma_{\tau_e}^2}{\partial \lambda_1} \right) \left( \frac{\partial \sigma_{\tau_e}^2}{\partial \lambda_2} \right) + \left( \frac{\partial \tilde{L}}{\partial \beta_{1,\tau_e}} \right) \left( \frac{\partial^2 \beta_{1,\tau_e}}{\partial \lambda_1^2} \right) + \left( \frac{\partial \tilde{L}}{\partial \sigma_{\tau_e}^2} \right) \left( \frac{\partial^2 \sigma_{\tau_e}^2}{\partial \lambda_1^2} \right)
\]

\[
\frac{\partial^2 \tilde{L}}{\partial \lambda_2^2} = \left( \frac{\partial^2 \tilde{L}}{\partial (\sigma_{\tau_e}^2)^2} \right) \left( \frac{\partial \sigma_{\tau_e}^2}{\partial \lambda_2} \right)^2 + \left( \frac{\partial \tilde{L}}{\partial \sigma_{\tau_e}^2} \right) \left( \frac{\partial^2 \sigma_{\tau_e}^2}{\partial \lambda_2^2} \right).
\]

The derivatives of \(\tilde{L}\) with respect to \(\beta_{1,\tau_e}\) and \(\sigma_{\tau_e}^2\) are given by (2.10) and (2.12). The derivatives of \(\beta_{1,\tau_e}\) and \(\sigma_{\tau_e}^2\) with respect to \(\lambda_1\) are given by

\[
\frac{\partial \beta_{1,\tau_e}}{\partial \lambda_1} = \frac{1}{\sqrt{g(\tau_e)}}
\]

\[
\frac{\partial \sigma_{\tau_e}^2}{\partial \lambda_1} = \frac{\lambda_1 z_\epsilon^2 \left( \bar{x}^* \sqrt{2N} + \lambda_2 \sqrt{g(\tau_e)} \right)^2}{N(2 + z_\epsilon^2) g(\tau_e)}
\]

\[
\frac{\partial^2 \sigma_{\tau_e}^2}{\partial \lambda_1^2} = \frac{\lambda_1^2 z_\epsilon^2}{N(2 + z_\epsilon^2)}
\]

\[
\frac{\partial^2 \sigma_{\tau_e}^2}{\partial \lambda_1 \partial \lambda_2} = \frac{2 \lambda_1 z_\epsilon^2 \left( \bar{x}^* \sqrt{2N} + \lambda_2 \sqrt{g(\tau_e)} \right)}{N(2 + z_\epsilon^2) g(\tau_e)}
\]

\[
\frac{\partial^2 \sigma_{\tau_e}^2}{\partial \lambda_2^2} = \frac{\lambda_1^2 z_\epsilon^2}{N(2 + z_\epsilon^2)}
\]

35
The generalised likelihood ratio statistic based on the conditional PL is given by

\[ w_c(\tau_\epsilon) = 2 \left( \tilde{L}_c(\hat{\tau}_\epsilon; y) - \tilde{L}_c(\tau_\epsilon; y) \right), \]

which is distributed approximately \( \chi^2_1 \), where \( \hat{\tau}_\epsilon \) is obtained by maximizing (2.24). Consequently, the label shelf-life, based on the 100(1 - \( \alpha \))\% lower confidence bound, is given by

\[ \hat{\tau}_{c,L(\alpha)} = \inf \left\{ t : (t \leq \hat{\tau}_\epsilon) \cap (w_c(t) \leq \chi^2_{1,1-2\alpha}) \right\}. \] (2.25)

Figure 2.2 shows an example of the three constrained profile likelihood approaches detailed above. The graph also shows the 95% lower confidence bounds for the true shelf-life for each approach based on \( N = 8 \). From the plot it can be seen that the modified and conditional PL approaches give similar profile likelihoods. This may suggest that the constrained parameters \( \beta_1, \tau_\epsilon \) and \( \sigma^2_{\tau_\epsilon} \) are almost orthogonal, and hence the determinant of the Jacobian in (2.22) is close to 1 (which has been verified numerically). Both approaches also result in lower label shelf-life estimates than the basic profile likelihood, which is reflected in the coverage probabilities in the simulation study presented in Section 2.3.
2.2.3 Constrained Profile Likelihood (CPL) Based Approaches

The approaches presented in Section 2.2.2 depend on the elimination of $\beta_0$. This may not be possible in a more general setting. However, as will be shown in this section, the constrained profile likelihood provides an alternative, yet equivalent, approach.

Again, consider the general likelihood problem where the parameters $\theta^T = (\beta_0, \beta_1, \sigma^2)$ need to be estimated. However, the parameters are now required to satisfy (2.7) for a given value of $\tau_\epsilon$. Consequently, denote these constrained parameters by $\theta_{\tau_\epsilon}$ and write the constraint as

$$g(\theta_{\tau_\epsilon}; \tau_\epsilon) = \beta_0,\tau_\epsilon + \beta_1,\tau_\epsilon \tau_\epsilon + z_\epsilon \sigma_{\tau_\epsilon} - \kappa = 0.$$  

(2.26)

One way of estimating maximum likelihood parameters under a constraint is via Lagrangian multipliers, that is, by maximizing the function

$$L(\theta_{\tau_\epsilon}; y) - \lambda g(\theta_{\tau_\epsilon}; \tau_\epsilon),$$  

(2.27)

where $\lambda$ is a Lagrangian multiplier. This can be done by differentiating with respect to $\theta_{\tau_\epsilon}$ and $\lambda$, and employing an iterative maximization scheme. If the constraint is linear in the parameters $\theta_{\tau_\epsilon}$, then the second derivative of $g(\theta_{\tau_\epsilon}; \tau_\epsilon)$ with respect to $\theta_{\tau_\epsilon}$ is zero, and hence does not contribute to the information matrix. If, however, the constraint is not linear in some parameters, then the second derivative will not equal zero. In particular, this second derivative can be very complicated, making estimation more difficult.

An alternative approach is based on an augmented form of Lagrangian multipliers, referred to as the \textit{Powell-Hestenes method} (Osborne, 2000). The advantage of this alternative is that the second derivative of the constraint is not required. Osborne suggests minimizing an objective function which is a scaled version of the negative log-likelihood — this is equivalent to maximizing the log-likelihood. The Powell-Hestenes method is used only for estimation of the parameters, given a value of $\tau_\epsilon$. The profile likelihood for $\tau_\epsilon$ is then obtained by evaluating (2.8) at $(\hat{\beta}_0,\tau_\epsilon, \hat{\beta}_1,\tau_\epsilon, \sigma^2_{\tau_\epsilon})$.

The objective function, based on the log-likelihood (2.8) and the constraint (2.26), takes
the form
\[ H(\theta_{\tau_e}; \tau_e, y) = \frac{1}{N} \left[ -L(\theta_{\tau_e}) \right] + \omega \left[ g(\theta_{\tau_e}; \tau_e) + \psi_{\tau_e} \right]^2, \]  

(2.28)

where \( \omega \) governs the importance placed on the constraint at each iteration. Generally, \( \omega \) is fixed and chosen to be of order \( O(\sqrt{N}) \) and \( \psi_{\tau_e} \) is an ancillary parameter, similar to the Lagrangian multiplier \( \lambda \). Osborne (2000) shows that the update for \( \psi_{\tau_e} \) at the \( m \)-th iteration is given by
\[ \psi_{\tau_e}^{(m+1)} = \psi_{\tau_e}^{(m)} + g(\theta_{\tau_e}^{(m)}; \tau_e). \]

The scores with respect to \( \beta_{\tau_e} \) and \( \sigma_{\tau_e}^2 \) are
\[ \frac{\partial H}{\partial \beta_{\tau_e}} = \frac{1}{N} \left[ -\frac{1}{\sigma_{\tau_e}^2} X^T (y - X\beta_{\tau_e}) \right] + 2\omega \left[ g(\theta_{\tau_e}; \tau_e) + \psi_{\tau_e} \right] x_{\tau_e}, \]
\[ \frac{\partial H}{\partial \sigma_{\tau_e}^2} = \frac{1}{N} \left[ \frac{N}{2\sigma_{\tau_e}^2} - \frac{1}{2\sigma_{\tau_e}^4} (y - X\beta_{\tau_e})^T (y - X\beta_{\tau_e}) \right] + \frac{\omega z_{\tau_e}}{\sigma_{\tau_e}} \left[ g(\theta_{\tau_e}; \tau_e) + \psi_{\tau_e} \right]. \]

In the constrained case, the second derivatives of (2.28) with respect to the parameters give the elements of an equivalent to the observed information matrix which is denoted by \( I_o' \) to distinguish it from the actual information matrix, \( I_o \), which is based on the log-likelihood (2.8),
\[ I_o'(\beta_{\tau_e}, \beta_{\tau_e}) = \frac{\partial^2 H}{\partial \beta_{\tau_e} \partial \beta_{\tau_e}^T} = \frac{1}{N} \left[ \frac{X^T X}{\sigma_{\tau_e}^2} \right] + 2\omega X_{\tau_e} x_{\tau_e}^T, \]
\[ I_o'(\beta_{\tau_e}, \sigma_{\tau_e}^2) = \frac{\partial^2 H}{\partial \beta_{\tau_e} \partial \sigma_{\tau_e}^2} = \frac{1}{N} \left[ \frac{1}{\sigma_{\tau_e}^4} X^T (y - X\beta_{\tau_e}) \right] + \frac{\omega z_{\tau_e}}{\sigma_{\tau_e}} x_{\tau_e}, \]
\[ I_o'(\sigma_{\tau_e}^2, \sigma_{\tau_e}^2) = \frac{\partial^2 H}{\partial \sigma_{\tau_e}^2 \partial \sigma_{\tau_e}^2} = \frac{1}{N} \left[ -\frac{N}{2\sigma_{\tau_e}^4} + \frac{1}{\sigma_{\tau_e}^6} (y - X\beta_{\tau_e})^T (y - X\beta_{\tau_e}) \right] + \frac{\omega z_{\tau_e}^2}{2\sigma_{\tau_e}^2}. \]

Note that components containing the second derivative of \( g(\theta_{\tau_e}; \tau_e) \) have been omitted based on the argument given by Osborne (2000).

Taking expectations gives components of a matrix which is equivalent to the expected
2.2. ESTIMATION OF THE LABEL SHELF-LIFE

information matrix, denoted by \( I_e' \). They are given by

\[
I_e'(\beta_{\tau_e}, \beta_{\tau_e}) = \frac{1}{N} \left[ \frac{X^T X}{\sigma^2_{\tau_e}} \right] + 2\omega x_{\tau_e} x_{\tau_e}^T
\]

\[
I_e'(\beta_{\tau_e}, \sigma^2_{\tau_e}) = 0 + \frac{\omega z_{\tau_e}}{\sigma_{\tau_e}} x_{\tau_e}
\]

\[
I_e'(\sigma^2_{\tau_e}, \sigma^2_{\tau_e}) = \frac{1}{N} \left[ \frac{N}{2\sigma^4_{\tau_e}} \right] + \frac{\omega z^2_{\tau_e}}{2\sigma^2_{\tau_e}}
\]

Note that the relationship between \( I'_o \) and \( I_o \) is

\[
I'_o = \frac{1}{N} I_o + 2\omega C_{\tau_e} C_{\tau_e}^T,
\]

where \( C_{\tau_e} \) is the \( p \times 1 \) vector of derivatives of the constraint such that

\[
C_{\tau_e} = \frac{\partial g(\theta_{\tau_e}; \tau_e)}{\partial \theta_{\tau_e}}.
\]

The same relationship also holds for \( I'_e \) and \( I_e \).

Estimates for \( \beta_{\tau_e} \) and \( \sigma^2_{\tau_e} \) can then be found using Fisher scoring. These estimates can be substituted into the likelihood and the likelihood can be profiled over \( \tau_e \), resulting in

\[
L(\tau_e; \hat{\beta}_{0,\tau_e}, \hat{\beta}_{1,\tau_e}, \hat{\sigma}^2_{\tau_e}, y) = -\frac{N}{2} \log 2\pi - \frac{N}{2} \log \hat{\sigma}^2_{\tau_e} - \frac{1}{2\hat{\sigma}^2_{\tau_e}} \sum_{i=1}^{N} (y_i - \hat{\beta}_{0,\tau_e} - \hat{\beta}_{1,\tau_e} x_i)^2. \tag{2.29}
\]

The maximum likelihood estimate of \( \tau_e \) is the value which minimises this profile likelihood. The unconstrained parameter estimates of \( \beta \) and \( \sigma^2 \) can be found by setting \( \omega = 0 \) and solving for \( \theta = (\beta, \sigma^2) \).

The constrained version of (2.20) is given by

\[
w(\tau_e) = 2 \left( L(\hat{\beta}_{\tau_e}, \hat{\sigma}^2_{\tau_e}; \hat{\tau}_e, y) - L(\hat{\beta}_{\tau_e}, \hat{\sigma}^2_{\tau_e}; \tau_e, y) \right),
\]

which is distributed approximately \( \chi^2_1 \), where the estimates of \( \beta_{\tau_e} \) and \( \sigma^2_{\tau_e} \) are obtained by maximizing (2.28), and the estimate of \( \tau_e \) is obtained by solving \( g(\hat{\theta}; \tau_e) = 0 \) for \( \tau_e \). Consequently, the label shelf-life, based on the 100(1 - \( \alpha \))% lower confidence bound, is given by

\[
\hat{\tau}_{e,L(\alpha)} = \inf \left\{ t : (t \leq \hat{\tau}_e) \& \ (w(t) \leq \chi^2_{1,1-2\alpha}) \right\}. \tag{2.30}
\]
The label shelf-life obtained using this method should be identical to that obtained using the basic profile likelihood in Section 2.2.2. This has been verified numerically.

Comments made about the large sample properties of the profile likelihood also apply here. Consequently, a modification to the CPL, similar to that suggested by Barndorff-Nielsen (1983), is desirable. Such a modification is developed below.

The Truncated Modified Constrained Profile Likelihood (TMCPL) Approach

The previous section showed that the constrained parameters cannot be expressed explicitly in terms of the unconstrained maximum likelihood estimates. Consequently, a Jacobian term analogous to the one used in (2.22) cannot be obtained. One option is to ignore the Jacobian term (Barndorff-Nielsen, 1983) and adjust the profile likelihood (2.29) purely with the log determinant of the appropriate information matrix. This approach is adopted here and the resulting method shall be termed truncated modified constrained profile likelihood (TMCPL) to distinguish it from the previously mentioned truncated modified profile likelihood.

To adjust the profile likelihood under the general constrained estimation approach presented in Section 2.2.3 the appropriate information matrix is required. The observed information matrix $I_o(\hat{\theta}_{\tau_\epsilon})$ based on (2.29) can easily be calculated and evaluated at the parameter estimates $\hat{\theta}_{\tau_\epsilon}$. However due to the constraint the parameters are linearly dependent. This means that $I_o(\hat{\theta}_{\tau_\epsilon})$ is not of full column rank, and has a zero determinant. Consequently, the adjustment which is based on the log determinant cannot be made. However, the approach suggested below approximates the “correct” information matrix used in (2.22).

This can be done by recognizing that the basic profile likelihood (2.19) and the constrained profile likelihood (2.29) must be the same for any given value of $\tau_\epsilon$. In addition, since $g(\hat{\theta}_{\tau_\epsilon}; \tau_\epsilon) = 0$, the equality can be written as

$$\tilde{L}(\tau_\epsilon; \vartheta_{\tau_\epsilon}) = \tilde{L}(\tau_\epsilon; \theta_{\tau_\epsilon}) + \lambda g(\theta_{\tau_\epsilon}; \tau_\epsilon),$$

(2.31)
where $\boldsymbol{\theta}^T_{\tau} = (\beta_{1,\tau}, \sigma^2_{\tau})$ denote the parameters under the basic profile likelihood model, $\boldsymbol{\hat{\theta}}^T_{\tau} = (\beta_{0,\tau}, \beta_{1,\tau}, \sigma^2_{\tau})$ denote the parameters under the constrained profile likelihood model, $\lambda$ is a Lagrangian multiplier, and the last term on the right hand side is zero when the constraint is satisfied.

The constraint $g(\theta_{\tau}; \tau)$ can be approximated using a Taylor Series expansion about $\theta_{\tau} = \theta_{0,\tau}$, say, that is,

$$g(\theta_{\tau}; \tau) \approx g(\theta_{0,\tau}; \tau) + \frac{\partial g(\theta_{\tau}; \tau)}{\partial \theta_{\tau}} |_{\theta_{\tau} = \theta_{0,\tau}} (\theta_{\tau} - \theta_{0,\tau}) .$$

Since both $\theta_{\tau}$ and $\theta_{0,\tau}$ must satisfy the constraint, it follows that

$$0 \approx 0 + C_0 (\theta_{\tau} - \theta_{0,\tau}) ,$$

where $C_0$ is the $1 \times p$ matrix of derivatives evaluated at $\theta_{0,\tau}$. Hence, let $C_0 \theta_{0,\tau} = c_0$, say. Substituting into (2.31) gives

$$\tilde{L}(\tau; \theta_{\tau}) \approx \tilde{L}(\tau; \theta_{0,\tau}) + \lambda \left[ C_0 \theta_{\tau} - c_0 \right] . \quad (2.32)$$

Let $L_0$ be the orthonormal matrix of size $(p - 1) \times p$, orthogonal to $C_0$, that is,

$$C_0 L_0^T = 0 \quad \text{and} \quad L_0 C_0^T = I_{p-1} ,$$

and write

$$T_0 \theta_{\tau} = \begin{bmatrix} L_0 \\ C_0 \end{bmatrix} \theta_{\tau} = \begin{bmatrix} \hat{\theta}_{\tau} \\ c_0 \end{bmatrix} . \quad (2.33)$$

Differentiating (2.32) with respect to $\vartheta_{j,\tau}$ gives

$$\frac{\partial \tilde{L}(\tau; \vartheta_{\tau})}{\partial \vartheta_{j,\tau}} \approx \frac{\partial L(\theta_{\tau})}{\partial \vartheta_{j,\tau}} \frac{\partial \theta_{\tau}}{\partial \vartheta_{j,\tau}} + \lambda C_0 \frac{\partial \theta_{\tau}}{\partial \vartheta_{j,\tau}} ,$$

and differentiating again with respect to $\vartheta_{k,\tau}$ gives

$$\frac{\partial^2 \tilde{L}(\theta_{\tau})}{\partial \vartheta_{j,\tau} \partial \vartheta_{k,\tau}} \approx \left( \frac{\partial \theta_{\tau}}{\partial \vartheta_{k,\tau}} \right) \left( \frac{\partial^2 L(\theta_{\tau})}{\partial \vartheta_{j,\tau} \partial \theta_{\tau}} \right) \left( \frac{\partial \theta_{\tau}}{\partial \vartheta_{j,\tau}} \right) + \left( \frac{\partial L(\theta_{\tau})}{\partial \theta_{\tau}} \right) \left( \frac{\partial^2 \theta_{\tau}}{\partial \vartheta_{j,\tau} \partial \vartheta_{k,\tau}} \right) + \lambda C_0 \frac{\partial \theta_{\tau}}{\partial \vartheta_{j,\tau}} \frac{\partial^2 \theta_{\tau}}{\partial \vartheta_{j,\tau} \partial \vartheta_{k,\tau}} .$$
From (2.33) it follows that

\[ T_0 \frac{\partial \theta_{\tau_e}}{\partial j_{\tau_e}} = e_j, \]

where \( e_j \) is a vector of zeros, except for a 1 in the \( j \)-th position. Now, differentiating again gives

\[ T_0 \frac{\partial^2 \theta_{\tau_e}}{\partial j_{\tau_e} \partial k_{\tau_e}} = 0, \]

which, since \( T_0 \) is of full column rank, means that \( \frac{\partial^2 \theta_{\tau_e}}{\partial j_{\tau_e} \partial k_{\tau_e}} = 0 \). Hence

\[
\frac{\partial^2 \tilde{L}(\theta_{\tau_e})}{\partial j_{\tau_e} \partial k_{\tau_e}} \approx \left( \frac{\partial^2 L(\theta_{\tau_e})}{\partial \theta_{\tau_e}^T} \right) \left( \frac{\partial^2 L(\theta_{\tau_e})}{\partial \theta_{\tau_e} \partial \theta_{\tau_e}^T} \right) \left( \frac{\partial \theta_{\tau_e}}{\partial j_{\tau_e}} \right) \]

Since \( \frac{\partial \theta_{\tau_e}}{\partial j_{\tau_e}} = T_0^{-1} e_j \) it follows that

\[
\frac{\partial^2 \tilde{L}(\theta_{\tau_e})}{\partial \theta_{\tau_e} \partial \theta_{\tau_e}^T} \approx e_j^T T_0^{-T} \left( \frac{\partial^2 L(\theta_{\tau_e})}{\partial \theta_{\tau_e} \partial \theta_{\tau_e}^T} \right) T_0^{-1} e_j,
\]

where \( T_0^{-T} = (T_0^{-1})^T \). Consequently, the equivalent to the observed information matrix of \( \theta_{\tau_e} \) is given by

\[
I_o'(\theta_{\tau_e}) \approx \left[ T_0^{-T} I_o(\theta_{\tau_e}) T_0^{-1} \right] \text{(Upper left } p-1 \text{ matrix)}.
\]

The inverse of \( T_0 \) is given by

\[
T_0^{-1} = \left[ L_0^T \left( L_0 L_0^T \right)^{-1} C_0^T \left( C_0 C_0^T \right)^{-1} \right] = \left[ L_0^T C_0^T \left( C_0 C_0^T \right)^{-1} \right],
\]

such that

\[
T_0^{-T} I_o(\theta_{\tau_e}) T_0^{-1} = \begin{bmatrix}
L_0 I_o(\theta_{\tau_e}) L_0^T & L_0 I_o(\theta_{\tau_e}) C_0^T \left( C_0 C_0^T \right)^{-1} \\
(C_0 C_0^T)^{-1} C_0 I_o(\theta_{\tau_e}) L_0^T & (C_0 C_0^T)^{-1} C_0 I_o(\theta_{\tau_e}) C_0^T \left( C_0 C_0^T \right)^{-1}
\end{bmatrix},
\]

Consequently,

\[
I_o'(\theta_{\tau_e}) \approx L_0 I_o(\theta_{\tau_e}) L_0^T, \tag{2.34}
\]

where \( I_o'(\theta_{\tau_e}) \) is the appropriate equivalent to the information matrix for use in the modified profile likelihood.

On the other hand,

\[
\det \left\{ T_0^{-T} I_o(\theta_{\tau_e}) T_0^{-1} \right\} = \left[ \det \left\{ T_0 T_0^T \right\} \right]^{-1} \det \left\{ I_o(\theta_{\tau_e}) \right\} = \left[ \det \left\{ C_0 C_0^T \right\} \right]^{-1} \det \left\{ I_o(\theta_{\tau_e}) \right\}, \tag{2.35}
\]

42
2.2. ESTIMATION OF THE LABEL SHELF-LIFE

An obvious choice for \( \theta \) since

\[
\text{det} \left\{ T_0 T_0^T \right\} = \text{det} \left[ \begin{array}{cc} L_0 L_0^T & 0 \\ 0 & C_0 C_0^T \end{array} \right] = \text{det} \left\{ C_0 C_0^T \right\}.
\]

But also

\[
\text{det} \left\{ T_0^{-T} I_o(\theta_{\tau_e}) T_0^{-1} \right\} = \text{det} \left\{ L_0 I_o(\theta_{\tau_e}) L_0^T \right\} \text{det} \left\{ (C_0 C_0^T)^{-1} C_0 I_o(\theta_{\tau_e}) C_0^T (C_0 C_0^T)^{-1} \right\} - \left( C_0 C_0^T \right)^{-1} C_0 I_o(\theta_{\tau_e}) L_0^T \left( L_0 I_o(\theta_{\tau_e}) L_0^T \right)^{-1} L_0 I_o(\theta_{\tau_e}) C_0^T (C_0 C_0^T)^{-1} \right\}.
\]

The second determinant on the right hand side can now be written as

\[
\text{det} \left\{ (C_0 C_0^T)^{-1} C_0 \left[ I_o(\theta_{\tau_e}) - I_o(\theta_{\tau_e}) L_0^T \left( L_0 I_o(\theta_{\tau_e}) L_0^T \right)^{-1} L_0 I_o(\theta_{\tau_e}) \right] C_0^T \right\} \left( C_0 C_0^T \right)^{-1}
\]

\[
= \text{det} \left\{ (C_0 C_0^T)^{-1} C_0 C_0^T \left( C_0 I_o(\theta_{\tau_e}) \right)^{-1} C_0^T \right\} \left( C_0 C_0^T \right)^{-1}
\]

\[
= \left[ \text{det} \left\{ C_0 I_o(\theta_{\tau_e})^{-1} C_0^T \right\} \right]^{-1}.
\]

Consequently,

\[
\text{det} \left\{ T_0^{-T} I_o(\theta_{\tau_e}) T_0^{-1} \right\} = \text{det} \left\{ L_0 I_o(\theta_{\tau_e}) L_0^T \right\} \left[ \text{det} \left\{ C_0 I_o(\theta_{\tau_e})^{-1} C_0^T \right\} \right]^{-1}. \quad (2.36)
\]

Combining (2.35) and (2.36) gives

\[
\text{det} \left\{ L_0 I_o(\theta_{\tau_e}) L_0^T \right\} = \text{det} \left\{ C_0 I_o(\theta_{\tau_e})^{-1} C_0^T \right\} \left[ \text{det} \left\{ C_0 C_0^T \right\} \right]^{-1} \text{det} \left\{ I_o(\theta_{\tau_e}) \right\}. \quad (2.37)
\]

This means that, as long as \( C_0 \) is known, there is no need to know the form of \( L_0 \).

The form of \( C_0 \) can be obtained by differentiating (2.26), that is,

\[
\frac{\partial g(\theta_{\tau_e}, \tau_e)}{\partial \theta_{\tau_e}^T \tau_e} = \left[ 1 \quad \tau_e \quad \frac{\xi_{\tau_e}}{2\sigma_{\tau_e}} \right]_{\theta_{\tau_e} = \theta_{0,\tau_e}}.
\]

An obvious choice for \( \theta_{0,\tau_e} \) is the constrained maximum likelihood estimate \( \hat{\theta}_{\tau_e} \), such that

\[
C_0 = \left[ 1 \quad \tau_e \quad \frac{\xi_{\tau_e}}{2\sigma_{\tau_e}} \right].
\]

The approximate modified profile likelihood is now given by

\[
\tilde{L}_t(\tau_e; y) = L(\hat{\theta}_{\tau_e}; y) - \frac{1}{2} \log \left[ \text{det} \left\{ C_0 I_o(\hat{\theta}_{\tau_e})^{-1} C_0^T \right\} \left[ \text{det} \left\{ C_0 C_0^T \right\} \right]^{-1} \text{det} \left\{ I_o(\hat{\theta}_{\tau_e}) \right\} \right]
\]

43
from which \( \hat{\tau}_e \) can be estimated.

The generalised likelihood ratio statistic is given by

\[
  w_t(\tau_e) = 2 \left( \tilde{L}_t(\hat{\tau}_e; y) - \tilde{L}_t(\tau_e; y) \right),
\]

which is distributed approximately \( \chi^2_1 \). Consequently, the label shelf-life, based on the 100(1 - \( \alpha \))% lower confidence bound, is given by

\[
  \hat{\tau}_{e,L(\alpha)} = \inf \left\{ t : (t \leq \hat{\tau}_e) \& (w_t(t) \leq \chi^2_{1,1-2\alpha}) \right\}. \tag{2.38}
\]

Again, the expected information matrix can be used instead of the observed information matrix.

However, it does not matter which information matrix is used, as both give the same adjustment to the likelihood. This is the case if it can be shown that the observed information matrix can be expressed in a form that does not involve the data. To do this, first consider the score vector for the constrained problem which equals zero at the constrained maximum likelihood estimate. Hence it follows that the elements of the score can be written as

\[
  \frac{1}{\hat{\sigma}^2_{\tau_e}} X^T(y - X\hat{\beta}_{\tau_e}) = 2N\omega \left[ g(\hat{\theta}_{\tau_e}; \tau_e) + \hat{\psi}_{\tau_e} \right] x_{\tau_e}
\]

and

\[
  \frac{1}{2\hat{\sigma}^4_{\tau_e}} (y - X\hat{\beta}_{\tau_e})^T(y - X\hat{\beta}_{\tau_e}) = \frac{N}{2\hat{\sigma}^2_{\tau_e}} + \frac{N\omega z_{\tau_e}}{\hat{\sigma}_{\tau_e}} \left[ g(\hat{\theta}_{\tau_e}; \tau_e) + \hat{\psi}_{\tau_e} \right].
\]

The left hand sides of these two equations now also feature in the equivalents of the observed information matrix elements \( I_o'(\beta_{\tau_e}, \sigma^2_{\tau_e}) \) and \( I_o'(\sigma^2_{\tau_e}, \sigma^2_{\tau_e}) \). Hence these elements can be written at the constrained maximum likelihood estimates as

\[
  I_o'(\hat{\beta}_{\tau_e}, \hat{\sigma}^2_{\tau_e}) = \frac{2\omega}{\hat{\sigma}^2_{\tau_e}} \left[ g(\hat{\theta}_{\tau_e}; \tau_e) + \hat{\psi}_{\tau_e} \right] x_{\tau_e} + \frac{\omega z_{\tau_e}}{\hat{\sigma}_{\tau_e}} x_{\tau_e}
\]

\[
  I_o'(\hat{\sigma}^2_{\tau_e}, \hat{\sigma}^2_{\tau_e}) = \frac{N}{2\hat{\sigma}^4_{\tau_e}} + \frac{2N z_{\tau_e} \omega}{\hat{\sigma}^3_{\tau_e}} \left[ g(\hat{\theta}_{\tau_e}; \tau_e) + \hat{\psi}_{\tau_e} \right].
\]

This now proves that the Powell-Hestenes equivalent of the observed information matrix can be expressed in terms that do not involve the data. Thus, taking expectations will have no effect on the adjustment term.
### 2.2.4 Non-central t-distribution (NCT) Approach

As pointed out in Section 2.1, this approach was suggested by Easterling (1969). The exact approach given here serves as a gold standard against which the other methods given in this chapter can be compared. However, it is shown in Chapter 3 that this approach cannot generally be used in the case when multiple batches are used in the estimation of the label shelf-life. Thus it is specifically for the single batch case, and for multiple batches another approach is necessary.

The distribution of the least squares estimate of $\beta$ is given by

$$\hat{\beta} \sim N\{\beta, \sigma^2 (X^T X)^{-1}\} .$$

Consequently, the distribution of the estimated mean response at time $t$, $\hat{\mu}_t = x^T_t \hat{\beta}$, is

$$\hat{\mu}_t \sim N\{\mu_t, \sigma^2 x^T_t (X^T X)^{-1} x_t\} .$$

Standardizing $\hat{\mu}_t$ and using (2.1) to replace the mean $\mu_t$ by $q_{\tau,\epsilon}$, gives

$$\frac{\hat{\mu}_t - q_{\tau,\epsilon} + z_\epsilon \sigma}{\sigma \sqrt{x^T_t (X^T X)^{-1} x_t}} \sim N(0, 1) ,$$

or equivalently,

$$\frac{\hat{\mu}_t - q_{\tau,\epsilon}}{\sigma \sqrt{x^T_t (X^T X)^{-1} x_t}} \sim N\left\{\frac{-z_\epsilon}{\sqrt{x^T_t (X^T X)^{-1} x_t}}, 1\right\} .$$

Since the mean of this distribution does not involve $\sigma$, replacing $\sigma$ by $s$ results in

$$\frac{\hat{\mu}_t - q_{\tau,\epsilon}}{s \sqrt{x^T_t (X^T X)^{-1} x_t}} \sim t'_{N-2}(\delta) ,$$

where $t'_{N-2}(\delta)$ denotes a non-central t-distribution with $N - 2$ degrees of freedom and non-centrality parameter $\delta = -z_\epsilon / \sqrt{x^T_t (X^T X)^{-1} x_t}$.

An estimate of the shelf-life is given by

$$\hat{\tau}_\epsilon = \inf \{t : \hat{\mu}_t + z_\epsilon s \leq \kappa\} ,$$

and the label shelf-life is given by the intersection of the exact 100(1-\alpha)% lower confidence bound and the specification limit, that is,

$$\hat{\tau}_{\epsilon, L(\alpha)} = \inf \left\{t : \hat{\mu}_t + Q_\alpha s \sqrt{x^T_t (X^T X)^{-1} x_t} \leq \kappa\right\} .$$
2.3 Simulation Results

Since real pharmaceutical stability data is generally commercially sensitive, the approaches detailed in Section 2.2 have been evaluated with the help of a simulation study using the software package R (Ihaka and Gentleman, 1996).

The simulations were conducted as follows. A random sample of size $r$ was drawn at each of $n$ times $x_i = 0, 3, 6, 9, 12, 18, 24,$ and $36$ months from $N(x_i^T \beta, \sigma^2)$, where $x_i^T = (1, x_i)$ and $\beta^T = (100, \beta_1)$. This can be done without loss of generality as the approaches, except the approximate conditional profile likelihood approach, are invariant under linear transformations. The total number of observations available for analysis is $N = nr = 8r$. Values of $\beta_1 = -0.2, -0.3$ and $-0.4$, $\sigma = 0.5, 1$ and $2$, and $r = 1, 2, 5, 10, 20,$ and $50$ were used. For each combination of $\beta_1$, $\sigma$ and $r$, 1000 simulations were performed and the proportion of label shelf-lives less than the true shelf-life was determined for each method. The true mean degradation lines for each $\beta_1$ are shown in Figure 2.3.

Assuming that $\kappa = 90$ and $\epsilon = 0.01$, the true shelf-life (in months) is given by $\tau_{0.01} = (90 - \beta_0 - z_{0.01} \sigma)/\beta_1$. The true shelf-lives for each parameter combination of $\beta_1$ and $\sigma$ are given in Table 2.1. For each simulation the label shelf-life $\hat{\tau}_{0.01,L(0.05)}$, based on a significance level of $\alpha = 0.05$, was calculated.

For each approach the proportion of label shelf-lives that were less than the true shelf-life is tabled in Tables 2.2–2.10. In addition, the bias and the Mean Square Error (MSE) were calculated and tabulated in Tables 2.11–2.19 and Tables 2.20–2.28, respectively.
Figure 2.3: At each point shown, \( r \) values are drawn randomly from a normal distribution with these means and with a standard deviation of \( \sigma = 0.5, 1 \) and \( 2 \).

<table>
<thead>
<tr>
<th align="left">( \beta_1 )</th>
<th>( \sigma = 0.5 )</th>
<th>( \sigma = 1.0 )</th>
<th>( \sigma = 2.0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td align="left">( -0.4 )</td>
<td>22.1</td>
<td>19.2</td>
<td>13.4</td>
</tr>
<tr>
<td align="left">( -0.3 )</td>
<td>29.5</td>
<td>25.6</td>
<td>17.8</td>
</tr>
<tr>
<td align="left">( -0.2 )</td>
<td>44.2</td>
<td>38.4</td>
<td>26.7</td>
</tr>
</tbody>
</table>

Table 2.1: True shelf-lives for each parameter combination of \( \beta_1 \) and \( \sigma \).

The bias is calculated as the difference between the average of the label shelf-lives and the true shelf-life, and hence is in months. This “bias” is not a true bias in the strict sense as the label shelf-life is a lower confidence bound for, rather than an estimate of, the true shelf-life. However, it does indicate how conservative a particular method is. That is, a negative bias implies that the label shelf-lives are, on average, shorter than the true shelf-life, and hence that the corresponding approach is conservative, on average.

The MSE is calculated according to

\[
\text{MSE} = \text{Var}[\text{Label Shelf-life}] + \text{Bias}^2.
\]

These three measures can be used to assess the various approaches. Firstly, the coverage probability should be as close to the desired 0.95 as possible. Methods with similar
coverage probabilities can then be compared using the bias and the MSE. That is, a method with a small bias or a small MSE will be better than one with a larger bias or MSE, respectively.

For each simulation the label shelf-life can be either greater than the true shelf-life (failure), or less than or equal to the true shelf-life (success). Consequently, the number of successes are distributed as a binomial distribution $B(1000, 1 - \alpha)$. Using the normal approximation to the binomial, an approximate 95% confidence interval for the true coverage probability $p$ is given by $\hat{p} \pm 0.0135$, where $\hat{p}$ is given in the cells of Tables 2.2–2.10.

The results in Tables 2.2–2.10, Tables 2.11–2.19 and Tables 2.20 are graphically summarised in Figures 2.5–2.7, and a discussion of the results can be found in the following Section. Note that the methods are abbreviated in order to reduce the size of the tables and graphs. The following key can be used to relate the abbreviations to the appropriate method; some of these have already been introduced in the previous sections.

- $z.z$ and $z.t$ denote the normal approximation approaches (2.6), using $\delta$ given in (2.3). The confidence bounds are calculated using the normal or $t$ distribution, respectively.
- ML.$z$ and ML.$t$ denote the normal approximation approach (2.6), using $\delta$ given in (2.4). The confidence bounds are calculated using the normal or $t$ distribution, respectively.
- U.$z$ and U.$t$ denote the normal approximation approach (2.6), using $\delta$ given in (2.5). The confidence bounds are calculated using the normal or $t$ distribution, respectively.
- PL denotes the profile likelihood approach (2.21)
- MPL.$o$ and MPL.$e$ denote the modified profile likelihood approach (2.23) using the observed and expected information matrix, respectively.
- TMPL.$o$ and TMPL.$e$ denote the truncated modified profile likelihood approach using the observed and expected information matrix, respectively.
• ACPL.o and ACPL.e denote the approximate conditional profile likelihood approach (2.25) using the observed and expected information matrix, respectively.

• TMCPL denotes the truncated modified constrained profile likelihood approach (2.38).

• NCT denotes the non-central t-distribution approach (3.17).

Note that zero shelf-lives can result from each method. It was noted that some parameter combinations resulted in many more zero shelf-lives than others. For example, Figure 2.4 shows combinations of \( \hat{\beta}_0, \hat{\beta}_1 \) and \( \hat{\sigma} \) for the simulation results of the ACPL approach using \( I_e \) when \( \beta_1 = -0.4 \) and \( \sigma = 2 \). It identifies the combinations of \( \hat{\beta}_0, \hat{\beta}_1 \) and \( \hat{\sigma} \) which resulted in the approach giving a zero shelf-life with a (+). This happened more frequently as the variability increased, starting first for only low intercepts and small slopes, but extending to the range of intercepts and slopes. Similar patterns are observed for all other approaches, and in general, larger underlying variation and quicker degradation result in more zero shelf-lives. This is not surprising as it suggests that the product has already more than 100\% of product fall below the specification limit with 100(1 − α)% confidence at the time of manufacture.
2.3. SIMULATION RESULTS

Figure 2.4: Graph of the Slope versus the Intercept, split on the size of the standard deviation, indicating when the ACPL approach resulted in a zero shelf-life (+) and a non-zero shelf-life (o).

Table 2.2: Proportion of label shelf-lives that were smaller than the true shelf-life for parameter values $\beta_1 = -0.4$, $\sigma = 0.5$, $\epsilon = 0.01$ and $\alpha = 0.05$.  

<table>
<thead>
<tr>
<th></th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.z</td>
<td>0.846</td>
<td>0.895</td>
<td>0.913</td>
<td>0.912</td>
<td>0.939</td>
<td>0.939</td>
</tr>
<tr>
<td>z.t</td>
<td>0.878</td>
<td>0.908</td>
<td>0.919</td>
<td>0.916</td>
<td>0.940</td>
<td>0.939</td>
</tr>
<tr>
<td>ML.z</td>
<td>0.695</td>
<td>0.838</td>
<td>0.880</td>
<td>0.899</td>
<td>0.923</td>
<td>0.930</td>
</tr>
<tr>
<td>ML.t</td>
<td>0.756</td>
<td>0.848</td>
<td>0.887</td>
<td>0.900</td>
<td>0.924</td>
<td>0.930</td>
</tr>
<tr>
<td>U.z</td>
<td>0.861</td>
<td>0.902</td>
<td>0.922</td>
<td>0.917</td>
<td>0.940</td>
<td>0.940</td>
</tr>
<tr>
<td>U.t</td>
<td>0.892</td>
<td>0.922</td>
<td>0.928</td>
<td>0.918</td>
<td>0.940</td>
<td>0.940</td>
</tr>
<tr>
<td>PL</td>
<td>0.812</td>
<td>0.884</td>
<td>0.912</td>
<td>0.913</td>
<td>0.940</td>
<td>0.940</td>
</tr>
<tr>
<td>MPL.o</td>
<td>0.977</td>
<td>0.975</td>
<td>0.963</td>
<td>0.945</td>
<td>0.956</td>
<td>0.950</td>
</tr>
<tr>
<td>MPL.e</td>
<td>0.972</td>
<td>0.969</td>
<td>0.959</td>
<td>0.939</td>
<td>0.954</td>
<td>0.948</td>
</tr>
<tr>
<td>TMPL.o</td>
<td>0.976</td>
<td>0.974</td>
<td>0.963</td>
<td>0.944</td>
<td>0.956</td>
<td>0.950</td>
</tr>
<tr>
<td>TMPL.e</td>
<td>0.972</td>
<td>0.967</td>
<td>0.959</td>
<td>0.939</td>
<td>0.952</td>
<td>0.948</td>
</tr>
<tr>
<td>TMCPL</td>
<td>0.968</td>
<td>0.965</td>
<td>0.957</td>
<td>0.938</td>
<td>0.950</td>
<td>0.947</td>
</tr>
<tr>
<td>ACPL.o</td>
<td>0.941</td>
<td>0.958</td>
<td>0.952</td>
<td>0.936</td>
<td>0.948</td>
<td>0.946</td>
</tr>
<tr>
<td>ACPL.e</td>
<td>0.921</td>
<td>0.950</td>
<td>0.945</td>
<td>0.931</td>
<td>0.946</td>
<td>0.945</td>
</tr>
<tr>
<td>NCT</td>
<td>0.953</td>
<td>0.959</td>
<td>0.952</td>
<td>0.936</td>
<td>0.948</td>
<td>0.946</td>
</tr>
</tbody>
</table>
Table 2.3: Proportion of label shelf-lives that were smaller than the true shelf-life for parameter values $\beta_1 = -0.4$, $\sigma = 1.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>N</th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.z</td>
<td>0.842</td>
<td>0.899</td>
<td>0.910</td>
<td>0.931</td>
<td>0.935</td>
<td>0.938</td>
</tr>
<tr>
<td>z.t</td>
<td>0.876</td>
<td>0.911</td>
<td>0.917</td>
<td>0.936</td>
<td>0.937</td>
<td>0.938</td>
</tr>
<tr>
<td>ML.z</td>
<td>0.713</td>
<td>0.817</td>
<td>0.874</td>
<td>0.918</td>
<td>0.921</td>
<td>0.931</td>
</tr>
<tr>
<td>ML.t</td>
<td>0.764</td>
<td>0.843</td>
<td>0.882</td>
<td>0.919</td>
<td>0.921</td>
<td>0.931</td>
</tr>
<tr>
<td>U.z</td>
<td>0.864</td>
<td>0.910</td>
<td>0.918</td>
<td>0.936</td>
<td>0.939</td>
<td>0.938</td>
</tr>
<tr>
<td>U.t</td>
<td>0.890</td>
<td>0.916</td>
<td>0.922</td>
<td>0.939</td>
<td>0.940</td>
<td>0.938</td>
</tr>
<tr>
<td>PL</td>
<td>0.811</td>
<td>0.891</td>
<td>0.910</td>
<td>0.931</td>
<td>0.936</td>
<td>0.938</td>
</tr>
<tr>
<td>MPL.o</td>
<td>0.961</td>
<td>0.974</td>
<td>0.954</td>
<td>0.961</td>
<td>0.957</td>
<td>0.950</td>
</tr>
<tr>
<td>MPL.e</td>
<td>0.951</td>
<td>0.969</td>
<td>0.949</td>
<td>0.959</td>
<td>0.955</td>
<td>0.947</td>
</tr>
<tr>
<td>TMPL.o</td>
<td>0.962</td>
<td>0.974</td>
<td>0.954</td>
<td>0.960</td>
<td>0.957</td>
<td>0.949</td>
</tr>
<tr>
<td>TMPL.e</td>
<td>0.955</td>
<td>0.969</td>
<td>0.949</td>
<td>0.959</td>
<td>0.955</td>
<td>0.947</td>
</tr>
<tr>
<td>TMCPL</td>
<td>0.945</td>
<td>0.967</td>
<td>0.944</td>
<td>0.952</td>
<td>0.955</td>
<td>0.945</td>
</tr>
<tr>
<td>ACPL.o</td>
<td>0.922</td>
<td>0.963</td>
<td>0.944</td>
<td>0.950</td>
<td>0.954</td>
<td>0.945</td>
</tr>
<tr>
<td>ACPL.e</td>
<td>0.915</td>
<td>0.952</td>
<td>0.933</td>
<td>0.947</td>
<td>0.952</td>
<td>0.942</td>
</tr>
<tr>
<td>NCT</td>
<td>0.932</td>
<td>0.964</td>
<td>0.944</td>
<td>0.950</td>
<td>0.954</td>
<td>0.945</td>
</tr>
</tbody>
</table>

Table 2.4: Proportion of label shelf-lives that were smaller than the true shelf-life for parameter values $\beta_1 = -0.4$, $\sigma = 2.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>N</th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.z</td>
<td>0.853</td>
<td>0.897</td>
<td>0.921</td>
<td>0.918</td>
<td>0.946</td>
<td>0.949</td>
</tr>
<tr>
<td>z.t</td>
<td>0.883</td>
<td>0.908</td>
<td>0.924</td>
<td>0.920</td>
<td>0.948</td>
<td>0.949</td>
</tr>
<tr>
<td>ML.z</td>
<td>0.711</td>
<td>0.819</td>
<td>0.882</td>
<td>0.899</td>
<td>0.930</td>
<td>0.944</td>
</tr>
<tr>
<td>ML.t</td>
<td>0.770</td>
<td>0.835</td>
<td>0.890</td>
<td>0.902</td>
<td>0.931</td>
<td>0.944</td>
</tr>
<tr>
<td>U.z</td>
<td>0.872</td>
<td>0.907</td>
<td>0.925</td>
<td>0.923</td>
<td>0.949</td>
<td>0.949</td>
</tr>
<tr>
<td>U.t</td>
<td>0.905</td>
<td>0.919</td>
<td>0.929</td>
<td>0.923</td>
<td>0.951</td>
<td>0.950</td>
</tr>
<tr>
<td>PL</td>
<td>0.838</td>
<td>0.888</td>
<td>0.921</td>
<td>0.919</td>
<td>0.948</td>
<td>0.949</td>
</tr>
<tr>
<td>MPL.o</td>
<td>0.979</td>
<td>0.969</td>
<td>0.963</td>
<td>0.954</td>
<td>0.967</td>
<td>0.955</td>
</tr>
<tr>
<td>MPL.e</td>
<td>0.973</td>
<td>0.968</td>
<td>0.961</td>
<td>0.950</td>
<td>0.966</td>
<td>0.955</td>
</tr>
<tr>
<td>TMPL.o</td>
<td>0.981</td>
<td>0.969</td>
<td>0.963</td>
<td>0.953</td>
<td>0.966</td>
<td>0.955</td>
</tr>
<tr>
<td>TMPL.e</td>
<td>0.975</td>
<td>0.967</td>
<td>0.961</td>
<td>0.949</td>
<td>0.965</td>
<td>0.954</td>
</tr>
<tr>
<td>TMCPL</td>
<td>0.957</td>
<td>0.946</td>
<td>0.951</td>
<td>0.941</td>
<td>0.959</td>
<td>0.952</td>
</tr>
<tr>
<td>ACPL.o</td>
<td>0.948</td>
<td>0.949</td>
<td>0.957</td>
<td>0.944</td>
<td>0.964</td>
<td>0.953</td>
</tr>
<tr>
<td>ACPL.e</td>
<td>0.938</td>
<td>0.941</td>
<td>0.950</td>
<td>0.941</td>
<td>0.959</td>
<td>0.952</td>
</tr>
<tr>
<td>NCT</td>
<td>0.963</td>
<td>0.956</td>
<td>0.957</td>
<td>0.946</td>
<td>0.964</td>
<td>0.953</td>
</tr>
</tbody>
</table>
### Table 2.5: Proportion of label shelf-lives that were smaller than the true shelf-life for parameter values $\beta_1 = -0.3$, $\sigma = 0.5$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th></th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.z</td>
<td>0.848</td>
<td>0.890</td>
<td>0.926</td>
<td>0.927</td>
<td>0.942</td>
<td>0.944</td>
</tr>
<tr>
<td>z.t</td>
<td>0.880</td>
<td>0.903</td>
<td>0.931</td>
<td>0.928</td>
<td>0.943</td>
<td>0.944</td>
</tr>
<tr>
<td>ML.z</td>
<td>0.725</td>
<td>0.830</td>
<td>0.898</td>
<td>0.909</td>
<td>0.924</td>
<td>0.940</td>
</tr>
<tr>
<td>ML.t</td>
<td>0.768</td>
<td>0.853</td>
<td>0.904</td>
<td>0.914</td>
<td>0.925</td>
<td>0.940</td>
</tr>
<tr>
<td>U.z</td>
<td>0.863</td>
<td>0.900</td>
<td>0.931</td>
<td>0.929</td>
<td>0.945</td>
<td>0.944</td>
</tr>
<tr>
<td>U.t</td>
<td>0.890</td>
<td>0.913</td>
<td>0.937</td>
<td>0.929</td>
<td>0.945</td>
<td>0.946</td>
</tr>
<tr>
<td>PL</td>
<td>0.822</td>
<td>0.886</td>
<td>0.928</td>
<td>0.928</td>
<td>0.944</td>
<td>0.944</td>
</tr>
<tr>
<td>MPL.o</td>
<td>0.969</td>
<td>0.968</td>
<td>0.971</td>
<td>0.952</td>
<td>0.960</td>
<td>0.950</td>
</tr>
<tr>
<td>MPL.e</td>
<td>0.961</td>
<td>0.961</td>
<td>0.967</td>
<td>0.949</td>
<td>0.957</td>
<td>0.949</td>
</tr>
<tr>
<td>TMPL.o</td>
<td>0.971</td>
<td>0.969</td>
<td>0.971</td>
<td>0.952</td>
<td>0.960</td>
<td>0.950</td>
</tr>
<tr>
<td>TMPL.e</td>
<td>0.961</td>
<td>0.961</td>
<td>0.966</td>
<td>0.949</td>
<td>0.957</td>
<td>0.949</td>
</tr>
<tr>
<td>TMCPL</td>
<td>0.958</td>
<td>0.950</td>
<td>0.962</td>
<td>0.947</td>
<td>0.957</td>
<td>0.949</td>
</tr>
<tr>
<td>ACPL.o</td>
<td>0.938</td>
<td>0.954</td>
<td>0.961</td>
<td>0.947</td>
<td>0.957</td>
<td>0.949</td>
</tr>
<tr>
<td>ACPL.e</td>
<td>0.927</td>
<td>0.933</td>
<td>0.949</td>
<td>0.943</td>
<td>0.953</td>
<td>0.948</td>
</tr>
<tr>
<td>NCT</td>
<td>0.938</td>
<td>0.951</td>
<td>0.957</td>
<td>0.945</td>
<td>0.957</td>
<td>0.948</td>
</tr>
</tbody>
</table>

### Table 2.6: Proportion of label shelf-lives that were smaller than the true shelf-life for parameter values $\beta_1 = -0.3$, $\sigma = 1.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th></th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.z</td>
<td>0.850</td>
<td>0.890</td>
<td>0.912</td>
<td>0.933</td>
<td>0.921</td>
<td>0.947</td>
</tr>
<tr>
<td>z.t</td>
<td>0.880</td>
<td>0.911</td>
<td>0.917</td>
<td>0.934</td>
<td>0.922</td>
<td>0.947</td>
</tr>
<tr>
<td>ML.z</td>
<td>0.707</td>
<td>0.833</td>
<td>0.886</td>
<td>0.909</td>
<td>0.910</td>
<td>0.940</td>
</tr>
<tr>
<td>ML.t</td>
<td>0.759</td>
<td>0.852</td>
<td>0.890</td>
<td>0.910</td>
<td>0.911</td>
<td>0.940</td>
</tr>
<tr>
<td>U.z</td>
<td>0.870</td>
<td>0.905</td>
<td>0.918</td>
<td>0.935</td>
<td>0.922</td>
<td>0.947</td>
</tr>
<tr>
<td>U.t</td>
<td>0.894</td>
<td>0.922</td>
<td>0.922</td>
<td>0.936</td>
<td>0.923</td>
<td>0.947</td>
</tr>
<tr>
<td>PL</td>
<td>0.829</td>
<td>0.885</td>
<td>0.912</td>
<td>0.934</td>
<td>0.922</td>
<td>0.947</td>
</tr>
<tr>
<td>MPL.o</td>
<td>0.971</td>
<td>0.973</td>
<td>0.963</td>
<td>0.954</td>
<td>0.951</td>
<td>0.963</td>
</tr>
<tr>
<td>MPL.e</td>
<td>0.965</td>
<td>0.963</td>
<td>0.954</td>
<td>0.948</td>
<td>0.942</td>
<td>0.960</td>
</tr>
<tr>
<td>TMPL.o</td>
<td>0.977</td>
<td>0.976</td>
<td>0.966</td>
<td>0.955</td>
<td>0.952</td>
<td>0.963</td>
</tr>
<tr>
<td>TMPL.e</td>
<td>0.969</td>
<td>0.967</td>
<td>0.955</td>
<td>0.949</td>
<td>0.944</td>
<td>0.961</td>
</tr>
<tr>
<td>TMCPL</td>
<td>0.962</td>
<td>0.950</td>
<td>0.950</td>
<td>0.944</td>
<td>0.941</td>
<td>0.959</td>
</tr>
<tr>
<td>ACPL.o</td>
<td>0.941</td>
<td>0.958</td>
<td>0.950</td>
<td>0.945</td>
<td>0.941</td>
<td>0.959</td>
</tr>
<tr>
<td>ACPL.e</td>
<td>0.924</td>
<td>0.939</td>
<td>0.940</td>
<td>0.941</td>
<td>0.936</td>
<td>0.954</td>
</tr>
<tr>
<td>NCT</td>
<td>0.941</td>
<td>0.951</td>
<td>0.947</td>
<td>0.944</td>
<td>0.941</td>
<td>0.958</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>16</td>
<td>40</td>
<td>80</td>
<td>160</td>
<td>400</td>
</tr>
<tr>
<td>---</td>
<td>------</td>
<td>------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>z.t</td>
<td>0.773</td>
<td>0.819</td>
<td>0.881</td>
<td>0.911</td>
<td>0.907</td>
<td>0.951</td>
</tr>
<tr>
<td>U.t</td>
<td>0.873</td>
<td>0.889</td>
<td>0.915</td>
<td>0.931</td>
<td>0.927</td>
<td>0.961</td>
</tr>
<tr>
<td>PL</td>
<td>0.902</td>
<td>0.908</td>
<td>0.920</td>
<td>0.933</td>
<td>0.930</td>
<td>0.961</td>
</tr>
<tr>
<td>MPL.o</td>
<td>0.967</td>
<td>0.958</td>
<td>0.956</td>
<td>0.959</td>
<td>0.948</td>
<td>0.965</td>
</tr>
<tr>
<td>MPL.e</td>
<td>0.963</td>
<td>0.953</td>
<td>0.951</td>
<td>0.953</td>
<td>0.945</td>
<td>0.964</td>
</tr>
<tr>
<td>TMPL.o</td>
<td>0.976</td>
<td>0.961</td>
<td>0.962</td>
<td>0.959</td>
<td>0.950</td>
<td>0.965</td>
</tr>
<tr>
<td>TMPL.e</td>
<td>0.972</td>
<td>0.958</td>
<td>0.954</td>
<td>0.954</td>
<td>0.946</td>
<td>0.965</td>
</tr>
<tr>
<td>TMCPL</td>
<td>0.958</td>
<td>0.941</td>
<td>0.941</td>
<td>0.945</td>
<td>0.937</td>
<td>0.963</td>
</tr>
<tr>
<td>ACPL.o</td>
<td>0.950</td>
<td>0.943</td>
<td>0.943</td>
<td>0.949</td>
<td>0.939</td>
<td>0.963</td>
</tr>
<tr>
<td>ACPL.e</td>
<td>0.945</td>
<td>0.938</td>
<td>0.940</td>
<td>0.944</td>
<td>0.937</td>
<td>0.963</td>
</tr>
<tr>
<td>NCT</td>
<td>0.957</td>
<td>0.943</td>
<td>0.943</td>
<td>0.949</td>
<td>0.939</td>
<td>0.963</td>
</tr>
</tbody>
</table>

Table 2.7: Proportion of label shelf-lives that were smaller than the true shelf-life for parameter values $\beta_1 = -0.3$, $\sigma = 2.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th></th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.t</td>
<td>0.877</td>
<td>0.901</td>
<td>0.922</td>
<td>0.941</td>
<td>0.938</td>
<td>0.951</td>
</tr>
<tr>
<td>U.t</td>
<td>0.909</td>
<td>0.917</td>
<td>0.929</td>
<td>0.944</td>
<td>0.938</td>
<td>0.952</td>
</tr>
<tr>
<td>PL</td>
<td>0.786</td>
<td>0.862</td>
<td>0.903</td>
<td>0.929</td>
<td>0.933</td>
<td>0.946</td>
</tr>
<tr>
<td>MPL.o</td>
<td>0.839</td>
<td>0.877</td>
<td>0.907</td>
<td>0.932</td>
<td>0.934</td>
<td>0.946</td>
</tr>
<tr>
<td>MPL.e</td>
<td>0.891</td>
<td>0.909</td>
<td>0.925</td>
<td>0.944</td>
<td>0.941</td>
<td>0.952</td>
</tr>
<tr>
<td>TMPL.o</td>
<td>0.920</td>
<td>0.923</td>
<td>0.932</td>
<td>0.945</td>
<td>0.943</td>
<td>0.953</td>
</tr>
<tr>
<td>TMPL.e</td>
<td>0.859</td>
<td>0.897</td>
<td>0.922</td>
<td>0.943</td>
<td>0.938</td>
<td>0.952</td>
</tr>
<tr>
<td>TMCPL</td>
<td>0.976</td>
<td>0.972</td>
<td>0.955</td>
<td>0.963</td>
<td>0.958</td>
<td>0.961</td>
</tr>
<tr>
<td>ACPL.o</td>
<td>0.965</td>
<td>0.956</td>
<td>0.950</td>
<td>0.959</td>
<td>0.953</td>
<td>0.959</td>
</tr>
<tr>
<td>ACPL.e</td>
<td>0.960</td>
<td>0.975</td>
<td>0.957</td>
<td>0.963</td>
<td>0.959</td>
<td>0.961</td>
</tr>
<tr>
<td>NCT</td>
<td>0.951</td>
<td>0.950</td>
<td>0.941</td>
<td>0.955</td>
<td>0.951</td>
<td>0.959</td>
</tr>
</tbody>
</table>

Table 2.8: Proportion of label shelf-lives that were smaller than the true shelf-life for parameter values $\beta_1 = -0.2$, $\sigma = 0.5$, $\epsilon = 0.01$ and $\alpha = 0.05$. 
2.5. SIMULATION RESULTS

Table 2.9: Proportion of label shelf-lives that were smaller than the true shelf-life for parameter values $\beta_1 = -0.2$, $\sigma = 1.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th></th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.z</td>
<td>0.848</td>
<td>0.898</td>
<td>0.925</td>
<td>0.934</td>
<td>0.939</td>
<td>0.946</td>
</tr>
<tr>
<td>z.t</td>
<td>0.882</td>
<td>0.910</td>
<td>0.932</td>
<td>0.936</td>
<td>0.939</td>
<td>0.946</td>
</tr>
<tr>
<td>ML.z</td>
<td>0.759</td>
<td>0.850</td>
<td>0.902</td>
<td>0.920</td>
<td>0.928</td>
<td>0.942</td>
</tr>
<tr>
<td>ML.t</td>
<td>0.811</td>
<td>0.862</td>
<td>0.908</td>
<td>0.921</td>
<td>0.929</td>
<td>0.942</td>
</tr>
<tr>
<td>U.z</td>
<td>0.869</td>
<td>0.907</td>
<td>0.931</td>
<td>0.936</td>
<td>0.939</td>
<td>0.946</td>
</tr>
<tr>
<td>U.t</td>
<td>0.896</td>
<td>0.917</td>
<td>0.936</td>
<td>0.937</td>
<td>0.939</td>
<td>0.946</td>
</tr>
<tr>
<td>PL</td>
<td>0.835</td>
<td>0.897</td>
<td>0.925</td>
<td>0.936</td>
<td>0.939</td>
<td>0.946</td>
</tr>
<tr>
<td>MPL.o</td>
<td>0.959</td>
<td>0.960</td>
<td>0.965</td>
<td>0.955</td>
<td>0.951</td>
<td>0.954</td>
</tr>
<tr>
<td>MPL.e</td>
<td>0.944</td>
<td>0.945</td>
<td>0.955</td>
<td>0.947</td>
<td>0.946</td>
<td>0.951</td>
</tr>
<tr>
<td>TMPL.o</td>
<td>0.972</td>
<td>0.970</td>
<td>0.968</td>
<td>0.961</td>
<td>0.955</td>
<td>0.956</td>
</tr>
<tr>
<td>TMPL.e</td>
<td>0.961</td>
<td>0.959</td>
<td>0.965</td>
<td>0.955</td>
<td>0.950</td>
<td>0.954</td>
</tr>
<tr>
<td>TMCPL</td>
<td>0.948</td>
<td>0.948</td>
<td>0.958</td>
<td>0.947</td>
<td>0.946</td>
<td>0.951</td>
</tr>
<tr>
<td>ACPL.o</td>
<td>0.945</td>
<td>0.948</td>
<td>0.960</td>
<td>0.951</td>
<td>0.948</td>
<td>0.952</td>
</tr>
<tr>
<td>ACPL.e</td>
<td>0.922</td>
<td>0.935</td>
<td>0.949</td>
<td>0.943</td>
<td>0.944</td>
<td>0.949</td>
</tr>
<tr>
<td>NCT</td>
<td>0.940</td>
<td>0.944</td>
<td>0.954</td>
<td>0.946</td>
<td>0.945</td>
<td>0.949</td>
</tr>
</tbody>
</table>

Table 2.10: Proportion of label shelf-lives that were smaller than the true shelf-life for parameter values $\beta_1 = -0.2$, $\sigma = 2.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th></th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.z</td>
<td>0.855</td>
<td>0.893</td>
<td>0.915</td>
<td>0.932</td>
<td>0.923</td>
<td>0.937</td>
</tr>
<tr>
<td>z.t</td>
<td>0.886</td>
<td>0.911</td>
<td>0.924</td>
<td>0.933</td>
<td>0.925</td>
<td>0.937</td>
</tr>
<tr>
<td>ML.z</td>
<td>0.741</td>
<td>0.825</td>
<td>0.895</td>
<td>0.914</td>
<td>0.909</td>
<td>0.932</td>
</tr>
<tr>
<td>ML.t</td>
<td>0.791</td>
<td>0.843</td>
<td>0.899</td>
<td>0.917</td>
<td>0.912</td>
<td>0.932</td>
</tr>
<tr>
<td>U.z</td>
<td>0.876</td>
<td>0.907</td>
<td>0.926</td>
<td>0.933</td>
<td>0.928</td>
<td>0.937</td>
</tr>
<tr>
<td>U.t</td>
<td>0.898</td>
<td>0.921</td>
<td>0.933</td>
<td>0.939</td>
<td>0.928</td>
<td>0.937</td>
</tr>
<tr>
<td>PL</td>
<td>0.832</td>
<td>0.891</td>
<td>0.916</td>
<td>0.933</td>
<td>0.925</td>
<td>0.937</td>
</tr>
<tr>
<td>MPL.o</td>
<td>0.958</td>
<td>0.959</td>
<td>0.955</td>
<td>0.953</td>
<td>0.950</td>
<td>0.946</td>
</tr>
<tr>
<td>MPL.e</td>
<td>0.945</td>
<td>0.951</td>
<td>0.948</td>
<td>0.949</td>
<td>0.942</td>
<td>0.943</td>
</tr>
<tr>
<td>TMPL.o</td>
<td>0.985</td>
<td>0.976</td>
<td>0.970</td>
<td>0.959</td>
<td>0.956</td>
<td>0.951</td>
</tr>
<tr>
<td>TMPL.e</td>
<td>0.980</td>
<td>0.972</td>
<td>0.959</td>
<td>0.956</td>
<td>0.955</td>
<td>0.948</td>
</tr>
<tr>
<td>TMCPL</td>
<td>0.959</td>
<td>0.954</td>
<td>0.949</td>
<td>0.949</td>
<td>0.940</td>
<td>0.942</td>
</tr>
<tr>
<td>ACPL.o</td>
<td>0.962</td>
<td>0.960</td>
<td>0.953</td>
<td>0.951</td>
<td>0.945</td>
<td>0.944</td>
</tr>
<tr>
<td>ACPL.e</td>
<td>0.951</td>
<td>0.950</td>
<td>0.946</td>
<td>0.948</td>
<td>0.937</td>
<td>0.942</td>
</tr>
<tr>
<td>NCT</td>
<td>0.948</td>
<td>0.953</td>
<td>0.949</td>
<td>0.949</td>
<td>0.942</td>
<td>0.943</td>
</tr>
</tbody>
</table>
Figure 2.5: Coverage probabilities plotted against $1/\sqrt{N}$ for each method and combination of parameters.
### Table 2.11: Bias (in months) of label shelf-lives for parameter values $\beta_1 = -0.4$, $\sigma = 0.5$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th></th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z.z$</td>
<td>-1.36</td>
<td>-1.06</td>
<td>-0.65</td>
<td>-0.45</td>
<td>-0.34</td>
<td>-0.21</td>
</tr>
<tr>
<td>$z.t$</td>
<td>-1.64</td>
<td>-1.14</td>
<td>-0.67</td>
<td>-0.46</td>
<td>-0.34</td>
<td>-0.21</td>
</tr>
<tr>
<td>ML.z</td>
<td>-0.67</td>
<td>-0.77</td>
<td>-0.55</td>
<td>-0.41</td>
<td>-0.32</td>
<td>-0.20</td>
</tr>
<tr>
<td>ML.t</td>
<td>-0.89</td>
<td>-0.84</td>
<td>-0.57</td>
<td>-0.41</td>
<td>-0.32</td>
<td>-0.20</td>
</tr>
<tr>
<td>U.z</td>
<td>-1.53</td>
<td>-1.13</td>
<td>-0.67</td>
<td>-0.46</td>
<td>-0.34</td>
<td>-0.21</td>
</tr>
<tr>
<td>U.t</td>
<td>-1.81</td>
<td>-1.20</td>
<td>-0.69</td>
<td>-0.47</td>
<td>-0.34</td>
<td>-0.21</td>
</tr>
<tr>
<td>PL</td>
<td>-1.19</td>
<td>-1.03</td>
<td>-0.65</td>
<td>-0.46</td>
<td>-0.34</td>
<td>-0.21</td>
</tr>
<tr>
<td>MPL.o</td>
<td>-3.29</td>
<td>-1.72</td>
<td>-0.86</td>
<td>-0.55</td>
<td>-0.38</td>
<td>-0.23</td>
</tr>
<tr>
<td>MPL.e</td>
<td>-3.08</td>
<td>-1.62</td>
<td>-0.82</td>
<td>-0.53</td>
<td>-0.37</td>
<td>-0.23</td>
</tr>
<tr>
<td>TMPL.o</td>
<td>-3.35</td>
<td>-1.71</td>
<td>-0.85</td>
<td>-0.55</td>
<td>-0.38</td>
<td>-0.23</td>
</tr>
<tr>
<td>TMPL.e</td>
<td>-3.12</td>
<td>-1.62</td>
<td>-0.82</td>
<td>-0.53</td>
<td>-0.37</td>
<td>-0.22</td>
</tr>
<tr>
<td>TMCPL</td>
<td>-2.85</td>
<td>-1.56</td>
<td>-0.80</td>
<td>-0.52</td>
<td>-0.37</td>
<td>-0.22</td>
</tr>
<tr>
<td>ACPL.o</td>
<td>-2.43</td>
<td>-1.48</td>
<td>-0.79</td>
<td>-0.52</td>
<td>-0.37</td>
<td>-0.22</td>
</tr>
<tr>
<td>ACPL.e</td>
<td>-2.23</td>
<td>-1.39</td>
<td>-0.76</td>
<td>-0.50</td>
<td>-0.36</td>
<td>-0.22</td>
</tr>
<tr>
<td>NCT</td>
<td>-2.61</td>
<td>-1.50</td>
<td>-0.79</td>
<td>-0.52</td>
<td>-0.37</td>
<td>-0.22</td>
</tr>
</tbody>
</table>

### Table 2.12: Bias (in months) of label shelf-lives for parameter values $\beta_1 = -0.4$, $\sigma = 1.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th></th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z.z$</td>
<td>-2.80</td>
<td>-1.99</td>
<td>-1.23</td>
<td>-0.91</td>
<td>-0.64</td>
<td>-0.40</td>
</tr>
<tr>
<td>$z.t$</td>
<td>-3.35</td>
<td>-2.14</td>
<td>-1.26</td>
<td>-0.92</td>
<td>-0.64</td>
<td>-0.40</td>
</tr>
<tr>
<td>ML.z</td>
<td>-1.39</td>
<td>-1.40</td>
<td>-1.03</td>
<td>-0.82</td>
<td>-0.60</td>
<td>-0.39</td>
</tr>
<tr>
<td>ML.t</td>
<td>-1.81</td>
<td>-1.53</td>
<td>-1.06</td>
<td>-0.83</td>
<td>-0.60</td>
<td>-0.39</td>
</tr>
<tr>
<td>U.z</td>
<td>-3.15</td>
<td>-2.12</td>
<td>-1.28</td>
<td>-0.93</td>
<td>-0.65</td>
<td>-0.41</td>
</tr>
<tr>
<td>U.t</td>
<td>-3.71</td>
<td>-2.27</td>
<td>-1.31</td>
<td>-0.94</td>
<td>-0.65</td>
<td>-0.41</td>
</tr>
<tr>
<td>PL</td>
<td>-2.43</td>
<td>-1.91</td>
<td>-1.23</td>
<td>-0.91</td>
<td>-0.64</td>
<td>-0.40</td>
</tr>
<tr>
<td>MPL.o</td>
<td>-6.92</td>
<td>-3.26</td>
<td>-1.63</td>
<td>-1.09</td>
<td>-0.73</td>
<td>-0.44</td>
</tr>
<tr>
<td>MPL.e</td>
<td>-6.41</td>
<td>-3.10</td>
<td>-1.57</td>
<td>-1.07</td>
<td>-0.71</td>
<td>-0.43</td>
</tr>
<tr>
<td>TMPL.o</td>
<td>-6.73</td>
<td>-3.26</td>
<td>-1.62</td>
<td>-1.09</td>
<td>-0.72</td>
<td>-0.43</td>
</tr>
<tr>
<td>TMPL.e</td>
<td>-6.28</td>
<td>-3.10</td>
<td>-1.57</td>
<td>-1.06</td>
<td>-0.71</td>
<td>-0.43</td>
</tr>
<tr>
<td>TMCPL</td>
<td>-5.08</td>
<td>-2.86</td>
<td>-1.51</td>
<td>-1.04</td>
<td>-0.70</td>
<td>-0.43</td>
</tr>
<tr>
<td>ACPL.o</td>
<td>-5.00</td>
<td>-2.80</td>
<td>-1.50</td>
<td>-1.04</td>
<td>-0.70</td>
<td>-0.43</td>
</tr>
<tr>
<td>ACPL.e</td>
<td>-4.62</td>
<td>-2.65</td>
<td>-1.44</td>
<td>-1.01</td>
<td>-0.69</td>
<td>-0.42</td>
</tr>
<tr>
<td>NCT</td>
<td>-5.37</td>
<td>-2.87</td>
<td>-1.51</td>
<td>-1.04</td>
<td>-0.70</td>
<td>-0.43</td>
</tr>
</tbody>
</table>
### Table 2.13: Bias (in months) of label shelf-lives for parameter values $\beta_1 = -0.4$, $\sigma = 2.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th></th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.z</td>
<td>-5.83</td>
<td>-4.06</td>
<td>-2.49</td>
<td>-1.73</td>
<td>-1.29</td>
<td>-0.83</td>
</tr>
<tr>
<td>z.t</td>
<td>-6.78</td>
<td>-4.37</td>
<td>-2.56</td>
<td>-1.75</td>
<td>-1.30</td>
<td>-0.83</td>
</tr>
<tr>
<td>ML.z</td>
<td>-3.13</td>
<td>-2.82</td>
<td>-2.08</td>
<td>-1.54</td>
<td>-1.20</td>
<td>-0.79</td>
</tr>
<tr>
<td>ML.t</td>
<td>-4.01</td>
<td>-3.10</td>
<td>-2.14</td>
<td>-1.56</td>
<td>-1.21</td>
<td>-0.79</td>
</tr>
<tr>
<td>U.z</td>
<td>-6.41</td>
<td>-4.33</td>
<td>-2.58</td>
<td>-1.77</td>
<td>-1.31</td>
<td>-0.83</td>
</tr>
<tr>
<td>U.t</td>
<td>-7.34</td>
<td>-4.64</td>
<td>-2.65</td>
<td>-1.79</td>
<td>-1.32</td>
<td>-0.84</td>
</tr>
<tr>
<td>PL</td>
<td>-5.19</td>
<td>-3.90</td>
<td>-2.48</td>
<td>-1.73</td>
<td>-1.29</td>
<td>-0.83</td>
</tr>
<tr>
<td>MPL.o</td>
<td>-11.00</td>
<td>-6.99</td>
<td>-3.36</td>
<td>-2.11</td>
<td>-1.46</td>
<td>-0.89</td>
</tr>
<tr>
<td>MPL.e</td>
<td>-10.51</td>
<td>-6.54</td>
<td>-3.24</td>
<td>-2.06</td>
<td>-1.44</td>
<td>-0.88</td>
</tr>
<tr>
<td>TMPL.o</td>
<td>-10.45</td>
<td>-6.49</td>
<td>-3.26</td>
<td>-2.07</td>
<td>-1.45</td>
<td>-0.89</td>
</tr>
<tr>
<td>TMPL.e</td>
<td>-9.90</td>
<td>-6.08</td>
<td>-3.14</td>
<td>-2.02</td>
<td>-1.43</td>
<td>-0.88</td>
</tr>
<tr>
<td>TMCPL</td>
<td>-6.45</td>
<td>-4.59</td>
<td>-2.79</td>
<td>-1.89</td>
<td>-1.37</td>
<td>-0.86</td>
</tr>
<tr>
<td>ACPL.o</td>
<td>-9.26</td>
<td>-5.83</td>
<td>-3.05</td>
<td>-1.98</td>
<td>-1.41</td>
<td>-0.87</td>
</tr>
<tr>
<td>ACPL.e</td>
<td>-8.65</td>
<td>-5.44</td>
<td>-2.94</td>
<td>-1.93</td>
<td>-1.38</td>
<td>-0.86</td>
</tr>
<tr>
<td>NCT</td>
<td>-9.40</td>
<td>-5.86</td>
<td>-3.07</td>
<td>-1.99</td>
<td>-1.41</td>
<td>-0.87</td>
</tr>
</tbody>
</table>

### Table 2.14: Bias (in months) of label shelf-lives for parameter values $\beta_1 = -0.3$, $\sigma = 0.5$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th></th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.z</td>
<td>-2.00</td>
<td>-1.48</td>
<td>-0.99</td>
<td>-0.69</td>
<td>-0.50</td>
<td>-0.33</td>
</tr>
<tr>
<td>z.t</td>
<td>-2.39</td>
<td>-1.59</td>
<td>-1.02</td>
<td>-0.70</td>
<td>-0.50</td>
<td>-0.33</td>
</tr>
<tr>
<td>ML.z</td>
<td>-1.08</td>
<td>-1.09</td>
<td>-0.86</td>
<td>-0.63</td>
<td>-0.47</td>
<td>-0.32</td>
</tr>
<tr>
<td>ML.t</td>
<td>-1.40</td>
<td>-1.19</td>
<td>-0.88</td>
<td>-0.64</td>
<td>-0.47</td>
<td>-0.32</td>
</tr>
<tr>
<td>U.z</td>
<td>-2.21</td>
<td>-1.56</td>
<td>-1.02</td>
<td>-0.71</td>
<td>-0.51</td>
<td>-0.33</td>
</tr>
<tr>
<td>U.t</td>
<td>-2.61</td>
<td>-1.67</td>
<td>-1.05</td>
<td>-0.72</td>
<td>-0.51</td>
<td>-0.33</td>
</tr>
<tr>
<td>PL</td>
<td>-1.78</td>
<td>-1.44</td>
<td>-0.99</td>
<td>-0.70</td>
<td>-0.50</td>
<td>-0.33</td>
</tr>
<tr>
<td>MPL.o</td>
<td>-4.63</td>
<td>-2.41</td>
<td>-1.29</td>
<td>-0.83</td>
<td>-0.56</td>
<td>-0.35</td>
</tr>
<tr>
<td>MPL.e</td>
<td>-4.23</td>
<td>-2.23</td>
<td>-1.23</td>
<td>-0.80</td>
<td>-0.55</td>
<td>-0.35</td>
</tr>
<tr>
<td>TMPL.o</td>
<td>-4.93</td>
<td>-2.44</td>
<td>-1.29</td>
<td>-0.83</td>
<td>-0.56</td>
<td>-0.35</td>
</tr>
<tr>
<td>TMPL.e</td>
<td>-4.49</td>
<td>-2.25</td>
<td>-1.23</td>
<td>-0.80</td>
<td>-0.55</td>
<td>-0.35</td>
</tr>
<tr>
<td>TMCPL</td>
<td>-4.09</td>
<td>-2.17</td>
<td>-1.21</td>
<td>-0.79</td>
<td>-0.55</td>
<td>-0.34</td>
</tr>
<tr>
<td>ACPL.o</td>
<td>-3.57</td>
<td>-2.10</td>
<td>-1.20</td>
<td>-0.79</td>
<td>-0.55</td>
<td>-0.34</td>
</tr>
<tr>
<td>ACPL.e</td>
<td>-3.19</td>
<td>-1.92</td>
<td>-1.14</td>
<td>-0.76</td>
<td>-0.53</td>
<td>-0.34</td>
</tr>
<tr>
<td>NCT</td>
<td>-3.64</td>
<td>-2.06</td>
<td>-1.18</td>
<td>-0.78</td>
<td>-0.54</td>
<td>-0.34</td>
</tr>
</tbody>
</table>

57
### 2.3. Simulation Results

<table>
<thead>
<tr>
<th></th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.z</td>
<td>-3.62</td>
<td>-2.81</td>
<td>-1.80</td>
<td>-1.27</td>
<td>-0.90</td>
<td>-0.61</td>
</tr>
<tr>
<td>z.t</td>
<td>-4.34</td>
<td>-3.00</td>
<td>-1.85</td>
<td>-1.28</td>
<td>-0.91</td>
<td>-0.61</td>
</tr>
<tr>
<td>ML.z</td>
<td>-1.83</td>
<td>-2.04</td>
<td>-1.54</td>
<td>-1.15</td>
<td>-0.84</td>
<td>-0.59</td>
</tr>
<tr>
<td>ML.t</td>
<td>-2.41</td>
<td>-2.22</td>
<td>-1.59</td>
<td>-1.16</td>
<td>-0.85</td>
<td>-0.59</td>
</tr>
<tr>
<td>U.z</td>
<td>-4.05</td>
<td>-2.97</td>
<td>-1.86</td>
<td>-1.29</td>
<td>-0.91</td>
<td>-0.61</td>
</tr>
<tr>
<td>U.t</td>
<td>-4.80</td>
<td>-3.17</td>
<td>-1.90</td>
<td>-1.31</td>
<td>-0.92</td>
<td>-0.61</td>
</tr>
<tr>
<td>PL</td>
<td>-3.17</td>
<td>-2.71</td>
<td>-1.80</td>
<td>-1.27</td>
<td>-0.91</td>
<td>-0.61</td>
</tr>
<tr>
<td>MPL.o</td>
<td>-8.53</td>
<td>-4.45</td>
<td>-2.34</td>
<td>-1.52</td>
<td>-1.02</td>
<td>-0.65</td>
</tr>
<tr>
<td>MPL.e</td>
<td>-7.98</td>
<td>-4.20</td>
<td>-2.24</td>
<td>-1.47</td>
<td>-1.00</td>
<td>-0.65</td>
</tr>
<tr>
<td>TMPL.o</td>
<td>-9.11</td>
<td>-4.61</td>
<td>-2.37</td>
<td>-1.53</td>
<td>-1.03</td>
<td>-0.66</td>
</tr>
<tr>
<td>TMPL.e</td>
<td>-8.59</td>
<td>-4.35</td>
<td>-2.27</td>
<td>-1.48</td>
<td>-1.00</td>
<td>-0.65</td>
</tr>
<tr>
<td>TMCPL</td>
<td>-6.93</td>
<td>-4.01</td>
<td>-2.19</td>
<td>-1.45</td>
<td>-0.99</td>
<td>-0.64</td>
</tr>
<tr>
<td>ACPL.o</td>
<td>-6.66</td>
<td>-3.94</td>
<td>-2.19</td>
<td>-1.45</td>
<td>-0.99</td>
<td>-0.64</td>
</tr>
<tr>
<td>ACPL.e</td>
<td>-6.13</td>
<td>-3.69</td>
<td>-2.09</td>
<td>-1.40</td>
<td>-0.97</td>
<td>-0.63</td>
</tr>
<tr>
<td>NCT</td>
<td>-6.90</td>
<td>-3.95</td>
<td>-2.17</td>
<td>-1.44</td>
<td>-0.98</td>
<td>-0.64</td>
</tr>
</tbody>
</table>

Table 2.15: Bias (in months) of label shelf-lives for parameter values $\beta_1 = -0.3$, $\sigma = 1.0$, $\epsilon = 0.01$ and $\alpha = 0.05$. 

<table>
<thead>
<tr>
<th></th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.z</td>
<td>-7.24</td>
<td>-5.08</td>
<td>-3.29</td>
<td>-2.39</td>
<td>-1.64</td>
<td>-1.11</td>
</tr>
<tr>
<td>z.t</td>
<td>-8.43</td>
<td>-5.48</td>
<td>-3.37</td>
<td>-2.42</td>
<td>-1.65</td>
<td>-1.11</td>
</tr>
<tr>
<td>ML.z</td>
<td>-3.65</td>
<td>-3.47</td>
<td>-2.75</td>
<td>-2.14</td>
<td>-1.53</td>
<td>-1.07</td>
</tr>
<tr>
<td>ML.t</td>
<td>-4.78</td>
<td>-3.83</td>
<td>-2.83</td>
<td>-2.17</td>
<td>-1.54</td>
<td>-1.07</td>
</tr>
<tr>
<td>U.z</td>
<td>-8.03</td>
<td>-5.43</td>
<td>-3.41</td>
<td>-2.44</td>
<td>-1.67</td>
<td>-1.12</td>
</tr>
<tr>
<td>U.t</td>
<td>-9.19</td>
<td>-5.84</td>
<td>-3.49</td>
<td>-2.47</td>
<td>-1.68</td>
<td>-1.12</td>
</tr>
<tr>
<td>PL</td>
<td>-6.38</td>
<td>-4.86</td>
<td>-3.27</td>
<td>-2.40</td>
<td>-1.65</td>
<td>-1.11</td>
</tr>
<tr>
<td>MPL.o</td>
<td>-14.37</td>
<td>-8.73</td>
<td>-4.36</td>
<td>-2.86</td>
<td>-1.86</td>
<td>-1.19</td>
</tr>
<tr>
<td>MPL.e</td>
<td>-13.77</td>
<td>-8.21</td>
<td>-4.21</td>
<td>-2.80</td>
<td>-1.83</td>
<td>-1.18</td>
</tr>
<tr>
<td>TMPL.o</td>
<td>-14.07</td>
<td>-8.40</td>
<td>-4.34</td>
<td>-2.86</td>
<td>-1.87</td>
<td>-1.20</td>
</tr>
<tr>
<td>TMPL.e</td>
<td>-13.37</td>
<td>-7.92</td>
<td>-4.20</td>
<td>-2.80</td>
<td>-1.84</td>
<td>-1.18</td>
</tr>
<tr>
<td>TMCPL</td>
<td>-8.58</td>
<td>-5.97</td>
<td>-3.74</td>
<td>-2.62</td>
<td>-1.76</td>
<td>-1.16</td>
</tr>
<tr>
<td>ACPL.o</td>
<td>-12.58</td>
<td>-7.55</td>
<td>-4.04</td>
<td>-2.73</td>
<td>-1.80</td>
<td>-1.17</td>
</tr>
<tr>
<td>ACPL.e</td>
<td>-11.75</td>
<td>-7.06</td>
<td>-3.90</td>
<td>-2.66</td>
<td>-1.77</td>
<td>-1.16</td>
</tr>
<tr>
<td>NCT</td>
<td>-12.29</td>
<td>-7.46</td>
<td>-4.04</td>
<td>-2.73</td>
<td>-1.80</td>
<td>-1.17</td>
</tr>
</tbody>
</table>

Table 2.16: Bias (in months) of label shelf-lives for parameter values $\beta_1 = -0.3$, $\sigma = 2.0$, $\epsilon = 0.01$ and $\alpha = 0.05$. 

58
### Table 2.17: Bias (in months) of label shelf-lives for parameter values $\beta_1 = -0.2$, $\sigma = 0.5$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>N</th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.z</td>
<td>-4.01</td>
<td>-3.05</td>
<td>-1.90</td>
<td>-1.47</td>
<td>-1.05</td>
<td>-0.70</td>
</tr>
<tr>
<td>z.t</td>
<td>-4.73</td>
<td>-3.26</td>
<td>-1.95</td>
<td>-1.48</td>
<td>-1.05</td>
<td>-0.70</td>
</tr>
<tr>
<td>ML.z</td>
<td>-2.60</td>
<td>-2.45</td>
<td>-1.69</td>
<td>-1.37</td>
<td>-1.00</td>
<td>-0.68</td>
</tr>
<tr>
<td>ML.t</td>
<td>-3.23</td>
<td>-2.65</td>
<td>-1.74</td>
<td>-1.39</td>
<td>-1.01</td>
<td>-0.68</td>
</tr>
<tr>
<td>U.z</td>
<td>-4.29</td>
<td>-3.16</td>
<td>-1.94</td>
<td>-1.49</td>
<td>-1.06</td>
<td>-0.70</td>
</tr>
<tr>
<td>U.t</td>
<td>-5.03</td>
<td>-3.37</td>
<td>-1.99</td>
<td>-1.50</td>
<td>-1.06</td>
<td>-0.70</td>
</tr>
<tr>
<td>PL</td>
<td>-3.72</td>
<td>-3.00</td>
<td>-1.91</td>
<td>-1.48</td>
<td>-1.05</td>
<td>-0.70</td>
</tr>
<tr>
<td>MPL.o</td>
<td>-8.35</td>
<td>-4.58</td>
<td>-2.39</td>
<td>-1.69</td>
<td>-1.16</td>
<td>-0.74</td>
</tr>
<tr>
<td>MPL.e</td>
<td>-7.32</td>
<td>-4.16</td>
<td>-2.26</td>
<td>-1.63</td>
<td>-1.12</td>
<td>-0.73</td>
</tr>
<tr>
<td>TMPL.o</td>
<td>-9.52</td>
<td>-4.78</td>
<td>-2.44</td>
<td>-1.71</td>
<td>-1.16</td>
<td>-0.74</td>
</tr>
<tr>
<td>TMPL.e</td>
<td>-8.31</td>
<td>-4.34</td>
<td>-2.30</td>
<td>-1.65</td>
<td>-1.13</td>
<td>-0.73</td>
</tr>
<tr>
<td>TMCPL</td>
<td>-7.48</td>
<td>-4.16</td>
<td>-2.25</td>
<td>-1.62</td>
<td>-1.12</td>
<td>-0.73</td>
</tr>
<tr>
<td>ACPL.o</td>
<td>-6.89</td>
<td>-4.14</td>
<td>-2.27</td>
<td>-1.64</td>
<td>-1.13</td>
<td>-0.73</td>
</tr>
<tr>
<td>ACPL.e</td>
<td>-5.94</td>
<td>-3.75</td>
<td>-2.13</td>
<td>-1.58</td>
<td>-1.10</td>
<td>-0.72</td>
</tr>
<tr>
<td>NCT</td>
<td>-6.52</td>
<td>-3.94</td>
<td>-2.19</td>
<td>-1.60</td>
<td>-1.11</td>
<td>-0.72</td>
</tr>
</tbody>
</table>

### Table 2.18: Bias (in months) of label shelf-lives for parameter values $\beta_1 = -0.2$, $\sigma = 1.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>N</th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.z</td>
<td>-6.31</td>
<td>-4.87</td>
<td>-3.36</td>
<td>-2.49</td>
<td>-1.73</td>
<td>-1.14</td>
</tr>
<tr>
<td>z.t</td>
<td>-7.48</td>
<td>-5.21</td>
<td>-3.44</td>
<td>-2.52</td>
<td>-1.74</td>
<td>-1.14</td>
</tr>
<tr>
<td>ML.z</td>
<td>-3.73</td>
<td>-3.78</td>
<td>-2.98</td>
<td>-2.31</td>
<td>-1.64</td>
<td>-1.11</td>
</tr>
<tr>
<td>ML.t</td>
<td>-4.73</td>
<td>-4.09</td>
<td>-3.06</td>
<td>-2.34</td>
<td>-1.65</td>
<td>-1.11</td>
</tr>
<tr>
<td>U.z</td>
<td>-6.89</td>
<td>-5.09</td>
<td>-3.44</td>
<td>-2.53</td>
<td>-1.75</td>
<td>-1.15</td>
</tr>
<tr>
<td>U.t</td>
<td>-8.07</td>
<td>-5.44</td>
<td>-3.52</td>
<td>-2.56</td>
<td>-1.76</td>
<td>-1.15</td>
</tr>
<tr>
<td>PL</td>
<td>-5.73</td>
<td>-4.77</td>
<td>-3.37</td>
<td>-2.51</td>
<td>-1.74</td>
<td>-1.15</td>
</tr>
<tr>
<td>MPL.o</td>
<td>-12.94</td>
<td>-7.33</td>
<td>-4.20</td>
<td>-2.88</td>
<td>-1.91</td>
<td>-1.21</td>
</tr>
<tr>
<td>MPL.e</td>
<td>-11.82</td>
<td>-6.72</td>
<td>-3.97</td>
<td>-2.77</td>
<td>-1.86</td>
<td>-1.19</td>
</tr>
<tr>
<td>TMPL.o</td>
<td>-16.14</td>
<td>-8.09</td>
<td>-4.40</td>
<td>-2.97</td>
<td>-1.95</td>
<td>-1.23</td>
</tr>
<tr>
<td>TMPL.e</td>
<td>-15.11</td>
<td>-7.45</td>
<td>-4.17</td>
<td>-2.86</td>
<td>-1.90</td>
<td>-1.21</td>
</tr>
<tr>
<td>TMCPL</td>
<td>-11.95</td>
<td>-6.82</td>
<td>-3.99</td>
<td>-2.78</td>
<td>-1.86</td>
<td>-1.19</td>
</tr>
<tr>
<td>ACPL.o</td>
<td>-11.70</td>
<td>-6.83</td>
<td>-4.04</td>
<td>-2.81</td>
<td>-1.88</td>
<td>-1.20</td>
</tr>
<tr>
<td>ACPL.e</td>
<td>-10.38</td>
<td>-6.23</td>
<td>-3.82</td>
<td>-2.70</td>
<td>-1.83</td>
<td>-1.18</td>
</tr>
<tr>
<td>NCT</td>
<td>-10.97</td>
<td>-6.51</td>
<td>-3.91</td>
<td>-2.75</td>
<td>-1.85</td>
<td>-1.19</td>
</tr>
</tbody>
</table>
### Table 2.19: Bias (in months) of label shelf-lives for parameter values $\beta_1 = -0.2$, $\sigma = 2.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th></th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.z</td>
<td>-10.88</td>
<td>-7.94</td>
<td>-5.09</td>
<td>-3.74</td>
<td>-2.69</td>
<td>-1.79</td>
</tr>
<tr>
<td>z.t</td>
<td>-12.65</td>
<td>-8.55</td>
<td>-5.21</td>
<td>-3.79</td>
<td>-2.70</td>
<td>-1.79</td>
</tr>
<tr>
<td>ML.z</td>
<td>-5.67</td>
<td>-5.60</td>
<td>-4.32</td>
<td>-3.39</td>
<td>-2.52</td>
<td>-1.72</td>
</tr>
<tr>
<td>ML.t</td>
<td>-7.81</td>
<td>-6.14</td>
<td>-4.45</td>
<td>-3.44</td>
<td>-2.54</td>
<td>-1.73</td>
</tr>
<tr>
<td>U.z</td>
<td>-11.88</td>
<td>-8.46</td>
<td>-5.25</td>
<td>-3.82</td>
<td>-2.73</td>
<td>-1.80</td>
</tr>
<tr>
<td>U.t</td>
<td>-13.53</td>
<td>-9.08</td>
<td>-5.38</td>
<td>-3.87</td>
<td>-2.74</td>
<td>-1.80</td>
</tr>
<tr>
<td>PL</td>
<td>-9.25</td>
<td>-7.56</td>
<td>-5.08</td>
<td>-3.76</td>
<td>-2.70</td>
<td>-1.79</td>
</tr>
<tr>
<td>MPL.o</td>
<td>-19.81</td>
<td>-12.60</td>
<td>-6.44</td>
<td>-4.37</td>
<td>-2.99</td>
<td>-1.90</td>
</tr>
<tr>
<td>MPL.e</td>
<td>-18.82</td>
<td>-11.84</td>
<td>-6.19</td>
<td>-4.24</td>
<td>-2.93</td>
<td>-1.88</td>
</tr>
<tr>
<td>TMPL.o</td>
<td>-21.87</td>
<td>-13.71</td>
<td>-6.89</td>
<td>-4.57</td>
<td>-3.09</td>
<td>-1.94</td>
</tr>
<tr>
<td>TMPL.e</td>
<td>-21.25</td>
<td>-13.07</td>
<td>-6.66</td>
<td>-4.45</td>
<td>-3.03</td>
<td>-1.92</td>
</tr>
<tr>
<td>TMCPL</td>
<td>-13.86</td>
<td>-9.76</td>
<td>-5.92</td>
<td>-4.15</td>
<td>-2.89</td>
<td>-1.87</td>
</tr>
<tr>
<td>ACPL.o</td>
<td>-20.50</td>
<td>-12.59</td>
<td>-6.29</td>
<td>-4.29</td>
<td>-2.95</td>
<td>-1.89</td>
</tr>
<tr>
<td>ACPL.e</td>
<td>-19.58</td>
<td>-11.79</td>
<td>-6.03</td>
<td>-4.16</td>
<td>-2.88</td>
<td>-1.86</td>
</tr>
<tr>
<td>NCT</td>
<td>-17.97</td>
<td>-11.45</td>
<td>-6.17</td>
<td>-4.24</td>
<td>-2.92</td>
<td>-1.88</td>
</tr>
</tbody>
</table>
Figure 2.6: Bias (in months) plotted against $1/\sqrt{N}$ for each method and combination of parameters.
## 2.3. Simulation Results

<table>
<thead>
<tr>
<th>N</th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.z</td>
<td>3.66</td>
<td>1.81</td>
<td>0.64</td>
<td>0.31</td>
<td>0.16</td>
<td>0.06</td>
</tr>
<tr>
<td>z.t</td>
<td>4.67</td>
<td>2.00</td>
<td>0.66</td>
<td>0.32</td>
<td>0.17</td>
<td>0.06</td>
</tr>
<tr>
<td>ML.z</td>
<td>1.80</td>
<td>1.20</td>
<td>0.51</td>
<td>0.27</td>
<td>0.15</td>
<td>0.06</td>
</tr>
<tr>
<td>ML.t</td>
<td>2.26</td>
<td>1.32</td>
<td>0.53</td>
<td>0.28</td>
<td>0.15</td>
<td>0.06</td>
</tr>
<tr>
<td>U.z</td>
<td>4.26</td>
<td>1.97</td>
<td>0.67</td>
<td>0.32</td>
<td>0.17</td>
<td>0.06</td>
</tr>
<tr>
<td>U.t</td>
<td>5.41</td>
<td>2.17</td>
<td>0.70</td>
<td>0.33</td>
<td>0.17</td>
<td>0.06</td>
</tr>
<tr>
<td>PL</td>
<td>3.08</td>
<td>1.72</td>
<td>0.64</td>
<td>0.32</td>
<td>0.17</td>
<td>0.06</td>
</tr>
<tr>
<td>MPL.o</td>
<td>14.16</td>
<td>3.81</td>
<td>0.97</td>
<td>0.41</td>
<td>0.20</td>
<td>0.07</td>
</tr>
<tr>
<td>MPL.e</td>
<td>12.69</td>
<td>3.47</td>
<td>0.91</td>
<td>0.40</td>
<td>0.19</td>
<td>0.07</td>
</tr>
<tr>
<td>TMPL.o</td>
<td>14.64</td>
<td>3.80</td>
<td>0.96</td>
<td>0.41</td>
<td>0.20</td>
<td>0.07</td>
</tr>
<tr>
<td>TMPL.e</td>
<td>13.07</td>
<td>3.45</td>
<td>0.90</td>
<td>0.39</td>
<td>0.19</td>
<td>0.07</td>
</tr>
<tr>
<td>TMCPL</td>
<td>10.96</td>
<td>3.25</td>
<td>0.88</td>
<td>0.39</td>
<td>0.19</td>
<td>0.07</td>
</tr>
<tr>
<td>ACPL.o</td>
<td>8.48</td>
<td>2.97</td>
<td>0.85</td>
<td>0.38</td>
<td>0.19</td>
<td>0.07</td>
</tr>
<tr>
<td>ACPL.e</td>
<td>7.45</td>
<td>2.69</td>
<td>0.80</td>
<td>0.36</td>
<td>0.18</td>
<td>0.07</td>
</tr>
<tr>
<td>NCT</td>
<td>9.64</td>
<td>3.06</td>
<td>0.86</td>
<td>0.38</td>
<td>0.19</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Table 2.20: Mean Square Error of label shelf-lives for parameter values $\beta_1 = -0.4$, $\sigma = 0.5$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>N</th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.z</td>
<td>15.70</td>
<td>6.55</td>
<td>2.33</td>
<td>1.17</td>
<td>0.59</td>
<td>0.23</td>
</tr>
<tr>
<td>z.t</td>
<td>20.00</td>
<td>7.23</td>
<td>2.42</td>
<td>1.19</td>
<td>0.59</td>
<td>0.23</td>
</tr>
<tr>
<td>ML.z</td>
<td>7.66</td>
<td>4.25</td>
<td>1.85</td>
<td>1.00</td>
<td>0.53</td>
<td>0.22</td>
</tr>
<tr>
<td>ML.t</td>
<td>9.56</td>
<td>4.69</td>
<td>1.91</td>
<td>1.02</td>
<td>0.54</td>
<td>0.22</td>
</tr>
<tr>
<td>U.z</td>
<td>18.34</td>
<td>7.16</td>
<td>2.45</td>
<td>1.21</td>
<td>0.60</td>
<td>0.23</td>
</tr>
<tr>
<td>U.t</td>
<td>23.26</td>
<td>7.89</td>
<td>2.54</td>
<td>1.23</td>
<td>0.61</td>
<td>0.23</td>
</tr>
<tr>
<td>PL</td>
<td>13.14</td>
<td>6.19</td>
<td>2.32</td>
<td>1.17</td>
<td>0.59</td>
<td>0.23</td>
</tr>
<tr>
<td>MPL.o</td>
<td>66.91</td>
<td>13.94</td>
<td>3.53</td>
<td>1.54</td>
<td>0.71</td>
<td>0.26</td>
</tr>
<tr>
<td>MPL.e</td>
<td>57.61</td>
<td>12.86</td>
<td>3.35</td>
<td>1.49</td>
<td>0.69</td>
<td>0.25</td>
</tr>
<tr>
<td>TMPL.o</td>
<td>60.90</td>
<td>13.85</td>
<td>3.51</td>
<td>1.54</td>
<td>0.71</td>
<td>0.26</td>
</tr>
<tr>
<td>TMPL.e</td>
<td>53.48</td>
<td>12.77</td>
<td>3.33</td>
<td>1.48</td>
<td>0.69</td>
<td>0.25</td>
</tr>
<tr>
<td>TMCPL</td>
<td>34.87</td>
<td>11.02</td>
<td>3.14</td>
<td>1.43</td>
<td>0.67</td>
<td>0.25</td>
</tr>
<tr>
<td>ACPL.o</td>
<td>37.53</td>
<td>10.83</td>
<td>3.10</td>
<td>1.42</td>
<td>0.67</td>
<td>0.25</td>
</tr>
<tr>
<td>ACPL.e</td>
<td>32.84</td>
<td>9.94</td>
<td>2.94</td>
<td>1.36</td>
<td>0.65</td>
<td>0.24</td>
</tr>
<tr>
<td>NCT</td>
<td>41.88</td>
<td>11.32</td>
<td>3.14</td>
<td>1.43</td>
<td>0.67</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Table 2.21: Mean Square Error of label shelf-lives for parameter values $\beta_1 = -0.4$, $\sigma = 1.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.  

62
### 2.3. Simulation Results

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.z</td>
<td>60.67</td>
<td>28.38</td>
<td>9.43</td>
<td>4.50</td>
<td>2.37</td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td>z.t</td>
<td>72.68</td>
<td>31.43</td>
<td>9.79</td>
<td>4.58</td>
<td>2.39</td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td>ML.z</td>
<td>32.83</td>
<td>18.28</td>
<td>7.37</td>
<td>3.86</td>
<td>2.15</td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td>ML.t</td>
<td>41.13</td>
<td>20.33</td>
<td>7.66</td>
<td>3.93</td>
<td>2.17</td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td>U.z</td>
<td>67.62</td>
<td>30.98</td>
<td>9.93</td>
<td>4.66</td>
<td>2.43</td>
<td>0.94</td>
<td></td>
</tr>
<tr>
<td>U.t</td>
<td>80.19</td>
<td>34.24</td>
<td>10.30</td>
<td>4.74</td>
<td>2.45</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>PL</td>
<td>53.40</td>
<td>26.95</td>
<td>9.37</td>
<td>4.52</td>
<td>2.39</td>
<td>0.94</td>
<td></td>
</tr>
<tr>
<td>MPL.o</td>
<td>137.46</td>
<td>64.65</td>
<td>15.00</td>
<td>6.04</td>
<td>2.88</td>
<td>1.05</td>
<td></td>
</tr>
<tr>
<td>MPL.e</td>
<td>128.83</td>
<td>57.84</td>
<td>14.11</td>
<td>5.82</td>
<td>2.81</td>
<td>1.03</td>
<td></td>
</tr>
<tr>
<td>TMPL.o</td>
<td>126.03</td>
<td>55.90</td>
<td>14.17</td>
<td>5.87</td>
<td>2.83</td>
<td>1.04</td>
<td></td>
</tr>
<tr>
<td>TMPL.e</td>
<td>116.23</td>
<td>50.10</td>
<td>13.34</td>
<td>5.65</td>
<td>2.76</td>
<td>1.02</td>
<td></td>
</tr>
<tr>
<td>TMCPL</td>
<td>53.56</td>
<td>29.18</td>
<td>10.79</td>
<td>5.04</td>
<td>2.58</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td>ACPL.o</td>
<td>109.37</td>
<td>48.34</td>
<td>12.82</td>
<td>5.49</td>
<td>2.71</td>
<td>1.01</td>
<td></td>
</tr>
<tr>
<td>ACPL.e</td>
<td>99.21</td>
<td>43.11</td>
<td>12.06</td>
<td>5.28</td>
<td>2.64</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>NCT</td>
<td>110.23</td>
<td>48.15</td>
<td>12.89</td>
<td>5.52</td>
<td>2.72</td>
<td>1.01</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.22: Mean Square Error of label shelf-lives for parameter values $\beta_1 = -0.4$, $\sigma = 2.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.z</td>
<td>7.88</td>
<td>3.64</td>
<td>1.47</td>
<td>0.69</td>
<td>0.35</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>z.t</td>
<td>9.93</td>
<td>4.00</td>
<td>1.52</td>
<td>0.70</td>
<td>0.36</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>ML.z</td>
<td>4.30</td>
<td>2.52</td>
<td>1.21</td>
<td>0.60</td>
<td>0.32</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>ML.t</td>
<td>5.31</td>
<td>2.77</td>
<td>1.25</td>
<td>0.61</td>
<td>0.33</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>U.z</td>
<td>8.98</td>
<td>3.92</td>
<td>1.53</td>
<td>0.71</td>
<td>0.36</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>U.t</td>
<td>11.27</td>
<td>4.31</td>
<td>1.58</td>
<td>0.72</td>
<td>0.36</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>PL</td>
<td>6.86</td>
<td>3.50</td>
<td>1.47</td>
<td>0.69</td>
<td>0.36</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>MPL.o</td>
<td>27.71</td>
<td>7.54</td>
<td>2.18</td>
<td>0.90</td>
<td>0.42</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>MPL.e</td>
<td>24.06</td>
<td>6.64</td>
<td>2.02</td>
<td>0.85</td>
<td>0.41</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>TMPL.o</td>
<td>31.27</td>
<td>7.73</td>
<td>2.18</td>
<td>0.90</td>
<td>0.42</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>TMPL.e</td>
<td>27.08</td>
<td>6.79</td>
<td>2.02</td>
<td>0.85</td>
<td>0.41</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>TMCPL</td>
<td>22.59</td>
<td>6.37</td>
<td>1.96</td>
<td>0.84</td>
<td>0.40</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>ACPL.o</td>
<td>18.12</td>
<td>6.04</td>
<td>1.95</td>
<td>0.84</td>
<td>0.40</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>ACPL.e</td>
<td>15.33</td>
<td>5.29</td>
<td>1.80</td>
<td>0.79</td>
<td>0.39</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>NCT</td>
<td>18.92</td>
<td>5.90</td>
<td>1.90</td>
<td>0.82</td>
<td>0.40</td>
<td>0.16</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.23: Mean Square Error of label shelf-lives for parameter values $\beta_1 = -0.3$, $\sigma = 0.5$, $\epsilon = 0.01$ and $\alpha = 0.05$.  
63
\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline
N   & 8   & 16  & 40  & 80  & 160 & 400 \\
\hline
z.z & 26.38 & 12.83 & 4.88 & 2.38 & 1.18 & 0.52 \\
\hline
z.t & 33.36 & 14.10 & 5.05 & 2.42 & 1.19 & 0.52 \\
\hline
ML.z & 13.58 & 8.65 & 3.96 & 2.08 & 1.08 & 0.49 \\
\hline
ML.t & 16.72 & 9.50 & 4.10 & 2.11 & 1.09 & 0.49 \\
\hline
U.z & 30.58 & 13.92 & 5.10 & 2.45 & 1.21 & 0.52 \\
\hline
U.t & 38.55 & 15.27 & 5.28 & 2.49 & 1.22 & 0.52 \\
\hline
PL   & 22.38 & 12.24 & 4.87 & 2.40 & 1.19 & 0.52 \\
\hline
MPL.o & 99.98 & 25.72 & 7.19 & 3.09 & 1.42 & 0.57 \\
\hline
MPL.e & 88.87 & 23.54 & 6.74 & 2.95 & 1.37 & 0.56 \\
\hline
TMPL.o & 108.04 & 27.27 & 7.35 & 3.13 & 1.43 & 0.58 \\
\hline
TMPL.e & 97.49 & 24.97 & 6.89 & 2.99 & 1.38 & 0.56 \\
\hline
TMCPL & 63.49 & 21.56 & 6.49 & 2.88 & 1.35 & 0.56 \\
\hline
ACPL.o & 64.93 & 21.14 & 6.48 & 2.88 & 1.35 & 0.56 \\
\hline
ACPL.e & 57.15 & 19.20 & 6.06 & 2.75 & 1.31 & 0.55 \\
\hline
NCT    & 68.50 & 21.35 & 6.42 & 2.85 & 1.34 & 0.55 \\
\hline
\end{tabular}
\caption{Mean Square Error of label shelf-lives for parameter values $\beta_1 = -0.3$, $\sigma = 1.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.}
\end{table}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline
N   & 8   & 16  & 40  & 80  & 160 & 400 \\
\hline
z.z & 97.92 & 46.93 & 16.79 & 8.23 & 3.99 & 1.67 \\
\hline
z.t & 119.34 & 51.92 & 17.40 & 8.38 & 4.03 & 1.68 \\
\hline
ML.z & 52.04 & 30.38 & 13.24 & 7.06 & 3.62 & 1.58 \\
\hline
ML.t & 67.91 & 33.65 & 13.74 & 7.18 & 3.65 & 1.58 \\
\hline
U.z & 110.05 & 51.28 & 17.65 & 8.51 & 4.08 & 1.70 \\
\hline
U.t & 132.84 & 56.63 & 18.30 & 8.66 & 4.12 & 1.70 \\
\hline
PL   & 85.48 & 44.46 & 16.67 & 8.26 & 4.02 & 1.68 \\
\hline
MPL.o & 239.72 & 105.45 & 25.79 & 10.85 & 4.80 & 1.87 \\
\hline
MPL.e & 225.07 & 95.30 & 24.43 & 10.48 & 4.68 & 1.84 \\
\hline
TMPL.o & 226.46 & 94.86 & 25.26 & 10.82 & 4.80 & 1.87 \\
\hline
TMPL.e & 208.33 & 85.74 & 23.96 & 10.45 & 4.68 & 1.84 \\
\hline
TMCPL & 92.73 & 50.11 & 19.55 & 9.33 & 4.38 & 1.78 \\
\hline
ACPL.o & 199.47 & 83.45 & 22.73 & 10.04 & 4.56 & 1.81 \\
\hline
ACPL.e & 180.27 & 74.82 & 21.51 & 9.68 & 4.45 & 1.79 \\
\hline
NCT    & 189.55 & 80.54 & 22.77 & 10.07 & 4.57 & 1.82 \\
\hline
\end{tabular}
\caption{Mean Square Error of label shelf-lives for parameter values $\beta_1 = -0.3$, $\sigma = 2.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.}
\end{table}
### Table 2.26: Mean Square Error of label shelf-lives for parameter values $\beta_1 = -0.2, \sigma = 0.5, \epsilon = 0.01$ and $\alpha = 0.05$. 

<table>
<thead>
<tr>
<th>N</th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.z</td>
<td>28.43</td>
<td>14.58</td>
<td>5.33</td>
<td>3.02</td>
<td>1.56</td>
<td>0.67</td>
</tr>
<tr>
<td>z.t</td>
<td>35.30</td>
<td>15.95</td>
<td>5.53</td>
<td>3.07</td>
<td>1.57</td>
<td>0.67</td>
</tr>
<tr>
<td>ML.z</td>
<td>17.79</td>
<td>11.10</td>
<td>4.58</td>
<td>2.75</td>
<td>1.47</td>
<td>0.65</td>
</tr>
<tr>
<td>ML.t</td>
<td>21.76</td>
<td>12.14</td>
<td>4.75</td>
<td>2.80</td>
<td>1.48</td>
<td>0.65</td>
</tr>
<tr>
<td>U.z</td>
<td>31.21</td>
<td>15.34</td>
<td>5.49</td>
<td>3.08</td>
<td>1.58</td>
<td>0.67</td>
</tr>
<tr>
<td>U.t</td>
<td>38.63</td>
<td>16.77</td>
<td>5.69</td>
<td>3.13</td>
<td>1.59</td>
<td>0.68</td>
</tr>
<tr>
<td>PL</td>
<td>25.85</td>
<td>14.28</td>
<td>5.37</td>
<td>3.05</td>
<td>1.57</td>
<td>0.67</td>
</tr>
<tr>
<td>MPL.o</td>
<td>86.32</td>
<td>26.80</td>
<td>7.52</td>
<td>3.75</td>
<td>1.80</td>
<td>0.73</td>
</tr>
<tr>
<td>MPL.e</td>
<td>69.45</td>
<td>23.05</td>
<td>6.86</td>
<td>3.53</td>
<td>1.73</td>
<td>0.71</td>
</tr>
<tr>
<td>TMPL.o</td>
<td>110.97</td>
<td>28.92</td>
<td>7.74</td>
<td>3.81</td>
<td>1.82</td>
<td>0.73</td>
</tr>
<tr>
<td>TMPL.e</td>
<td>88.59</td>
<td>24.76</td>
<td>7.06</td>
<td>3.59</td>
<td>1.75</td>
<td>0.72</td>
</tr>
<tr>
<td>TMCPL</td>
<td>72.59</td>
<td>23.04</td>
<td>6.81</td>
<td>3.51</td>
<td>1.72</td>
<td>0.71</td>
</tr>
<tr>
<td>ACPL.o</td>
<td>63.24</td>
<td>22.89</td>
<td>6.92</td>
<td>3.56</td>
<td>1.74</td>
<td>0.71</td>
</tr>
<tr>
<td>ACPL.e</td>
<td>50.26</td>
<td>19.68</td>
<td>6.31</td>
<td>3.36</td>
<td>1.67</td>
<td>0.70</td>
</tr>
<tr>
<td>NCT</td>
<td>58.00</td>
<td>21.18</td>
<td>6.56</td>
<td>3.44</td>
<td>1.70</td>
<td>0.70</td>
</tr>
</tbody>
</table>

### Table 2.27: Mean Square Error of label shelf-lives for parameter values $\beta_1 = -0.2, \sigma = 1.0, \epsilon = 0.01$ and $\alpha = 0.05$. 

<table>
<thead>
<tr>
<th>N</th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.z</td>
<td>79.33</td>
<td>38.36</td>
<td>16.29</td>
<td>8.76</td>
<td>4.24</td>
<td>1.80</td>
</tr>
<tr>
<td>z.t</td>
<td>96.56</td>
<td>41.80</td>
<td>16.84</td>
<td>8.90</td>
<td>4.27</td>
<td>1.81</td>
</tr>
<tr>
<td>ML.z</td>
<td>48.72</td>
<td>28.32</td>
<td>13.78</td>
<td>7.89</td>
<td>3.95</td>
<td>1.73</td>
</tr>
<tr>
<td>ML.t</td>
<td>57.44</td>
<td>30.80</td>
<td>14.26</td>
<td>8.02</td>
<td>3.98</td>
<td>1.73</td>
</tr>
<tr>
<td>U.z</td>
<td>88.62</td>
<td>40.76</td>
<td>16.85</td>
<td>8.95</td>
<td>4.30</td>
<td>1.82</td>
</tr>
<tr>
<td>U.t</td>
<td>107.65</td>
<td>44.38</td>
<td>17.42</td>
<td>9.10</td>
<td>4.34</td>
<td>1.82</td>
</tr>
<tr>
<td>PL</td>
<td>70.73</td>
<td>37.28</td>
<td>16.35</td>
<td>8.84</td>
<td>4.28</td>
<td>1.81</td>
</tr>
<tr>
<td>MPL.o</td>
<td>218.23</td>
<td>69.01</td>
<td>22.67</td>
<td>10.86</td>
<td>4.92</td>
<td>1.97</td>
</tr>
<tr>
<td>MPL.e</td>
<td>193.42</td>
<td>60.78</td>
<td>20.83</td>
<td>10.25</td>
<td>4.72</td>
<td>1.92</td>
</tr>
<tr>
<td>TMPL.o</td>
<td>320.67</td>
<td>81.30</td>
<td>24.46</td>
<td>11.37</td>
<td>5.08</td>
<td>2.01</td>
</tr>
<tr>
<td>TMPL.e</td>
<td>294.06</td>
<td>71.81</td>
<td>22.49</td>
<td>10.74</td>
<td>4.87</td>
<td>1.96</td>
</tr>
<tr>
<td>TMCPL</td>
<td>187.02</td>
<td>61.75</td>
<td>20.97</td>
<td>10.29</td>
<td>4.73</td>
<td>1.92</td>
</tr>
<tr>
<td>ACPL.o</td>
<td>194.05</td>
<td>61.98</td>
<td>21.36</td>
<td>10.45</td>
<td>4.79</td>
<td>1.94</td>
</tr>
<tr>
<td>ACPL.e</td>
<td>163.16</td>
<td>54.32</td>
<td>19.60</td>
<td>9.86</td>
<td>4.60</td>
<td>1.89</td>
</tr>
<tr>
<td>NCT</td>
<td>172.06</td>
<td>58.09</td>
<td>20.38</td>
<td>10.11</td>
<td>4.68</td>
<td>1.91</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>8</td>
<td>16</td>
<td>40</td>
<td>80</td>
<td>160</td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>z.z</td>
<td>229.19</td>
<td>109.68</td>
<td>38.85</td>
<td>20.48</td>
<td>10.61</td>
<td>4.44</td>
</tr>
<tr>
<td>z.t</td>
<td>285.03</td>
<td>121.22</td>
<td>40.21</td>
<td>20.81</td>
<td>10.70</td>
<td>4.46</td>
</tr>
<tr>
<td>ML.z</td>
<td>136.69</td>
<td>72.67</td>
<td>31.25</td>
<td>17.91</td>
<td>9.73</td>
<td>4.22</td>
</tr>
<tr>
<td>ML.t</td>
<td>163.77</td>
<td>79.54</td>
<td>32.35</td>
<td>18.21</td>
<td>9.81</td>
<td>4.23</td>
</tr>
<tr>
<td>U.z</td>
<td>254.93</td>
<td>119.64</td>
<td>40.70</td>
<td>21.08</td>
<td>10.82</td>
<td>4.49</td>
</tr>
<tr>
<td>U.t</td>
<td>300.82</td>
<td>132.08</td>
<td>42.10</td>
<td>21.42</td>
<td>10.91</td>
<td>4.51</td>
</tr>
<tr>
<td>PL</td>
<td>191.46</td>
<td>102.79</td>
<td>38.67</td>
<td>20.59</td>
<td>10.70</td>
<td>4.47</td>
</tr>
<tr>
<td>MPL.o</td>
<td>485.86</td>
<td>224.90</td>
<td>55.40</td>
<td>25.63</td>
<td>12.34</td>
<td>4.88</td>
</tr>
<tr>
<td>MPL.e</td>
<td>457.77</td>
<td>205.54</td>
<td>52.43</td>
<td>24.62</td>
<td>11.98</td>
<td>4.78</td>
</tr>
<tr>
<td>TMPL.o</td>
<td>532.13</td>
<td>239.20</td>
<td>60.64</td>
<td>27.32</td>
<td>12.92</td>
<td>5.02</td>
</tr>
<tr>
<td>TMPL.e</td>
<td>508.88</td>
<td>220.87</td>
<td>57.60</td>
<td>26.27</td>
<td>12.54</td>
<td>4.92</td>
</tr>
<tr>
<td>TMCPL</td>
<td>234.61</td>
<td>126.65</td>
<td>47.21</td>
<td>23.56</td>
<td>11.71</td>
<td>4.73</td>
</tr>
<tr>
<td>ACPL.o</td>
<td>506.72</td>
<td>224.52</td>
<td>52.99</td>
<td>24.84</td>
<td>12.08</td>
<td>4.81</td>
</tr>
<tr>
<td>ACPL.e</td>
<td>479.83</td>
<td>203.40</td>
<td>50.05</td>
<td>23.84</td>
<td>11.72</td>
<td>4.72</td>
</tr>
<tr>
<td>NCT</td>
<td>418.66</td>
<td>187.49</td>
<td>51.66</td>
<td>24.50</td>
<td>11.96</td>
<td>4.78</td>
</tr>
</tbody>
</table>

Table 2.28: Mean Square Error of label shelf-lives for parameter values $\beta_1 = -0.2$, $\sigma = 2.0$, $\epsilon = 0.01$ and $\alpha = 0.05$. 
Figure 2.7: Mean Square Error plotted against $1/N$ for each method and combination of parameters.
As the simulations results were obtained in what can be considered a designed experiment, it is natural to fit models to the coverage probability, bias and MSE to summarize the above tables succinctly. It should however be noted that because the models are fitted to results of simulations they are subject to error. Consequently, the results from these models should not be viewed as absolute, but as a guide together with the previous figures.

From Figure 2.5 it can be seen that the coverage probability tends to 0.95, the desired confidence level, in a near linear fashion as \(1/\sqrt{N}\) approaches 0 (\(N\) approaches infinity). The rate of this convergence clearly depends on the method used, but possibly also on the slope of the degradation curve \(\beta_1\) and the variability \(\sigma\). Consequently, a possible full model for the coverage probability (CP) of method \(i\) may be

\[
CP_i = \alpha_{i0} + \alpha_{i1} \frac{1}{\sqrt{N}} + \alpha_{i2} \beta_1^* + \alpha_{i3} \sigma^* \\
+ \alpha_{i4} \left( \frac{\beta_1^*}{\sqrt{N}} \right) + \alpha_{i5} \left( \frac{\sigma^*}{\sqrt{N}} \right) + \alpha_{i6} \left( \beta_1^* \sigma^* \right) \\
+ \alpha_{i7} \left( \frac{\beta_1^* \sigma^*}{\sqrt{N}} \right) + e_i ,
\]

where \(\beta_1^* = \beta_1 - 0.3\), \(\sigma^* = \sigma - 1\), and \(e_i\) is a random error which represents the sampling variability inherent in the simulations. Standard model reduction was conducted by testing the higher order interactions, observing marginality, in turn (assuming a significance level of 1\%). Table 2.29 gives the parameter estimates for all significant effects (there was no significant three-way interaction and consequently the corresponding column has been omitted from the table). These are discussed further in Section 2.4.

With respect to the bias of the various methods, a similar modelling approach was performed on the simulation results. However, from Figure 2.6 it is to be expected that \(1/\sqrt{N}\), \(\beta_1\) and \(\sigma\) are likely to be involved in a three-way interaction. In an attempt to overcome the complexity of such a model, the bias relative to the shelf-life was used as the response. That is, the full model for method \(i\) is of the form

\[
\frac{\text{Bias}_{i \text{shelf-life}}}{\text{Shelf-life}} = \alpha_{i0} + \alpha_{i1} \frac{1}{\sqrt{N}} + \alpha_{i2} \beta_1^* + \alpha_{i3} \sigma^* \\
+ \alpha_{i4} \left( \frac{\beta_1^*}{\sqrt{N}} \right) + \alpha_{i5} \left( \frac{\sigma^*}{\sqrt{N}} \right) + \alpha_{i6} \left( \beta_1^* \sigma^* \right) \\
+ \alpha_{i7} \left( \frac{\beta_1^* \sigma^*}{\sqrt{N}} \right) + e_i ,
\]

68
where $\beta_1^*$, $\sigma^*$, and $e_i$ is as above. The parameter estimates for each method are given in Table 2.30. These are discussed further in Section 2.4.

Since the bias also plays a significant role in the MSE, and considering the patterns in Figure 2.7, it is immediately obvious that a model for the MSE will also involve the sample size, slope of the degradation curve and the product variability. Due to the difficulty of interpretation of the estimates in Table 2.30, there is little use in attempting to fit a model for the MSE — the patterns are simply obtained from Figure 2.7.

<table>
<thead>
<tr>
<th>Method</th>
<th>Int.</th>
<th>$\beta_1^*$</th>
<th>$\sigma^*$</th>
<th>$\beta_1^* \sigma^*$</th>
<th>$\frac{1}{\sqrt{N}}$</th>
<th>$\frac{\beta_1^*}{\sqrt{N}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.z</td>
<td>0.961</td>
<td>0.036</td>
<td>-0.002</td>
<td>-0.062</td>
<td>-0.293</td>
<td></td>
</tr>
<tr>
<td>z.t</td>
<td>0.954</td>
<td>0.038</td>
<td>-0.002</td>
<td>-0.055</td>
<td>-0.192</td>
<td></td>
</tr>
<tr>
<td>ML.z</td>
<td>0.982</td>
<td>-0.036</td>
<td></td>
<td></td>
<td>-0.675</td>
<td>0.8</td>
</tr>
<tr>
<td>ML.t</td>
<td>0.969</td>
<td>-0.012</td>
<td>-0.006</td>
<td>0.703</td>
<td>-0.509</td>
<td></td>
</tr>
<tr>
<td>U.z</td>
<td>0.958</td>
<td></td>
<td></td>
<td></td>
<td>-0.236</td>
<td></td>
</tr>
<tr>
<td>U.t</td>
<td>0.951</td>
<td></td>
<td></td>
<td></td>
<td>-0.145</td>
<td></td>
</tr>
<tr>
<td>PL</td>
<td>0.969</td>
<td>0.052</td>
<td>-0.001</td>
<td>-0.077</td>
<td>-0.368</td>
<td></td>
</tr>
<tr>
<td>MPL.o</td>
<td>0.952</td>
<td>-0.009</td>
<td>-0.002</td>
<td>-0.047</td>
<td>0.053</td>
<td></td>
</tr>
<tr>
<td>MPL.e</td>
<td>0.950</td>
<td>-0.029</td>
<td>-0.001</td>
<td>-0.052</td>
<td>0.032</td>
<td></td>
</tr>
<tr>
<td>TMPL.o</td>
<td>0.951</td>
<td></td>
<td></td>
<td></td>
<td>0.073</td>
<td></td>
</tr>
<tr>
<td>TMPL.e</td>
<td>0.949</td>
<td></td>
<td></td>
<td></td>
<td>0.056</td>
<td></td>
</tr>
<tr>
<td>ACPL.o</td>
<td>0.950</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACPL.e</td>
<td>0.951</td>
<td></td>
<td></td>
<td></td>
<td>-0.052</td>
<td></td>
</tr>
<tr>
<td>NCT</td>
<td>0.949</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMCPL</td>
<td>0.947</td>
<td></td>
<td></td>
<td></td>
<td>0.030</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.29: For each method the parameter estimates for the model of coverage probability on $1/\sqrt{N}$, $\beta_1^* = \beta_1 + 0.3$ and $\sigma^* = \sigma - 1$ are given. Blank values indicate that the parameter was not significant at the 1% level and the model was re-fitted without the corresponding term(s).
2.4 Discussion

From Table 2.29 it can be seen that there are essentially two groups — the methods above and including MPL.e and those below and including TMPL.o. The first group depends on the slope of the degradation curve and the product variability, while the second group does not. Clearly, the second group is superior as these methods depend less on the nature of the underlying degradation parameters. In addition, these methods also tend to be conservative, as can be seen from Figure 2.5 and the positive coefficients of $1/\sqrt{N}$ (except ACPL.e).

It should be noted that based on 1000 simulations and a desired confidence level of 0.95 the likely range of coverage probabilities is about 0.95 ± 0.014, as shown in Section 2.3. Consequently, if the coverage probability lies outside the range (0.936, 0.964) in the lim-
iting case as $N \to \infty$, then the true limiting coverage probability is likely to differ from 0.95. However, for some methods the intercept may lie outside these limit because higher order terms, such as $1/N^{3/2}$, were omitted from the already complex models. Including such a term may reduce the intercept to a level that it is not significantly different from 0.95.

From the tables and graphs in Section 2.3 the following general statements can be made.

- The coverage probability tends toward 0.95 as $N \to \infty$ for all methods. Some methods appear to depend on the degradation parameters more than others;

- The bias and MSE tend toward zero as $N \to \infty$. The rate at which different methods reduce in bias depends on the sample size, the slope of the degradation curve and the product variability, as can be seen from Figure 2.6 and Table 2.30;

- Smaller coefficients of $1/\sqrt{N}$ indicate that the corresponding method works better for small sample sizes;

- Larger variability, as indicated by larger values of $\sigma$, tends to result in larger bias and MSE;

- Slower degradation, as indicated by smaller (absolute) values of $\beta_1$, also tends to result in larger bias and MSE.

- Methods with lower coverage probabilities tend to also exhibit smaller bias and MSE. This is not surprising as all three measures suggest that the label shelf-life is too large. However, coverage probability is the most important of these three performance criteria. Methods with roughly equal coverage can then be further compared by their bias and MSE. Consequently, unless mentioned otherwise, performance will hereafter refer to coverage probability.

The approaches based on the normal approximation under-performed with respect to coverage probability for all values of $\beta_1$ and $\sigma$ that were chosen for the simulations, except for very large $N$. Correspondingly, the normal approximations result in a bias and
MSE that are smaller than most other methods. Of the three choices for \( \delta \), the unbiased choice gave the best results in terms of coverage probability, followed closely by the choice \( \delta = z_c \). The maximum likelihood choice gave the worst coverage probabilities, performing particularly badly for very small \( N \). The choice of \( Q_\alpha = t_{\alpha,N-2} \) consistently gave better results than \( Q_\alpha = z_\alpha \), with the difference being negligible for large \( N \), as expected. A likely reason why these normal approximations do not work very well is that \( q_t,\epsilon \) is not independent of \( \text{Var}[q_t,\epsilon] - s \) features in both of them. Further work is still required to investigate this issue.

The basic profile likelihood approach performed the worst of all the profile likelihood based approaches. It also performed worse than the unbiased normal approximations in all cases. Correspondingly, the bias and MSE are generally smaller. This approach performed badly for small \( N \) and reasonably well for large \( N \), which is probably because the profile likelihood is more like a quadratic for large \( N \) (see Figure 2.2), and hence the \( \chi^2 \) approximation is better.

The MPL had coverage probabilities that were much better than those obtained using the basic profile likelihood, especially for smaller \( N \). Likewise, the bias and MSE were larger. The use of the observed information matrix resulted in higher coverage probabilities than the expected information matrix, which tended to be closer to the desired 0.95.

The coverage probability for all the above methods appears to be affected by the sample size, the rate of degradation as well as the product variability. On the other hand, the following methods, except ACPL.o and NCT, tend to be affected only by the sample size. All of the following methods are likely to perform well if they can be adapted to the multi-batch case.

The TMPL gave coverage probabilities that were higher than those obtained using the MPL for almost all cases of \( \beta_1 \) and \( \sigma \). Again, the use of the observed information matrix gave more conservative coverage probabilities than the expected information matrix.

The TMCPL also resulted in coverage probabilities that were very close to 0.95 for all values of \( \beta_1, \sigma \) and \( N \), and for small values of \( N \) this approach tended to be conservative.
As pointed out previously, the use of the Powell-Hestenes equivalents of observed and expected information matrix results in identical estimates of shelf-life. Furthermore, this approach resulted in values of bias and MSE that were very close to those of the exact approach; for small $N$ and large $\sigma$ they were, at times, considerably smaller.

The ACPL resulted in coverage probabilities that were acceptably close to 0.95, except for small $N$. Like the modified profile likelihood approach, use of the observed information resulted in higher coverage probabilities than the expected information. Also, the bias and MSE are correspondingly larger.

The non-central t-distribution approach resulted in very good coverage probabilities for all $N$. This is expected since the confidence bounds for the $\epsilon$ quantile are based on exact results.

In conclusion, the NCT, TMCPL, ACPL and TMPL approaches are about equally as good with respect to coverage, bias and MSE. Clearly, in the single batch case the NCT approach is the one that should be used as it is based on exact results. However, the other approaches also offer good alternatives, but TMCPL is considered to be the second best choice as it is conservative for all $N$, but less so than TMPL. The normal approximations and the basic PL approach cannot recommended, except for very large sample sizes, which are unlikely to be encountered in practice.
CHAPTER 3

Estimation of Shelf-Life:
The Multi-Batch Case

In this chapter an attempt is made to overcome the shortcomings of the current regulatory guidelines by extending the definitions proposed in Chapter 2 and by generalizing the methods of estimation to the batch case.

3.1 A New Definition

A lot of effort has so far been expended on developing valid tests for batch-to-batch variability, as can be seen from the number of papers discussed in Section 1.2. Unlike the current regulatory approach, many of these testing procedures tend to assume that batches are random effects. However, basing the data analysis on preliminary tests has been shown to be potentially misleading, not to mention inefficient. Consequently, such an approach is not endorsed here.

The current regulatory approach becomes more difficult and cumbersome as more batches are included in the analysis. This is because the probability of batches being found to be different increases as more batches are used in the analysis. A separate shelf-life for
3.1 A NEW DEFINITION

each batch may consequently need to be calculated and the minimum of these shelf-lives found. This suggests that bad designs with few batches, and possibly few observations per batch, reap the benefit of requiring a simpler analysis (Ruberg and Hsu, 1990).

In addition, Shao and Chow (1994) pointed out that there is no statistical justification for the “minimum of three” approach. Consequently, they developed an approach with more statistical merit. In particular, using more batches with their approach has the effect of increasing the label shelf-life (1.10) by decreasing the variability in \( x_t^T \bar{b} \). However, it heavily relies on the assumption that \( \sigma^2 \rightarrow 0 \), or at least that \( \sigma^2 \) is small compared to the batch variability \( D \).

However, if instead the batch-variability approaches zero, then (1.10) approaches

\[
\bar{\tau}_t = \inf \left\{ t : \mathbf{x}_t^T \bar{b} - t'_{1-\alpha} (k-1, -z_{\epsilon} \sqrt{k}) s \sqrt{\mathbf{x}_t^T (X^TX)^{-1} \mathbf{x}_t} \leq \eta \right\},
\]

which is not consistent with the current regulatory approach and (1.3). This illustrates that different values for the label shelf-life can be obtained based on which definition is used. This in turn depends on the outcome of the test for significance of batch-to-batch variability, which gives even more reason not to use this two stage approach.

Instead, good definitions of true, estimated and label shelf-life for batches should have the following three properties.

1. They should be consistent with the corresponding definitions in the single batch case.

2. The bias of the label shelf-life should decrease as more batches are used in the analysis.

3. The MSE of the label shelf-life should decrease as more batches are used in the analysis.

These properties lead to the following set of new definitions in the batch-to-batch case.
Definition 3 Under model (1.6), the true shelf-life, denoted by $\tau_\epsilon$, is the minimum time at which $100\epsilon\%$ of drug product from all future batches has an activity level that is less than or equal to the predetermined specification limit, $\kappa$, that is,

$$\tau_\epsilon = \inf \left\{ t : x_t^T \beta + z_\epsilon \sqrt{\sigma^2 + x_t^T D x_t} \leq \kappa \right\}.$$ 

The estimated shelf-life, $\hat{\tau}_\epsilon$, is then defined as the time at which the estimated $\epsilon$ quantile of $Y$ intersects the specification limit.

Definition 4 Based on Definition 3, the label shelf-life, denoted by $\hat{\tau}_{\epsilon,L(\alpha)}$, is the $100(1-\alpha)\%$ lower confidence bound for the true shelf-life.

Note that $\tau_{0.5}$ again represents the true shelf-life currently used by regulatory bodies, but it differs from Shao and Chow’s definition of $\tau_\epsilon$ in (1.9) through the inclusion of $\sigma^2$ under the square root. This is because this definition is based on the unconditional distribution of an observation $Y_{ij}$, and hence is influenced by both the within and between batch variability. This clearly makes sense from a quality control and quality assurance point of view — the shelf-life can be extended by reducing the variability between tablets within a batch, as well as by reducing the variability between batches. The effects which these two sources of variability have on the true shelf-life are as follows:

1. If batch-to-batch variability becomes very small in comparison to $\sigma^2$, then Definition 3 becomes Definition 1. Hence each batch will have $100\epsilon\%$ of product fall below the specification limit by the time the shelf-life has expired.

2. If batch-to-batch variability becomes very large compared to $\sigma^2$, then $100\epsilon\%$ of batches will have $100\%$ of product fail, and the remaining $100(1-\epsilon)\%$ of batches will have none of the product fail by the time the shelf-life has expired.

This is illustrated as follows. Denote the activity level of a unit $i$ of drug product from batch $j$, when the true shelf-life has expired, as $Y_{ij}$ and denoted it’s mean by $\mu_j = \ldots$
$x_{ij}(\tau_e)^T b_j$, where $\mu_j \sim N(\mu, \sigma^2_B)$ and $\sigma^2_B = x_{ij}(\tau_e)^T D x_{ij}(\tau_e)$. The distribution of $Y_{ij}$ is given by

$$Y_{ij}|\mu_j \sim N(\mu_j, \sigma^2),$$

and unconditionally

$$Y_{ij} \sim N(\mu, \sigma^2 + \sigma^2_B).$$

A unit of drug product can now be classified as defective if it has an activity less than the specification limit. Hence, the probability of a drug product from batch $j$ falling below $\kappa$ is

$$P_Y(\text{Defective}|\mu_j) = P_Y(Y_{ij}|\mu_j) = \Phi\left(\frac{\kappa - \mu_j}{\sigma}\right),$$

where $P_Y$ indicates the probability with respect to $Y_{ij}$. Consequently, the probability over all batches that $P_Y(\text{Defective}|\mu_j)$ is less than or equal to $\epsilon$ is given by

$$P_\mu\{P_Y(\text{defective}|\mu_j) \leq \epsilon\} = P_\mu\left\{\Phi\left(\frac{\kappa - \mu_j}{\sigma}\right) \leq \epsilon\right\}
= 1 - P_\mu\left\{\mu_j \leq \kappa - \Phi^{-1}(\epsilon)\sigma\right\}
= 1 - \Phi\left\{\frac{\mu - \Phi^{-1}(\epsilon)\sigma - \kappa}{\sigma_B}\right\}
= \Phi\left\{\frac{\Phi^{-1}(\epsilon)\sigma - \kappa}{\sigma_B}\right\}
= \Phi\left\{\frac{\Phi^{-1}(\epsilon)\sigma - \kappa}{\sigma_B}\right\}.$$

A graphical interpretation of this is given in Figure 3.1, where $\epsilon$ denotes the proportion of defectives per batch, and $F(\epsilon)$ denotes the proportion of batches with a proportion of $\epsilon$ defectives. As previously indicated, this graph shows the following when the overall rate of defects is 0.10:

- When $\sigma^2$ is large compared to $\sigma^2_B$, almost all batches have 100\% defective units.
- When $\sigma^2$ is small compared to $\sigma^2_B$, almost 100(1 - $\epsilon$)\% of batches have zero defects, while 100\% of batches have 100\% defects.
- For ratios of $\sigma^2_B$ to $\sigma^2$ within those extremes, batches vary from no defectives to 100\% defectives.
As in the single batch case, there are several ways of estimating the label shelf-life, and they are detailed in the following sections.

### 3.2 Estimation of the Label Shelf-Life

Interest now lies in the $\epsilon$ quantile of the unconditional distribution of the response given $t$. This is in comparison to the current regulatory approach which focuses on the mean response. Assuming linear degradation, the $\epsilon$ quantile at time $t$ is given by

$$q_{t,\epsilon} = x_t^T \beta + z_\epsilon \sqrt{\sigma^2 + x_t^T D x_t} = x_t^T \beta + z_\epsilon A^{1/2} , \quad (3.1)$$

assuming that neither the within nor the between batch variability depends on $t$.

In Chapter 2 several approaches for estimating the shelf-life and obtaining a label shelf-life were investigated. These approaches are developed for the multi-batch case in the following sections.
3.2.1 Normal Approximation

An estimate for \(q_{t,\epsilon}\) can be obtained by replacing \(\beta\), \(\sigma^2\) and \(D\) in (3.1) with appropriate estimates. As in Chapter 2, it will be assumed that this estimate is approximately normally distributed, and confidence bounds will be calculated.

The ordinary least squares estimates of \(b_j\) are obtained by treating batches as fixed, such that the estimates are given by \(\hat{b}_j = (X_j^T X_j)^{-1} X_j^T y_j\), with distribution \(\hat{b}_j \sim N(\beta, \Omega_j)\), where \(\Omega_j = D + \sigma^2 (X_j^T X_j)^{-1}\). Carter and Yang (1986) take a weighted average of these \(\hat{b}_j\) as an estimate of \(\beta\), that is,

\[
\hat{\beta}(\gamma) = \Omega \left( \frac{1}{k} \sum_{j=1}^{k} \Omega_j^{-1} \hat{b}_j \right)
\]

(3.2)

where \(\gamma\) is the \(1 + \frac{p(1+p)}{2}\) column vector containing \(\sigma^2\) and the upper-diagonal elements of \(D\) and \(\Omega\) is the harmonic mean of the \(\Omega_j\), that is,

\[
\Omega = \left( \frac{1}{k} \sum_{j=1}^{k} \Omega_j^{-1} \right)^{-1}.
\]

It is straightforward to show that (3.2) is equal to the generalised least squares estimator of \(\beta\) under (1.7), given by

\[
\hat{\beta}(\gamma) = \left( X^T \Sigma^{-1} X \right)^{-1} \left( X^T \Sigma^{-1} y \right)
\]

\[
= \left( \sum_{j=1}^{k} X_j^T \Sigma_j^{-1} X_j \right)^{-1} \left( \sum_{j=1}^{k} X_j \Sigma_j^{-1} y_j \right),
\]

by noting that

\[
\Sigma_j^{-1} = \frac{1}{\sigma^2} \left[ I - X_j (\sigma^2 D^{-1} + X_j^T X_j)^{-1} X_j^T \right]
\]

\[
\Omega_j^{-1} = \frac{1}{\sigma^2} X_j^T \left[ I - X_j (\sigma^2 D^{-1} + X_j^T X_j)^{-1} X_j^T \right] X_j,
\]

and hence \(X_j^T \Sigma_j^{-1} X_j \equiv (X_j^T X_j)^{-1} X_j^T = X_j^T \Sigma_j^{-1}\) and \(\Omega_j^{-1} = X_j^T \Sigma_j^{-1} X_j\). The distribution of the generalised least squares estimate is given by

\[
\hat{\beta}(\gamma) \sim N(\beta, \frac{1}{k} \Omega).
\]
An estimate of $\sigma^2$ is obtained by pooling the residuals across all batches, that is,

$$s^2 = \frac{1}{\nu} \sum_{j=1}^{k} (y_j - X_j b_j)^T (y_j - X_j b_j),$$

where $\nu = N - kp$. The distribution of $\nu s^2 / \sigma^2$ is $\chi^2_{\nu}$.

An estimate of $\Omega$ is given by

$$(k - 1) \hat{\Omega} = \sum_{j=1}^{k} (b_i - \bar{b})(b_i - \bar{b})^T,$$

where $\bar{b} = \frac{1}{k} \sum_{j=1}^{k} b_j$. The distribution of $(k - 1) \hat{\Omega}$ is Wishart with $k - 1$ degrees of freedom, that is,

$$(k - 1) \hat{\Omega} \sim W(k - 1, \Omega).$$

The Wishart distribution is a multivariate generalization of the $\chi^2$ distribution and it should be noted that for any $p \times 1$ vector $x$

$$(k - 1) \frac{x^T \hat{\Omega} x}{x^T \Omega x} \sim \chi^2_{k-1}.$$

An estimate of $D$ can then be obtained using

$$\hat{D} = \hat{\Omega} - s^2 \frac{k}{k} \sum_{j=1}^{k} (X_j^T X_j)^{-1},$$

which can potentially be non-positive definite. In such a case it may be reasonable to use the estimator suggested by Carter and Yang (1986).

A general form of the estimate of (3.1) can then be written as

$$\hat{q}_{t,\epsilon} = x_t^T \hat{\beta}(\hat{\gamma}) + \sqrt{\delta_W^2 s^2 G + \delta_B^2 x_t^T \hat{\Omega} x_t},$$

where $G = 1 - x_t^T \left[ \frac{1}{k} \sum_{j=1}^{k} (X_j^T X_j)^{-1} \right] x_t$, and possible choices for $\delta_W^2$ and $\delta_B^2$ include:

1. $\delta_W^2 = \delta_B^2 = z_{\epsilon^2}$,

2. $\sqrt{\delta_W^2 s^2 G + \delta_B^2 x_t^T \hat{\Omega} x_t}$ is maximum likelihood for $z_{\epsilon^2 A^{1/2}}$, and

3. $\sqrt{\delta_W^2 s^2 G + \delta_B^2 x_t^T \hat{\Omega} x_t}$ is unbiased for $z_{\epsilon^2 A^{1/2}}$. 

80
As before, the performance of the normal approximation relies on the choice of \( \delta_W^2 \) and \( \delta_B^2 \) and the closeness to normality of \( x^T \beta(\hat{\gamma}) \) and \( \sqrt{\delta_W^2 s^2 G + \delta_B^2 x^T \hat{\Omega} x} \). Note that \( s^2 \) and \( \hat{\Omega} \) are the maximum likelihood estimates for \( \sigma^2 \) and \( \Omega \), which can easily be obtained from the joint likelihood of \( s^2 \) and \( \hat{\Omega} \). It then follows that (1) and (2) above are identical.

Satterthwaite (1946) suggested a method of determining an approximate distribution of a linear combination of mean squares, like that under the square root. The Satterthwaite approximation states that when \( m^2 \) is a linear combination of mean squares, that is, \( m^2 = \sum c_i m_i^2 \), where the \( c_i \) are known constants and \( m_i^2 \) is estimated on \( f_i \) degrees of freedom, then \( m^2 \) is distributed approximately as \( \sigma^2 m^2 / f \), where \( \mathbb{E}[m^2] = \sigma^2 m^2 \) and the approximate degrees of freedom, \( f \), are given by

\[
f = \frac{\left(\sum c_i m_i^2\right)^2}{\sum c_i m_i^2}.
\]

Consequently,

\[
\frac{\delta_W^2 s^2 G + \delta_B^2 x^T \hat{\Omega} x}{\delta_W^2 \sigma^2 G + \delta_B^2 x^T \hat{\Omega} x} \sim \frac{\chi_f^2}{f},
\]

where the Satterthwaite degrees of freedom are given by

\[
f = \frac{\left[\delta_W^2 s^2 G + \delta_B^2 x^T \hat{\Omega} x\right]^2}{\sum_{i=1}^{k}(n_i-2) + \frac{\left[\delta_B^2 x^T \hat{\Omega} x\right]^2}{k-1}}.
\]

As a result, the expected value and variance of the square root are given by

\[
\mathbb{E}\left[\sqrt{\delta_W^2 s^2 G + \delta_B^2 x^T \hat{\Omega} x}\right] = C_f \sqrt{\delta_W^2 \sigma^2 G + \delta_B^2 x^T \hat{\Omega} x}
\]

and

\[
\text{Var}\left[\sqrt{\delta_W^2 s^2 G + \delta_B^2 x^T \hat{\Omega} x}\right] = (1 - C_f^2) \left[\delta_W^2 \sigma^2 G + \delta_B^2 x^T \hat{\Omega} x\right].
\]

Hence, the mean and variance of \( \hat{q}_{t,\epsilon} \) are given by

\[
\mathbb{E}[\hat{q}_{t,\epsilon}] = x^T \beta + C_f \sqrt{\delta_W^2 s^2 G + \delta_B^2 x^T \hat{\Omega} x}
\]

and

\[
\text{Var}[\hat{q}_{t,\epsilon}] = \frac{1}{k} x^T \hat{\Omega} x + (1 - C_f^2) \left[\delta_W^2 \sigma^2 G + \delta_B^2 x^T \hat{\Omega} x\right].
\]
From the expected value it follows that the three choices for $\delta^2_W$ and $\delta^2_B$ are

Estimate (1) and (2): $\delta^2_W = \delta^2_B = z^2_\epsilon$ \hspace{1cm} (3.5)

Estimate (3): $\delta^2_W = \delta^2_B = \frac{z^2_\epsilon}{C_\nu^2}$, \hspace{1cm} (3.6)

and an estimate of the variance can be obtained by substituting $s^2$ and $\hat{\Omega}$ for $\sigma^2$ and $\Omega$ in the variance for $\hat{q}_{t,\epsilon}$. Thus, the estimated shelf-life is given by

$$\hat{\tau}_\epsilon = \inf \{ t : \hat{q}_{t,\epsilon} \leq \kappa \} ,$$

and the label shelf-life is the time at which the 100$(1 - \alpha)\%$ lower confidence bound for $q_{t,\epsilon}$ intersects the specification limit, that is,

$$\hat{\tau}_{\epsilon, L(\alpha)} = \inf \left\{ t : \hat{q}_{t,\epsilon} + t_{\alpha, f} \sqrt{\frac{1}{s^2} \hat{\Omega} x_t + (1 - C^2)} \left[ \delta^2_W s^2 G + \delta^2_B x_t^T \hat{\Omega} x_t \right] \leq \kappa \right\}. \hspace{1cm} (3.7)$$

Notice that (3.7) must be solved iteratively for $t$, as $t$ is used under the square root sign as well as in the calculation of $t_{\alpha, f}$. Also, $z_\alpha$ could be used instead of $t_{\alpha, f}$, but the simulations in Chapter 2 have shown that the normal quantile performed worse than the appropriate quantile from the t-distribution, especially for small samples.

### 3.2.2 Various Profile Likelihood Based Approaches

This section details methods analogous to those in 2.2.2 and 2.2.3 for the situation when multiple batches are analysed together.

The label shelf-life $\hat{\tau}_{\epsilon, L(\alpha)}$ is again the smallest time $\tau$ for which the null hypothesis

$$H_0 : \beta_0 + \beta_1 \tau + z_\epsilon \sqrt{\sigma + x_t^T D x_t} = \kappa$$ \hspace{1cm} (3.8)

is retained on a one-sided 100$\alpha\%$ level test. Both $\kappa$ and $z_\epsilon$ are known constants, and thus the likelihood can be reformulated as a function of the parameters $(\beta_1, \sigma^2, D, \tau)$, by eliminating $\beta_0$. Hypotheses about $\tau$ can then be tested using the profile likelihood, which is obtained by maximizing the likelihood over $(\beta_1, \sigma, D)$ for given values of $\tau$. 
The Basic Profile Likelihood

The log-likelihood based on (1.7) is given by
\[
L(\beta, \sigma, D; y) = -\frac{1}{2} \log |\Sigma| \quad - \frac{1}{2} \left( y - X\beta \right)^T \Sigma^{-1} \left( y - X\beta \right)
\]
\[
= -\frac{1}{2} \sum_{j=1}^{k} \log |\Sigma_j| \quad - \frac{1}{2} \sum_{j=1}^{k} r_j^T \Sigma_j^{-1} r_j ,
\]
where \( r_j = y_j - X_j\beta \).

The maximum likelihood estimates of the parameters are found by differentiating (3.9) with respect to \( \beta \), \( \sigma^2 \) and the elements \( d_l \) of \( D \), and equating the derivatives to zero. That is,
\[
\frac{\partial L}{\partial \beta} = \sum_{j=1}^{k} X_j^T \Sigma_j^{-1} r_j
\]
\[
\frac{\partial L}{\partial \sigma^2} = -\frac{1}{2} \sum_{j=1}^{k} \text{tr} \{ \Sigma_j^{-1} \} \quad + \frac{1}{2} \sum_{j=1}^{k} r_j^T \Sigma_j^{-2} r_j
\]
\[
\frac{\partial L}{\partial d_l} = -\frac{1}{2} \sum_{j=1}^{k} \text{tr} \left\{ \Sigma_j^{-1} X_j \frac{\partial D}{\partial d_l} X_j^T \right\} \quad + \frac{1}{2} \sum_{j=1}^{k} r_j^T \Sigma_j^{-1} X_j \frac{\partial D}{\partial d_l} X_j^T \Sigma_j^{-1} r_j ,
\]
which do not have explicit solutions. However, Fisher scoring can be employed to find the maximum likelihood estimates.

The elements of the observed information matrix \( I_o \) are given by
\[
I_o(\beta, \beta) = \sum_{j=1}^{k} X_j^T \Sigma_j^{-1} X_j
\]
\[
I_o(\beta, \sigma^2) = \sum_{j=1}^{k} X_j^T \Sigma_j^{-2} r_j
\]
\[
I_o(\beta, d_l) = \sum_{j=1}^{k} X_j^T \Sigma_j^{-1} X_j \frac{\partial D}{\partial d_l} X_j^T \Sigma_j^{-1} r_j
\]
\[
I_o(\sigma^2, \sigma^2) = -\frac{1}{2} \sum_{j=1}^{k} \text{tr} \{ \Sigma_j^{-2} \} \quad + \frac{1}{2} \sum_{j=1}^{k} r_j^T \Sigma_j^{-3} r_j
\]
\[
I_o(\sigma^2, d_l) = -\frac{1}{2} \sum_{j=1}^{k} \text{tr} \left\{ \Sigma_j^{-1} X_j \frac{\partial D}{\partial d_l} X_j^T \Sigma_j^{-1} \right\} \quad + \sum_{j=1}^{k} r_j^T \Sigma_j^{-2} X_j \frac{\partial D}{\partial d_l} X_j^T \Sigma_j^{-1} r_j
\]
3.2. ESTIMATION OF THE LABEL SHELF-LIFE

\[
I_o(d_l, d_m) = -\frac{1}{2} \sum_{j=1}^{k} \text{tr} \left\{ \Sigma_j^{-1} X_j \frac{\partial D}{\partial d_l} X_j^T \Sigma_j^{-1} X_j \frac{\partial D}{\partial d_m} X_j^T \right\} \\
+ \sum_{j=1}^{k} \frac{r_j^T \Sigma_j^{-1} X_j \frac{\partial D}{\partial d_l} X_j^T \Sigma_j^{-1} X_j \frac{\partial D}{\partial d_m} X_j^T \Sigma_j^{-1} r_j}{2}.
\]

Hence, the elements of the expected information matrix are

\[
I_e(\beta, \beta) = \sum_{j=1}^{k} X_j^T \Sigma_j^{-1} X_j \\
I_e(\beta, \sigma^2) = 0 \\
I_e(\beta, d_l) = 0 \\
I_e(\sigma^2, \beta) = \frac{1}{2} \sum_{j=1}^{k} \text{tr} \left\{ \Sigma_j^{-2} \right\} \\
I_e(\sigma^2, d_l) = \frac{1}{2} \sum_{j=1}^{k} \text{tr} \left\{ \Sigma_j^{-1} X_j^T \frac{\partial D}{\partial d_l} X_j^T \Sigma_j^{-1} \right\} \\
I_e(d_l, d_m) = \frac{1}{2} \sum_{j=1}^{k} \text{tr} \left\{ \Sigma_j^{-1} X_j^T \frac{\partial D}{\partial d_l} X_j^T \Sigma_j^{-1} X_j^T \frac{\partial D}{\partial d_m} X_j^T \right\}.
\]

Denoting the maximum likelihood estimate of \( \beta, \sigma^2 \) and \( D \) by \( \hat{\beta}, \hat{\sigma}^2 \) and \( \hat{D} \), respectively, it is easy to find the maximum likelihood estimate of \( \tau_e \), that is,

\[
\hat{\tau}_e = \inf \left\{ t : x^T_t \hat{\beta} + z_t \sqrt{\hat{\sigma}^2 + x^T_t \hat{D} x_t} \leq \kappa \right\}.
\]

The profile likelihood of \( \tau_e \) is obtained by eliminating \( \beta_0 \) in (3.9) using the null hypothesis (3.8), and subscripting the parameters with \( \tau_e \) to indicate their dependence on \( \tau_e \), that is,

\[
\bar{L}(\tau_e; y) = -\frac{1}{2} \sum_{j=1}^{k} \log |\Sigma_j, \tau_e| - \frac{1}{2} \sum_{j=1}^{k} r^*_j \Sigma_j^{-1} r^*_j,
\]

where

\[
r^*_j = y^*_j - X^*_j \beta_{1, \tau_e} + 1_j z \sqrt{\sigma^2 + x^T_{\tau_e} D_{\tau_e} x_{\tau_e}},
\]

and \( 1_j \) is a \( N_j \times 1 \) vector of 1's, \( y^*_j = y_j - 1_j \kappa \), \( X^*_j = X(j) - 1_j \tau_e \), \( X(j) \) is the model matrix \( X_j \) without the column corresponding to \( \beta_0 \), and \( A^{1/2} = \sqrt{\sigma^2 + x^T_{\tau_e} D_{\tau_e} x_{\tau_e}} \), where \( x_{\tau_e} = x(\tau_e) \).
Differentiating the profile likelihood with respect to \( \beta_{1,\tau_c}, \sigma_{\tau_c}^2 \) and \( d_{l,\tau_c} \) gives

\[
\frac{\partial \tilde{L}}{\partial \beta_{1,\tau_c}} = \sum_{j=1}^{k} X_j^T \Sigma_{j,\tau_c}^{-1} r_j^* \\
\frac{\partial \tilde{L}}{\partial \sigma_{\tau_c}^2} = -\frac{1}{2} \sum_{j=1}^{k} \text{tr} \left\{ \Sigma_{j,\tau_c}^{-1} \right\} + \frac{1}{2} \sum_{j=1}^{k} r_j^T \Sigma_{j,\tau_c} r_j^* - \frac{z_\epsilon}{2A^{1/2}} \sum_{j=1}^{k} 1_j^T \Sigma_{j,\tau_c}^{-1} r_j^* \\
\frac{\partial \tilde{L}}{\partial d_{l,\tau_c}} = -\frac{1}{2} \sum_{j=1}^{k} \text{tr} \left\{ \Sigma_{j,\tau_c}^{-1} X_j \frac{\partial D_{\tau_c}}{\partial d_{l,\tau_c}} X_j^T \right\} + \frac{1}{2} \sum_{j=1}^{k} r_j^T \Sigma_{j,\tau_c} r_j^* - \frac{z_\epsilon}{2A^{1/2}} \left( x_{\tau_c}^T \frac{\partial D_{\tau_c}}{\partial d_{l,\tau_c}} x_{\tau_c} \right) \sum_{j=1}^{k} 1_j^T \Sigma_{j,\tau_c}^{-1} r_j^*.
\]

The elements of the observed information matrix are consequently given by

\[
I_o(\beta_{1,\tau_c}, \beta_{1,\tau_c}) = \sum_{j=1}^{k} X_j^T \Sigma_{j,\tau_c}^{-1} X_j^* \\
I_o(\beta_{1,\tau_c}, \sigma_{\tau_c}^2) = \sum_{j=1}^{k} X_j^T \Sigma_{j,\tau_c}^{-1} r_j^* - \frac{z_\epsilon}{2A^{1/2}} \sum_{j=1}^{k} X_j^T \Sigma_{j,\tau_c}^{-1} 1_j \\
I_o(\beta_{1,\tau_c}, d_{l,\tau_c}) = \sum_{j=1}^{k} X_j^T \Sigma_{j,\tau_c}^{-1} X_j \frac{\partial D_{\tau_c}}{\partial d_{l,\tau_c}} X_j^T \Sigma_{j,\tau_c}^{-1} r_j^* - \frac{z_\epsilon}{2A^{1/2}} \left( x_{\tau_c}^T \frac{\partial D_{\tau_c}}{\partial d_{l,\tau_c}} x_{\tau_c} \right) \sum_{j=1}^{k} 1_j^T \Sigma_{j,\tau_c}^{-1} 1_j \\
I_o(\sigma_{\tau_c}^2, \sigma_{\tau_c}^2) = -\frac{1}{2} \sum_{j=1}^{k} \text{tr} \left\{ \Sigma_{j,\tau_c}^{-1} \right\} + \frac{1}{2} \sum_{j=1}^{k} r_j^T \Sigma_{j,\tau_c}^{-3} r_j^* + \frac{z_\epsilon^2}{4A} \sum_{j=1}^{k} 1_j^T \Sigma_{j,\tau_c}^{-1} 1_j \\
I_o(\sigma_{\tau_c}^2, d_{l,\tau_c}) = -\frac{1}{2} \sum_{j=1}^{k} \text{tr} \left\{ \Sigma_{j,\tau_c}^{-1} X_j \frac{\partial D_{\tau_c}}{\partial d_{l,\tau_c}} X_j^T \Sigma_{j,\tau_c}^{-1} \right\} + \frac{1}{2} \sum_{j=1}^{k} r_j^T \Sigma_{j,\tau_c}^{-2} X_j \frac{\partial D_{\tau_c}}{\partial d_{l,\tau_c}} X_j^T \Sigma_{j,\tau_c}^{-1} r_j^* + \frac{z_\epsilon^2}{4A} \left( x_{\tau_c}^T \frac{\partial D_{\tau_c}}{\partial d_{l,\tau_c}} x_{\tau_c} \right) \sum_{j=1}^{k} 1_j^T \Sigma_{j,\tau_c}^{-2} r_j^* \\
- \frac{z_\epsilon^2}{4A^{3/2}} \left( x_{\tau_c}^T \frac{\partial D_{\tau_c}}{\partial d_{l,\tau_c}} x_{\tau_c} \right) \sum_{j=1}^{k} 1_j^T \Sigma_{j,\tau_c}^{-1} r_j^* \\
I_o(d_{l,\tau_c}, d_{m,\tau_c}) = -\frac{1}{2} \sum_{j=1}^{k} \text{tr} \left\{ \Sigma_{j,\tau_c}^{-1} X_j \frac{\partial D_{\tau_c}}{\partial d_{m,\tau_c}} X_j^T \Sigma_{j,\tau_c}^{-1} X_j \frac{\partial D_{\tau_c}}{\partial d_{l,\tau_c}} X_j^T \right\}
\]

85
Taking expectations gives the elements of the expected information matrix,

\[
I_e(\beta, \sigma^2) = \sum_{j=1}^{k} X_j^T \Sigma_{j, \tau}^{-1} X_j
\]

\[
I_e(\beta, \sigma^2, d) = \frac{z_\epsilon}{2A^{1/2}} \sum_{j=1}^{k} X_j^T \Sigma_{j, \tau}^{-1} X_j
\]

\[
I_e(\sigma^2, \sigma^2, d) = \frac{1}{2} \sum_{j=1}^{k} \text{tr}\left(\Sigma_{j, \tau}^{-1}\right) + \frac{z_\epsilon^2}{4A} \sum_{j=1}^{k} 1_j^T \Sigma_{j, \tau}^{-1} 1_j
\]

\[
I_e(d, m) = \frac{1}{2} \sum_{j=1}^{k} \text{tr}\left(\Sigma_{j, \tau}^{-1}\right) d_j + \frac{z_\epsilon^2}{4A} \sum_{j=1}^{k} 1_j^T \Sigma_{j, \tau}^{-1} 1_j
\]

Consequently, Fisher scoring can be implemented to obtain estimates \(\hat{\beta}_{1, \tau}, \hat{\sigma}_{\tau}^2\) and \(\hat{D}_{\tau}\). As before, the likelihood ratio test can be used to obtain the label shelf-life from a 100\((1-\alpha)\)% lower confidence bound for \(\tau\). Let \(\hat{L}(\tau; y)\) be the likelihood evaluated at the maximum likelihood estimates and let \(\tilde{L}(\tau; y)\) be the likelihood evaluated for a given value of \(\tau\).
Then, asymptotically, the generalised likelihood ratio statistic is given by

\[ w(\tau) = 2 \left( \tilde{L}(\hat{\tau}_e; y) - \tilde{L} (\tau_e; y) \right), \]

which is distributed approximately $\chi^2_1$ when (3.8) is true. Consequently, the label shelf-life is given by

\[ \hat{\tau}_{e,L(\alpha)} = \inf \left\{ t : (t \leq \hat{\tau}_e) \& (w(t) \leq \chi^2_{1,1-2\alpha}) \right\}. \tag{3.10} \]

**The Truncated Modified Profile Likelihood (TMPL) Approach**

Unlike in the single batch case of Chapter 2 parameter estimates cannot be obtained analytically. Consequently, the Jacobian term used in the modified profile likelihood (2.22) cannot be calculated. Nevertheless, it is possible to calculate the observed and expected truncated modified profile likelihoods.

The observed modified profile log-likelihood for $\tau_e$ is given by

\[ \tilde{L}_o^e (\tau_e; y) = \tilde{L} (\tau_e; y) - \frac{1}{2} \log \left| o_{\beta_1, \tau_e, \sigma^2_{\tau_e}, \check{D}_{\tau}} \right| - \log \left| \partial \left( \hat{\beta}_{1, \tau_e}, \hat{\sigma}^2_{\tau_e}, \hat{D}_{\tau} \right) \right|. \]

Consequently, the observed truncated modified profile likelihood is obtained by ignoring the Jacobian term, that is,

\[ \tilde{L}_t^o (\tau_e; y) = \tilde{L} (\tau_e; y) + \frac{1}{2} \log \left| o_{\beta_1, \tau_e, \sigma^2_{\tau_e}, \check{D}_{\tau}} \right|. \]

The generalised likelihood ratio statistic is then given by

\[ w_t(\tau) = 2 \left( \tilde{L}_t^o (\hat{\tau}_e; y) - \tilde{L}_t^o (\tau_e; y) \right), \]

which is distributed approximately $\chi^2_1$ when (3.8) is true. Consequently, the label shelf-life is given by

\[ \hat{\tau}_{e,L(\alpha)} = \inf \left\{ t : (t \leq \hat{\tau}_e) \& (w_t(t) \leq \chi^2_{1,1-2\alpha}) \right\}. \tag{3.11} \]

Similarly to Chapter 2, the expected information matrix can be used, resulting in $\tilde{L}_t^o$ being replaced by $\tilde{L}_t^e$. 

87
The Conditional Profile Likelihood Approach

The conditional profile likelihood (2.24) depends on the orthogonality of the parameters to $\psi = \tau_\epsilon$. Consequently, the original parameters $\phi^T = (\beta_{1,\tau_\epsilon}, \sigma^2_{\tau_\epsilon}, d_{1,\tau_\epsilon}, d_{2,\tau_\epsilon}, d_{3,\tau_\epsilon})$ must be transformed to give a new set of parameters $\lambda^T = (\lambda_1, \ldots, \lambda_5)$, which are orthogonal to $\psi = \tau_\epsilon$. This is done by solving the differential equations

$$I_e(\phi, \phi) \frac{\partial \phi}{\partial \psi} = -I_e(\phi, \psi),$$

Note that the information matrix $I_e(\phi, \phi)$ is a $5 \times 5$ matrix, which makes the inversion, and consequent integration, analytically intractable. Hence, the conditional profile likelihood does not offer a usable alternative in the multi-batch case.

### 3.2.3 Constrained Profile Likelihood Based Approaches

It has been shown in Section 2.2.3 that the constrained profile likelihood, based on the Powell-Hestenes approach advocated by Osborne (2000), can be used in place of the basic profile likelihood approach. As expected, the results were identical numerically. In addition, an approximation to the modified profile likelihood was obtained even though the Jacobian term could not be approximated. This section will detail both methods when applied to the batch case.

The objective function, based on (3.9), is given by

$$H(\beta_{\tau_\epsilon}, \sigma^2, D, \tau_\epsilon; y) = \frac{1}{N} \left[ -L(\theta_{\tau_\epsilon}) \right] + \omega \left[ g(\theta_{\tau_\epsilon}; \tau_\epsilon) + \psi \right]^2,$$

where $\theta^T_{\tau_\epsilon} = (\beta_{\tau_\epsilon}, \sigma^2_{\tau_\epsilon}, d_{1,\tau_\epsilon}, d_{2,\tau_\epsilon}, d_{3,\tau_\epsilon})$ and

$$g(\theta_{\tau_\epsilon}; \tau_\epsilon) = x^T_{\tau_\epsilon} \beta_{\tau_\epsilon} + z_{\epsilon} A^{1/2}_{\tau_\epsilon} - \kappa,$$

where $A^{1/2}_{\tau_\epsilon} = \sqrt{\sigma^2_{\tau_\epsilon} + x^T_{\tau_\epsilon} D_{\tau_\epsilon} x_{\tau_\epsilon}}$. The aim is to minimise the objective function subject to $g(\theta_{\tau_\epsilon}; \tau_\epsilon) = 0$. This can again be achieved with the help of Fisher Scoring.
Letting \( r_j = y_j - X_j \beta_{\tau_c} \), the elements of the score vector are given by

\[
\frac{\partial H}{\partial \beta_{\tau_c}} = \frac{1}{N} \left[ -\sum_{j=1}^{k} X_j^T \Sigma_{j,\tau_c}^{-1} r_j \right] + 2\omega \left[ g(\theta_{\tau_c}; \tau_c) + \psi \right] x_{\tau_c}
\]

\[
\frac{\partial H}{\partial \sigma_{\tau_c}^2} = \frac{1}{2N} \left[ \sum_{j=1}^{k} \text{tr} \left\{ \Sigma_{j,\tau_c}^{-1} \right\} - \sum_{j=1}^{k} r_j^T \Sigma_{j,\tau_c}^{-2} r_j \right] + \frac{\omega z_{\tau_c}}{A_{\tau_c}^{1/2}} \left[ g(\theta_{\tau_c}; \tau_c) + \psi \right]
\]

\[
\frac{\partial H}{\partial d_{l,\tau_c}} = \frac{1}{2N} \left[ \sum_{j=1}^{k} \text{tr} \left\{ \Sigma_{j,\tau_c}^{-1} X_j \frac{\partial D_{\tau_c}}{\partial d_{l,\tau_c}} X_j^T \right\} - \sum_{j=1}^{k} r_j^T \Sigma_{j,\tau_c}^{-1} X_j \frac{\partial D_{\tau_c}}{\partial d_{l,\tau_c}} X_j^T \Sigma_{j,\tau_c}^{-1} r_j \right] + \frac{\omega z_{\tau_c}}{A_{\tau_c}^{1/2}} \left( x_{\tau_c}^T \frac{\partial D_{\tau_c}}{\partial d_{l,\tau_c}} x_{\tau_c} \right) \left[ g(\theta_{\tau_c}; \tau_c) + \psi \right]
\]

Note that the unconstrained estimates can be obtained by ignoring the terms involving the constraint, which is equivalent to setting \( \omega = 0 \). It is easy to see that this yields the same maximum likelihood estimates that were obtained in Section 3.2.2.

The elements of the Powell-Hestenes equivalents of the observed information matrix, ignoring terms involving the second derivative of \( g(\theta_{\tau_c}; \tau_c) \), are then given by

\[
I_o'(\beta_{\tau_c}, \beta_{\tau_c}) = \frac{1}{N} \left[ \sum_{j=1}^{k} X_j^T \Sigma_{j,\tau_c}^{-1} X_j \right] + 2\omega x_{\tau_c} x_{\tau_c}^T
\]

\[
I_o'(\beta_{\tau_c}, \sigma_{\tau_c}^2) = \frac{1}{N} \left[ \sum_{j=1}^{k} X_j^T \Sigma_{j,\tau_c}^{-2} r_j \right] + \frac{\omega z_{\tau_c}^2}{2A_{\tau_c}} \left( x_{\tau_c}^T \frac{\partial D_{\tau_c}}{\partial d_{l,\tau_c}} x_{\tau_c} \right)
\]

\[
I_o'(\beta_{\tau_c}, d_{l,\tau_c}) = \frac{1}{N} \left[ \sum_{j=1}^{k} X_j^T \Sigma_{j,\tau_c}^{-1} X_j \frac{\partial D_{\tau_c}}{\partial d_{l,\tau_c}} X_j^T \Sigma_{j,\tau_c}^{-1} r_j \right] + \frac{\omega z_{\tau_c}^2}{2A_{\tau_c}} \left( x_{\tau_c}^T \frac{\partial D_{\tau_c}}{\partial d_{l,\tau_c}} x_{\tau_c} \right)
\]

\[
I_o'(\sigma_{\tau_c}^2, \sigma_{\tau_c}^2) = \frac{1}{N} \left[ -\frac{1}{2} \sum_{j=1}^{k} \text{tr} \left\{ \Sigma_{j,\tau_c}^{-2} \right\} + \sum_{j=1}^{k} r_j^T \Sigma_{j,\tau_c}^{-3} r_j \right] + \frac{\omega z_{\tau_c}^2}{2A_{\tau_c}} \left( x_{\tau_c}^T \frac{\partial D_{\tau_c}}{\partial d_{l,\tau_c}} x_{\tau_c} \right)
\]

\[
I_o'(\sigma_{\tau_c}^2, d_{l,\tau_c}) = \frac{1}{N} \left[ \sum_{j=1}^{k} X_j^T \Sigma_{j,\tau_c}^{-2} X_j \frac{\partial D_{\tau_c}}{\partial d_{l,\tau_c}} X_j^T \Sigma_{j,\tau_c}^{-1} r_j \right] + \frac{\omega z_{\tau_c}^2}{2A_{\tau_c}} \left( x_{\tau_c}^T \frac{\partial D_{\tau_c}}{\partial d_{l,\tau_c}} x_{\tau_c} \right)
\]

\[
I_o'(d_{l,\tau_c}, d_{m,\tau_c}) = \frac{1}{N} \left[ \sum_{j=1}^{k} X_j^T \Sigma_{j,\tau_c}^{-1} X_j \frac{\partial D_{\tau_c}}{\partial d_{m,\tau_c}} X_j^T \Sigma_{j,\tau_c}^{-1} r_j \right] + \frac{\omega z_{\tau_c}^2}{2A_{\tau_c}} \left( x_{\tau_c}^T \frac{\partial D_{\tau_c}}{\partial d_{m,\tau_c}} x_{\tau_c} \right)
\]
Taking expectations gives the elements of the equivalent versions of the expected information matrix

\[ I_e' (\beta_{\tau_e}, \beta_{\tau_e}) = \frac{1}{N} \left[ \sum_{j=1}^{k} X_j^T \Sigma^{-1}_{j,\tau_e} X_j \right] + 2\omega \times x_{\tau_e} x_{\tau_e}^T \]

\[ I_e' (\beta_{\tau_e}, \sigma^2_{\tau_e}) = 0 + \frac{\omega_z}{A_{\tau_e}^{1/2}} x_{\tau_e} \]

\[ I_e' (\beta_{\tau_e}, d_{l,\tau_e}) = 0 + \frac{\omega_z}{A_{\tau_e}^{1/2}} \left( x_{\tau_e}^T \frac{\partial D_{\tau_e}}{\partial d_{l,\tau_e}} x_{\tau_e} \right) \]

\[ I_e' (\sigma^2_{\tau_e}, \sigma^2_{\tau_e}) = \frac{1}{2N} \sum_{j=1}^{k} \text{tr} \left( \Sigma^{-2}_{j,\tau_e} \right) + \frac{\omega_z^2}{2A_{\tau_e}} \]

\[ I_e' (\sigma^2_{\tau_e}, d_{l,\tau_e}) = \frac{1}{2N} \sum_{j=1}^{k} \text{tr} \left( \Sigma^{-1}_{j,\tau_e} X_j \frac{\partial D_{\tau_e}}{\partial d_{l,\tau_e}} X_j^T \Sigma^{-1}_{j,\tau_e} \right) + \frac{\omega_z^2}{2A_{\tau_e}} \left( x_{\tau_e}^T \frac{\partial D_{\tau_e}}{\partial d_{l,\tau_e}} x_{\tau_e} \right) \]

As before, Fisher scoring can now be used to obtain estimates \( \hat{\beta}_{\tau_e}, \hat{\sigma}^2_{\tau_e} \) and \( \hat{D}_{\tau_e} \). As before, the likelihood ratio test can be used to obtain the label shelf-life from the 100(1 - \( \alpha \))% lower confidence bound for \( \tau_e \). Let \( \tilde{L}(\hat{\tau}_e; y) \) be the likelihood evaluated at the maximum likelihood estimates and let \( \tilde{L}(\tau_e; y) \) be the likelihood evaluated for a given value of \( \tau_e \).

Then, asymptotically, the generalised likelihood ratio statistic is given by

\[ w(\tau) = 2 \left( \tilde{L}(\hat{\tau}_e; y) - \tilde{L}(\tau_e; y) \right), \]

which is distributed approximately \( \chi^2_1 \) when (3.8) is true. Consequently, the label shelf-life is given by

\[ \hat{\tau}_{e,L(\alpha)} = \inf \left\{ t : (t \leq \hat{\tau}_e) \& (w(t) \leq \chi^2_{1,1-2\alpha}) \right\}. \]  

(3.14)

**The Modified Constrained Profile Likelihood**

In Section 3.2.2 it was shown that a modified profile likelihood approach was not possible when \( \beta_1 \) was eliminated from the profile likelihood. Nevertheless, the truncated modified profile likelihood provides a usable adjustment when the modified profile likelihood is
not possible. In addition, an approximation to the truncated modified profile likelihood, was developed in Section 2.2.3. This approach has performed well in the simulations of Chapter 2 and this Section will detail its application in the multi-batch scenario.

The approximation detailed in Section 2.2.3 was very general in terms of the number of parameters that are estimated. As such, the approximation can very easily be applied to the multi-batch scenario, since (2.34) and (2.37) still apply here, with \( p = 6 \).

The derivative of the constraint (3.13) in the batch case gives

\[
\frac{\partial g(\theta_\tau, \tau_\epsilon)}{\partial \theta_\tau^T} = \left[ 1 \quad \tau_\epsilon \quad \frac{z_1\tau_\epsilon}{2A_{\tau_\epsilon}^{1/2}} \quad \frac{z_2\tau_\epsilon}{2A_{\tau_\epsilon}^{1/2}} \quad \frac{z_3\tau_\epsilon^2}{2A_{\tau_\epsilon}^{1/2}} \right],
\]

where the obvious choice for \( \theta_0, \tau_\epsilon \) is again the constrained maximum likelihood estimate \( \hat{\theta}_{\tau_\epsilon} \), giving

\[
C_0 = \left[ 1 \quad \tau_\epsilon \quad \frac{z_1}{2A_{\tau_\epsilon}^{1/2}} \quad \frac{z_2}{2A_{\tau_\epsilon}^{1/2}} \quad \frac{z_3\tau_\epsilon}{2A_{\tau_\epsilon}^{1/2}} \right].
\]

The approximate modified profile likelihood is now given by

\[
\tilde{L}_a(\tau_\epsilon; y) = L(\hat{\theta}_{\tau_\epsilon}; y) - \frac{1}{2} \log \left[ \text{det} \left\{ C_0 \Omega(\hat{\theta}_{\tau_\epsilon})^{-1} C_0^T \right\} \right] - \frac{1}{2} \log \left( \text{det} \left\{ I_o(\hat{\theta}_{\tau_\epsilon}) \right\} \right)
\]

from which \( \hat{\tau}_\epsilon \) can be estimated. The generalised likelihood ratio statistic is given by

\[
w_a(\tau_\epsilon) = 2 \left( \tilde{L}_a(\hat{\tau}_\epsilon; y) - \tilde{L}_a(\tau_\epsilon; y) \right),
\]

which is distributed approximately \( \chi_1^2 \) under (2.7). Consequently, the label shelf-life, based on the 100(1 - \( \alpha \))% lower confidence bound, is given by

\[
\hat{\tau}_{\epsilon,L(\alpha)} = \inf \left\{ t : (t \leq \hat{\tau}_\epsilon) \cap (w_a(t) \leq \chi_{1,1-2\alpha}^2) \right\}.
\]

\[
3.2.4 \quad \text{Non-central t-distribution (NCT) Approach}
\]

In Section 2.2.4 an exact lower confidence bound for \( \tau_\epsilon \) was found based on the distribution of \( \hat{\beta} \). In the batch case, the distribution of \( \hat{\beta} \) is given by

\[
\hat{\beta} \sim N \left( \beta, \frac{1}{\hat{\tau}} \Omega \right).
\]
3.2. ESTIMATION OF THE LABEL SHELF-LIFE

as described in Section 3.2.1. The mean response at time \( t \) is then given by \( x_t^T \hat{\beta} \), with distribution

\[
x_t^T \hat{\beta} \sim N \left( x_t^T \beta, \frac{1}{k} x_t^T \Omega x_t \right).
\]

From (3.1) it follows that \( x_t^T \beta = q_{t, \epsilon} - z_{\epsilon} A^{1/2} \), and hence standardizing \( x_t^T \hat{\beta} \) yields

\[
\begin{align*}
\frac{\left( x_t^T \hat{\beta} - q_{t, \epsilon} + z_{\epsilon} A^{1/2} \right) \sqrt{k}}{\sqrt{x_t^T \Omega x_t}} &\sim N(0, 1).
\end{align*}
\]

Alternatively, this can be written as

\[
\frac{(x_t^T \hat{\beta} - q_{t, \epsilon}) \sqrt{k}}{\sqrt{x_t^T \Omega x_t}} \sim N \left\{ -z_{\epsilon} A^{1/2} \sqrt{k} \frac{1}{\sqrt{x_t^T \Omega x_t}}, 1 \right\}.
\]

Note that \( A^{1/2} = \sqrt{\sigma^2 + x_t^T D x_t} \) which does not equal \( \sqrt{x_t^T \Omega x_t} \). This can easily be seen from the case when all \( X_j \) are equal to \( X \), say, in which case \( \Omega = D + \sigma^2 (X^T X)^{-1} \) and \( x_t^T \Omega x_t = x_t^T D x_t + \sigma^2 x_t^T (X^T X)^{-1} x_t \). Consequently, unlike Section 2.2.4, these elements do not cancel from the mean of the normal distribution.

However, when between batch variability is large compared to within batch variability, or more precisely, when \( x_t^T D x_t \) is large compared to \( \sigma^2 \), then \( A^{1/2} \approx \sqrt{x_t^T \Omega x_t} \approx \sqrt{x_t^T D x_t} \), in which case

\[
\frac{(x_t^T \hat{\beta} - q_{t, \epsilon}) \sqrt{k}}{\sqrt{x_t^T \Omega x_t}} \sim N \left\{ -z_{\epsilon} \sqrt{k} \frac{1}{\sqrt{x_t^T \Omega x_t}}, 1 \right\}.
\]

Consequently, replacing \( x_t^T \Omega x_t \) by its estimate \( x_t^T \hat{\Omega} x_t \) results in a non-central t-distribution, that is,

\[
\frac{(x_t^T \hat{\beta} - q_{t, \epsilon}) \sqrt{k}}{\sqrt{x_t^T \hat{\Omega} x_t}} \sim t' \left( k - 1, \delta \right),
\]

where \( \delta = -z_{\epsilon} \sqrt{k} \) denotes the non-centrality parameter. The label shelf-life is given by the intersection of the 100(1 - \( \alpha \))% lower confidence bound and the specification limit, that is,

\[
\hat{\tau}_{\epsilon, L(\alpha)} = \inf \left\{ t : x_t^T \hat{\beta} + Q_\alpha \sqrt{x_t^T \hat{\Omega} x_t} \leq \kappa \right\},
\]

where \( Q_\alpha = t'_{\alpha} (k - 1, -\delta) \) is the lower \( \alpha \) quantile of the non-central t-distribution with \( k - 1 \) degrees of freedom and non-centrality parameter \( \delta = -z_{\epsilon} \sqrt{k} \). Note that this approximation is what Shao and Chow (1994) rely on for their definition of shelf-life in
3.2. ESTIMATION OF THE LABEL SHELF-LIFE

the batch case. In fact, (3.17) is equivalent to the label shelf-life of Shao and Chow, with the above derivation being alternative to that provided by Shao and Chow.

It should be noted that in the limiting case as \( k \to \infty \), the label shelf-life (3.17) becomes

\[
\hat{\tau}_{\epsilon,L(\alpha)} = \inf \left\{ t : x_t^T \beta + z_\epsilon \sqrt{\sigma^2 x_t^T (X^T X)^{-1} x_t + x_t^T D x_t} \leq \kappa \right\},
\]

since \( t'_\alpha(k-1,-\delta) \to (z_\alpha + z_\epsilon \sqrt{k}) \). Generally \( x_t (X^T X)^{-1} x_t < 1 \), and hence it follows that

\[
z_\epsilon \sqrt{\sigma^2 x_t^T (X^T X)^{-1} x_t + x_t^T D x_t} \leq z_\epsilon \sqrt{\sigma^2 + x_t^T D x_t}.
\]

Consequently, using this method will result in a label shelf-life that is greater than the true shelf-life. In addition, since \( (X^T X)^{-1} \to 0 \) as \( N \to \infty \) it follows that this approximation gets worse as the sample size increases.

However, Shao and Chow (1994) use the following definition of the true shelf-life

\[
\tau^*_\epsilon = \inf \left\{ t : x_t^T \beta + z_\epsilon \sqrt{x_t^T D x_t} \leq \kappa \right\}.
\] (3.18)

Thus, Shao and Chow’s label shelf-life is the 100(1 - \( \alpha \))% lower confidence bound for the time at which the \( \epsilon \) quantile of the distribution of \( x_t^T b_j \) intersects the specification limit.

In this case, (3.17) will be conservative, as was observed by Shao and Chow.

When \( \sigma^2 \) is not small compared to \( x_t^T D x_t \) then an approximate answer may be obtained by replacing \( \sigma^2 \), \( D \) and \( \Omega \) in the mean of the distribution in (3.16) by their respective estimates and considering the mean as fixed. In this case, (3.17) still holds, but the non-centrality parameter takes the form

\[
\delta = -z_\epsilon \sqrt{k \left( s^2 + x_t^T \hat{D} x_t \right) / x_t^T \hat{\Omega} x_t}.
\] (3.19)

Even though these two NCT approaches are approximate only, they will be further evaluated in Section 3.3.
3.3 Simulation Results

As in the single batch case, a simulation study is used to evaluate the approaches proposed in this chapter. These simulations were again performed with the help of the software R (Ihaka and Gentleman, 1996). Compared to the single batch case, simulations in the batch case are very much more computationally intensive.

The simulations were structured as follows. A random sample of batch effects of size \( k = 3, 5, 10 \) and 20 was drawn from the \( b_j \sim N(\beta, D) \). The mean was chosen to be \( \beta^T = (100, -0.3) \), while

\[
D = \begin{bmatrix}
1 & d_2 \\
d_2 & 0.003
\end{bmatrix}
\]

for \( d_2 = -0.03, 0 \) or 0.03. Then a random sample of responses of size \( r = 1, 2 \) and 5 was drawn from each conditional distribution \( Y_{ij} | b_j \sim N(x_i^T b_j, \sigma^2) \), where \( x_i = c(1, x_i) \), the \( x_i \) are as before 0, 3, 6, 9, 12, 18, 24 and 36 months, and \( \sigma = 0.5, 1 \) and 2. The total number of observations used in the analysis is then \( N = knr \), where \( n \) denotes the number of times \( (n = 8) \). Note that, without loss of generality, the same times \( x_i \) were used for each batch. This was done for two reasons: firstly, the coding and analysis are easier and less time consuming, and secondly, these times are those suggested by the regulatory authorities (even though other times may be more appropriate, c.f. Section 1.1).

Assuming that the specification limit was again set at \( \kappa = 90 \) and that \( \epsilon = 0.01 \), the true shelf-life (in months) is given by the smaller root to \( q_{t, \epsilon} = \kappa \). For each combination of \( k, r, d_2 \) and \( \sigma \), 1000 simulations were performed and for each simulation the data were analysed using each method, using a 95% lower confidence bound (\( \alpha = 0.05 \)). For each method the proportion of label shelf-lives that were smaller than the true shelf-life was calculated. These proportions can be found in Tables 3.1–3.9. Approximate confidence bounds for the true proportions can again be obtained using the normal approximation to the binomial described in Section 2.3.

The simulation results are summarised in the following tables and graphs. The coverage probabilities are displayed in Tables 3.1–3.9 and Figure 3.2; the bias is summarised in
Tables 3.10–3.18 and Figure 3.3; and the mean square error is summarised in Tables 3.19–3.27 and Figure 3.4. Again, the “bias” is not a bias in the true sense, but serves as an indicator of how conservative a particular method is (c.f. Section 2.3). As in Chapter 2 the methods are abbreviated, and the key is as follows.

- z.t and U.t denote the normal approximation approaches (3.7), using δ given in (3.5) and (3.6), respectively. The confidence bounds are calculated using $t$ distribution.
- CPL denotes the constrained profile likelihood approach (3.14)
- TMCPL denotes the truncated modified constrained profile likelihood approach (3.15).
- NCT1 and NCT2 denote the non-central t-distribution approach (3.17), using the unmodified δ and the modified δ given in (3.19).

Similar to Chapter 2, these tables can be summarized further by fitting appropriate models. In the case of the coverage probability it can be seen from Tables 3.1-3.9 and Figure 3.2 that the coverage probability does not appear to depend appreciably on the number of assays per batch. A similar full model for the coverage probability to that of Chapter 2 was fitted here. However, the full model involved $1/k, 1/\sqrt{N}, \sigma^*$ and $d_*^2$ as explanatory variables. Again, standard model reduction was conducted by testing higher order terms in turn and observing marginality (assuming a significance level of 1%). Table 3.28 gives the parameter estimates for all significant effects (interactions that were not significant under any method are omitted from the table). Even though only one value of $\beta_1$ was used in these simulations, it is not unreasonable to expect that the methods behave in a similar way to those in Chapter 2.

A similar modeling approach to that in Chapter 2 was also undertaken for the bias, and the results of the significant parameter estimates (at the 1% significance level) are presented in Table 3.29. The results are discussed in Section 3.4.
<table>
<thead>
<tr>
<th>(k)</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>(r)</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>(3.7) using (3.5)</td>
<td>0.991</td>
<td>0.981</td>
<td>0.973</td>
<td>0.990</td>
<td>0.993</td>
<td>0.988</td>
<td>1.000</td>
<td>0.999</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>(3.7) using (3.6)</td>
<td>0.993</td>
<td>0.984</td>
<td>0.976</td>
<td>0.990</td>
<td>0.994</td>
<td>0.990</td>
<td>1.000</td>
<td>0.999</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>TMPL (3.11)</td>
<td>1.000</td>
<td>0.998</td>
<td>0.998</td>
<td>0.998</td>
<td>1.000</td>
<td>1.000</td>
<td>0.999</td>
<td>0.996</td>
<td>0.933</td>
<td>0.97</td>
<td>0.996</td>
<td>0.986</td>
</tr>
<tr>
<td>CPL (3.14)</td>
<td>0.853</td>
<td>0.837</td>
<td>0.866</td>
<td>0.863</td>
<td>0.880</td>
<td>0.877</td>
<td>0.904</td>
<td>0.903</td>
<td>0.899</td>
<td>0.934</td>
<td>0.920</td>
<td>0.915</td>
</tr>
<tr>
<td>TMCPL (3.15)</td>
<td>0.990</td>
<td>0.994</td>
<td>0.981</td>
<td>0.968</td>
<td>0.971</td>
<td>0.971</td>
<td>0.955</td>
<td>0.945</td>
<td>0.956</td>
<td>0.963</td>
<td>0.945</td>
<td>0.949</td>
</tr>
<tr>
<td>NCT (3.17)</td>
<td>0.961</td>
<td>0.961</td>
<td>0.956</td>
<td>0.943</td>
<td>0.945</td>
<td>0.945</td>
<td>0.942</td>
<td>0.922</td>
<td>0.939</td>
<td>0.948</td>
<td>0.921</td>
<td>0.918</td>
</tr>
<tr>
<td>NCT (3.17) using (3.19)</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>0.974</td>
<td>0.979</td>
<td>0.981</td>
<td>0.968</td>
<td>0.957</td>
<td>0.978</td>
<td>0.976</td>
<td>0.959</td>
<td>0.965</td>
</tr>
</tbody>
</table>

Table 3.1: Proportion of label shelf-lives that were smaller than the true shelf-life for parameter values \(\beta_1 = -0.3\), \(d_2 = 0.00\), \(\sigma = 0.5\), \(\epsilon = 0.01\) and \(\alpha = 0.05\).

<table>
<thead>
<tr>
<th>(k)</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>(r)</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>(3.7) using (3.5)</td>
<td>1.000</td>
<td>0.998</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>(3.7) using (3.6)</td>
<td>1.000</td>
<td>0.998</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>TMPL (3.11)</td>
<td>0.997</td>
<td>0.999</td>
<td>0.999</td>
<td>1.000</td>
<td>0.999</td>
<td>1.000</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>0.995</td>
<td>0.986</td>
</tr>
<tr>
<td>CPL (3.14)</td>
<td>0.875</td>
<td>0.861</td>
<td>0.858</td>
<td>0.885</td>
<td>0.870</td>
<td>0.877</td>
<td>0.902</td>
<td>0.898</td>
<td>0.897</td>
<td>0.917</td>
<td>0.922</td>
<td>0.926</td>
</tr>
<tr>
<td>TMCPL (3.15)</td>
<td>0.997</td>
<td>1.000</td>
<td>0.993</td>
<td>0.987</td>
<td>0.981</td>
<td>0.984</td>
<td>0.960</td>
<td>0.954</td>
<td>0.956</td>
<td>0.951</td>
<td>0.950</td>
<td>0.963</td>
</tr>
<tr>
<td>NCT (3.17)</td>
<td>0.957</td>
<td>0.959</td>
<td>0.939</td>
<td>0.924</td>
<td>0.916</td>
<td>0.917</td>
<td>0.876</td>
<td>0.876</td>
<td>0.854</td>
<td>0.809</td>
<td>0.800</td>
<td>0.794</td>
</tr>
<tr>
<td>NCT (3.17) using (3.19)</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>0.994</td>
<td>0.994</td>
<td>0.993</td>
<td>0.989</td>
<td>0.988</td>
<td>0.996</td>
</tr>
</tbody>
</table>

Table 3.2: Proportion of label shelf-lives that were smaller than the true shelf-life for parameter values \(\beta_1 = -0.3\), \(d_2 = 0.00\), \(\sigma = 1.0\), \(\epsilon = 0.01\) and \(\alpha = 0.05\).
### Table 3.3: Proportion of label shelf-lives that were smaller than the true shelf-life for parameter values $\beta_1 = -0.3$, $d_2 = 0.00$, $\sigma = 2.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>$k$</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r$</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(3.7) using (3.5)</th>
<th>1.00</th>
<th>1.00</th>
<th>1.00</th>
<th>1.00</th>
<th>1.00</th>
<th>1.00</th>
<th>1.00</th>
<th>1.00</th>
<th>1.00</th>
<th>1.00</th>
<th>1.00</th>
<th>1.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>(3.7) using (3.6)</td>
<td>0.914</td>
<td>0.897</td>
<td>0.865</td>
<td>0.901</td>
<td>0.891</td>
<td>0.893</td>
<td>0.911</td>
<td>0.910</td>
<td>0.927</td>
<td>0.918</td>
<td>0.924</td>
<td>0.933</td>
</tr>
<tr>
<td>TMPL (3.11)</td>
<td>0.987</td>
<td>0.998</td>
<td>0.999</td>
<td>0.985</td>
<td>0.996</td>
<td>0.994</td>
<td>0.958</td>
<td>0.961</td>
<td>0.971</td>
<td>0.948</td>
<td>0.954</td>
<td>0.953</td>
</tr>
<tr>
<td>CPL (3.14)</td>
<td>0.949</td>
<td>0.919</td>
<td>0.901</td>
<td>0.782</td>
<td>0.734</td>
<td>0.703</td>
<td>0.528</td>
<td>0.405</td>
<td>0.364</td>
<td>0.169</td>
<td>0.095</td>
<td>0.065</td>
</tr>
<tr>
<td>NCT (3.17)</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>NCT (3.17) using (3.19)</td>
<td>0.978</td>
<td>0.975</td>
<td>0.963</td>
<td>0.987</td>
<td>0.987</td>
<td>0.986</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>(3.7) using (3.5)</td>
<td>0.983</td>
<td>0.979</td>
<td>0.968</td>
<td>0.988</td>
<td>0.988</td>
<td>0.989</td>
<td>1.000</td>
<td>0.999</td>
<td>0.999</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>(3.7) using (3.6)</td>
<td>0.992</td>
<td>0.996</td>
<td>1.000</td>
<td>0.998</td>
<td>1.000</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>0.987</td>
<td>0.990</td>
<td>0.972</td>
<td>0.965</td>
</tr>
<tr>
<td>TMPL (3.11)</td>
<td>0.798</td>
<td>0.813</td>
<td>0.817</td>
<td>0.866</td>
<td>0.854</td>
<td>0.839</td>
<td>0.894</td>
<td>0.907</td>
<td>0.878</td>
<td>0.916</td>
<td>0.911</td>
<td>0.924</td>
</tr>
<tr>
<td>CPL (3.14)</td>
<td>0.987</td>
<td>0.982</td>
<td>0.968</td>
<td>0.964</td>
<td>0.951</td>
<td>0.945</td>
<td>0.948</td>
<td>0.960</td>
<td>0.938</td>
<td>0.953</td>
<td>0.944</td>
<td>0.951</td>
</tr>
<tr>
<td>TMCPL (3.15)</td>
<td>0.957</td>
<td>0.952</td>
<td>0.941</td>
<td>0.951</td>
<td>0.941</td>
<td>0.936</td>
<td>0.945</td>
<td>0.954</td>
<td>0.932</td>
<td>0.935</td>
<td>0.928</td>
<td>0.941</td>
</tr>
<tr>
<td>NCT (3.17)</td>
<td>0.991</td>
<td>0.994</td>
<td>0.998</td>
<td>0.970</td>
<td>0.971</td>
<td>0.971</td>
<td>0.961</td>
<td>0.972</td>
<td>0.959</td>
<td>0.967</td>
<td>0.957</td>
<td>0.970</td>
</tr>
</tbody>
</table>

### Table 3.4: Proportion of label shelf-lives that were smaller than the true shelf-life for parameter values $\beta_1 = -0.3$, $d_2 = 0.03$, $\sigma = 0.5$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>$k$</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r$</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(3.7) using (3.5)</th>
<th>0.978</th>
<th>0.975</th>
<th>0.963</th>
<th>0.987</th>
<th>0.987</th>
<th>0.986</th>
<th>0.999</th>
<th>0.999</th>
<th>0.999</th>
<th>1.000</th>
<th>1.000</th>
<th>1.000</th>
</tr>
</thead>
<tbody>
<tr>
<td>(3.7) using (3.6)</td>
<td>0.983</td>
<td>0.979</td>
<td>0.968</td>
<td>0.988</td>
<td>0.988</td>
<td>0.989</td>
<td>1.000</td>
<td>0.999</td>
<td>0.999</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>TMPL (3.11)</td>
<td>0.992</td>
<td>0.996</td>
<td>1.000</td>
<td>0.998</td>
<td>1.000</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>0.987</td>
<td>0.990</td>
<td>0.972</td>
<td>0.965</td>
</tr>
<tr>
<td>CPL (3.14)</td>
<td>0.798</td>
<td>0.813</td>
<td>0.817</td>
<td>0.866</td>
<td>0.854</td>
<td>0.839</td>
<td>0.894</td>
<td>0.907</td>
<td>0.878</td>
<td>0.916</td>
<td>0.911</td>
<td>0.924</td>
</tr>
<tr>
<td>TMCPL (3.15)</td>
<td>0.987</td>
<td>0.982</td>
<td>0.968</td>
<td>0.964</td>
<td>0.951</td>
<td>0.945</td>
<td>0.948</td>
<td>0.960</td>
<td>0.938</td>
<td>0.953</td>
<td>0.944</td>
<td>0.951</td>
</tr>
<tr>
<td>NCT (3.17)</td>
<td>0.957</td>
<td>0.952</td>
<td>0.941</td>
<td>0.951</td>
<td>0.941</td>
<td>0.936</td>
<td>0.945</td>
<td>0.954</td>
<td>0.932</td>
<td>0.935</td>
<td>0.928</td>
<td>0.941</td>
</tr>
<tr>
<td>NCT (3.17) using (3.19)</td>
<td>0.991</td>
<td>0.994</td>
<td>0.998</td>
<td>0.970</td>
<td>0.971</td>
<td>0.971</td>
<td>0.961</td>
<td>0.972</td>
<td>0.959</td>
<td>0.967</td>
<td>0.957</td>
<td>0.970</td>
</tr>
</tbody>
</table>
### Table 3.5: Proportion of label shelf-lives that were smaller than the true shelf-life for parameter values $\beta_1 = -0.3$, $d_2 = 0.03$, $\sigma = 1.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.  

<table>
<thead>
<tr>
<th>k</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>(3.7) using (3.5)</td>
<td>0.998</td>
<td>0.996</td>
<td>0.997</td>
<td>0.999</td>
<td>0.998</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>(3.7) using (3.6)</td>
<td>0.999</td>
<td>0.996</td>
<td>0.997</td>
<td>0.999</td>
<td>0.998</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>TMPL (3.11)</td>
<td>0.996</td>
<td>0.996</td>
<td>0.999</td>
<td>0.997</td>
<td>0.998</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>0.994</td>
<td>0.999</td>
<td>0.988</td>
</tr>
<tr>
<td>CPL (3.14)</td>
<td>0.820</td>
<td>0.842</td>
<td>0.827</td>
<td>0.848</td>
<td>0.876</td>
<td>0.869</td>
<td>0.896</td>
<td>0.894</td>
<td>0.908</td>
<td>0.922</td>
<td>0.908</td>
<td>0.926</td>
</tr>
<tr>
<td>TMCPL (3.15)</td>
<td>0.997</td>
<td>0.997</td>
<td>0.985</td>
<td>0.965</td>
<td>0.970</td>
<td>0.974</td>
<td>0.951</td>
<td>0.946</td>
<td>0.955</td>
<td>0.949</td>
<td>0.947</td>
<td>0.958</td>
</tr>
<tr>
<td>NCT (3.17)</td>
<td>0.952</td>
<td>0.939</td>
<td>0.938</td>
<td>0.913</td>
<td>0.935</td>
<td>0.937</td>
<td>0.896</td>
<td>0.895</td>
<td>0.901</td>
<td>0.854</td>
<td>0.825</td>
<td>0.855</td>
</tr>
<tr>
<td>NCT (3.17) using (3.19)</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>0.993</td>
<td>0.997</td>
<td>1.000</td>
<td>0.986</td>
<td>0.987</td>
<td>0.989</td>
<td>0.984</td>
<td>0.978</td>
<td>0.989</td>
</tr>
</tbody>
</table>

### Table 3.6: Proportion of label shelf-lives that were smaller than the true shelf-life for parameter values $\beta_1 = -0.3$, $d_2 = 0.03$, $\sigma = 2.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.  

<table>
<thead>
<tr>
<th>k</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>(3.7) using (3.5)</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>(3.7) using (3.6)</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>TMPL (3.11)</td>
<td>0.997</td>
<td>0.998</td>
<td>1.000</td>
<td>0.999</td>
<td>1.000</td>
<td>0.999</td>
<td>0.989</td>
<td>0.984</td>
<td>0.993</td>
<td>0.984</td>
<td>0.985</td>
<td>0.977</td>
</tr>
<tr>
<td>CPL (3.14)</td>
<td>0.904</td>
<td>0.851</td>
<td>0.830</td>
<td>0.893</td>
<td>0.882</td>
<td>0.887</td>
<td>0.914</td>
<td>0.909</td>
<td>0.899</td>
<td>0.945</td>
<td>0.925</td>
<td>0.928</td>
</tr>
<tr>
<td>TMCPL (3.15)</td>
<td>0.991</td>
<td>0.993</td>
<td>0.995</td>
<td>0.983</td>
<td>0.993</td>
<td>0.982</td>
<td>0.948</td>
<td>0.954</td>
<td>0.945</td>
<td>0.964</td>
<td>0.954</td>
<td>0.947</td>
</tr>
<tr>
<td>NCT (3.17)</td>
<td>0.944</td>
<td>0.909</td>
<td>0.887</td>
<td>0.822</td>
<td>0.808</td>
<td>0.809</td>
<td>0.656</td>
<td>0.594</td>
<td>0.530</td>
<td>0.346</td>
<td>0.252</td>
<td>0.207</td>
</tr>
<tr>
<td>NCT (3.17) using (3.19)</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>0.999</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>
Table 3.7: Proportion of label shelf-lives that were smaller than the true shelf-life for parameter values $\beta_1 = -0.3$, $d_2 = -0.03$, $\sigma = 0.5$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>k</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>(3.7) using (3.5)</td>
<td>0.991</td>
<td>0.991</td>
<td>0.989</td>
<td>0.997</td>
<td>0.995</td>
<td>0.996</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>(3.7) using (3.6)</td>
<td>0.993</td>
<td>0.994</td>
<td>0.991</td>
<td>0.997</td>
<td>0.995</td>
<td>0.997</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>TMPL (3.11)</td>
<td>0.997</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>0.994</td>
<td>0.992</td>
<td>0.993</td>
<td>0.991</td>
<td>0.987</td>
<td>0.982</td>
<td></td>
</tr>
<tr>
<td>CPL (3.14)</td>
<td>0.910</td>
<td>0.909</td>
<td>0.914</td>
<td>0.883</td>
<td>0.885</td>
<td>0.872</td>
<td>0.889</td>
<td>0.903</td>
<td>0.919</td>
<td>0.926</td>
<td>0.920</td>
<td>0.912</td>
</tr>
<tr>
<td>TMCPL (3.15)</td>
<td>1.000</td>
<td>1.000</td>
<td>0.996</td>
<td>0.992</td>
<td>0.980</td>
<td>0.982</td>
<td>0.961</td>
<td>0.952</td>
<td>0.980</td>
<td>0.951</td>
<td>0.957</td>
<td>0.948</td>
</tr>
<tr>
<td>NCT (3.17)</td>
<td>0.970</td>
<td>0.973</td>
<td>0.957</td>
<td>0.934</td>
<td>0.935</td>
<td>0.933</td>
<td>0.932</td>
<td>0.920</td>
<td>0.943</td>
<td>0.916</td>
<td>0.904</td>
<td>0.882</td>
</tr>
<tr>
<td>NCT (3.17) using (3.19)</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>0.982</td>
<td>0.987</td>
<td>0.993</td>
<td>0.971</td>
<td>0.973</td>
<td>0.989</td>
<td>0.967</td>
<td>0.979</td>
<td>0.978</td>
</tr>
</tbody>
</table>

Table 3.8: Proportion of label shelf-lives that were smaller than the true shelf-life for parameter values $\beta_1 = -0.3$, $d_2 = -0.03$, $\sigma = 1.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.
### Table 3.9: Proportion of label shelf-lives that were smaller than the true shelf-life for parameter values $\beta_1 = -0.3$, $d_2 = -0.03$, $\sigma = 2.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.  

<table>
<thead>
<tr>
<th>k</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>(3.7) using (3.5)</td>
<td>1.00</td>
<td>1.00</td>
<td>0.98</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>(3.7) using (3.6)</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>TMPL (3.11)</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td>CPL (3.14)</td>
<td>0.93</td>
<td>0.94</td>
<td>0.91</td>
<td>0.93</td>
<td>0.93</td>
<td>0.93</td>
<td>0.93</td>
<td>0.93</td>
<td>0.93</td>
<td>0.93</td>
<td>0.93</td>
</tr>
<tr>
<td>TMCPL (3.15)</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>1.00</td>
<td>0.96</td>
<td>0.97</td>
<td>0.97</td>
<td>0.95</td>
<td>0.95</td>
<td>0.96</td>
</tr>
<tr>
<td>NCT (3.17)</td>
<td>0.92</td>
<td>0.89</td>
<td>0.86</td>
<td>0.72</td>
<td>0.59</td>
<td>0.51</td>
<td>0.28</td>
<td>0.16</td>
<td>0.07</td>
<td>0.03</td>
<td>0.00</td>
</tr>
<tr>
<td>NCT (3.17) using (3.19)</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Figure 3.2: Coverage Probability plotted against $1/\sqrt{N}$ for each method and combination of parameters.
<table>
<thead>
<tr>
<th></th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>(3.7) using (3.5)</td>
<td>-17.3</td>
<td>-16.3</td>
<td>-15.3</td>
<td>-13.9</td>
<td>-13.7</td>
<td>-13.3</td>
<td>-12.6</td>
<td>-12.6</td>
<td>-12.6</td>
<td>-12.0</td>
<td>-11.9</td>
<td>-11.9</td>
</tr>
<tr>
<td>(3.7) using (3.6)</td>
<td>-17.8</td>
<td>-16.9</td>
<td>-15.8</td>
<td>-14.4</td>
<td>-14.2</td>
<td>-13.9</td>
<td>-12.9</td>
<td>-12.9</td>
<td>-12.9</td>
<td>-12.2</td>
<td>-12.1</td>
<td>-12.0</td>
</tr>
<tr>
<td>TMPL (3.11)</td>
<td>-20.8</td>
<td>-20.7</td>
<td>-20.9</td>
<td>-20.8</td>
<td>-20.9</td>
<td>-17.9</td>
<td>-17.9</td>
<td>-15.2</td>
<td>-14.0</td>
<td>-13.7</td>
<td>-13.5</td>
<td></td>
</tr>
<tr>
<td>CPL (3.14)</td>
<td>-7.7</td>
<td>-7.2</td>
<td>-7.5</td>
<td>-4.9</td>
<td>-5.1</td>
<td>-5.1</td>
<td>-3.6</td>
<td>-3.7</td>
<td>-3.6</td>
<td>-2.6</td>
<td>-2.6</td>
<td>-2.6</td>
</tr>
<tr>
<td>TMCPL (3.15)</td>
<td>-13.0</td>
<td>-12.5</td>
<td>-12.3</td>
<td>-7.8</td>
<td>-7.9</td>
<td>-7.7</td>
<td>-4.7</td>
<td>-4.7</td>
<td>-4.5</td>
<td>-3.1</td>
<td>-3.1</td>
<td>-2.9</td>
</tr>
<tr>
<td>NCT (3.17)</td>
<td>-15.4</td>
<td>-15.0</td>
<td>-14.8</td>
<td>-8.6</td>
<td>-9.0</td>
<td>-8.7</td>
<td>-4.7</td>
<td>-4.7</td>
<td>-4.6</td>
<td>-2.9</td>
<td>-2.8</td>
<td>-2.7</td>
</tr>
<tr>
<td>NCT (3.17) using (3.19)</td>
<td>-17.6</td>
<td>-17.2</td>
<td>-17.2</td>
<td>-9.8</td>
<td>-10.3</td>
<td>-10.1</td>
<td>-5.4</td>
<td>-5.5</td>
<td>-5.5</td>
<td>-3.4</td>
<td>-3.4</td>
<td>-3.4</td>
</tr>
</tbody>
</table>

Table 3.10: Bias of label shelf-lives for parameter values $\beta_1 = -0.3$, $d_2 = 0.00$, $\sigma = 0.5$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th></th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>(3.7) using (3.5)</td>
<td>-18.5</td>
<td>-17.5</td>
<td>-16.8</td>
<td>-17.0</td>
<td>-15.9</td>
<td>-15.4</td>
<td>-15.5</td>
<td>-15.0</td>
<td>-14.6</td>
<td>-14.7</td>
<td>-14.3</td>
<td>-14.3</td>
</tr>
<tr>
<td>(3.7) using (3.6)</td>
<td>-18.6</td>
<td>-17.7</td>
<td>-17.0</td>
<td>-17.2</td>
<td>-16.1</td>
<td>-15.7</td>
<td>-15.7</td>
<td>-15.1</td>
<td>-14.7</td>
<td>-14.8</td>
<td>-14.4</td>
<td>-14.4</td>
</tr>
<tr>
<td>TMPL (3.11)</td>
<td>-18.6</td>
<td>-19.0</td>
<td>-19.4</td>
<td>-19.0</td>
<td>-19.3</td>
<td>-19.4</td>
<td>-15.3</td>
<td>-14.5</td>
<td>-13.2</td>
<td>-13.2</td>
<td>-11.3</td>
<td>-10.4</td>
</tr>
<tr>
<td>CPL (3.14)</td>
<td>-7.7</td>
<td>-6.8</td>
<td>-6.4</td>
<td>-5.3</td>
<td>-4.8</td>
<td>-4.9</td>
<td>-3.5</td>
<td>-3.4</td>
<td>-3.3</td>
<td>-2.5</td>
<td>-2.4</td>
<td>-2.5</td>
</tr>
<tr>
<td>TMCPL (3.15)</td>
<td>-13.3</td>
<td>-13.2</td>
<td>-12.4</td>
<td>-7.8</td>
<td>-7.6</td>
<td>-7.6</td>
<td>-4.6</td>
<td>-4.5</td>
<td>-4.3</td>
<td>-3.0</td>
<td>-2.9</td>
<td>-3.0</td>
</tr>
<tr>
<td>NCT (3.17)</td>
<td>-14.7</td>
<td>-14.4</td>
<td>-13.7</td>
<td>-8.3</td>
<td>-7.6</td>
<td>-7.7</td>
<td>-3.7</td>
<td>-3.5</td>
<td>-3.3</td>
<td>-1.8</td>
<td>-1.6</td>
<td>-1.6</td>
</tr>
<tr>
<td>NCT (3.17) using (3.19)</td>
<td>-19.1</td>
<td>-19.3</td>
<td>-19.5</td>
<td>-12.9</td>
<td>-12.6</td>
<td>-13.0</td>
<td>-6.5</td>
<td>-6.5</td>
<td>-6.4</td>
<td>-4.0</td>
<td>-3.9</td>
<td>-4.1</td>
</tr>
</tbody>
</table>

Table 3.11: Bias of label shelf-lives for parameter values $\beta_1 = -0.3$, $d_2 = 0.00$, $\sigma = 1.0$, $\epsilon = 0.01$ and $\alpha = 0.05$. 
### Table 3.12: Bias of label shelf-lives for parameter values $\beta_1 = -0.3$, $d_2 = 0.00$, $\sigma = 2.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>$k$</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r$</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>TMPL (3.11)</td>
<td>$-14.4$</td>
<td>$-14.7$</td>
<td>$-14.7$</td>
<td>$-14.7$</td>
<td>$-14.5$</td>
<td>$-14.1$</td>
<td>$-8.3$</td>
<td>$-7.5$</td>
<td>$-6.2$</td>
<td>$-4.7$</td>
<td>$-4.7$</td>
<td>$-3.2$</td>
</tr>
<tr>
<td>CPL (3.14)</td>
<td>$-8.1$</td>
<td>$-7.0$</td>
<td>$-5.6$</td>
<td>$-5.6$</td>
<td>$-4.9$</td>
<td>$-4.1$</td>
<td>$-3.8$</td>
<td>$-3.1$</td>
<td>$-2.9$</td>
<td>$-2.5$</td>
<td>$-2.2$</td>
<td>$-2.1$</td>
</tr>
<tr>
<td>TMCPL (3.15)</td>
<td>$-10.5$</td>
<td>$-10.9$</td>
<td>$-9.6$</td>
<td>$-7.0$</td>
<td>$-6.6$</td>
<td>$-6.1$</td>
<td>$-4.3$</td>
<td>$-3.8$</td>
<td>$-3.7$</td>
<td>$-2.8$</td>
<td>$-2.5$</td>
<td>$-2.4$</td>
</tr>
<tr>
<td>NCT (3.17)</td>
<td>$-12.1$</td>
<td>$-10.8$</td>
<td>$-9.7$</td>
<td>$-5.9$</td>
<td>$-4.4$</td>
<td>$-3.2$</td>
<td>$-0.5$</td>
<td>$0.7$</td>
<td>$1.0$</td>
<td>$1.9$</td>
<td>$2.6$</td>
<td>$2.9$</td>
</tr>
<tr>
<td>NCT (3.17) using (3.19)</td>
<td>$-14.9$</td>
<td>$-14.9$</td>
<td>$-14.9$</td>
<td>$-14.6$</td>
<td>$-14.8$</td>
<td>$-14.9$</td>
<td>$-10.6$</td>
<td>$-10.4$</td>
<td>$-10.5$</td>
<td>$-6.1$</td>
<td>$-6.0$</td>
<td>$-6.1$</td>
</tr>
</tbody>
</table>

### Table 3.13: Bias of label shelf-lives for parameter values $\beta_1 = -0.3$, $d_2 = 0.03$, $\sigma = 0.5$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>$k$</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r$</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>(3.7) using (3.5)</td>
<td>$-16.2$</td>
<td>$-15.4$</td>
<td>$-14.3$</td>
<td>$-14.2$</td>
<td>$-13.4$</td>
<td>$-12.9$</td>
<td>$-12.8$</td>
<td>$-12.7$</td>
<td>$-12.4$</td>
<td>$-12.2$</td>
<td>$-12.1$</td>
<td>$-12.2$</td>
</tr>
<tr>
<td>(3.7) using (3.6)</td>
<td>$-16.5$</td>
<td>$-15.8$</td>
<td>$-14.8$</td>
<td>$-14.7$</td>
<td>$-13.9$</td>
<td>$-13.3$</td>
<td>$-13.1$</td>
<td>$-13.0$</td>
<td>$-12.7$</td>
<td>$-12.4$</td>
<td>$-12.2$</td>
<td>$-12.3$</td>
</tr>
<tr>
<td>TMPL (3.11)</td>
<td>$-18.7$</td>
<td>$-18.8$</td>
<td>$-18.9$</td>
<td>$-18.9$</td>
<td>$-18.9$</td>
<td>$-18.9$</td>
<td>$-18.5$</td>
<td>$-18.5$</td>
<td>$-14.7$</td>
<td>$-5.3$</td>
<td>$-4.6$</td>
<td>$-3.8$</td>
</tr>
<tr>
<td>CPL (3.14)</td>
<td>$-6.3$</td>
<td>$-6.5$</td>
<td>$-6.5$</td>
<td>$-5.4$</td>
<td>$-5.3$</td>
<td>$-5.0$</td>
<td>$-4.0$</td>
<td>$-4.1$</td>
<td>$-3.7$</td>
<td>$-2.9$</td>
<td>$-2.8$</td>
<td>$-2.9$</td>
</tr>
<tr>
<td>TMCPL (3.15)</td>
<td>$-12.1$</td>
<td>$-11.7$</td>
<td>$-11.4$</td>
<td>$-8.0$</td>
<td>$-7.5$</td>
<td>$-7.2$</td>
<td>$-4.9$</td>
<td>$-5.0$</td>
<td>$-4.6$</td>
<td>$-3.4$</td>
<td>$-3.2$</td>
<td>$-3.3$</td>
</tr>
<tr>
<td>NCT (3.17)</td>
<td>$-13.8$</td>
<td>$-13.9$</td>
<td>$-13.6$</td>
<td>$-9.4$</td>
<td>$-9.2$</td>
<td>$-8.9$</td>
<td>$-5.2$</td>
<td>$-5.3$</td>
<td>$-4.9$</td>
<td>$-3.3$</td>
<td>$-3.1$</td>
<td>$-3.2$</td>
</tr>
<tr>
<td>NCT (3.17) using (3.19)</td>
<td>$-15.7$</td>
<td>$-15.8$</td>
<td>$-15.8$</td>
<td>$-10.3$</td>
<td>$-10.3$</td>
<td>$-10.0$</td>
<td>$-5.8$</td>
<td>$-5.9$</td>
<td>$-5.6$</td>
<td>$-3.8$</td>
<td>$-3.6$</td>
<td>$-3.7$</td>
</tr>
</tbody>
</table>
### Table 3.14: Bias of label shelf-lives for parameter values $\beta_1 = -0.3$, $d_2 = 0.03$, $\sigma = 1.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>k</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
</tbody>
</table>

- **(3.7) using (3.5)**: $-16.8$ $-16.5$ $-15.8$ $-16.1$ $-15.5$ $-14.6$ $-15.2$ $-14.3$ $-14.2$ $-14.2$ $-13.8$ $-13.8$
- **(3.7) using (3.6)**: $-16.9$ $-16.6$ $-15.9$ $-16.2$ $-15.7$ $-14.8$ $-15.3$ $-14.4$ $-14.3$ $-14.3$ $-13.8$ $-13.8$
- **TMPL (3.11)**: $-16.7$ $-17.5$ $-17.6$ $-17.3$ $-17.6$ $-17.7$ $-15.5$ $-15.1$ $-15.3$ $-9.7$ $-6.3$ $-4.8$
- **CPL (3.14)**: $-6.4$ $-6.5$ $-6.3$ $-5.2$ $-5.2$ $-5.2$ $-3.9$ $-3.5$ $-3.9$ $-2.8$ $-2.6$ $-2.7$
- **TMCPL (3.15)**: $-12.1$ $-12.3$ $-12.0$ $-7.4$ $-7.5$ $-7.6$ $-4.8$ $-4.5$ $-4.8$ $-3.2$ $-3.1$ $-3.2$
- **NCT (3.17)**: $-13.3$ $-13.3$ $-13.0$ $-8.5$ $-8.3$ $-8.2$ $-4.4$ $-3.9$ $-4.2$ $-2.3$ $-2.1$ $-2.1$
- **NCT (3.17) using (3.19)**: $-17.2$ $-17.5$ $-17.7$ $-12.1$ $-12.2$ $-12.2$ $-6.7$ $-6.4$ $-6.8$ $-4.2$ $-4.1$ $-4.2$

### Table 3.15: Bias of label shelf-lives for parameter values $\beta_1 = -0.3$, $d_2 = 0.03$, $\sigma = 2.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>k</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
</tbody>
</table>

- **TMPL (3.11)**: $-13.2$ $-13.5$ $-13.6$ $-13.5$ $-13.4$ $-13.2$ $-8.4$ $-7.4$ $-6.6$ $-4.7$ $-4.5$ $-3.5$
- **CPL (3.14)**: $-7.0$ $-6.1$ $-5.0$ $-5.5$ $-4.9$ $-4.5$ $-3.9$ $-3.4$ $-3.1$ $-2.8$ $-2.5$ $-2.3$
- **TMCPL (3.15)**: $-9.8$ $-9.9$ $-8.7$ $-6.6$ $-6.4$ $-6.1$ $-4.3$ $-4.0$ $-3.8$ $-3.0$ $-2.7$ $-2.6$
- **NCT (3.17)**: $-10.9$ $-9.8$ $-8.9$ $-6.3$ $-5.1$ $-4.5$ $-1.5$ $-0.7$ $-0.1$ $0.8$ $1.4$ $1.8$
<table>
<thead>
<tr>
<th>k</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>(3.7) using (3.5)</td>
<td>-18.6</td>
<td>-17.4</td>
<td>-16.5</td>
<td>-13.9</td>
<td>-13.0</td>
<td>-12.7</td>
<td>-11.5</td>
<td>-11.2</td>
<td>-11.3</td>
<td>-10.4</td>
<td>-10.2</td>
<td>-10.2</td>
</tr>
<tr>
<td>(3.7) using (3.6)</td>
<td>-19.3</td>
<td>-18.3</td>
<td>-17.2</td>
<td>-14.7</td>
<td>-13.7</td>
<td>-13.5</td>
<td>-11.9</td>
<td>-11.6</td>
<td>-11.7</td>
<td>-10.6</td>
<td>-10.4</td>
<td>-10.3</td>
</tr>
<tr>
<td>TMPL (3.11)</td>
<td>-23.2</td>
<td>-23.7</td>
<td>-23.8</td>
<td>-23.7</td>
<td>-23.7</td>
<td>-23.7</td>
<td>-16.8</td>
<td>-16.2</td>
<td>-15.8</td>
<td>-15.4</td>
<td>-14.5</td>
<td>-14.5</td>
</tr>
<tr>
<td>CPL (3.14)</td>
<td>-8.7</td>
<td>-8.4</td>
<td>-8.9</td>
<td>-4.8</td>
<td>-4.6</td>
<td>-4.6</td>
<td>-2.9</td>
<td>-2.9</td>
<td>-3.0</td>
<td>-2.1</td>
<td>-2.0</td>
<td>-2.0</td>
</tr>
<tr>
<td>TMCPL (3.15)</td>
<td>-15.5</td>
<td>-15.1</td>
<td>-14.8</td>
<td>-8.4</td>
<td>-8.2</td>
<td>-8.1</td>
<td>-4.0</td>
<td>-4.0</td>
<td>-4.2</td>
<td>-2.6</td>
<td>-2.4</td>
<td>-2.4</td>
</tr>
<tr>
<td>NCT (3.17)</td>
<td>-16.5</td>
<td>-16.1</td>
<td>-16.1</td>
<td>-7.4</td>
<td>-7.2</td>
<td>-6.9</td>
<td>-3.6</td>
<td>-3.5</td>
<td>-3.6</td>
<td>-2.1</td>
<td>-1.9</td>
<td>-1.9</td>
</tr>
<tr>
<td>NCT (3.17) using (3.19)</td>
<td>-19.5</td>
<td>-19.2</td>
<td>-19.5</td>
<td>-9.2</td>
<td>-9.2</td>
<td>-9.0</td>
<td>-4.5</td>
<td>-4.5</td>
<td>-4.7</td>
<td>-2.9</td>
<td>-2.7</td>
<td>-2.8</td>
</tr>
</tbody>
</table>

Table 3.16: Bias of label shelf-lives for parameter values \( \beta_1 = -0.3, d_2 = -0.03, \sigma = 0.5, \epsilon = 0.01 \) and \( \alpha = 0.05 \).

<table>
<thead>
<tr>
<th>k</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>(3.7) using (3.5)</td>
<td>-20.6</td>
<td>-19.8</td>
<td>-18.7</td>
<td>-18.3</td>
<td>-17.2</td>
<td>-15.9</td>
<td>-15.8</td>
<td>-15.1</td>
<td>-15.0</td>
<td>-14.6</td>
<td>-14.1</td>
<td>-14.2</td>
</tr>
<tr>
<td>(3.7) using (3.6)</td>
<td>-20.8</td>
<td>-20.0</td>
<td>-18.9</td>
<td>-18.6</td>
<td>-17.5</td>
<td>-16.1</td>
<td>-16.0</td>
<td>-15.3</td>
<td>-15.1</td>
<td>-14.7</td>
<td>-14.2</td>
<td>-14.3</td>
</tr>
<tr>
<td>TMPL (3.11)</td>
<td>-21.0</td>
<td>-21.3</td>
<td>-21.6</td>
<td>-20.7</td>
<td>-21.2</td>
<td>-21.8</td>
<td>-14.7</td>
<td>-13.8</td>
<td>-11.3</td>
<td>-13.5</td>
<td>-12.3</td>
<td>-6.5</td>
</tr>
<tr>
<td>CPL (3.14)</td>
<td>-8.9</td>
<td>-8.9</td>
<td>-7.8</td>
<td>-5.2</td>
<td>-4.6</td>
<td>-4.4</td>
<td>-2.7</td>
<td>-2.7</td>
<td>-2.7</td>
<td>-2.0</td>
<td>-1.9</td>
<td>-1.8</td>
</tr>
<tr>
<td>TMCPL (3.15)</td>
<td>-15.6</td>
<td>-15.8</td>
<td>-14.6</td>
<td>-8.3</td>
<td>-7.9</td>
<td>-8.0</td>
<td>-3.9</td>
<td>-3.8</td>
<td>-3.9</td>
<td>-2.5</td>
<td>-2.3</td>
<td>-2.2</td>
</tr>
<tr>
<td>NCT (3.17)</td>
<td>-16.2</td>
<td>-15.2</td>
<td>-14.8</td>
<td>-6.8</td>
<td>-5.9</td>
<td>-5.7</td>
<td>-2.2</td>
<td>-2.0</td>
<td>-1.8</td>
<td>-0.8</td>
<td>-0.5</td>
<td>-0.2</td>
</tr>
</tbody>
</table>

Table 3.17: Bias of label shelf-lives for parameter values \( \beta_1 = -0.3, d_2 = -0.03, \sigma = 1.0, \epsilon = 0.01 \) and \( \alpha = 0.05 \).
<table>
<thead>
<tr>
<th>k</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>(3.7) using (3.5)</td>
<td>-16.2</td>
<td>-16.3</td>
<td>-16.3</td>
<td>-16.3</td>
<td>-16.3</td>
<td>-16.3</td>
<td>-16.3</td>
<td>-16.3</td>
<td>-16.3</td>
<td>-16.3</td>
<td>-16.3</td>
<td>-16.3</td>
</tr>
<tr>
<td>(3.7) using (3.6)</td>
<td>-16.2</td>
<td>-16.3</td>
<td>-16.3</td>
<td>-16.3</td>
<td>-16.3</td>
<td>-16.3</td>
<td>-16.3</td>
<td>-16.3</td>
<td>-16.3</td>
<td>-16.3</td>
<td>-16.3</td>
<td>-16.3</td>
</tr>
<tr>
<td>TMPL (3.11)</td>
<td>-15.7</td>
<td>-16.0</td>
<td>-16.2</td>
<td>-16.0</td>
<td>-15.9</td>
<td>-15.3</td>
<td>-7.6</td>
<td>-6.2</td>
<td>-4.9</td>
<td>-4.5</td>
<td>-3.9</td>
<td>-2.4</td>
</tr>
<tr>
<td>CPL (3.14)</td>
<td>-8.3</td>
<td>-7.6</td>
<td>-5.9</td>
<td>-6.4</td>
<td>-4.8</td>
<td>-4.0</td>
<td>-3.5</td>
<td>-2.8</td>
<td>-2.3</td>
<td>-2.2</td>
<td>-1.8</td>
<td>-1.5</td>
</tr>
<tr>
<td>TMCPL (3.15)</td>
<td>-11.8</td>
<td>-12.2</td>
<td>-10.6</td>
<td>-8.1</td>
<td>-7.2</td>
<td>-6.4</td>
<td>-4.1</td>
<td>-3.5</td>
<td>-3.1</td>
<td>-2.6</td>
<td>-2.1</td>
<td>-1.8</td>
</tr>
<tr>
<td>NCT (3.17)</td>
<td>-12.3</td>
<td>-11.0</td>
<td>-9.7</td>
<td>-5.7</td>
<td>-2.6</td>
<td>-0.9</td>
<td>1.5</td>
<td>2.6</td>
<td>3.5</td>
<td>3.3</td>
<td>4.3</td>
<td>5.1</td>
</tr>
<tr>
<td>NCT (3.17) using (3.19)</td>
<td>-16.3</td>
<td>-16.3</td>
<td>-16.3</td>
<td>-16.0</td>
<td>-16.2</td>
<td>-16.3</td>
<td>-11.4</td>
<td>-11.3</td>
<td>-11.2</td>
<td>-6.3</td>
<td>-6.2</td>
<td>-6.2</td>
</tr>
</tbody>
</table>

Table 3.18: Bias of label shelf-lives for parameter values $\beta_1 = -0.3$, $d_2 = -0.03$, $\sigma = 2.0$, $\epsilon = 0.01$ and $\alpha = 0.05$. 
Figure 3.3: Bias plotted against $1/\sqrt{N}$ for each method and combination of parameters.
<table>
<thead>
<tr>
<th>k</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>(3.7) using (3.5)</td>
<td>333.8</td>
<td>308.7</td>
<td>279.7</td>
<td>226.8</td>
<td>219.3</td>
<td>208.7</td>
<td>173.6</td>
<td>172.8</td>
<td>172.9</td>
<td>150.3</td>
<td>149.1</td>
<td>146.8</td>
</tr>
<tr>
<td>(3.7) using (3.6)</td>
<td>348.0</td>
<td>323.1</td>
<td>294.4</td>
<td>242.5</td>
<td>234.7</td>
<td>224.5</td>
<td>183.0</td>
<td>181.9</td>
<td>181.4</td>
<td>154.4</td>
<td>153.0</td>
<td>150.6</td>
</tr>
<tr>
<td>TMPL (3.11)</td>
<td>434.0</td>
<td>433.4</td>
<td>436.8</td>
<td>434.9</td>
<td>437.3</td>
<td>436.0</td>
<td>353.5</td>
<td>354.7</td>
<td>283.8</td>
<td>262.6</td>
<td>255.5</td>
<td>252.9</td>
</tr>
<tr>
<td>CPL (3.14)</td>
<td>111.7</td>
<td>103.6</td>
<td>101.2</td>
<td>45.3</td>
<td>47.0</td>
<td>47.5</td>
<td>20.2</td>
<td>21.5</td>
<td>20.6</td>
<td>10.1</td>
<td>10.2</td>
<td>10.0</td>
</tr>
<tr>
<td>TMCPL (3.15)</td>
<td>195.6</td>
<td>178.5</td>
<td>178.1</td>
<td>76.9</td>
<td>78.0</td>
<td>75.7</td>
<td>28.9</td>
<td>29.9</td>
<td>27.3</td>
<td>12.8</td>
<td>12.8</td>
<td>12.0</td>
</tr>
<tr>
<td>NCT (3.17)</td>
<td>283.5</td>
<td>274.5</td>
<td>271.2</td>
<td>104.4</td>
<td>110.2</td>
<td>106.5</td>
<td>30.4</td>
<td>31.8</td>
<td>30.5</td>
<td>11.6</td>
<td>11.6</td>
<td>11.3</td>
</tr>
<tr>
<td>NCT (3.17) using (3.19)</td>
<td>331.2</td>
<td>320.3</td>
<td>321.8</td>
<td>123.5</td>
<td>131.6</td>
<td>128.0</td>
<td>37.1</td>
<td>38.9</td>
<td>37.9</td>
<td>14.9</td>
<td>15.1</td>
<td>14.9</td>
</tr>
</tbody>
</table>

Table 3.19: Mean Square Error of label shelf-lives for parameter values $\beta_1 = -0.3$, $d_2 = 0.00$, $\sigma = 0.5$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>k</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>(3.7) using (3.5)</td>
<td>352.2</td>
<td>322.6</td>
<td>302.2</td>
<td>303.2</td>
<td>269.6</td>
<td>254.8</td>
<td>252.4</td>
<td>234.9</td>
<td>221.8</td>
<td>221.8</td>
<td>209.5</td>
<td>208.7</td>
</tr>
<tr>
<td>(3.7) using (3.6)</td>
<td>355.7</td>
<td>327.9</td>
<td>307.5</td>
<td>309.5</td>
<td>275.3</td>
<td>261.3</td>
<td>258.0</td>
<td>240.0</td>
<td>226.3</td>
<td>225.2</td>
<td>212.4</td>
<td>211.2</td>
</tr>
<tr>
<td>TMPL (3.11)</td>
<td>352.4</td>
<td>368.4</td>
<td>378.3</td>
<td>367.3</td>
<td>376.5</td>
<td>379.1</td>
<td>267.9</td>
<td>248.5</td>
<td>222.6</td>
<td>231.4</td>
<td>189.4</td>
<td>169.5</td>
</tr>
<tr>
<td>CPL (3.14)</td>
<td>104.9</td>
<td>87.3</td>
<td>79.5</td>
<td>49.9</td>
<td>42.3</td>
<td>43.0</td>
<td>19.4</td>
<td>18.7</td>
<td>17.5</td>
<td>9.4</td>
<td>8.8</td>
<td>9.1</td>
</tr>
<tr>
<td>TMCPL (3.15)</td>
<td>195.5</td>
<td>191.7</td>
<td>177.7</td>
<td>73.9</td>
<td>70.8</td>
<td>70.1</td>
<td>27.0</td>
<td>26.3</td>
<td>24.9</td>
<td>11.8</td>
<td>11.1</td>
<td>11.5</td>
</tr>
<tr>
<td>NCT (3.17)</td>
<td>260.5</td>
<td>253.3</td>
<td>238.9</td>
<td>103.3</td>
<td>89.6</td>
<td>89.6</td>
<td>22.8</td>
<td>21.3</td>
<td>19.7</td>
<td>6.9</td>
<td>6.1</td>
<td>6.0</td>
</tr>
<tr>
<td>NCT (3.17) using (3.19)</td>
<td>366.0</td>
<td>374.6</td>
<td>379.8</td>
<td>185.4</td>
<td>176.9</td>
<td>185.3</td>
<td>49.7</td>
<td>49.0</td>
<td>47.4</td>
<td>18.7</td>
<td>18.3</td>
<td>19.1</td>
</tr>
</tbody>
</table>

Table 3.20: Mean Square Error of label shelf-lives for parameter values $\beta_1 = -0.3$, $d_2 = 0.00$, $\sigma = 1.0$, $\epsilon = 0.01$ and $\alpha = 0.05$. 
### Table 3.21: Mean Square Error of label shelf-lives for parameter values $\beta_1 = -0.3$, $d_2 = 0.00$, $\sigma = 2.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>k</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>$(3.7)$ using $(3.5)$</td>
<td>220.5</td>
<td>221.2</td>
<td>221.2</td>
<td>221.0</td>
<td>221.3</td>
<td>221.3</td>
<td>221.1</td>
<td>221.4</td>
<td>221.4</td>
<td>221.4</td>
<td>221.4</td>
<td>221.4</td>
</tr>
<tr>
<td>$(3.7)$ using $(3.6)$</td>
<td>220.6</td>
<td>221.2</td>
<td>221.2</td>
<td>221.0</td>
<td>221.3</td>
<td>221.3</td>
<td>221.2</td>
<td>221.4</td>
<td>221.4</td>
<td>221.4</td>
<td>221.4</td>
<td>221.4</td>
</tr>
<tr>
<td>TMPL $(3.11)$</td>
<td>210.2</td>
<td>217.1</td>
<td>217.6</td>
<td>216.4</td>
<td>211.3</td>
<td>204.2</td>
<td>85.6</td>
<td>74.5</td>
<td>50.9</td>
<td>36.2</td>
<td>39.8</td>
<td>15.7</td>
</tr>
<tr>
<td>CPL $(3.14)$</td>
<td>97.9</td>
<td>81.9</td>
<td>59.2</td>
<td>53.6</td>
<td>42.4</td>
<td>28.8</td>
<td>22.7</td>
<td>15.1</td>
<td>13.2</td>
<td>9.3</td>
<td>7.3</td>
<td>6.3</td>
</tr>
<tr>
<td>TMCPL $(3.15)$</td>
<td>122.4</td>
<td>126.5</td>
<td>98.2</td>
<td>57.1</td>
<td>50.5</td>
<td>42.9</td>
<td>23.7</td>
<td>18.8</td>
<td>17.1</td>
<td>10.6</td>
<td>8.6</td>
<td>7.6</td>
</tr>
<tr>
<td>NCT $(3.17)$</td>
<td>173.3</td>
<td>158.6</td>
<td>140.9</td>
<td>83.0</td>
<td>61.9</td>
<td>42.7</td>
<td>13.8</td>
<td>11.0</td>
<td>10.4</td>
<td>7.8</td>
<td>11.1</td>
<td>12.1</td>
</tr>
<tr>
<td>NCT $(3.17)$ using $(3.19)$</td>
<td>221.3</td>
<td>221.4</td>
<td>221.4</td>
<td>213.5</td>
<td>219.0</td>
<td>220.9</td>
<td>121.1</td>
<td>113.6</td>
<td>113.2</td>
<td>40.1</td>
<td>38.3</td>
<td>38.4</td>
</tr>
</tbody>
</table>

### Table 3.22: Mean Square Error of label shelf-lives for parameter values $\beta_1 = -0.3$, $d_2 = 0.03$, $\sigma = 0.5$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>k</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>$(3.7)$ using $(3.5)$</td>
<td>291.3</td>
<td>271.1</td>
<td>244.8</td>
<td>228.2</td>
<td>206.3</td>
<td>192.0</td>
<td>176.0</td>
<td>173.4</td>
<td>164.7</td>
<td>155.2</td>
<td>151.0</td>
<td>153.0</td>
</tr>
<tr>
<td>$(3.7)$ using $(3.6)$</td>
<td>299.1</td>
<td>280.8</td>
<td>255.6</td>
<td>239.2</td>
<td>218.4</td>
<td>204.2</td>
<td>183.6</td>
<td>181.0</td>
<td>171.9</td>
<td>159.0</td>
<td>154.6</td>
<td>156.6</td>
</tr>
<tr>
<td>TMPL $(3.11)$</td>
<td>352.7</td>
<td>356.0</td>
<td>357.4</td>
<td>356.7</td>
<td>357.7</td>
<td>356.7</td>
<td>347.8</td>
<td>347.1</td>
<td>260.4</td>
<td>41.6</td>
<td>31.8</td>
<td>21.4</td>
</tr>
<tr>
<td>CPL $(3.14)$</td>
<td>88.4</td>
<td>91.9</td>
<td>87.2</td>
<td>53.9</td>
<td>51.8</td>
<td>50.4</td>
<td>25.5</td>
<td>25.3</td>
<td>23.7</td>
<td>12.6</td>
<td>11.9</td>
<td>12.3</td>
</tr>
<tr>
<td>TMCPL $(3.15)$</td>
<td>170.6</td>
<td>158.9</td>
<td>155.3</td>
<td>82.0</td>
<td>72.3</td>
<td>70.2</td>
<td>33.0</td>
<td>32.5</td>
<td>29.9</td>
<td>15.2</td>
<td>14.2</td>
<td>14.2</td>
</tr>
<tr>
<td>NCT $(3.17)$</td>
<td>232.0</td>
<td>238.2</td>
<td>233.5</td>
<td>116.8</td>
<td>114.1</td>
<td>109.5</td>
<td>37.8</td>
<td>37.6</td>
<td>35.1</td>
<td>15.0</td>
<td>14.0</td>
<td>14.4</td>
</tr>
<tr>
<td>NCT $(3.17)$ using $(3.19)$</td>
<td>266.8</td>
<td>271.3</td>
<td>269.6</td>
<td>131.0</td>
<td>130.1</td>
<td>125.4</td>
<td>43.5</td>
<td>43.9</td>
<td>41.3</td>
<td>18.0</td>
<td>17.1</td>
<td>17.7</td>
</tr>
<tr>
<td>k</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>----</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>r</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>(3.7) using (3.5)</td>
<td>292.5</td>
<td>282.4</td>
<td>263.8</td>
<td>269.2</td>
<td>252.6</td>
<td>227.5</td>
<td>239.8</td>
<td>211.6</td>
<td>208.2</td>
<td>208.6</td>
<td>193.7</td>
<td>192.8</td>
</tr>
<tr>
<td>(3.7) using (3.6)</td>
<td>294.9</td>
<td>285.4</td>
<td>267.1</td>
<td>273.2</td>
<td>256.5</td>
<td>231.9</td>
<td>243.4</td>
<td>215.5</td>
<td>211.9</td>
<td>211.3</td>
<td>196.1</td>
<td>194.9</td>
</tr>
<tr>
<td>TMPL (3.11)</td>
<td>285.4</td>
<td>308.9</td>
<td>313.3</td>
<td>305.7</td>
<td>312.0</td>
<td>314.4</td>
<td>259.0</td>
<td>250.7</td>
<td>258.9</td>
<td>133.9</td>
<td>65.1</td>
<td>35.8</td>
</tr>
<tr>
<td>CPL (3.14)</td>
<td>86.1</td>
<td>82.3</td>
<td>79.0</td>
<td>52.4</td>
<td>47.9</td>
<td>47.0</td>
<td>24.1</td>
<td>20.6</td>
<td>23.3</td>
<td>11.8</td>
<td>10.8</td>
<td>10.9</td>
</tr>
<tr>
<td>TMCPL (3.15)</td>
<td>165.3</td>
<td>166.1</td>
<td>166.5</td>
<td>70.1</td>
<td>70.8</td>
<td>71.4</td>
<td>30.6</td>
<td>27.0</td>
<td>30.0</td>
<td>14.2</td>
<td>13.0</td>
<td>13.2</td>
</tr>
<tr>
<td>NCT (3.17)</td>
<td>215.5</td>
<td>218.0</td>
<td>212.0</td>
<td>106.5</td>
<td>96.5</td>
<td>95.6</td>
<td>30.3</td>
<td>25.5</td>
<td>28.1</td>
<td>10.1</td>
<td>8.8</td>
<td>8.6</td>
</tr>
<tr>
<td>NCT (3.17) using (3.19)</td>
<td>298.3</td>
<td>308.4</td>
<td>313.6</td>
<td>165.3</td>
<td>162.9</td>
<td>164.1</td>
<td>53.2</td>
<td>48.5</td>
<td>53.4</td>
<td>21.1</td>
<td>20.0</td>
<td>20.5</td>
</tr>
</tbody>
</table>

Table 3.23: Mean Square Error of label shelf-lives for parameter values $\beta_1 = -0.3$, $d_2 = 0.03$, $\sigma = 1.0$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>k</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>(3.7) using (3.5)</td>
<td>187.7</td>
<td>187.9</td>
<td>188.4</td>
<td>188.5</td>
<td>188.6</td>
<td>188.5</td>
<td>188.5</td>
<td>188.6</td>
<td>188.6</td>
<td>188.6</td>
<td>188.6</td>
<td>188.6</td>
</tr>
<tr>
<td>(3.7) using (3.6)</td>
<td>187.8</td>
<td>187.9</td>
<td>188.4</td>
<td>188.5</td>
<td>188.6</td>
<td>188.5</td>
<td>188.5</td>
<td>188.6</td>
<td>188.6</td>
<td>188.6</td>
<td>188.6</td>
<td>188.6</td>
</tr>
<tr>
<td>TMPL (3.11)</td>
<td>177.5</td>
<td>183.5</td>
<td>185.8</td>
<td>184.7</td>
<td>181.0</td>
<td>177.0</td>
<td>84.4</td>
<td>66.5</td>
<td>53.8</td>
<td>31.6</td>
<td>30.3</td>
<td>17.3</td>
</tr>
<tr>
<td>CPL (3.14)</td>
<td>80.3</td>
<td>68.1</td>
<td>53.2</td>
<td>50.8</td>
<td>40.5</td>
<td>34.7</td>
<td>24.3</td>
<td>18.6</td>
<td>15.7</td>
<td>10.9</td>
<td>8.9</td>
<td>7.7</td>
</tr>
<tr>
<td>TMCPL (3.15)</td>
<td>105.9</td>
<td>105.0</td>
<td>82.3</td>
<td>50.4</td>
<td>47.4</td>
<td>43.5</td>
<td>24.6</td>
<td>21.5</td>
<td>19.0</td>
<td>11.8</td>
<td>10.0</td>
<td>9.0</td>
</tr>
<tr>
<td>NCT (3.17)</td>
<td>146.8</td>
<td>134.2</td>
<td>121.9</td>
<td>79.2</td>
<td>59.5</td>
<td>50.1</td>
<td>16.9</td>
<td>12.4</td>
<td>11.0</td>
<td>5.5</td>
<td>6.9</td>
<td>8.0</td>
</tr>
<tr>
<td>NCT (3.17) using (3.19)</td>
<td>188.6</td>
<td>188.6</td>
<td>188.6</td>
<td>182.8</td>
<td>186.6</td>
<td>188.2</td>
<td>104.5</td>
<td>100.4</td>
<td>99.8</td>
<td>38.8</td>
<td>36.8</td>
<td>35.4</td>
</tr>
</tbody>
</table>

Table 3.24: Mean Square Error of label shelf-lives for parameter values $\beta_1 = -0.3$, $d_2 = 0.03$, $\sigma = 2.0$, $\epsilon = 0.01$ and $\alpha = 0.05$. 
## 3. SIMULATION RESULTS

<table>
<thead>
<tr>
<th>k</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>(3.7) using (3.5)</td>
<td>400.2</td>
<td>361.0</td>
<td>332.2</td>
<td>240.4</td>
<td>212.4</td>
<td>202.2</td>
<td>150.4</td>
<td>141.7</td>
<td>142.9</td>
<td>114.6</td>
<td>110.5</td>
<td>108.9</td>
</tr>
<tr>
<td>(3.7) using (3.6)</td>
<td>422.6</td>
<td>387.9</td>
<td>356.7</td>
<td>265.1</td>
<td>233.0</td>
<td>226.3</td>
<td>162.1</td>
<td>152.6</td>
<td>154.0</td>
<td>117.8</td>
<td>113.5</td>
<td>111.7</td>
</tr>
<tr>
<td>TMPL (3.11)</td>
<td>548.6</td>
<td>563.1</td>
<td>564.7</td>
<td>563.1</td>
<td>563.7</td>
<td>561.1</td>
<td>359.5</td>
<td>345.7</td>
<td>336.2</td>
<td>338.1</td>
<td>318.5</td>
<td>322.8</td>
</tr>
<tr>
<td>CPL (3.14)</td>
<td>143.8</td>
<td>131.8</td>
<td>142.3</td>
<td>47.8</td>
<td>44.0</td>
<td>46.7</td>
<td>13.6</td>
<td>13.8</td>
<td>14.9</td>
<td>6.6</td>
<td>5.7</td>
<td>5.8</td>
</tr>
<tr>
<td>TMCPL (3.15)</td>
<td>273.1</td>
<td>258.3</td>
<td>255.3</td>
<td>95.5</td>
<td>89.0</td>
<td>92.9</td>
<td>21.8</td>
<td>22.2</td>
<td>23.4</td>
<td>9.0</td>
<td>7.8</td>
<td>7.5</td>
</tr>
<tr>
<td>NCT (3.17)</td>
<td>337.4</td>
<td>323.3</td>
<td>328.8</td>
<td>84.0</td>
<td>78.5</td>
<td>74.8</td>
<td>18.7</td>
<td>17.8</td>
<td>18.5</td>
<td>6.9</td>
<td>5.8</td>
<td>5.8</td>
</tr>
<tr>
<td>NCT (3.17) using (3.19)</td>
<td>411.5</td>
<td>400.1</td>
<td>410.7</td>
<td>111.6</td>
<td>108.1</td>
<td>105.1</td>
<td>25.9</td>
<td>25.5</td>
<td>26.9</td>
<td>10.4</td>
<td>9.4</td>
<td>9.5</td>
</tr>
</tbody>
</table>

Table 3.25: Mean Square Error of label shelf-lives for parameter values $\beta_1 = -0.3$, $d_2 = -0.03$, $\sigma = 0.5$, $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>k</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>(3.7) using (3.5)</td>
<td>436.9</td>
<td>409.7</td>
<td>374.2</td>
<td>357.3</td>
<td>320.9</td>
<td>276.9</td>
<td>270.1</td>
<td>243.2</td>
<td>238.4</td>
<td>223.4</td>
<td>206.0</td>
<td>208.4</td>
</tr>
<tr>
<td>(3.7) using (3.6)</td>
<td>443.6</td>
<td>416.4</td>
<td>380.7</td>
<td>367.1</td>
<td>330.9</td>
<td>284.8</td>
<td>277.3</td>
<td>249.8</td>
<td>244.1</td>
<td>227.6</td>
<td>209.2</td>
<td>211.1</td>
</tr>
<tr>
<td>TMPL (3.11)</td>
<td>450.2</td>
<td>463.9</td>
<td>472.5</td>
<td>444.9</td>
<td>460.0</td>
<td>477.3</td>
<td>279.1</td>
<td>259.5</td>
<td>198.1</td>
<td>261.0</td>
<td>242.4</td>
<td>105.9</td>
</tr>
<tr>
<td>CPL (3.14)</td>
<td>136.2</td>
<td>135.1</td>
<td>108.6</td>
<td>55.9</td>
<td>43.4</td>
<td>38.3</td>
<td>13.2</td>
<td>12.4</td>
<td>13.7</td>
<td>5.9</td>
<td>5.5</td>
<td>4.8</td>
</tr>
<tr>
<td>TMCPL (3.15)</td>
<td>269.7</td>
<td>273.1</td>
<td>244.9</td>
<td>89.8</td>
<td>82.5</td>
<td>83.1</td>
<td>20.6</td>
<td>19.5</td>
<td>21.8</td>
<td>8.1</td>
<td>7.4</td>
<td>7.0</td>
</tr>
<tr>
<td>NCT (3.17)</td>
<td>321.6</td>
<td>296.7</td>
<td>278.7</td>
<td>82.8</td>
<td>67.0</td>
<td>60.5</td>
<td>11.3</td>
<td>9.8</td>
<td>8.9</td>
<td>3.1</td>
<td>2.9</td>
<td>2.4</td>
</tr>
<tr>
<td>NCT (3.17) using (3.19)</td>
<td>462.2</td>
<td>475.6</td>
<td>479.5</td>
<td>202.7</td>
<td>195.6</td>
<td>190.5</td>
<td>39.9</td>
<td>39.4</td>
<td>39.4</td>
<td>14.6</td>
<td>14.4</td>
<td>13.9</td>
</tr>
</tbody>
</table>

Table 3.26: Mean Square Error of label shelf-lives for parameter values $\beta_1 = -0.3$, $d_2 = -0.03$, $\sigma = 1.0$, $\epsilon = 0.01$ and $\alpha = 0.05$. 
3.3. Simulation Results

<table>
<thead>
<tr>
<th></th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>10</th>
<th>20</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>k</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>(3.7) using (3.5)</td>
<td>263.5</td>
<td>265.5</td>
<td>265.9</td>
<td>265.4</td>
<td>265.8</td>
<td>266.0</td>
<td>266.0</td>
<td>266.0</td>
<td>266.0</td>
<td>266.0</td>
<td>266.0</td>
</tr>
<tr>
<td>(3.7) using (3.6)</td>
<td>263.7</td>
<td>265.6</td>
<td>265.9</td>
<td>265.4</td>
<td>265.9</td>
<td>266.0</td>
<td>266.0</td>
<td>266.0</td>
<td>266.0</td>
<td>266.0</td>
<td>266.0</td>
</tr>
<tr>
<td>TMPL (3.11)</td>
<td>251.0</td>
<td>257.2</td>
<td>262.2</td>
<td>258.3</td>
<td>255.1</td>
<td>239.8</td>
<td>77.0</td>
<td>55.5</td>
<td>36.2</td>
<td>38.2</td>
<td>34.4</td>
</tr>
<tr>
<td>CPL (3.14)</td>
<td>104.1</td>
<td>93.1</td>
<td>61.7</td>
<td>65.5</td>
<td>42.8</td>
<td>28.7</td>
<td>20.4</td>
<td>12.4</td>
<td>8.3</td>
<td>7.4</td>
<td>5.0</td>
</tr>
<tr>
<td>TMCPL (3.15)</td>
<td>154.4</td>
<td>158.9</td>
<td>119.6</td>
<td>75.3</td>
<td>61.1</td>
<td>47.2</td>
<td>23.5</td>
<td>16.6</td>
<td>12.3</td>
<td>8.7</td>
<td>6.1</td>
</tr>
<tr>
<td>NCT (3.17)</td>
<td>195.8</td>
<td>174.9</td>
<td>156.4</td>
<td>90.7</td>
<td>56.3</td>
<td>35.7</td>
<td>12.3</td>
<td>14.3</td>
<td>17.8</td>
<td>13.4</td>
<td>21.2</td>
</tr>
<tr>
<td>NCT (3.17) using (3.19)</td>
<td>266.0</td>
<td>266.0</td>
<td>266.0</td>
<td>257.6</td>
<td>263.7</td>
<td>265.7</td>
<td>141.3</td>
<td>134.4</td>
<td>130.6</td>
<td>42.6</td>
<td>40.0</td>
</tr>
</tbody>
</table>

Table 3.27: Mean Square Error of label shelf-lives for parameter values \( \beta_1 = -0.3, \ d_2 = -0.03, \ \sigma = 2.0, \ \epsilon = 0.01 \) and \( \alpha = 0.05 \).
Figure 3.4: Mean Square Error plotted against $1/N$ for each method and combination of parameters.
Table 3.28: For each method the parameter estimates for the model of coverage probability on $1/k$, $1/\sqrt{N}$, $\sigma^*$ and $d_2$ are given. Blank values indicate that the parameter was not significant at the 1% level and the model was re-fitted without the corresponding term(s).

<table>
<thead>
<tr>
<th>Method</th>
<th>Int</th>
<th>$\sigma$</th>
<th>$d_2$</th>
<th>$\sigma^* d_2$</th>
<th>$1/k$</th>
<th>$\sigma^*/k$</th>
<th>$d_2/k$</th>
<th>$\sigma^*/\sqrt{N}$</th>
<th>$\sigma^*/k\sqrt{N}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.t</td>
<td>1.002</td>
<td>-0.003</td>
<td>0.037</td>
<td>-0.044</td>
<td>-0.031</td>
<td>0.040</td>
<td>-0.550</td>
<td>0.672</td>
<td></td>
</tr>
<tr>
<td>U.t</td>
<td>1.002</td>
<td>-0.002</td>
<td>0.033</td>
<td>-0.040</td>
<td>-0.026</td>
<td>0.034</td>
<td>-0.483</td>
<td>0.590</td>
<td></td>
</tr>
<tr>
<td>TMPL</td>
<td>0.973</td>
<td>-0.006</td>
<td></td>
<td></td>
<td>0.097</td>
<td>0.024</td>
<td></td>
<td>0.054</td>
<td>-0.216</td>
</tr>
<tr>
<td>CPL</td>
<td>0.928</td>
<td>0.024</td>
<td>0.345</td>
<td>-0.217</td>
<td>-0.129</td>
<td>-5.432</td>
<td>-0.018</td>
<td>-0.066</td>
<td>0.194</td>
</tr>
<tr>
<td>TMCPPL</td>
<td>0.945</td>
<td>0.006</td>
<td>-0.221</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCT1</td>
<td>0.684</td>
<td>-0.574</td>
<td>3.281</td>
<td>2.415</td>
<td>0.842</td>
<td>1.747</td>
<td>-11.078</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCT2</td>
<td>0.979</td>
<td>0.024</td>
<td>-0.122</td>
<td>0.124</td>
<td>0.058</td>
<td>-0.066</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.29: For each method the parameter estimates for the model of bias on $1/k$, $1/\sqrt{N}$, $\sigma^*$ and $d_2$ are given. Blank values indicate that the parameter was not significant at the 1% level and the model was re-fitted without the corresponding term(s).

<table>
<thead>
<tr>
<th>Method</th>
<th>Int</th>
<th>$\sigma$</th>
<th>$d_2$</th>
<th>$\sigma^* d_2$</th>
<th>$1/k$</th>
<th>$\sigma^*/k$</th>
<th>$d_2/k$</th>
<th>$\sigma^*/\sqrt{N}$</th>
<th>$\sigma^*/k\sqrt{N}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>z.t</td>
<td>0.20</td>
<td>0.03</td>
<td>0.58</td>
<td>-0.61</td>
<td>0.19</td>
<td>-0.19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U.t</td>
<td>0.19</td>
<td>0.03</td>
<td>0.86</td>
<td>-0.59</td>
<td>0.20</td>
<td>-0.20</td>
<td>-1.75</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>TMPL</td>
<td>0.15</td>
<td>-0.09</td>
<td></td>
<td></td>
<td>0.55</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPL</td>
<td>0.03</td>
<td>-0.01</td>
<td>0.55</td>
<td>-0.09</td>
<td>0.23</td>
<td>-0.11</td>
<td>-1.68</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>TMCPPL</td>
<td>0.01</td>
<td>0.00</td>
<td>0.52</td>
<td></td>
<td>0.56</td>
<td>-0.13</td>
<td>-1.68</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>NCT1</td>
<td>-0.04</td>
<td>-0.08</td>
<td>0.83</td>
<td>0.72</td>
<td></td>
<td></td>
<td>-1.22</td>
<td>0.06</td>
<td>0.10</td>
</tr>
<tr>
<td>NCT2</td>
<td>0.03</td>
<td>0.06</td>
<td>0.31</td>
<td>0.78</td>
<td></td>
<td></td>
<td>-0.29</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.4 Discussion

As in Chapter 2, the performance of a method is primarily indicated by how close its coverage probability is to the desired 0.95. The bias and MSE are secondary measures of goodness that can be used to choose between methods with a similar coverage probability.

Clearly, due to the number of effects and interactions involved in the models for coverage probability and bias, it is difficult to give a simple interpretation of them. However, the following general observations can be made from the tables and graphs in Section 3.3:

- Most methods resulted in conservative coverage probabilities;
- The coverage probability approaches the desired level of 0.95 as more batches are used. Smaller coefficients of $1/k$ indicate that the method works better for a small number of batches;
- For most methods, the coverage probability does not appear to depend on the number of assays per batch;
- A positive covariance $d_2$ tends to result in lower coverage probabilities than a negative covariance.

However, for large numbers of batches, i.e. $k \geq 10$, positive values of $d_2$ tend to result in larger values of bias. A possible reason for this has been suggested in Chapter 1. That is, the variability in the response at time $t$ due to batches is $d_1 + 2d_2t + d_3t^2$, which is larger when $d_2$ is positive. Consequently, the label shelf-life may be smaller, and hence the bias larger, due to the increased uncertainty.

The two normal approximations perform about equally for all practical purposes, as expected. Both give conservative coverage probabilities for all parameter combinations. In terms of bias and MSE both perform better for large numbers of batches and small within-batch variability. However, it should be noted that both methods often resulted
in a zero shelf-life for small numbers of batches, and hence neither is of much practical value.

The TMPL approach was also conservative for all parameter combinations. However, its bias and MSE tended to be worse than the normal approaches when the within batch variability was small.

The CPL approach was, except for the NCT1 approach, the only approach that fell short of the desired coverage probability of 0.95. This approach also appears to be the only one largely affected by the covariance $d_2$ and the within batch variability $\sigma$.

The TMCPL performed the best of all these methods. It was conservative for small numbers of batches, which is a good property, but approached the desired coverage of 0.95 as the number of batches increased. It also appears to be only slightly affected by the covariance $d_2$ and within batch variability $\sigma$ — at least for the range of values used for these parameters. Furthermore, the bias and MSE for this method were smaller than most other methods, and both measures tended toward zero as the number of batches increased.

The NCT1 approach had reasonable coverage probabilities for small values of $\sigma$, but worked very badly for large values of $\sigma$. This is clearly because the approximation in Section 3.2.4 does not work well for these cases. This tends to get worse as $k$ and $N$ increase and for negative $d_2$. This is as expected.

These observations are different from the findings by Shao and Chow (1994), who suggest that this approach is conservative for large numbers of batches and large $\sigma$. Clearly, this is because Shao and Chow use (3.18) as the definition for the true shelf-life in the batch case, which is different to Definition 3.

The NCT2 approach was also conservative — much more so than TMCPL — across the range of parameter values. Consequently, the bias and MSE were also larger.

In conclusion, in the multi-batch case, the recommended method for obtaining a label shelf-life is the TMCPL method, followed by the NCT2. The normal approximations and
TMPL cannot be recommended due to their large bias and MSE, which is mainly due to many of the shelf-lives being estimated as zero, while the CPL approach cannot be recommended due to their low coverage probabilities. The NCT1 approach clearly offers a reasonable solution for small values of $\sigma$, but should not be used for large values.
CHAPTER 4

Estimation of Shelf-Life:
The GLM Single Batch Case

4.1 Introduction

The ICH guidelines state that

“regression analysis is considered an appropriate approach to evaluating the stability data for a quantitative attribute”

and that

“the nature of the degradation relationship will determine whether the data should be transformed for linear regression analysis. Usually the relationship can be represented by a linear, quadratic or cubic function on an arithmetic or logarithmic scale”.

These statements implicitly assume that the residuals from fitting an appropriate model are normally distributed. This assumption has been used extensively in the literature, as well as in the development of the new definitions of shelf-life in Chapters 2 and 3.
4.1 INTRODUCTION

The relationship between potency and time depends on the order of the kinetics involved in the degradation process. A linear relationship between potency and time is referred to as a zero order kinetic relationship, while a linear relationship between the logarithm of potency and time is referred to as a first order kinetic relationship (Jerne and Perry, 1956; Kirkwood, 1977; Carstensen and Rhodes, 2000).

The possible requirement for a transformation of the response raises several questions: Is a linear relationship between the response and time the most appropriate; does the assumption of normality hold; and, in particular, is the assumption of constancy of variance realistic? Rather than transforming the data, generalised linear models (GLMs) may offer a more appropriate alternative (McCullagh and Nelder, 1989).

The link function used in GLMs allows for different relationships between the mean response and the explanatory variables. Distributions other than the Normal can be fitted and heteroscedasticity can be accommodated using mean-variance relationships. For example, using the Gamma distribution, the variance of the response $Y$ can depend on the mean according to the relationship \( \text{Var}[Y] \propto E[Y]^2 \), and the mean can be allowed to decay or grow exponentially. It is not difficult to imagine a process in which the response increases or decreases over time in such a fashion. Such processes are not covered in the existing literature on stability analysis, but they will be the focus of this chapter.

Assume that the assay result $y$ is a realisation of the random variable $Y$, which has a distribution that belongs to the exponential family of distributions. The probability density function (pdf) of $Y$ is

$$f(y; \mu, \phi) = \exp \left\{ \frac{A}{\phi} [ya(\mu) - b(a(\mu))] + c(y, \phi) \right\},$$

where

- $A$ is a prior known weight,
- $\phi > 0$ is the dispersion parameter,
- the mean $\mu$ is generally a function of the covariates,
• \( a = a(\mu) \) is an increasing and invertible function of \( \mu \) that does vary with the observations and is independent of \( \phi \),

• \( b(a) = b(a(\mu)) \) does not vary with the observations and is independent of \( \phi \), and

• \( c(y, \phi) \) is a function of the response and \( \phi \) that does not involve \( \mu \).

Further assume that the linear predictor is \( \eta = x^T \beta = l(\mu) \) where \( l(\cdot) \) is the link function. Consequently, \( \mu = l^{-1}(\eta) = m(\eta) \), say.

It can be shown (McCullagh and Nelder, 1989) that

\[
E[Y] = \mu = \frac{\partial b}{\partial a}, \\
\text{Var}[Y] = \phi V(\mu) = \frac{\phi \partial \mu}{A} 
\]

where \( V(\mu) \) denotes the variance function which does not depend on the scale parameter \( \phi \).

Assuming that the scale parameter does not depend on \( x \), the probability density function at time \( t \) is denoted by \( f(y; \mu_t, \phi) \). Similarly, the distribution function, which gives the probability \( P(Y \leq y) \) at time \( t \), shall be denoted by \( F(y; \mu_t, \phi) \).

Similar to Chapter 2, interest lies in the time at which 100\(\epsilon\)% of the drug product falls below the specification limit. Hence, the following definition extends Definition 1 to the exponential family of distributions.

**Definition 5** Assume that the amount of active ingredient \( y \) at time \( t \) comes from a distribution that belongs to the exponential family of distributions with pdf given in (4.1) and mean \( \mu_t \). Then, the true shelf-life, denoted by \( \tau_\epsilon \), is the minimum time at which 100\(\epsilon\)% of the drug product from the current batch has an activity level that is less than or equal to the pre-determined specification limit \( \kappa \), that is,

\[
\tau_\epsilon = \inf \{ t : F(\kappa; \mu_t, \phi) \leq \epsilon \},
\]

where \( F(\cdot) \) denotes the distribution function of \( y \).
In Chapters 2 and 3 it was shown that the normal approximation to the lower \( \epsilon \) quantile was not very satisfactory. Since the parameters are not guaranteed to be normally distributed, the normal approximation is even less likely to hold here. Furthermore, the exact approach is no longer applicable for the same reason. However, a natural way to estimate parameters in the GLM setting is via likelihood methods. In general, \( \beta_0 \) cannot be eliminated via substitution, and consequently the basic profile likelihood will not provide a workable approach. However the constrained profile likelihood gives a suitable approach due to the simple penalizing of the likelihood function. This approach is investigated in the next Section.

### 4.2 Likelihood based approaches

Let \( y = (y_1, \ldots, y_N)^T \) be a sample of the response with covariates \( x_{i1}, \ldots, x_{ip} \) corresponding to the \( i \)-th response. Then, the joint log likelihood of the parameters \( \beta \) and \( \phi \), based on (4.1), is given by

\[
L(\beta, \phi; y) = \sum_{i=1}^{N} \frac{A_i}{\phi} \left\{ y_i a_i - b_i(a_i) \right\} + c(y_i, \phi).
\]

The elements of the score are given by

\[
\frac{\partial L}{\partial \beta_i} = \sum_{i=1}^{N} \frac{A_i}{\phi} \left\{ (y_i - \frac{\partial b_i}{\partial a_i}) \frac{\partial a_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_i} \right\},
\]

\[
\frac{\partial L}{\partial \phi} = \sum_{i=1}^{N} \frac{A_i}{\phi^2} \left\{ y_i a_i - b_i(a_i) \right\} + \frac{\partial c_i}{\partial \phi},
\]

where \( c_i = c(y_i, \phi) \). The score of \( \beta_i \) can be written as

\[
\frac{\partial L}{\partial \beta_i} = \left( y_i - \mu_i \right) \left( \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_i} \right) x_{il} = e_i w_i x_{il}.
\]

Letting \( e = (e_1, \ldots, e_N)^T \) and \( W = \text{diag}(w_1, \ldots, w_N) \), the score of \( \beta \) can be written as

\[
\frac{\partial L}{\partial \beta} = X^T W e.
\]
The elements of the observed information matrix are given by

\[
I_o(\beta_l, \beta_m) = \sum_{i=1}^{N} \frac{A_i}{\phi} \left\{ \frac{\partial a_i}{\partial \mu_i} \left( \frac{\partial \mu_i}{\partial \eta_i} \right)^2 - (y_i - \mu_i) \right\} x_{il} x_{im}
\]

\[
I_o(\beta_l, \phi) = \frac{1}{\phi} \frac{\partial L}{\partial \beta_l}
\]

\[
I_o(\phi, \phi) = \sum_{i=1}^{N} -2 \frac{A_i}{\phi^3} \left\{ y_i a_i - b_i(a_i) \right\} - \frac{\partial^2 c_i}{\partial \phi^2},
\]

and the elements of the expected information matrix are

\[
I_e(\beta_l, \beta_m) = \sum_{i=1}^{N} \frac{A_i}{\phi} \frac{\partial a_i}{\partial \mu_i} \left( \frac{\partial \mu_i}{\partial \eta_i} \right)^2 x_{il} x_{im}
\]

\[
I_e(\beta, \phi) = 0
\]

\[
I_e(\phi, \phi) = \sum_{i=1}^{N} -2 \frac{A_i}{\phi^3} \left\{ \mu_i a_i - b_i(a_i) \right\} - E \left[ \frac{\partial^2 c_i}{\partial \phi^2} \right].
\]

Since \(I_e(\beta, \beta) = \text{Var} \left[ \frac{\partial L}{\partial \beta} \right]\) and \(\text{Var}[\mathbf{e}] = W^{-1}\) it follows that \(I_e(\beta, \beta) = X^T W X\).

Fisher scoring can then be used, with appropriate starting estimate, to find the maximum likelihood estimates for \(\beta\) and \(\phi\).

### 4.2.1 Constrained Profile Likelihood

As in Chapter 2 the constrained estimates are found by minimizing the objective function

\[
H(\theta_{\tau}; \tau, \mathbf{y}) = \frac{1}{N} \left[ -L(\theta_{\tau}) \right] + \omega \left[ g(\theta_{\tau}; \tau) + \psi \right]^2,
\]

where \(\theta_{\tau} = (\beta_{\tau}^T, \phi_{\tau})^T\), and \(g(\theta_{\tau}; \tau)\) can take the form

\[
g(\theta_{\tau}; \tau) = F(\kappa; \beta_{\tau}, \phi_{\tau}) - \epsilon
\]

\[
g(\theta_{\tau}; \tau) = F^{-1}(\epsilon; \beta_{\tau}, \phi_{\tau}) - \kappa
\]

where \(F^{-1}(\cdot)\) denotes the inverse of \(F(\cdot)\).
The score vector now has components

\[
\frac{\partial H}{\partial \beta_{\tau_e}} = \frac{1}{N} \left[ -X^T W e \right] + 2\omega \left[ g(\theta_{\tau_e}; \tau_e) + \psi \right] \frac{\partial g}{\partial \beta_{\tau_e}}
\]

\[
\frac{\partial H}{\partial \phi_{\tau_e}} = \frac{1}{N} \left[ -\sum_{i=1}^{N} \frac{A_i}{\phi_{\tau_e}^2} \left( y_i a_i + b_i(a_i) \right) + \frac{\partial c_i}{\partial \phi_{\tau_e}} \right] + 2\omega \left[ g(\theta_{\tau_e}; \tau_e) + \psi \right] \frac{\partial g}{\partial \phi_{\tau_e}}.
\]

Consequently, the components of the Powell-Hestenes equivalents of the observed information matrix are given by

\[
I_o'(\beta_{\tau_e}, \beta_{\tau_e}) = \frac{1}{N} \left[ X^T W X \right] + 2\omega \left( \frac{\partial g}{\partial \beta_{\tau_e}} \right) \left( \frac{\partial g}{\partial \beta_{\tau_e}^T} \right)
\]

\[
I_o'(\beta_{\tau_e}, \phi_{\tau_e}) = \frac{1}{N} \left[ \frac{1}{\phi_{\tau_e}} X^T W e \right] + 2\omega \left( \frac{\partial g}{\partial \beta_{\tau_e}} \right) \left( \frac{\partial g}{\partial \phi_{\tau_e}} \right)
\]

\[
I_o'(\phi_{\tau_e}, \phi_{\tau_e}) = \frac{1}{N} \left[ \sum_{i=1}^{N} -2A_i \phi_{\tau_e}^2 \left( y_i a_i - b_i(a_i) \right) - \frac{\partial^2 c_i}{\partial \phi_{\tau_e}^2} \right] + 2\omega \left( \frac{\partial^2 g}{\partial \phi_{\tau_e}^2} \right),
\]

and the elements of the equivalent of the expected information matrix are

\[
I_e'(\beta_{\tau_e}, \beta_{\tau_e}) = \frac{1}{N} \left[ X^T W X \right] + 2\omega \left( \frac{\partial g}{\partial \beta_{\tau_e}} \right) \left( \frac{\partial g}{\partial \beta_{\tau_e}^T} \right)
\]

\[
I_e'(\beta_{\tau_e}, \phi_{\tau_e}) = 0 + 2\omega \left( \frac{\partial g}{\partial \beta_{\tau_e}} \right) \left( \frac{\partial g}{\partial \phi_{\tau_e}} \right)
\]

\[
I_e'(\phi_{\tau_e}, \phi_{\tau_e}) = \frac{1}{N} \left[ \sum_{i=1}^{N} -2A_i \phi_{\tau_e}^3 \left( \mu_i a_i - b_i(a_i) \right) - E \left[ \frac{\partial^2 c_i}{\partial \phi_{\tau_e}^2} \right] \right] + 2\omega \left( \frac{\partial^2 g}{\partial \phi_{\tau_e}^2} \right).
\]

The form of the score and information matrix now depend on the derivatives of \( g(\cdot) \) which are investigated below.

First, however, it is worthwhile to note that the modification of the constrained profile likelihood developed in Chapter 2 can also be used here. The matrix \( C_0 \) is obtained by evaluating

\[
\frac{\partial g}{\partial \theta_{\tau_e}^T} = \left[ \frac{\partial g}{\partial \theta_{\tau_e}} \frac{\partial g}{\partial \beta_{\tau_e}} \frac{\partial g}{\partial \phi_{\tau_e}} \right]
\]

at \( \theta_{\tau_e} = \hat{\theta}_{\tau_e} \).
4.2. LIKELIHOOD BASED APPROACHES

Derivative of \( g(\theta_{\tau_e}; \tau_e) = F(\kappa; \beta_{\tau_e}, \phi_{\tau_e}) - \epsilon \)

In this case, \( \kappa \) and \( \epsilon \) are assumed fixed and differentiating \( g(\theta_{\tau_e}; \tau_e) \) with respect to \( \beta_{\tau_e} \) gives

\[
\frac{\partial g}{\partial \beta_{\tau_e}} = \frac{\partial F(\kappa)}{\partial \mu_{\tau_e}} \frac{\partial \mu_{\tau_e}}{\partial \eta_{\tau_e}} x_{\tau_e},
\]

where \( F(\kappa) = F(\kappa; \beta_{\tau_e}, \phi_{\tau_e}) \) for notational simplicity. The second derivative on the right hand side depends only on the link function, while the first element equals

\[
\frac{\partial F(\kappa)}{\partial \mu_{\tau_e}} = \int_{-\infty}^{\kappa} \frac{\partial f(z)}{\partial \mu_{\tau_e}} dz,
\]

where \( F(\kappa) = F(\kappa; \beta_{\tau_e}, \phi_{\tau_e}) \) and \( f(z) = f(z; \beta_{\tau_e}, \phi_{\tau_e}) \). This reduces to

\[
\frac{\partial F(\kappa)}{\partial \mu_{\tau_e}} = \int_{-\infty}^{\kappa} f(z) \frac{z - \mu_{\tau_e}}{\phi_{\tau_e} V(\mu_{\tau_e})} dz = \frac{1}{\phi_{\tau_e} V(\mu_{\tau_e})} \left[ \bar{\mu}_{\tau_e}(\kappa) - \mu_{\tau_e} \right] F(\kappa),
\]

where \( \bar{\mu}_{\tau_e}(\kappa) \) denotes the mean, given \( \tau_e \), of the distribution of \( Y \), right truncated at \( \kappa \).

Similarly, differentiating \( g(\theta_{\tau_e}; \tau_e) \) with respect to \( \phi_{\tau_e} \) gives

\[
\frac{\partial g}{\partial \phi_{\tau_e}} = \frac{\partial F(\kappa)}{\partial \phi_{\tau_e}} = M(\kappa; \mu_{\tau_e}, \phi_{\tau_e}) F(\kappa),
\]

where

\[
M(\kappa; \mu_{\tau_e}, \phi_{\tau_e}) = -\frac{A}{\phi_{\tau_e}^2} \left[ \bar{\mu}_{\tau_e}(\kappa)a_{\tau_e} - b_{\tau_e} \right] + \int_{-\infty}^{\kappa} \frac{\partial c(z, \phi_{\tau_e})}{\partial \phi_{\tau_e}} \frac{f(z)}{F(\kappa)} dz
\]

and \( a_{\tau_e} = a(\mu_{\tau_e}) \) and \( b_{\tau_e} = b(a(\mu_{\tau_e})) \). Whether the integral can be calculated analytically is unknown until the distribution, and hence the form of the pdf, is known.

Derivative of \( g(\theta_{\tau_e}; \tau_e) = F^{-1}(\epsilon; \beta_{\tau_e}, \phi_{\tau_e}) - \kappa \)

As above, \( \kappa \) and \( \epsilon \) are known and fixed in advance. However, at any given time \( \tau_e \) the value of \( F^{-1}(\epsilon) = F^{-1}(\epsilon; \beta_{\tau_e}, \phi_{\tau_e}) \) is not fixed as it depends on the value of the parameters. Hence, let \( F^{-1}(\epsilon) = y_{\tau_e}(\epsilon; \mu_{\tau_e}, \phi_{\tau_e}) = y_{\tau_e} \), which is the \( \epsilon \) quantile of the response at time \( \tau_e \). Then \( F(y_{\tau_e}) \) can be differentiated using Leibniz’ Theorem, that is,

\[
\frac{\partial F(y_{\tau_e})}{\partial \mu_{\tau_e}} = \left[ \int_{-\infty}^{y_{\tau_e}} \frac{\partial f(z)}{\partial \mu_{\tau_e}} dz \right] + f(y_{\tau_e}) \frac{\partial y_{\tau_e}}{\partial \mu_{\tau_e}} = \frac{1}{\phi_{\tau_e} V(\mu_{\tau_e})} \left[ \bar{\mu}_{\tau_e}(y_{\tau_e}) - \mu_{\tau_e} \right] F(y_{\tau_e}) + f(y_{\tau_e}) \frac{\partial y_{\tau_e}}{\partial \mu_{\tau_e}}.
\]
However, since $F(y_{\tau\epsilon}) = \epsilon$ is constant it follows that
\[
\frac{\partial F^{-1}(\epsilon)}{\partial \mu_{\tau\epsilon}} = -\frac{[\bar{\mu}_{\tau\epsilon}(F^{-1}(\epsilon)) - \mu_{\tau\epsilon}]}{\phi_{\tau\epsilon} V(\mu_{\tau\epsilon})} \frac{\epsilon}{f(F^{-1}(\epsilon))}.
\] (4.8)

The derivative of the constraint with respect to $\mu_{\tau\epsilon}$ is then given by
\[
\frac{\partial g}{\partial \mu_{\tau\epsilon}} = \frac{\partial F^{-1}(\epsilon)}{\partial \mu_{\tau\epsilon}} \frac{\epsilon}{f(F^{-1}(\epsilon))},
\] (4.9)

Using the same argument, it can be shown that
\[
\frac{\partial g}{\partial \phi_{\tau\epsilon}} = \frac{\partial F^{-1}(\epsilon)}{\partial \phi_{\tau\epsilon}} = -M(F^{-1}(\epsilon), \mu_{\tau\epsilon}, \phi_{\tau\epsilon}) \frac{\epsilon}{f(F^{-1}(\epsilon))},
\] (4.10)

where
\[
M(F^{-1}(\epsilon), \mu_{\tau\epsilon}, \phi_{\tau\epsilon}) = -\frac{A_{\phi_{\tau\epsilon}}^2}{\phi_{\tau\epsilon}^2} \left[ \bar{\mu}_{\tau\epsilon}(F^{-1}(\epsilon)) a_{\tau\epsilon} - b_{\tau\epsilon} \right] + \int_{-\infty}^{F^{-1}(\epsilon)} \frac{\partial c(z, \phi_{\tau\epsilon})}{\partial \phi_{\tau\epsilon}} f(z) \frac{\epsilon}{\epsilon} dz.
\]

### 4.2.2 Example: The Normal Distribution

In this section the constrained likelihood is applied to normally distributed data. The results can then be compared with the corresponding results in Chapter 2.

The likelihood, written in exponential family notation, is
\[
L(\beta, \phi; y) = \sum_{i=1}^{N} \frac{1}{\phi} \left[ y_i \mu_i - \frac{\mu_i^2}{2} \right] - \frac{1}{2} \left[ \frac{y_i^2}{\phi} + \log 2\pi + \log \phi \right],
\]
where $\phi = \sigma^2$ is the scale parameter, $A_i = 1$, $a(\mu_i) = \mu_i$ and $b_i(a_i) = a_i^2/2 = \mu_i^2/2$. Also, the mean and variance are given by $E[Y_i] = \mu_i$ and $\text{Var}[Y_i] = \phi$, such that the variance function is simply $V[\mu_i] = 1$. Possible choices for the link function are the identity, the canonical link, the inverse and the log link.

The objective function, for a given $\tau\epsilon$, is
\[
H(\theta_{\tau\epsilon}; y) = \frac{1}{N} \left[ -L(\theta_{\tau\epsilon}; y) \right] + \omega \left[ g(\theta_{\tau\epsilon}; \tau\epsilon) - \psi \right]^2,
\]
where $\theta_{\tau\epsilon}^T = (\beta_{\tau\epsilon}^T, \phi_{\tau\epsilon})$. The general forms of the score vector and information matrices can be obtained using the usual approach. Exact forms clearly depend on the choice of link function and the constraint $g(\cdot)$.
The elements of the score vector are now given by

\[
\frac{\partial H}{\partial \beta_{\tau\epsilon}} = -\frac{1}{N\phi_{\tau\epsilon}} \sum_{i=1}^{N} (y_i - \mu_{i,\tau\epsilon}) \frac{\partial \mu_{i,\tau\epsilon}}{\partial \eta_{i,\tau\epsilon}} x_i + 2\omega \left[ g(\theta_{\tau\epsilon}; \tau\epsilon) + \psi \right] \frac{\partial g}{\partial \beta_{\tau\epsilon}}
\]

\[
\frac{\partial H}{\partial \phi_{\tau\epsilon}} = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\phi_{i,\tau\epsilon}^2} \left( y_i \mu_{i,\tau\epsilon} - \frac{\mu_{i,\tau\epsilon}^2}{2} \right) - \frac{y_i^2}{2\phi_{\tau\epsilon}^2} + \frac{1}{2\phi_{\tau\epsilon}} \right] + 2\omega \left[ g(\theta_{\tau\epsilon}; \tau\epsilon) + \psi \right] \frac{\partial g}{\partial \phi_{\tau\epsilon}}
\]

where \(\mu_{i,\tau\epsilon}\) and \(\eta_{i,\tau\epsilon}\) denote the constrained mean and linear predictor corresponding to the \(i\)-th observation, respectively.

The elements of the Powell-Hestenes equivalent of the observed information matrix are

\[
I_o' (\beta_{\tau\epsilon}, \beta_{\tau\epsilon}) = -\frac{1}{N\phi_{\tau\epsilon}} \sum_{i=1}^{N} \left\{ (y_i - \mu_{i,\tau\epsilon}) \frac{\partial^2 \mu_{i,\tau\epsilon}}{\partial \eta_{i,\tau\epsilon}^2} - \left( \frac{\partial \mu_{i,\tau\epsilon}}{\partial \eta_{i,\tau\epsilon}} \right)^2 \right\} x_i x_i^T + 2\omega \frac{\partial g}{\partial \beta_{\tau\epsilon}} \frac{\partial g}{\partial \beta_{\tau\epsilon}^T}
\]

\[
I_o' (\beta_{\tau\epsilon}, \phi_{\tau\epsilon}) = \frac{1}{N\phi_{\tau\epsilon}^2} \sum_{i=1}^{N} (y_i - \mu_{i,\tau\epsilon}) \frac{\partial \mu_{i,\tau\epsilon}}{\partial \eta_{i,\tau\epsilon}} x_i + 2\omega \frac{\partial g}{\partial \beta_{\tau\epsilon}} \frac{\partial g}{\partial \phi_{\tau\epsilon}}
\]

\[
I_o' (\phi_{\tau\epsilon}, \phi_{\tau\epsilon}) = \frac{1}{N} \sum_{i=1}^{N} 2 \frac{\partial^2 \phi_{\tau\epsilon}}{\partial \phi_{\tau\epsilon}^2} \left( \mu_{i,\tau\epsilon}^2 - 2y_i \mu_{i,\tau\epsilon} + \frac{y_i^2}{\phi_{\tau\epsilon}^2} - \frac{1}{2\phi_{\tau\epsilon}^2} \right) + 2\omega \left( \frac{\partial g}{\partial \phi_{\tau\epsilon}} \right)^2
\]

Taking expectations gives the elements of the equivalent of the expected information matrix, that is,

\[
I_e' (\beta_{\tau\epsilon}, \beta_{\tau\epsilon}) = \frac{1}{N\phi_{\tau\epsilon}} \sum_{i=1}^{N} \left( \frac{\partial \mu_{i,\tau\epsilon}}{\partial \eta_{i,\tau\epsilon}} \right)^2 x_i x_i^T + 2\omega \frac{\partial g}{\partial \beta_{\tau\epsilon}} \frac{\partial g}{\partial \beta_{\tau\epsilon}^T}
\]

\[
I_e' (\beta_{\tau\epsilon}, \phi_{\tau\epsilon}) = 0 + 2\omega \frac{\partial g}{\partial \beta_{\tau\epsilon}} \frac{\partial g}{\partial \phi_{\tau\epsilon}}
\]

\[
I_e' (\phi_{\tau\epsilon}, \phi_{\tau\epsilon}) = \frac{1}{2\phi_{\tau\epsilon}^2} + 2\omega \left( \frac{\partial g}{\partial \phi_{\tau\epsilon}} \right)^2
\]

The score vector and the equivalent of the expected information matrix can then be used in Fisher scoring to obtain estimates for \(\beta_{\tau\epsilon}\) and \(\phi_{\tau\epsilon}\) for given values of \(\tau\epsilon\). Evaluating the likelihood at these constrained estimates gives the profile likelihood for \(\tau\epsilon\), \(\tilde{L}(\tau\epsilon)\), and the value of \(\tau\epsilon\) that maximises this profile likelihood is the maximum likelihood estimate of the shelf-life, \(\hat{\tau}\). The generalised likelihood ratio statistic is given by

\[
w(\tau\epsilon) = 2 \left( \tilde{L}(\hat{\tau}) - \tilde{L}(\tau\epsilon) \right)
\]
which is distributed approximately $\chi^2_1$. Consequently, the label shelf-life, based on the $100(1 - \alpha)$% lower confidence bound, is given by

\[
\hat{\tau}_{e,L(\alpha)} = \inf \left\{ t : (t \leq \hat{\tau}_e) \& (w(t) \leq \chi^2_{1,1-2\alpha}) \right\}.
\] (4.11)

As elucidated previously, the exact values of the score vector and information matrices rely partially on the form of the constraint. The two forms of constraint discussed in Section 4.2.1 will now be applied to the normal case. Before doing so, note that both forms involve the right truncated mean. Higher order moments of the truncated distribution may also arise depending on the form of $c(y, \phi)$.

For the normal distribution, the mean of the normal distribution of $Y$, truncated at $\kappa$ is given by (Patel and Read, 1982)

\[
\tilde{\mu}_\tau(\kappa) = \mu_\tau - \phi_\tau \frac{f(\kappa)}{F(\kappa)},
\]

and the variance is

\[
\tilde{\sigma}^2_\tau(\kappa) = \phi_\tau \left[ 1 - (\kappa - \mu_\tau) \frac{f(\kappa)}{F(\kappa)} \right] - \left[ \phi_\tau \frac{f(\kappa)}{F(\kappa)} \right]^2.
\]

**Constraint 1:** $g(\theta_\tau; \tau_e) = F(\kappa; \beta_\tau, \phi_\tau) - \epsilon$

The derivative of the constraint with respect to $\beta_\tau$ is

\[
\frac{\partial g}{\partial \beta_\tau} = \frac{1}{\phi_\tau} \left[ \tilde{\mu}_\tau(\kappa) - \mu_\tau \right] F(\kappa) x_\tau = -f(\kappa) x_\tau.
\]

Similarly, the derivative of the constraint with respect to $\phi_\tau$ is given by

\[
\frac{\partial g}{\partial \phi_\tau} = \frac{F(\kappa)}{\phi^2_\tau} \mu_\tau \left[ \frac{\mu_\tau}{2} - \tilde{\mu}_\tau \right] + \int_{-\infty}^{\kappa} \left[ \frac{t^2}{2\phi^2_\tau} - \frac{1}{2\phi_\tau} \right] f(t) . dt
\]

\[
= \frac{F(\kappa)}{\phi^2_\tau} \mu_\tau \left[ \frac{\mu_\tau}{2} - \tilde{\mu}_\tau \right] + \frac{F(\kappa)}{2\phi^2_\tau} \left[ \tilde{\sigma}^2_\tau(\kappa) + \tilde{\mu}^2_\tau \right] - \frac{F(\kappa)}{2\phi_\tau} = -\frac{\kappa - \mu_\tau}{2\phi_\tau} f(\kappa).
\]
4.2. LIKELIHOOD BASED APPROACHES

Constraint 2: \( g(\theta, \tau, \epsilon) = F^{-1}(\epsilon; \beta, \phi) - \kappa \)

The derivative of the constraint with respect to \( \beta \) is

\[
\frac{\partial g}{\partial \beta} = -\left[ \frac{\mu}{\phi} F^{-1}(\epsilon) - \mu \right] \epsilon \tau = x_\tau.
\]

Similarly, the derivative of the constraint with respect to \( \phi \) is given by

\[
\frac{\partial g}{\partial \phi} = -\left\{ \frac{-\mu_2^2}{2\phi_\tau} + \frac{\mu_2 \mu_\tau_\tau}{\epsilon \phi_\tau} - \frac{1}{2\phi_\tau} + \frac{\sigma^2_\tau}{2\phi_\tau^2} + \frac{\tilde{\mu}_2^2}{2\phi_\tau^2} \right\} \frac{\epsilon}{f(F^{-1}(\epsilon))},
\]

which reduces to

\[
\frac{\partial g}{\partial \phi} = \frac{F^{-1}(\epsilon) - \mu_\tau}{2\phi_\tau}.
\]

Notice the closeness of both derivatives to the derivatives of Constraint 1. Also note that this constraint is identical to the constraint used in Section 2.2.3.

4.2.3 Example: The Gamma Distribution

This section details the methods developed in Section 4.2.1 when applied to Gamma distributed data.

There are various parameterizations of the Gamma distribution available. For sake of consistency with (4.1), a modified form of the density used in McCullagh and Nelder (1989) is chosen here.

Let \( Y \) have a gamma distribution with mean and scale parameters \( \mu \) and \( \phi \), respectively. This distribution is denoted by \( \gamma(\mu, \phi) \). The density of \( Y \) is given by

\[
f(y) = \frac{y^{\frac{1}{\phi} - 1} e^{-\frac{y}{\phi\mu}}}{(\phi\mu)^{\frac{1}{\phi}} \Gamma\left(\frac{1}{\phi}\right)}.
\]
4.2. LIKELIHOOD BASED APPROACHES

Consequently, the joint log-likelihood, in exponential family notation, for a sample of size \( N \) is given by

\[
L(\beta, \phi; y) = \sum_{i=1}^{N} \frac{1}{\phi} \left\{ -\frac{y_i}{\mu_i} - \log \mu_i \right\} + \left\{ \left( \frac{1}{\phi} - 1 \right) \log \mu_i - \frac{1}{\phi} \log \phi - \log \Gamma \left( \frac{1}{\phi} \right) \right\}
\]

where \( \Gamma \) denotes the gamma function (Abramowitz and Stegun 1972, (6.1.1)). Hence, it follows that \( A_i = 1 \), \( a_i(\mu_i) = -\frac{1}{\mu_i} \) and \( b_i(a_i) = -\log(-a_i) = \log \mu_i \). Furthermore, \( E[Y_i] = \mu_i \) and \( \text{Var}[Y_i] = \phi \mu_i^2 \), and hence the variance function equals \( V(\mu_i) = \mu_i^2 \).

Possible choices for the link function are the inverse, which is the canonical link, the log, and the identity link.

As before, the objective function used to find the constrained maximum likelihood estimates is given by (4.2), and the general forms of the Powell-Hestenes equivalents of the score vector and information matrices can be found quite easily. However, the exact forms depend on the choice of link function and form of constraint \( g(\cdot) \).

The elements of the score vector are found as usual, that is,

\[
\frac{\partial H}{\partial \beta_{\tau_e}} = -\frac{1}{N\phi_{\tau_e}} \sum_{i=1}^{N} \left( \frac{y_i}{\mu_i^{2,\tau_e}} - \frac{1}{\mu_i^{2,\tau_e}} \right) \frac{\partial \mu_i,\tau_e}{\partial \eta_i,\tau_e} x_i + 2\omega \left[ g(\theta_{\tau_e}; \tau_e) + \psi \right] \frac{\partial g}{\partial \beta_{\tau_e}}
\]

\[
\frac{\partial H}{\partial \phi_{\tau_e}} = -\frac{1}{N\phi_{\tau_e}^2} \sum_{i=1}^{N} \left( \frac{y_i}{\mu_i,\tau_e} + \log \mu_i,\tau_e - \log y_i \right)
\]

\[
+ N(\log \phi_{\tau_e} - 1) + N\psi \left( \frac{1}{\phi_{\tau_e}} \right) + 2\omega \left[ g(\theta_{\tau_e}; \tau_e) + \psi \right] \frac{\partial g}{\partial \phi_{\tau_e}}
\]

where \( \mu_i,\tau_e \) and \( \eta_i,\tau_e \) again denote the constrained mean and linear predictor corresponding to the \( i \)-th observation, respectively, and \( \psi(\cdot) \) denotes the digamma function (Abramowitz and Stegun 1972, (6.3.1)).

The elements of the Powell-Hestenes equivalent of the observed information matrix are then

\[
I^\prime_{o}(\beta_{\tau_e}, \beta_{\tau_e}) = -\frac{1}{N\phi_{\tau_e}} \sum_{i=1}^{N} \left\{ \left( \frac{1}{\mu_i^{2,\tau_e}} - \frac{2y_i}{\mu_i^{3,\tau_e}} \right) \left( \frac{\partial \mu_i,\tau_e}{\partial \eta_i,\tau_e} \right)^2 \right\}
\]

\[
+ \left( \frac{y_i}{\mu_i^{2,\tau_e}} - \frac{1}{\mu_i^{1,\tau_e}} \right) \frac{\partial^2 \mu_i,\tau_e}{\partial \eta_i^{2,\tau_e}} x_i x_i^T + 2\omega \frac{\partial g}{\partial \beta_{\tau_e}} \frac{\partial g}{\partial \beta_{\tau_e}^T}
\]

129
4.2. LIKELIHOOD BASED APPROACHES

\[ I_o'(\beta_{\tau_\epsilon}, \phi_{\tau_\epsilon}) = \frac{1}{N\phi_{\tau_\epsilon}^2} \sum_{i=1}^{N} \left( \frac{y_i}{\mu_{i,\tau_\epsilon}} - \frac{1}{\mu_{i,\tau_\epsilon}} \right) \frac{\partial \mu_{i,\tau_\epsilon}}{\partial \eta_{i,\tau_\epsilon}} x_i + 2\omega \frac{\partial g}{\partial \beta_{\tau_\epsilon}} \frac{\partial g}{\partial \phi_{\tau_\epsilon}} \]

\[ I_o'(\phi_{\tau_\epsilon}, \phi_{\tau_\epsilon}) = \frac{1}{N\phi_{\tau_\epsilon}^3} \left\{ 2 \left[ \sum_{i=1}^{N} \left( \frac{y_i}{\mu_{i,\tau_\epsilon}} + \log \mu_{i,\tau_\epsilon} - \log y_i \right) + N(\log \phi_{\tau_\epsilon} - 1) + N\psi\left(\frac{1}{\phi_{\tau_\epsilon}}\right) \right] \right. \]

\[ \left. - N \left( 1 - \frac{\psi^{(1)}\left(\frac{1}{\phi_{\tau_\epsilon}}\right)}{\phi_{\tau_\epsilon}} \right) \right\} + 2\omega \left( \frac{\partial g}{\partial \phi_{\tau_\epsilon}} \right)^2, \]

where \( \psi^{(1)}(\cdot) \) is the trigamma function, the first derivative of the digamma function (Abramowitz and Stegun 1972, (6.3.1)), and the second derivatives of \( g(\cdot) \) have been ignored as before (Osborne, 2000).

Taking expectations, the equivalent of the expected information matrix is given by

\[ I_e'(\beta_{\tau_\epsilon}, \beta_{\tau_\epsilon}) = \frac{1}{N\phi_{\tau_\epsilon}} \sum_{i=1}^{N} \left\{ \frac{1}{\mu_{i,\tau_\epsilon}^2} \left( \frac{\partial \mu_{i,\tau_\epsilon}}{\partial \eta_{i,\tau_\epsilon}} \right)^2 \right\} x_i x_i^T + 2\omega \frac{\partial g}{\partial \beta_{\tau_\epsilon}} \frac{\partial g}{\partial \beta_{\tau_\epsilon}^T} \]

\[ I_e'(\beta_{\tau_\epsilon}, \phi_{\tau_\epsilon}) = 0 + 2\omega \frac{\partial g}{\partial \beta_{\tau_\epsilon}} \frac{\partial g}{\partial \phi_{\tau_\epsilon}} \]

\[ I_e'(\phi_{\tau_\epsilon}, \phi_{\tau_\epsilon}) = \frac{1}{\phi_{\tau_\epsilon}^3} \left\{ \psi^{(1)}\left(\frac{1}{\phi_{\tau_\epsilon}}\right) - 1 \right\} + 2\omega \left( \frac{\partial g}{\partial \phi_{\tau_\epsilon}} \right)^2, \]

Specific results for the link functions are obtained by substituting the appropriate derivative for \( \frac{\partial \mu_{i,\tau_\epsilon}}{\partial \eta_{i,\tau_\epsilon}} \).

Before discussing the two forms of constraints, it will be beneficial to give results for the left truncated mean for the gamma distribution.

The \( r \)-th moment of the left truncated gamma distribution is (Johnson et al., 1994a)

\[ \tilde{\mu}_{\tau_\epsilon}^{(r)}(\kappa) = (\phi_{\tau_\epsilon} \mu_{\tau_\epsilon})^r \frac{\Gamma_{\kappa^*} \left( \frac{1}{\phi_{\tau_\epsilon}} + r \right)}{\Gamma_{\kappa^*} \left( \frac{1}{\phi} \right)}, \]

where \( \kappa^* = \frac{\kappa}{\phi_{\tau_\epsilon} \mu_{\tau_\epsilon}} \) and

\[ \Gamma_x(\alpha) = \int_{0}^{x} t^{\alpha-1} e^{-t} dt \]

denotes the incomplete gamma function (Johnson et al. 1994, (17.4)). Consequently, the left truncated mean is given by

\[ \tilde{\mu}_{\tau_\epsilon}(\kappa) = \phi_{\tau_\epsilon} \mu_{\tau_\epsilon} \frac{\Gamma_{\kappa^*} \left( \frac{1}{\phi_{\tau_\epsilon}} + 1 \right)}{\Gamma_{\kappa^*} \left( \frac{1}{\phi} \right)}. \]
4.2. LIKELIHOOD BASED APPROACHES

**Constraint 1:** \( g(\theta_{\tau}; \tau_e) = F(\kappa; \beta_{\tau_e}, \phi_{\tau_e}) - \epsilon \)

In order to calculate (4.5), the derivative of \( F(\kappa) \) is required. The general form of this derivative is given by (4.6), which becomes here

\[
\frac{\partial F(\kappa)}{\partial \mu_{\tau_e}} = \frac{1}{\phi_{\tau_e} \mu_{\tau_e}^2} \left[ \frac{\phi_{\tau_e} \mu_{\tau_e} \Gamma(\frac{1}{\phi_{\tau_e}})(\frac{1}{\mu_{\tau_e}} + 1)}{\Gamma(\frac{1}{\phi_{\tau_e}})} - \mu_{\tau_e} \right] \left[ \frac{\Gamma(\frac{1}{\phi_{\tau_e}})}{\Gamma(\frac{1}{\phi_{\tau_e}})} \right] \\
= \frac{\phi_{\tau_e} \Gamma(\frac{1}{\phi_{\tau_e}})(\frac{1}{\mu_{\tau_e}} + 1) - \Gamma(\frac{1}{\phi_{\tau_e}})}{\phi_{\tau_e} \mu_{\tau_e} \Gamma(\frac{1}{\phi_{\tau_e}})},
\]

since \( F(\kappa) = \Gamma(\frac{1}{\phi_{\tau_e}})/\Gamma(\frac{1}{\phi_{\tau_e}}) \). This form can further be simplified (Abramowitz and Stegun 1972, (6.5.22)), that is

\[
\frac{\partial F(\kappa)}{\partial \mu_{\tau_e}} = - \left( \mu_{\tau_e} \Gamma(\frac{1}{\phi_{\tau_e}}) \right)^{-1} \left( \frac{\kappa}{\phi_{\tau_e} \mu_{\tau_e}} \right)^{\frac{1}{\phi_{\tau_e}}} \exp \left( - \frac{\kappa}{\phi_{\tau_e} \mu_{\tau_e}} \right). 
\]

Hence, the derivative of the constraint with respect to \( \beta_{\tau_e} \) is given by

\[
\frac{\partial g}{\partial \beta_{\tau_e}} = - \left( \mu_{\tau_e} \Gamma(\frac{1}{\phi_{\tau_e}}) \right)^{-1} \left( \frac{\kappa}{\phi_{\tau_e} \mu_{\tau_e}} \right)^{\frac{1}{\phi_{\tau_e}}} \exp \left( - \frac{\kappa}{\phi_{\tau_e} \mu_{\tau_e}} \right) \frac{\partial \mu_{\tau_e}}{\partial \eta_{\tau_e}} x_{\tau_e},
\]

with the exact form depending on the link function.

Similarly, the derivative of \( g(\cdot) \) with respect to \( \phi_{\tau_e} \) is found by substituting the relevant pieces into (4.7). This gives

\[
\frac{\partial g}{\partial \phi_{\tau_e}} = \frac{F(\kappa)}{\phi_{\tau_e}^2 \mu_{\tau_e}} \left\{ \tilde{\mu}_{\tau_e}(\kappa) \frac{\mu_{\tau_e}}{\mu_{\tau_e}} + \log \frac{\mu_{\tau_e}}{\phi_{\tau_e}} + \log \phi_{\tau_e} - 1 + \psi(\frac{1}{\phi_{\tau_e}}) \right\} - \frac{1}{F(\kappa)} \int_0^\kappa f(z) \log(z) \, dz.
\]

Unfortunately, there is no analytic solution for the integral. However, it can easily be evaluated numerically. The exact value again depends on the choice of link function.

**Constraint 2:** \( g(\theta_{\tau}; \tau_e) = F^{-1}(\epsilon; \beta_{\tau_e}, \phi_{\tau_e}) - \kappa \)

In the gamma case (4.8) becomes

\[
\frac{\partial F^{-1}(\epsilon)}{\partial \mu_{\tau_e}} = - \frac{1}{\phi_{\tau_e} \mu_{\tau_e}} \left[ \frac{\tilde{\mu}_{\tau_e}(F^{-1})}{\mu_{\tau_e}} - 1 \right] \frac{\epsilon}{f(F^{-1}(\epsilon))}.
\]

This can now be substituted into (4.9) and an exact form can be determined for the various choices of link function.
Similarly, the derivative of the constraint with respect to $\phi_{\tau_{c}}$ is given by

$$
\frac{\partial g}{\partial \phi_{\tau_{c}}} = -\epsilon \phi_{\tau_{c}} f(F^{-1}) \left\{ \frac{\mu_{\tau_{c}}(F^{-1})}{\varphi_{\tau_{c}}} \log \mu_{\tau_{c}} + \log \phi_{\tau_{c}} - 1 + \psi\left(\frac{1}{\varphi_{\tau_{c}}}\right) - \frac{1}{\epsilon} \int_{0}^{F^{-1}} f(z) \log(z) \, dz \right\},
$$

where $F^{-1} = F^{-1}(\epsilon)$. As in the previous section, the integral cannot be solved analytically, but it can be evaluated numerically.

### 4.3 Simulation Results

As in previous chapters, a simulation study is used to evaluate the approaches proposed in this chapter and the software used to perform these simulations was again R (Ihaka and Gentleman, 1996). The simulations were run for normally and gamma distributed data. Both of these are further discussed in the following sections.

It should be noted that Constraint 1 is on a probability scale, which is generally much smaller than the scale of the log-likelihood or objective function. Consequently, the Fisher scoring scheme may have difficulty converging to the solution quickly. This was numerically confirmed and hence this form of constraint was not used in the simulations.

#### 4.3.1 The Normal Distribution

The simulations were structured in the same way as those in Chapter 2. Furthermore, in order to compare the result more directly with those of Chapter 2 the same simulated data were used here. However, under the general GLM framework only the CPL and TMCPL methods are available.

As expected, the label shelf-lives obtained with both methods described in this chapter were identical to those obtained using the methods of Chapter 2. Consequently, as the coverage probability, bias and MSE are identical to those previously tabulated, no tables are given here, and the reader is referred to the tables and discussion in Chapter 2.
4.3.2 The Gamma Distribution

Gamma distributed data that can be used for GLM analysis exhibits a specific mean-variance relationship, that is, the variance is proportional to the square of the mean. As a consequence, the motivation for data generation that is used in the normal case is unlikely to be applicable here. Therefore, rather than considering the label shelf-life that is based on the percent label claim of active ingredient, a different type of process is considered here.

Consider a product which contains a very small initial amount of an unpleasant or toxic chemical, say. Now assume that the amount of toxin increases in an exponential manner over time, and that the variability in the toxicity from one unit to another is proportional to the mean level. Consequently, a gamma model may be reasonable. The true shelf-life is the time at which no more than 100(1 − ε)% of units have toxicity that is greater than the specification limit. Thus, interest now lies in the 100(1 − α)% upper confidence bound for the 1 − ε quantile.

The simulations in the gamma case were constructed as follows. A random sample of size \( r \) was drawn at each of \( n \) times \( x_i \) from \( \gamma(\mu, \phi) \), where \( \mu \) and \( \eta \) are related via the log-link. The linear predictor \( \eta \) was given by \( \log(10) + \beta_1 x_i \), such that the mean initial amount of toxicity equals 10. The \( x_i \) used in these simulations were taken to be \( x_i = 0, 1, \ldots, 12 \) days. The total number of observations available for analysis is \( N = 13r \). Values of \( \beta_1 = 0.5, 0.55 \) and 0.6, \( \phi = 0.5, 0.6, \) and 0.7, and \( r = 1, 2, 5, 10 \) and 20 were used. The values of \( \beta_1 \) and \( \phi \) were chosen as the corresponding growth curves exhibit a steady but not too rapid increased in toxicity. For each combination of \( \beta_1, \phi \) and \( r \), 1000 simulations were performed and the proportion of label shelf-lives less than the true shelf-life, based on a specification limit of \( \kappa = 1000 \), was determined for each method. The true mean degradation lines for each \( \beta_1 \) are shown in Figure 4.1, and the density of the response at time \( t = 10 \) days is displayed in Figure 4.2 for each parameter combination of \( \beta_1 \) and \( \phi \).

Assuming that \( \kappa = 1000 \), the true shelf-lives for each parameter combination of \( \beta_1 \) and \( \phi \) are given in Table 4.1. For each simulation, the label shelf-life \( \hat{\tau}_{0.99, U(0.05)} \), based on a
significance level of \( \alpha = 0.05 \) was calculated. Note that \( \hat{\tau}_{0.99,U(0.05)} \) denotes the 95% upper confidence bound, which is equivalent to \( \hat{\tau}_{0.99,L(0.95)} \).

<table>
<thead>
<tr>
<th></th>
<th>( \beta_1 = 0.5 )</th>
<th>( \beta_1 = 0.55 )</th>
<th>( \beta_1 = 0.6 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \phi = 0.5 )</td>
<td>11.42</td>
<td>10.38</td>
<td>9.51</td>
</tr>
<tr>
<td>( \phi = 0.6 )</td>
<td>11.25</td>
<td>10.23</td>
<td>9.38</td>
</tr>
<tr>
<td>( \phi = 0.7 )</td>
<td>11.11</td>
<td>10.10</td>
<td>9.26</td>
</tr>
</tbody>
</table>

Table 4.1: True shelf-lives (in days) for each parameter combination of \( \beta_1 \) and \( \phi \).

Figure 4.1: At each point shown, \( r \) values are drawn randomly drawn from a gamma distribution with these means and with a dispersion parameter of \( \phi = 0.5, 0.6 \) and 0.7.
Figure 4.2: The density of the response at time $t = 10$ days for each parameter combination of $\beta_1$ and $\phi$ – the orange shaded areas indicates a proportion of $\epsilon = 0.01$.

The simulation results are summarized in the following tables and graphs. Unlike in the previous chapters, there are only two methods that were used to estimate the label shelf-life. Consequently, the tables are structured differently to those in previous chapters. Tables 4.2 and 4.3 give the coverage probability, Tables 4.4 and 4.5 give the bias, and Tables 4.6 and 4.7 give the MSE for the CPL and TMCPL method, respectively.
### Table 4.2: Proportion of label shelf-lives estimated using the CPL method that were smaller than the true shelf-life for various parameter values of $\beta_1$ and $\phi$, and $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>$\beta_1$</th>
<th>$\phi$</th>
<th>1</th>
<th>2</th>
<th>5</th>
<th>10</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.50</td>
<td>0.5</td>
<td>0.876</td>
<td>0.915</td>
<td>0.915</td>
<td>0.952</td>
<td>0.948</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.889</td>
<td>0.907</td>
<td>0.925</td>
<td>0.939</td>
<td>0.939</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>0.883</td>
<td>0.927</td>
<td>0.930</td>
<td>0.930</td>
<td>0.939</td>
</tr>
<tr>
<td>0.55</td>
<td>0.5</td>
<td>0.882</td>
<td>0.913</td>
<td>0.936</td>
<td>0.947</td>
<td>0.944</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.882</td>
<td>0.930</td>
<td>0.924</td>
<td>0.939</td>
<td>0.928</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>0.890</td>
<td>0.902</td>
<td>0.929</td>
<td>0.935</td>
<td>0.949</td>
</tr>
<tr>
<td>0.60</td>
<td>0.5</td>
<td>0.873</td>
<td>0.911</td>
<td>0.917</td>
<td>0.949</td>
<td>0.947</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.881</td>
<td>0.906</td>
<td>0.930</td>
<td>0.940</td>
<td>0.942</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>0.879</td>
<td>0.919</td>
<td>0.929</td>
<td>0.929</td>
<td>0.938</td>
</tr>
</tbody>
</table>

### Table 4.3: Proportion of label shelf-lives estimated using the TMCPL method that were smaller than the true shelf-life for various parameter values of $\beta_1$ and $\phi$, and $\epsilon = 0.01$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>$\beta_1$</th>
<th>$\phi$</th>
<th>1</th>
<th>2</th>
<th>5</th>
<th>10</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.50</td>
<td>0.5</td>
<td>0.999</td>
<td>0.994</td>
<td>0.976</td>
<td>0.983</td>
<td>0.971</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.997</td>
<td>0.991</td>
<td>0.973</td>
<td>0.969</td>
<td>0.967</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>0.998</td>
<td>0.995</td>
<td>0.985</td>
<td>0.969</td>
<td>0.966</td>
</tr>
<tr>
<td>0.55</td>
<td>0.5</td>
<td>0.994</td>
<td>0.987</td>
<td>0.977</td>
<td>0.976</td>
<td>0.962</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.996</td>
<td>0.990</td>
<td>0.979</td>
<td>0.978</td>
<td>0.963</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>0.998</td>
<td>0.990</td>
<td>0.987</td>
<td>0.978</td>
<td>0.970</td>
</tr>
<tr>
<td>0.60</td>
<td>0.5</td>
<td>0.997</td>
<td>0.987</td>
<td>0.969</td>
<td>0.985</td>
<td>0.966</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.997</td>
<td>0.988</td>
<td>0.966</td>
<td>0.968</td>
<td>0.967</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>0.993</td>
<td>0.990</td>
<td>0.979</td>
<td>0.968</td>
<td>0.962</td>
</tr>
</tbody>
</table>
Figure 4.3: Coverage Probability plotted against $1/\sqrt{N}$ for each method and combination of parameters.
### Simulation Results

#### Table 4.4: Bias of label shelf-lives estimated using the CPL method for various parameter values of $\beta_1$ and $\phi$, and $\epsilon = 0.01$ and $\alpha = 0.05$.  

<table>
<thead>
<tr>
<th>$\beta_1$</th>
<th>$\phi$</th>
<th>1</th>
<th>2</th>
<th>5</th>
<th>10</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.50</td>
<td>0.5</td>
<td>-0.9</td>
<td>-0.7</td>
<td>-0.5</td>
<td>-0.4</td>
<td>-0.3</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>-1.0</td>
<td>-0.8</td>
<td>-0.6</td>
<td>-0.4</td>
<td>-0.3</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>-1.0</td>
<td>-0.8</td>
<td>-0.6</td>
<td>-0.4</td>
<td>-0.3</td>
</tr>
<tr>
<td>0.55</td>
<td>0.5</td>
<td>-0.8</td>
<td>-0.6</td>
<td>-0.4</td>
<td>-0.3</td>
<td>-0.2</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>-0.8</td>
<td>-0.7</td>
<td>-0.4</td>
<td>-0.3</td>
<td>-0.2</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>-0.9</td>
<td>-0.7</td>
<td>-0.5</td>
<td>-0.3</td>
<td>-0.2</td>
</tr>
<tr>
<td>0.60</td>
<td>0.5</td>
<td>-0.7</td>
<td>-0.5</td>
<td>-0.3</td>
<td>-0.3</td>
<td>-0.2</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>-0.7</td>
<td>-0.6</td>
<td>-0.4</td>
<td>-0.3</td>
<td>-0.2</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>-0.7</td>
<td>-0.6</td>
<td>-0.4</td>
<td>-0.3</td>
<td>-0.2</td>
</tr>
</tbody>
</table>

#### Table 4.5: Bias of label shelf-lives estimated using the TMCPL method for various parameter values of $\beta_1$ and $\phi$, and $\epsilon = 0.01$ and $\alpha = 0.05$.  

<table>
<thead>
<tr>
<th>$\beta_1$</th>
<th>$\phi$</th>
<th>1</th>
<th>2</th>
<th>5</th>
<th>10</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.50</td>
<td>0.5</td>
<td>-2.5</td>
<td>-1.4</td>
<td>-0.7</td>
<td>-0.5</td>
<td>-0.3</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>-2.6</td>
<td>-1.5</td>
<td>-0.8</td>
<td>-0.5</td>
<td>-0.3</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>-2.8</td>
<td>-1.6</td>
<td>-0.8</td>
<td>-0.5</td>
<td>-0.4</td>
</tr>
<tr>
<td>0.55</td>
<td>0.5</td>
<td>-2.0</td>
<td>-1.1</td>
<td>-0.6</td>
<td>-0.4</td>
<td>-0.3</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>-2.1</td>
<td>-1.2</td>
<td>-0.6</td>
<td>-0.4</td>
<td>-0.3</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>-2.3</td>
<td>-1.2</td>
<td>-0.7</td>
<td>-0.4</td>
<td>-0.3</td>
</tr>
<tr>
<td>0.60</td>
<td>0.5</td>
<td>-1.6</td>
<td>-0.9</td>
<td>-0.5</td>
<td>-0.3</td>
<td>-0.2</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>-1.8</td>
<td>-0.9</td>
<td>-0.5</td>
<td>-0.3</td>
<td>-0.2</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>-1.9</td>
<td>-1.0</td>
<td>-0.5</td>
<td>-0.3</td>
<td>-0.2</td>
</tr>
</tbody>
</table>
### 4.3 Simulation Results

Figure 4.4: Bias plotted against $1/\sqrt{N}$ for each method and combination of parameters.
### Table 4.6: MSE of label shelf-lives estimated using the CPL method for various parameter values of $\beta_1$ and $\phi$, and $\epsilon = 0.01$ and $\alpha = 0.05$. 

<table>
<thead>
<tr>
<th>$\beta_1$</th>
<th>$\phi$</th>
<th>1</th>
<th>2</th>
<th>5</th>
<th>10</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.50</td>
<td>0.5</td>
<td>1.5</td>
<td>0.8</td>
<td>0.4</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>1.6</td>
<td>0.9</td>
<td>0.4</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>1.8</td>
<td>1.0</td>
<td>0.4</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>0.55</td>
<td>0.5</td>
<td>1.1</td>
<td>0.6</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>1.1</td>
<td>0.7</td>
<td>0.3</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>1.3</td>
<td>0.7</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>0.60</td>
<td>0.5</td>
<td>0.8</td>
<td>0.4</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.8</td>
<td>0.4</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>0.9</td>
<td>0.5</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
</tr>
</tbody>
</table>

### Table 4.7: MSE of label shelf-lives estimated using the TMCPL method for various parameter values of $\beta_1$ and $\phi$, and $\epsilon = 0.01$ and $\alpha = 0.05$. 

<table>
<thead>
<tr>
<th>$\beta_1$</th>
<th>$\phi$</th>
<th>1</th>
<th>2</th>
<th>5</th>
<th>10</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.50</td>
<td>0.5</td>
<td>6.9</td>
<td>2.2</td>
<td>0.6</td>
<td>0.3</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>7.5</td>
<td>2.4</td>
<td>0.7</td>
<td>0.3</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>8.4</td>
<td>2.7</td>
<td>0.8</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>0.55</td>
<td>0.5</td>
<td>4.4</td>
<td>1.4</td>
<td>0.4</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>4.8</td>
<td>1.6</td>
<td>0.5</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>6.0</td>
<td>1.7</td>
<td>0.5</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>0.60</td>
<td>0.5</td>
<td>2.9</td>
<td>0.9</td>
<td>0.3</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>4.0</td>
<td>1.0</td>
<td>0.3</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>5.2</td>
<td>1.2</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
</tr>
</tbody>
</table>
Figure 4.5: Mean Square Error plotted against $1/N$ for each method and combination of parameters.
Similar to Chapter 2 these tables can be summarized further by fitting appropriate models. The full models for coverage probability for each method involved $1/\sqrt{N}$, $\beta^*_1 = \beta_1 - 0.55$ and $\phi^* = \phi - 0.6$. Standard model reduction was used to test higher order terms in turn using a significance level of 1% (observing marginality). Table 4.8 gives the parameter estimates for all significant effects. This modeling approach was repeated for the bias, and the parameter estimates are given in Table 4.9. The results from these models are discussed in Section 4.4.

<table>
<thead>
<tr>
<th>Method</th>
<th>(Intercept)</th>
<th>$\frac{1}{\sqrt{N}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPL</td>
<td>0.962</td>
<td>-0.275</td>
</tr>
<tr>
<td>TMCPL</td>
<td>0.960</td>
<td>0.138</td>
</tr>
</tbody>
</table>

Table 4.8: For each method the parameter estimates for the model of coverage probability on $1/\sqrt{N}$, $\beta^*_1 = \beta_1 - 0.55$ and $\phi^* = \phi - 1$ are given. Blank values indicate that the parameter was not significant at the 1% level and the model was re-fitted without the corresponding term(s).

<table>
<thead>
<tr>
<th>Method</th>
<th>(Intercept)</th>
<th>$\beta^*_1$</th>
<th>$\phi^*$</th>
<th>$\frac{1}{\sqrt{N}}$</th>
<th>$\frac{\beta_1}{\sqrt{N}}$</th>
<th>$\frac{\phi}{\sqrt{N}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPL</td>
<td>-0.009</td>
<td>0.085</td>
<td>-0.032</td>
<td>-0.268</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMCPL</td>
<td>0.037</td>
<td>-0.040</td>
<td>0.049</td>
<td>-0.850</td>
<td>1.806</td>
<td>-0.756</td>
</tr>
</tbody>
</table>

Table 4.9: For each method the parameter estimates for the model of bias on $1/\sqrt{N}$, $\beta^*_1 = \beta_1 - 0.55$ and $\phi^* = \phi - 0.6$ are given. Blank values indicate that the parameter was not significant at the 1% level and the model was re-fitted without the corresponding term(s).

### 4.4 Discussion

This chapter has extended the definition of label shelf-life and the methods for estimating a label shelf-life to the situation when generalized linear model assumptions are applicable. When the data are normally distributed then the comments of Chapter 2 apply as the
GLM based methods result in the same shelf-life as the corresponding methods in Chapter 2.

In the gamma case, the CPL method tends to give coverage probabilities that fall short of the desired 0.95 for small $N$, but it does tend toward 0.95 as $N \to \infty$. This is analogous to the observations made for the normal distribution in Chapter 2. However, the rate at which the coverage probability approaches 0.95 does appear to only depend on $N$. While this is different to the observations made in Chapter 2, this may be due to the choice of values for $\beta_1$ and $\phi$ that were used in the simulations. In addition, the bias of the CPL method also tends to be relatively small when compared to the bias of the TMCPL method. This is not surprising as the CPL tends to give coverage probabilities that fall short of the desired 0.95.

The TMCPL in the gamma case results in conservative coverage probabilities, as in the normal case, but here it appears to remain conservative for large $N$. Similar to Chapter 2, the rate at which the coverage probability approaches 0.95 depends only on the sample size $N$. The bias on the other hand appears to depend on $N$, $\beta_1$ and $\phi$, which is analogous to the observations made in Chapter 2.

In conclusion, the TMCPL can also be recommended in the GLM case, as the coverage probabilities obtain from this method tend to be conservative, but close to the desired level of 0.95. The CPL can only be recommended for very large sample size.
Several issues have been raised in Chapter 1 in relation to the design of stability trials and the analysis of stability data as currently outlined in the regulatory guidelines.

The current frequency and spacing of time points at which the product is assayed have been shown to be less than optimal. Consequently, it is suggested that the product be sampled more frequently near the shelf-life, instead of being sampled more frequently during the first year after manufacture. This has been shown to have immediate effects on the accuracy of the estimated shelf-life and thus can extend the label shelf-life. However, to sample very infrequently in the first year is of limited practical value as it is recognized that estimates of the label shelf-life are usually desirable early in the life of a drug product. Consequently, it is recommended that the product be assayed at equal time intervals throughout its life, instead of less frequently during the second and consecutive years of storage.

In respect to the assay result, the current guidelines assume that assay variability is the only source of variability affecting the assay result. However, this thesis argues that the variability in the drug product needs to be taken into account when determining a label shelf-life. The only way these sources of variation can be distinguished is by duplicating the assays at each time point. Assuming that the variability of the drug product can be identified, new precise definitions for the true and label shelf-life are suggested.
Another source of variability is due to batches. The current regulatory guidelines treat batches as fixed effects, which does not allow conclusions to be drawn about future batches. Assuming that batches are random effects has led many authors to develop statistical tests which can be used to test whether the intercept and slope have a non-zero variance and covariance, or whether they can be treated as fixed for all batches. It is pointed out here that basing the estimation on a test of significance has some inherent problems. Consequently, it is suggested that no such test be used, but that instead the model appropriately account for batch-to-batch variability.

Based on the new definitions for true shelf-life, several methods for determining a label shelf-life have been investigated. The results from simulation studies have shown that the truncated modified profile likelihood approach is the most suitable under a variety of circumstances, including cases where batch-to-batch variability or non-normally distributed data are present. Consequently, this method should be used in all situations except the most simple — linear regression. In this situation the non-central t-distribution provides an optimal solution.

While this thesis has mainly focused on the definition and estimation of a label shelf-life in the pharmaceutical industry, the methods developed here are not limited to this application. The following section gives some examples of where the use of the tolerance bound as suggested in this thesis may be useful.

## 5.1 Other Applications

This thesis is primarily concerned with estimating a lower $100(1 - \alpha)\%$ confidence bound for the time at which a proportion $\epsilon$ of product first falls below the specification limit. There are however, many more similar circumstances in which the new definitions and associated method of estimation can be useful. Some applications which come to mind are the following.

- Toxicity of a drug product: Some drug products may increase in toxicity over time.
The manufacturer may be interested in estimating the time at which the 95% upper confidence bound for the 0.99 quantile reaches the specification limit set for safe consumption.

- **Safety of food products**: This scenario is very similar to the toxicity of drug products *(c.f. Section 4.2.3)*.

- **Grow-out periods of farmed fish species**: Fish species may accumulate chemical residues during the time they are commercially farmed. Consequently, the producer may be interested in the time at which he can be 90% confident that no more than 5% of fish have a residue level greater than some pre-determined limit. In this case, animals may correspond to the batches, while herds may add yet another source of variation not yet investigated.

- **Withholding periods**: Drugs and antibiotics can often be detected in the meat, milk, etc. of farmed animals after being administered. The levels found in the produce generally degrade over time. The withholding period is the time period for which the produce should not be consumed (by humans) after the drug is administered. This withholding period would require the estimation of the time at which no more than 1% of animals show a residue level of more than the Maximum Residue Level (MRL) with 95% confidence.

- **Fuel tanks in cars**: The distance that can be travelled by car with a single tank of fuel will depend on many factors, including driver, size of fuel tank and probably others. Drivers are likely to be random effects, similar to batches, while the size of the fuel tank can be assumed to be the determining variable. Consequently, it should be possible to determine the size of fuel tank that would be required such that at most 5% of trips fall short of a desired target with 95% confidence. *(While this example is clearly fabricated, it illustrates the potential for the approach detailed in this thesis).*
5.2 Future Work

While this thesis has addressed the issues indicated in Chapter 1, it has also raised many more which will require further research. These include, but are not limited to, the following questions which are listed in no particular order.

- It has been assumed that assay-to-assay variability is either negligible or has been eliminated by taking assay averages, for example. How can assay variability be included in the modeling process?

- It has previously been mentioned that there are several possible sources of variability. One such source, which has been assumed to be negligible, is day-to-day variability. The question that arises is how such a source of variability can be included formally in the analysis? This is particularly of interest when batches are put on stability trials at different times, but are analysed at the same time. This situation is depicted in Figure 5.1.

- How do $\beta_1$, $\sigma$ and $d_2$ interact in the batch case? More simulations are required here.

- Random effects models in the GLM framework can already be fitted, but how can the methods of Chapter 4 be generalised to the batch case? There is already an extensive literature on fitting generalised linear mixed models (GLMM), e.g. Schall (1991). However, these methods are approximate, and “efficient ways to fit GLMMs are a research topic” (Venables and Ripley, 2002).

- Accelerated stability testing is a common approach to obtain estimates of the shelf-life sooner. How do the methods presented in this thesis generalise to the nonlinear models which are used in accelerated testing?

- How do the methods presented in this thesis perform when fractional factorial designs are used in the stability trial?

- Can non-parametric quantile regression methods (Koenker and Bassett Jr, 1978) be used in a similar way to those of Chapter 2, and can they be extended to the batch
5.2. FUTURE WORK

- How do these non-parametric methods compare to the methods presented here?

  - The non-normal approximations have been shown to be poor. Why is this the case? Is it because under the approximation the quantile $q_{t, \epsilon}$ and its variance $\text{Var} [ q_{t, \epsilon} ]$ are not really independent? Can this be remedied (easily)?

- Poisson count data may arise when observing the growth of pathogens in food products over time, e.g. Salmonella. How do the approaches in Chapter 4 perform under Poisson assumptions? Can the methods of Chapter 4.1 be generalised to other distributions?

- Bootstrap methods are often useful to estimate quantiles and their confidence bounds. How do various bootstrap confidence intervals perform in the determination of a label shelf-life?

- How do the TMCPL and other methods behave for different values of $\epsilon$ and $\alpha$?

![Figure 5.1: Illustration of a stability trial design where batches are added to the trial over time. At any particular point in time all batches are assayed.](image-url)
Bibliography


Appendix A

The following paper has been accepted for publication in

Pharmaceutical Statistics vol 3, 2004 pp: 3-11

Andreas Kiermeier, Richard G.Jarrett and Arunas P.Verbyla
A new approach to estimating shelf life.

NOTE: This publication is included in the print copy of the thesis.
It is also available online to authorised users at:

http://dx.doi.org/10.1002/pst.78
A new approach to estimating shelf-life  
This is a preprint of an article published in Pharmaceutical Statistics, 2004, 3, 3-11  
http://www.interscience.wiley.com/  
Andreas Kiermeier\textsuperscript{1}  
The University of Adelaide/  
SARDI  
Arūnas P. Verbyla  
The University of Adelaide/  
SARDI  
24 June 2003  

Abstract  
The current approach to the estimation of shelf-life and the determination of the label shelf-life as detailed in the ICH guidelines to industry is presented. The shortcomings of the status quo are explained and a possible solution is offered, which gives rise to a new definition of shelf-life. Several methods for calculating a label shelf-life are presented and investigated using a simulation study. Recommendations to adopt the new definition and to increase sample sizes are made.  
Keywords: Stability Analysis, Shelf-life, Quantile, Constraint, Modified Profile Likelihood  

1 Introduction  
Pharmaceutical companies are required by law to indicate the expiry date of each pharmaceutical product on the immediate container. This expiry date indicates the end of the period of time, known as shelf-life, during which the product can be expected to meet specifications. The guidelines which govern how the label  
\textsuperscript{1}Author to whom correspondence should be addressed: BiometricsSA, PMB 1, Glen Osmond, SA 5064, Australia, e-mail: Andreas.Kiermeier@adelaide.edu.au, Phone: +61 8 8303 6819, Facsimile: +61 8 8303 6761
shelf-life, and hence the expiry date, is determined are set out by the International Conference on Harmonisation (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use ([1], [2]).

The ICH states that "usually the relationship [between the sampling time and assay result] can be represented by a linear, quadratic or cubic function on an arithmetic or logarithmic scale". An appropriate statistical model for such a relationship might take the form

\[ y_i = x_i^T \beta + \epsilon_i, \]  

(1)

where \( y_i \) is the response from the \( i \)-th assay for \( i = 1, \ldots, N \), on the arithmetic or logarithmic scale, \( x_i \) is a \( p \)-dimensional predictor vector corresponding to the \( i \)-th response, \( \beta \) is a \( p \)-dimensional vector of fixed effects, and \( \epsilon_i \) are independently and identically distributed normal errors, with mean 0 and variance \( \sigma^2 \). In matrix notation (1) can be expressed as

\[ y = X\beta + \epsilon, \]

where \( X \) is the design matrix with \( i \)-th row \( x_i^T \). Variation can arise from two sources. In this paper we are concerned with the unit-to-unit variability, as exemplified by the variation in the level of active ingredient from one tablet to another, say. A second source is the variability associated with the measurement or assay system. We will suppose that either this is small relative to \( \sigma^2 \) or that it is separately estimated and eliminated from the estimate of \( \sigma^2 \).

The current definition of shelf-life ([1]) is

"the time interval that a drug product is expected to remain within the..."
approved shelf-life specification provided that it is stored under the conditions defined on the label in the proposed containers and closure.”

Assuming that the response decreases over time, this definition is equivalent to the mathematical definition for the true shelf-life of a single batch ([3]), namely the time \( \tau \) such that

\[
\tau = \inf \{ t : x(t)^T \beta \leq \kappa \}, \tag{2}
\]

where \( \inf \) denotes the infimum or greatest lower bound, \( \kappa \) is the specification limit (or acceptance criterion) and \( x(t) \) is the vector of regressors at time \( t \). A similar definition applies to responses that increase with time.

The guidelines also stipulate that the approach for determining the label shelf-life is

“to determine the time at which the 95% one-sided confidence limit for the mean curve intersects the acceptance criterion”.

This approach attempts to give an assurance that the label shelf-life is smaller than the true shelf-life with probability \( 1 - \alpha \). Hence, the label shelf-life is a \( 100(1 - \alpha)\% \) lower confidence bound for the true shelf-life.

Consequently, the label shelf-life, \( \hat{\tau}_{L(\alpha)} \), is given by the time at which the lower confidence bound for \( x(t)^T \beta \) intersects the specification limit ([4]),

\[
\hat{\tau}_{L(\alpha)} = \inf \{ t : x(t)^T \hat{\beta} + t_{\alpha,N-p} s \sqrt{x(t)^T (X^T X)^{-1} x(t)} \leq \kappa \},
\]

where \( t_{\alpha,N-p} \) is the \( \alpha \) quantile of the \( t \)-distribution with \( N - p \) degrees of freedom, and \( \hat{\beta} \) and \( s \) are the least squares estimates of \( \beta \) and \( \sigma \).
A different approach, which is based on the prediction interval of a future observation, was proposed in [5]. They define what will be referred to as a modified label shelf-life, given by the intersection of the 100(1 - \alpha)\% lower prediction bound and the specification limit, that is, the smallest value of \( t \) for which

\[ x(t)^T \hat{\beta} + t_{\alpha,N-p}s\sqrt{1 + x(t)^T(X^TX)^{-1}x(t)} \leq \kappa. \]  

(3)

This approach attempts to make a statement about the ability of an individual unit (a tablet, say) to meet the specifications at time \( t \), rather than the ability of the mean of all units of the batch to meet the specifications.

Which definition is used clearly has important implications for the consumer.

The implications and problems with the current methodology will be discussed in Section 2. This leads to a new formal definition of true shelf-life which addresses these problems. In Section 3 several approaches are given that provide an estimate of the label shelf-life corresponding to this new definition. These approaches are evaluated in a simulation study presented in Section 4. A discussion of the results and an indication of further work is given in Section 5.

2 An assessment of the various definitions

The current definition of the true shelf-life, \( \tau \), as given in (2), is based on the line of best fit. The estimated shelf-life, \( \hat{\tau} \), is given by the intersection of the mean response and the specification limit, i.e. it satisfies

\[ \hat{\tau} = \inf \{ t : x(t)^T \hat{\beta} \leq \kappa \}. \]

The use of the lower confidence bound for the mean response yields the label shelf-life, which is a lower confidence bound for the true shelf-life. As the sample size
increases the estimated shelf-life is determined more accurately. In the limiting case as \( N \to \infty \) the label shelf-life approaches the true shelf-life. However, in the limit, only 50\% of the product remains within specification at time \( \hat{\tau} \).

If the unit-to-unit variability, \( \sigma \), is small, this may not be a problem since the specification limit will not have been breached by a large amount. However, if \( \sigma \) is large, some products may fall far short of the specification, with potentially serious consequences.

Another undesirable property of the current regulatory approach is the lack of influence the product variability has on the estimation process. Increasing the number of samples can compensate for a manufacturing process with large variability \( \sigma \). This is often cheaper than attempting to improve the manufacturing process, but it does not put enough emphasis on good manufacturing practices and commitment to quality.

On the other hand, the modified label shelf-life does not approach the true shelf-life in large samples, but instead has the large sample property that ensures that a proportion \( \alpha \) of the drug product falls below the specification limit at the time the modified label shelf-life has expired. This follows from the fact that as \( N \to \infty \), equation (3) reduces to

\[
x(t)^T \beta + z_\alpha \sigma \leq \kappa ,
\]

where the left hand side is the \( \alpha \) quantile of the distribution of \( Y \) given \( x(t) \), and \( z_\alpha \) is the \( \alpha \) quantile of the standard normal distribution. We believe this to be a more appropriate basis for the true shelf-life.

To overcome the problems with the current regulatory approach and the pre-
diction bound approach, the following new definitions of true shelf-life, estimated shelf-life and label shelf-life are proposed.

**Definition 1** Under model (1), the true shelf-life, denoted by $\tau_\epsilon$, is the minimum time at which $100\epsilon\%$ of the drug product from the current batch has an activity level that is less than or equal to the pre-determined specification limit $\kappa$, that is,

$$\tau_\epsilon = \inf \left\{ t : x(t)^T \beta + z_\epsilon \sigma \leq \kappa \right\},$$

where $x(t)^T \beta$ is the mean activity level at time $t$ and $z_\epsilon$ is the $\epsilon$ quantile of the standard normal distribution.

Note that $\tau_{0.5}$ represents the true shelf-life currently used by regulatory bodies.

The estimated shelf-life, $\hat{\tau}_\epsilon$, is then defined as the time at which the estimated $\epsilon$ quantile of $Y$ intersects the specification limit.

**Definition 2** Based on Definition 1, the label shelf-life, denoted by $\hat{\tau}_{\epsilon, L(\alpha)}$, is the $100(1 - \alpha)\%$ lower confidence bound for the true shelf-life, $\tau_\epsilon$.

This is presented graphically in Figure 1. Similar definitions apply to responses that increase with time.

*[Figure 1 near here]*

Irrespective of the method of estimation that is used, several general points can be made about this definition. Firstly, with $100(1 - \alpha)\%$ confidence, at least $100(1 - \epsilon)\%$ of drug product will meet the specification when the label shelf-life has been reached. Secondly, the true shelf-life can be estimated more accurately by increasing sample size, and thirdly, reducing the variability in the manufacturing process is a good way to increase the label shelf-life.

Several possibilities for obtaining a label shelf-life are available. These are detailed in the next section for the case of linear degradation. Unless otherwise indicated, the terms *true shelf-life, estimated shelf-life* and *label shelf-life* will henceforth refer to the new definitions.
3 Estimation of the label shelf-life

Suppose the mean response is $\mu_t = x(t)^T \beta = \beta_0 + \beta_1 t$, where $t$ denotes time. Interest lies in the $\epsilon$ quantile of the response, given $t$, which under (1) is given by

$$q_{t,\epsilon} = \beta_0 + \beta_1 t + z_{\epsilon} \sigma,$$

assuming that the variability does not depend on the mean.

There are at least three possible approaches that can be used to estimate and find confidence bounds for the true shelf-life. These are based on a Normal Approximation approach, a variety of Profile Likelihood approaches, and a Noncentral $t$-distribution approach, and are discussed in the following sections.

3.1 Normal Approximation

There are various ways of estimating $q_{t,\epsilon}$. The general form of the estimate ([6]) is

$$\hat{q}_{t,\epsilon} = \hat{\beta}_0 + \hat{\beta}_1 t + \delta s,$$

where $\hat{\beta}_0$, $\hat{\beta}_1$ and $s$ denote the least squares estimates of $\beta_0$, $\beta_1$ and $\sigma$, and various choices for $\delta$ are available.

Independence of $(\hat{\beta}_0, \hat{\beta}_1)$ and $s^2$ leads to

$$E(\hat{q}_{t,\epsilon}) = x(t)^T \beta + \delta \sigma C_\nu$$

$$\text{Var}(\hat{q}_{t,\epsilon}) = \sigma^2 \left[ x(t)^T (X^T X)^{-1} x(t) + \delta^2 (1 - C_\nu^2) \right],$$

where $\nu = N - 2$ in the linear case,

$$C_\nu = \sqrt{\frac{2}{\nu}} \frac{\Gamma \left( \frac{\nu + 1}{2} \right)}{\Gamma \left( \frac{\nu}{2} \right)} \approx 1 - \frac{1}{4\nu},$$

and we assume that $\hat{q}_{t,\epsilon}$ is normally distributed with this mean and variance.
The performance of the normal approximation to \( \hat{q}_{t,\epsilon} \) relies on (a) the choice of \( \delta \), and (b) the closeness to normality of \( s \). By the Central Limit Theorem it is reasonable to expect the distribution of \( \hat{q}_{t,\epsilon} \) to be approximately normal as the sample size \( N \to \infty \).

Possible choices for \( \delta \), and their resulting values, are

- \( \delta = z_\epsilon \) \hspace{1cm} (6)
- \( \delta s \) is the maximum likelihood estimate for \( z_\epsilon \sigma \), i.e. \( \delta = z_\epsilon \sqrt{\frac{\nu}{N}} \) \hspace{1cm} (7)
- \( \delta s \) is the unbiased estimate for \( z_\epsilon \sigma \), i.e. \( \delta = \frac{z_\epsilon}{C_\nu} \) \hspace{1cm} (8)

An estimate of the variance of \( \hat{q}_{t,\epsilon} \) can be obtained by either substituting the least squares estimate for \( \sigma \), if using (6) or (8), or the maximum likelihood estimate for \( \sigma \), if using (7). The estimated shelf-life is then given by \( \hat{\tau}_\epsilon = \inf \{ t : \hat{q}_{t,\epsilon} \leq \kappa \} \).

The label shelf-life is the 100(1 - \( \alpha \))% lower confidence bound for the time at which \( q_{t,\epsilon} \) intersects the specification limit, that is,

\[
\hat{\tau}_{\epsilon,L(\alpha)} = \inf \left\{ t : \left( x(t)^T \hat{\beta} + \delta s \right) + q \hat{\sigma} \sqrt{x(t)^T (X^T X)^{-1} x(t)} + \delta^2 (1 - C_\nu^2) \leq \kappa \right\},
\]

where \( \hat{\sigma} = s \) for (6) and (8), and \( \hat{\sigma} = s \sqrt{\frac{\nu}{N}} \) for (7). In subsequent simulations we will try both \( z_\alpha \) and \( t_{\alpha,\nu} \) as possible values for \( q \).

### 3.2 Various Profile Likelihood Approaches

The label shelf-life \( \hat{\tau}_{\epsilon,L(\alpha)} \) is the smallest time for which the null hypothesis

\[
H_0 : \beta_0 + \beta_1 \tau + z_\epsilon \sigma = \kappa
\]

is not rejected on a one-sided 100\( \alpha \)% level test. Since \( \kappa \) and \( z_\epsilon \) are known, we can reformulate the likelihood as a function of the parameters \( \{ \beta_1, \sigma^2, \tau \} \), by substituting
\[ \beta_0 = \kappa - \beta_1 \tau - z_\sigma \sigma. \]
Hypotheses about \( \tau \) can then be tested using the profile likelihood for \( \tau \), obtained by maximizing the likelihood over \( \{ \beta_1, \sigma^2 \} \) for each value of \( \tau \).

Two adjustments to the profile likelihood, based on [7] and [8], are also investigated.

### 3.2.1 The Basic Profile Likelihood Approach

After substituting for \( \beta_0 \), the log-likelihood based on (1) is given by

\[ L(\beta_1, \sigma^2; \tau; y) = -\frac{N}{2} \log \sigma^2 \sigma^2 + \frac{1}{2\sigma^2} \sum_{i=1}^{N} (y_i^* - \beta_1 x_i^* + z_i \sigma)^2, \tag{9} \]

where \( y_i^* = y_i - \kappa \) and \( x_i^* = x_i - \tau \) for notational simplicity, and \( \beta_1, \tau \) and \( \sigma^2 \) denote the slope and variance for given \( \tau \). The maximum likelihood estimates of \( \beta_1, \tau \) and \( \sigma \) are

\[ \hat{\beta}_1, \tau = \hat{\beta}_1 + \frac{\bar{x}^*}{s_{xx} V} (\bar{y}^* - \hat{\beta}_1 \bar{x}^* + z_\epsilon \hat{\sigma}) \tag{10} \]
\[ \hat{\sigma}_\tau = \sqrt{\hat{\sigma}^2 + \frac{(\bar{y}^* - \hat{\beta}_1 \bar{x}^*)^2}{4N^2V^2} (z_\epsilon^2 + 4NV) + \frac{z_\epsilon(\bar{y}^* - \hat{\beta}_1 \bar{x}^*)}{2NV}}, \tag{11} \]

where \( \hat{\beta}_1 \) and \( \hat{\sigma}^2 \) are the unconstrained maximum likelihood estimates of \( \beta_1 \) and \( \sigma^2 \), respectively, \( \bar{x}^* = \bar{x} - \bar{t}, \bar{y}^* = \bar{y} - \kappa, s_{xx} = \sum_{i=1}^{N} (x_i - \bar{x})^2 \) and \( V = \frac{1}{N} + \frac{(\bar{x}^*)^2}{s_{xx}} \). Details of the derivation can be found in [9].

The maximum likelihood estimate of the shelf-life under this approach is given by \( \hat{\tau} = \inf \{ t : x(t)^T \hat{\beta} + z_\epsilon \hat{\sigma} \leq \kappa \} \), where \( \hat{\beta} \) and \( \hat{\sigma} \) are the unconstrained maximum likelihood estimates of \( \beta \) and \( \sigma \), respectively.

Let \( L(\hat{\beta}_1, \hat{\sigma}^2, \hat{\tau}; y) \) be the likelihood evaluated at the maximum likelihood estimates and let \( L(\hat{\beta}_1, \hat{\sigma}^2, \tau; y) \) be the likelihood evaluated, for a given value of \( \tau \), at the constrained maximum likelihood estimates. Then, asymptotically, the general-
ized likelihood ratio statistic is given by

$$w(\tau) = 2 \left( L(\hat{\beta}_1, \hat{\sigma}^2, \hat{\tau}; y) - L(\hat{\beta}_{1,\tau}, \hat{\sigma}^2_{\tau}, \tau; y) \right), \quad (12)$$

which is distributed approximately $\chi^2_1$. Consequently, the label shelf-life, based on the $100(1 - \alpha)\%$ lower confidence bound, is given by

$$\hat{\tau}_{\alpha, L(\alpha)} = \inf_t \{ t : (t \leq \hat{\tau}_e) \& (w(t) \leq \chi^2_{1,1-2\alpha}) \}. \quad (13)$$

### 3.2.2 The Modified Profile Likelihood Approach

We have a general likelihood problem in which we wish to estimate a scalar parameter $\tau$ in the presence of nuisance parameter vector $(\beta_1, \sigma^2)$. A modified profile likelihood, ([7]), appropriate for estimation of $\tau$ is given by

$$L_m(\tau; y) = L(\hat{\beta}_{1,\tau}, \hat{\sigma}^2_{\tau}, \tau; y) - \frac{1}{2} \log \left| I(\hat{\beta}_{1,\tau}, \hat{\sigma}^2_{\tau}) \right| - \log \left| \frac{\partial(\hat{\beta}_{1,\tau}, \hat{\sigma}^2_{\tau})}{\partial(\hat{\beta}_1, \hat{\sigma}^2)} \right|, \quad (14)$$

where $I(\cdot)$ denotes the observed or expected information matrix, based on (9). The elements of the the Jacobian, the last term in (14), are obtained by differentiating (10) and (11) with respect to the maximum likelihood estimates $\hat{\beta}_1$ and $\hat{\sigma}^2$.

The label shelf-life, based on the $100(1 - \alpha)\%$ lower confidence bound, is obtained from (12) and (13) by replacing $L(\cdot)$ by $L_m(\cdot)$, where $\hat{\tau}_e$ denotes the value of $\tau$ that maximizes $L_m(\cdot)$.

In general, there may not be explicit solutions of the constrained parameters in terms of the maximum likelihood parameters. Consequently, it may not be feasible to calculate the Jacobian. Barndorff-Nielsen ([7]) suggests that the Jacobian term be ignored in such instances. A potential alternative is given in the following section.
3.2.3 The Conditional Profile Likelihood Approach

The conditional profile likelihood ([8]) is

\[ L_c(\tau; y) = L(\tau, \hat{\lambda}_c; y) - \frac{1}{2} \log |I(\hat{\lambda}_c)|, \]  

(15)

where \( \lambda \) and \( \tau \) are orthogonal parameters. This conditional likelihood is very similar to the modified profile likelihood, except that the term involving the Jacobian is ignored, since it is of order \( 1/N \) for orthogonal parameters. The conditional profile likelihood is not invariant under transformations of \( \lambda \), but it has been argued ([8]) that the orthogonal parameterization reduces this lack of invariance.

In our case, one choice of the parameter vector \( \lambda^T = (\lambda_1, \lambda_2) \) is

\[ \lambda_1 = \beta_{1, \tau} \sqrt{g(\tau)} \quad \text{and} \quad \lambda_2 = \sqrt{2N \frac{\sigma(2 + z_{1}^2) - z_{1} \beta_{1, \tau} \bar{x}^*}{z_{1} \beta_{1, \tau} \sqrt{g(\tau)}}}, \]

where \( g(\tau) = s_{xx}(2 + z_{1}^2) + 2N \bar{x}^* \).

The elements of the observed or expected information matrix \( I(\hat{\lambda}_c) \) are found by differentiating \( L(\cdot) \) with respect to \( \lambda \).

Again, the label shelf-life, based on the 100(1 - \( \alpha \))% lower confidence bound, is obtained from (12) and (13), where \( L(\cdot) \) is replaced by \( L_c(\cdot) \), and \( \tau_c \) denotes the value of \( \tau \) that maximizes \( L_c(\cdot) \).

Figure 2 shows an example of the three constrained profile likelihood approaches detailed above. The graph also shows the 95% lower confidence bounds for the true shelf-life for each approach based on \( N = 8 \). From the plot it can be seen that the modified and conditional profile likelihood approaches give similar profile likelihoods. This suggests that the constrained parameters \( \beta_{1, \tau} \) and \( \sigma_{\tau}^2 \) are almost orthogonal, and hence the determinant of the Jacobian in (14) is close to 1 (which
has been verified numerically). Both approaches also result in lower label shelf-lives, which is reflected in the coverage probabilities discussed in Section 4.

[Figure 2 near here]

3.3 Noncentral t-distribution Approach

As the distribution of the estimated mean response $x(t)^T \hat{\beta}$ is normal, it can be shown that

$$
\frac{x(t)^T \hat{\beta} - q_{t,c}}{\sqrt{x(t)^T (X^TX)^{-1}x(t)}} \sim t'_{N-2}(\delta),
$$

where $t'_{N-2}(\delta)$ denotes a noncentral t-distribution with $N - 2$ degrees of freedom and noncentrality parameter $\delta = -z_c / \sqrt{x(t)^T (X^TX)^{-1}x(t)}$.

The label shelf-life is then given by the intersection of the exact $100(1 - \alpha)$% lower confidence bound and the specification limit, that is,

$$
\hat{t}_{\alpha,L(\alpha)} = \inf \left\{ t : x(t)^T \hat{\beta} + q \sqrt{x(t)^T (X^TX)^{-1}x(t)} \leq \kappa \right\},
$$

where $q = t'_{\alpha,N-2}(-\delta)$.

4 Simulation Results

Since real pharmaceutical stability data is generally commercially sensitive, we have evaluated the approaches detailed in Section 3 with the help of a simulation study.

The structure of the simulations was that a random sample of size $r$ was drawn at each of $n$ times $x_i = 0, 3, 6, 9, 12, 18, 24$, and 36 months from $N(x_i^T \beta, \sigma^2)$, where $x_i^T = (1, x_i)$, $\beta^T = (100, -0.4)$ and $\sigma^2 = 1$. This can be done without loss of generality as the approaches, except the conditional profile likelihood approach, are invariant under linear transformations. The total number of observations available
for analysis is $N = nr = 8r$. Values of $r=1, 2, 5, 10, 20,$ and $50$ were used. For each value of $r$, 1000 simulations were performed and the proportion of label shelf-lives covering the true shelf-life was determined for each method.

Assuming that $\kappa = 90$ and $\epsilon = 0.01$, the true shelf-life is equal to $\tau_{0.01} = 20.89$ months. For each simulation the label shelf-life $\hat{\tau}_{0.01, z(0.05)}$, based on $\alpha = 0.05$, was calculated and compared to $\tau_{0.01}$. For each approach the proportion of label shelf-lives that were less than the true shelf-life is tabulated in Table 1.

For each simulation the label shelf-life can be either greater than the true shelf-life (failure), or less than or equal to the true shelf-life (success). Consequently, the number of successes are distributed as a binomial distribution $B(1000, 1 - \alpha)$. Using the normal approximation to the binomial, an approximate confidence interval for the true coverage probability $p$ is given by $\hat{p} \pm 0.0135$, where $\hat{p}$ is given in the cells of Table 1.

[Table 1 near here]

The approaches based on the normal approximation under-performed with respect to coverage, except for very large $N$. Of the three choices for $\delta$, the unbiased choice gave the best results, followed closely by the choice $\delta = z_{\epsilon}$. The maximum likelihood choice gave the worst coverage probabilities, performing particularly badly for very small $N$. The choice of $q = t_{\alpha, N-2}$ consistently gave better results than $q = z_{\alpha}$, with the difference being negligible for large $N$ as expected.

The basic constrained profile likelihood approach performed badly for small $N$ and reasonably well for large $N$. This is probably because the profile likelihood is more like a quadratic for large $N$ (see Figure 2) and hence the $\chi^2$ approximation is

13
The modified profile likelihood approach, using the observed information matrix, performed well for all \( N \). However, the coverage probabilities were conservative, especially for small \( N \).

The conditional profile likelihood, using the observed information matrix, resulted in coverage probabilities that were not significantly different from 0.95 for values of \( N \geq 16 \). This approach gave coverage probabilities very close to those of the noncentral t-distribution approach.

The noncentral t-distribution approach performed very well for \( N \geq 16 \). This is hardly surprising since the confidence bounds are based on exact results.

Further simulations have been performed with values of \( \beta_1 = -0.4, -0.3 \) and \(-0.2 \) and \( \sigma = 0.5, 1 \) and \( 2 \) (the details can be found in an unpublished Ph.D. thesis by the first author). In general, the noncentral t-distribution approach resulted in coverage probabilities that were satisfactorily close to 0.95 for all values of \( N \).

5 Conclusions

We have proposed a new definition for the true shelf-life of a batch which is based on the \( \epsilon \) quantile of the distribution of the responses. We recommend that this new definition be adopted by regulatory authorities.

Irrespective of the approach used, we advocate that pharmaceutical companies use as large a number of assays as possible, because of the problems often experienced with small samples. In particular, we recommend that more time points be used during the second and consequent years of stability testing, which will reduce the influence that individual time points currently have on the estimation of the pa-
rameters. This will also provide more information near the true shelf-life, resulting in more accurate estimation. In particular, we suggest that products are tested at least every three months throughout the shelf-life period. We also suggest that at least two units are assayed at each time point and that each assay is repeated at least twice. This will enable estimation of the assay-to-assay variability as well as the unit-to-unit variability, allowing for more accurate shelf-life estimation.

Various approaches for the estimation of a lower confidence bound for this new definition have been described and evaluated with the help of a simulation study. The approach based on the noncentral t-distribution should be used as it is based on exact inferences and gives good coverage probabilities of the label shelf-life. However, the extensions to the multiple batch case which are discussed below are not amenable to an exact approach, hence the examination of alternatives.

The simulation results suggest that treating the shelf-life as approximately normal does not work well for small $N$. However, the assumption of approximate normality may offer a workable approach for very large $N$.

The basic profile likelihood approach does not work as well as the modified or the conditional profile likelihood approach, both of which work about as well as the exact approach. However, the modified profile likelihood approach tends to be more conservative.

In this paper the variation described is variation between units of a drug product, with assay variation assumed negligible. Where assay variation is not negligible, the $\epsilon$ quantile can only be determined by estimating the $\sigma$ related to unit-to-unit variation. As a referee has pointed out, variability between assays performed at
different times may be larger than the variability between assays performed at the same time. This source of variation has also been considered to be negligible. The inclusion of these sources of variation will be investigated in future research.

This work has not addressed the situation when a number of batches are used in the estimation process. The inclusion of batch-to-batch variability involves random effects in the model. It is no longer possible to use the noncentral t-distribution and one of the profile likelihood approaches may offer a suitable alternative. Research in this important area is currently being undertaken.

The assumption of linear degradation of the product may not hold in practice. While the current guidelines suggest that the observations may be analyzed on a transformed scale, in particular the log scale, there is no guarantee that the variance will be constant or that the error terms will be normally distributed. Further work is underway to investigate how the approaches detailed here perform when the degradation is not linear or when the variation requires generalized linear model assumptions.

References


<table>
<thead>
<tr>
<th>N</th>
<th>8</th>
<th>16</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>(5) using (6) and $q = z_\alpha$</td>
<td>0.842</td>
<td>0.899</td>
<td>0.910</td>
<td>0.931</td>
<td>0.935</td>
<td>0.938</td>
</tr>
<tr>
<td></td>
<td>$q = t_{\alpha,N-2}$</td>
<td>0.876</td>
<td>0.911</td>
<td>0.917</td>
<td>0.936</td>
<td>0.937</td>
</tr>
<tr>
<td>(5) using (7) and $q = z_\alpha$</td>
<td>0.713</td>
<td>0.817</td>
<td>0.874</td>
<td>0.918</td>
<td>0.921</td>
<td>0.931</td>
</tr>
<tr>
<td></td>
<td>$q = t_{\alpha,N-2}$</td>
<td>0.764</td>
<td>0.843</td>
<td>0.882</td>
<td>0.919</td>
<td>0.921</td>
</tr>
<tr>
<td>(5) using (8) and $q = z_\alpha$</td>
<td>0.864</td>
<td>0.910</td>
<td>0.918</td>
<td>0.936</td>
<td>0.939</td>
<td>0.938</td>
</tr>
<tr>
<td></td>
<td>$q = t_{\alpha,N-2}$</td>
<td>0.890</td>
<td>0.916</td>
<td>0.922</td>
<td>0.939</td>
<td>0.940</td>
</tr>
<tr>
<td>Basic PL</td>
<td>0.811</td>
<td>0.891</td>
<td>0.910</td>
<td>0.931</td>
<td>0.936</td>
<td>0.938</td>
</tr>
<tr>
<td>Mod. PL</td>
<td>0.974</td>
<td>0.974</td>
<td>0.954</td>
<td>0.961</td>
<td>0.957</td>
<td>0.950</td>
</tr>
<tr>
<td>Cond. PL</td>
<td>0.922</td>
<td>0.963</td>
<td>0.944</td>
<td>0.950</td>
<td>0.954</td>
<td>0.945</td>
</tr>
<tr>
<td>Non-central t-distribution</td>
<td>0.932</td>
<td>0.964</td>
<td>0.944</td>
<td>0.950</td>
<td>0.954</td>
<td>0.945</td>
</tr>
</tbody>
</table>

Table 1: Results from a simulation study (1000 iterations) of the different approaches to estimate the labeled shelf-life, based on the 95% lower confidence bound and a true shelf-life using $\epsilon = 0.01$. The numbers in the body of the table give the proportion of label shelf-lives that were smaller than the true shelf-life.
Figure 1: Estimating the shelf-life using the proposed definition with $\epsilon = 0.05$ and $\kappa = 90$. 
Figure 2: Example profile likelihood plots of the basic, modified and conditional profile likelihoods, with estimated label shelf-lives, based on the 95% lower confidence bound for the true shelf-life $\tau_{0.01}$, based on a sample of size $N = 8$ and the same $\beta_0$, $\beta_1$ and $\sigma$ as are used in the simulation study described in Section 4.