Chapter 3

PARAMETRIC MODEL FOR A SINGLE SAMPLE FROM A HOMOGENEOUS POPULATION

3.1 Introduction

Under a univariate model, a distribution is fitted to the lifetimes without using any covariates. The model must describe the basic underlying distribution of lifetimes.

Let T be a non-negative continuous random variable representing lifetime from a homogeneous population. $Y = \ln(T)$ is used to represent the log-lifetime.

The standard way of fitting parametric models to an observed set of survival data is to use the **method of maximum likelihood** (refer to [5, page 319-322]).

A new method of fitting parametric models to an observed set of survival data will be introduced in this chapter and is called **maximum likelihood estimation subject to constraints**.

3.2 Standard Method of Maximum Likelihood Estimation

3.2.1 Introduction

In the univariate case, a log-linear model (a linear model in log-lifetime) could be fitted to a survival data set. This model is of the form

$$\boxed{Y = \ln T = \mu + \sigma W}$$

where W is the error distribution, μ is the location parameter and σ is the scale parameter.

The standard way of fitting such a model to an observed set of survival data is to use the method of maximum likelihood.

3.2.2 Likelihood function for the linear model in log-time

The likelihood function for this linear model in log-time may be derived as follows.

Consider the n pairs (y_i, δ_i) $\quad i = 1, 2, ..., n$ in the data set with $y_i = ln(t_i)$.

The basic form of the likelihood function for **random right-censored continuous data** is, from Equation 2.6 .6, equal to

$$\begin{aligned} L(\mu,\sigma) &= \prod_{i=1}^{n} [f_Y(y_i)]^{\delta_i} \cdot [S_Y(y_i)]^{1-\delta_i} \\ &= \prod_{i=1}^{n} \left[f_W(\frac{y_i-\mu}{\sigma}) \cdot \frac{1}{\sigma} \right]^{\delta_i} \cdot \left[S_W(\frac{y_i-\mu}{\sigma}) \right]^{1-\delta_i} \qquad (3.2\ .1) \end{aligned}$$

The log-likelihood function for random right-censored continuous data is then

$$lnL(\mu,\sigma) = \sum \delta_i \cdot ln \left[f_W(\frac{y_i-\mu}{\sigma}) \cdot \frac{1}{\sigma} \right] + \sum (1-\delta_i) \cdot ln \left[S_W(\frac{y_i-\mu}{\sigma}) \right] \qquad (3.2\ .2)$$

with

the first sum over observed lifetimes (uncensored observations)
the second sum over right-censored observations.

The basic form of the likelihood function for **interval-censored data** follows from Equation 2.6 .9 as

$$L(\mu,\sigma) = \prod_{i=1}^{n} [f_Y(y_i)]^{\delta_i} \cdot [S_Y(y_i)]^{1-\delta_i} \cdot [1 - S_Y(y_i)]^{\delta_i} \cdot [S_Y(b_i) - S_Y(y_i)]^{1-\delta_i}$$

$$= \prod_{i=1}^{n} \left[f_W(\frac{y_i - \mu}{\sigma}) \cdot \frac{1}{\sigma} \right]^{\delta_i} \cdot \left[S_W(\frac{y_i - \mu}{\sigma}) \right]^{1-\delta_i} \cdot$$
$$\left[1 - S_W(\frac{y_i - \mu}{\sigma}) \right]^{\delta_i} \cdot \left[S_W(\frac{b_i - \mu}{\sigma}) - S_W(\frac{y_i - \mu}{\sigma}) \right]^{1-\delta_i} \quad (3.2\ .3)$$

with b_i the lower end of a censoring interval.

The log-likelihood function for interval-censored data is

$$\begin{aligned} lnL(\mu, \sigma) &= \sum \delta_i \cdot ln \left[f_W(\frac{y_i - \mu}{\sigma}) \cdot \frac{1}{\sigma} \right] + \sum (1 - \delta_i) \cdot ln \left[S_W(\frac{y_i - \mu}{\sigma}) \right] + \\ & \sum (\delta_i) \cdot ln \left[1 - S_W(\frac{y_i - \mu}{\sigma}) \right] + \sum (1 - \delta_i) \cdot ln \left[S_W(\frac{b_i - \mu}{\sigma}) - S_W(\frac{y_i - \mu}{\sigma}) \right] \end{aligned} \quad (3.2\ .4)$$

with

the first sum over observed lifetimes (uncensored observations)
the second sum over right-censored observations
the third sum over left-censored observations
the fourth sum over interval-censored observations.

3.2.3 Maximum likelihood estimators of the log-linear parameters

[5] shows how maximum likelihood estimators of the log-linear parameters μ and σ associated with

- the extreme value distribution, the error distribution for the Weibull model
- the logistic distribution, the error distribution for the log-logistic model
- the normal distribution, the error distribution for the lognormal model

can be found numerically by the Newton-Raphson procedure (refer to [21]). When the iterative procedure has converged, the variance-covariance matrix of the log-linear parameter estimates can be approximated by the inverse of the information matrix, evaluated at the parameter estimates. The square roots of the diagonal elements of this matrix are then the standard errors of the estimated values of the log-linear parameters.

[4] shows how the LIFEREG procedure of the SAS statistical package computes these maximum likelihood estimators of the log-linear parameters μ and σ and explains how the SAS

output must be interpreted. SAS allows for right-, left- and interval-censored data. The SAS programs appear in Appendix A. The variance-covariance matrix of the log-linear parameters μ and σ, obtained from the observed information matrix, are also available in the SAS package.

The invariance property of the maximum likelihood estimator provides that the maximum likelihood estimators of λ and α at the Weibull and log-logistic are then given by

$$\hat{\lambda} = \exp\left\{\frac{-\hat{\mu}}{\hat{\sigma}}\right\} \quad \text{and} \quad \hat{\alpha} = \frac{1}{\hat{\sigma}} \tag{3.2 .5}$$

An application of this standard technique to a reallife insurance company data set is done in chapter 5.

Applying the method of statistical differentials, also called delta method ([13, page 69-72]), leads to formulaes for the standard errors of the estimates and the covariance between the two estimates.

$$var(\hat{\lambda}) = \exp\left(\frac{-2\hat{\mu}}{\hat{\sigma}}\right) \cdot \left[\frac{var(\hat{\mu})}{\hat{\sigma}^2} + \hat{\mu}^2\frac{var(\hat{\sigma})}{\hat{\sigma}^4} - 2\hat{\mu}\frac{cov(\hat{\mu},\hat{\sigma})}{\hat{\sigma}^3}\right] \tag{3.2 .6}$$

$$var(\hat{\alpha}) = \frac{var(\hat{\sigma})}{\hat{\sigma}^4} \tag{3.2 .7}$$

$$cov(\hat{\lambda},\hat{\alpha}) = \exp\left(\frac{-\hat{\mu}}{\hat{\sigma}}\right) \cdot \left[\frac{cov(\hat{\mu},\hat{\sigma})}{\hat{\sigma}^3} - \hat{\mu}\frac{var(\hat{\sigma})}{\hat{\sigma}^4}\right] \tag{3.2 .8}$$

Once maximum likelihood estimates of the parameters μ and σ, or equivalently, λ and α are computed, estimates of the survivor function and the hazard function are available for the distribution of T (or $Y = \ln(T)$), that is the Weibull (or extreme value), log-logistic (or logistic) and lognormal (or normal).

3.3 MLE subject to Constraints - A Fixed Censoring Time

3.3.1 Introduction

Proposition 1, which is proved in [11], provides a method of finding the ML estimate for the mean vector of the exponential family, subject to certain constraints on the mean vector. From the estimate of the mean vector the estimates of the parameters in the model are computed.

Models can be easily formulated in terms of the implied constraints, which may be linear or non-linear in $\boldsymbol{\mu}$.

Proposition 1

Consider a random vector $\boldsymbol{y}$, with probability function belonging to the exponential family. Let $\boldsymbol{g}(\boldsymbol{\mu})$ be a continuous vector valued function of $\boldsymbol{\mu}$, for which the first order partial derivatives exist.
Let $\boldsymbol{G}_\mu = \dfrac{\partial \boldsymbol{g}(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}$ be the derivative of $\boldsymbol{g}(\boldsymbol{\mu})$ with respect to $\boldsymbol{\mu}$ and $\boldsymbol{G}_y = \left.\dfrac{\partial \boldsymbol{g}(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}\right|_{\boldsymbol{\mu}=\boldsymbol{y}}$.
The ML estimate of $\boldsymbol{\mu}$ subject to the constraints $\boldsymbol{g}(\boldsymbol{\mu}) = \boldsymbol{0}$, is given by

$$\widehat{\boldsymbol{\mu}}_c = \boldsymbol{y} - (\boldsymbol{G}_\mu \boldsymbol{V}_\mu)'(\boldsymbol{G}_y \boldsymbol{V}_\mu \boldsymbol{G}_\mu')^* \boldsymbol{g}(\boldsymbol{y}) + o(||\boldsymbol{y} - \boldsymbol{\mu}||) \qquad (3.3\ .1)$$

This result implies that the MLE of $\boldsymbol{\mu}$ must be obtained iteratively.

The variance-covariance matrix $\boldsymbol{V}$ could be known, or it could be some function of $\boldsymbol{\mu}$, say $\boldsymbol{V}_\mu$. The iterative use of the estimation procedure thus depends on the form of $\boldsymbol{G}_\mu$ and $\boldsymbol{V}_\mu$.

The matrix $\boldsymbol{G}_y \boldsymbol{V}_\mu \boldsymbol{G}_\mu'$ in Equation 3.3 .1 should be non-singular, and therefore the inverse, denoted by $*$, is any generalized inverse (refer to [31, page 123]).

An expression for the asymptotic variance-covariance matrix of the estimator $\widehat{\boldsymbol{\mu}}_c$ is given in Proposition 2.

Proposition 2

The asymptotic variance-covariance matrix of $\widehat{\boldsymbol{\mu}}_c$ is given by

$$cov\,(\widehat{\boldsymbol{\mu}}_c) = \boldsymbol{V}_\mu - (\boldsymbol{G}_\mu \boldsymbol{V}_\mu)'(\boldsymbol{G}_\mu \boldsymbol{V}_\mu \boldsymbol{G}_\mu')^* \boldsymbol{G}_\mu \boldsymbol{V}_\mu \qquad (3.3\ .2)$$

In [10] these propositions are applied to provide a method for fitting certain continuous probability distributions to an observed frequency distribution. The method requires that some function of the cumulative distribution function must be written as a **linear model**. The estimation algorithm described in [11] is applied to find the maximum likelihood estimates of the parameters in this linear model. From these estimates, the estimates of the parameters of the distribution can be found. This fitting method will be described in the notation of the survival analysis problem under consideration, regarding the lapses of policies.

3.3.2 Notation for a fixed censoring time

Consider the simple experimental design, as described in [38], where all policies enter at the same time with C the pre-assigned fixed censoring time. Instead of observing $X_1, X_2, ..., X_n$ only $T_1, T_2, ..., T_n$ are observed where

$$T_j = \begin{cases} X_j & \text{if} \quad X_j \leq C \\ C & \text{if} \quad X_j > C. \end{cases}$$

The survival data, based on a sample of size n, can then be represented by pairs of random variables (T_j, δ_j) where $T_1, T_2, ..., T_n$ are independent identically distributed random variables, each with distribution function F and density function f. δ_j is the survival status of the j^{th} policy and indicates whether the lifetime for the j^{th} policy corresponds to a lapse $(\delta_j = 1)$ or is censored $(\delta_j = 0)$.

A frequency distribution is formed when the observed values of the random variables $T_1, T_2, ..., T_n$ are grouped into k adjacent, non-overlapping fixed lifetime intervals $[x_{j-1}; x_j) \quad j = 1, 2, ..., k$ with $x_0 = 0$, $x_{k-1} = C$ and $x_k = \infty$, as shown in Table 3.1.

Table 3.1: **Relative frequency distribution of survival data - fixed censoring time**

Interval number	Lifetime Intervals	Frequency Vector $\boldsymbol{f}$	Relative Frequency Vector $\boldsymbol{p}$	Probability Vector $\boldsymbol{\pi}$	Vector of Upper Class Boundaries $\boldsymbol{x}$
first	$[0, x_1)$	f_1	p_1	π_1	x_1
second	$[x_1, x_2)$	f_2	p_2	π_2	x_2
third	$[x_2, x_3)$	f_3	p_3	π_3	x_3
...	...	...	...	...	...
...	...	...	...	...	...
$(k-1)^{th}$	$[x_{k-2}, x_{k-1})$	f_{k-1}	p_{k-1}	π_{k-1}	x_{k-1}
k^{th}	$[x_{k-1}, \infty)$	f_k	p_k	π_k	

In Table 3.1, the last interval in the second column is an open interval containing all the censored lifetimes. The $x_j's$, $j = 1, 2, \ldots, k-1$ represent the upper class boundaries and f_j denotes the observed frequency for the j^{th} lifetime interval with n the total number of observations $\quad j = 1, 2, \ldots, k$.

Define

$$\begin{aligned} \boldsymbol{x} &= (x_1, x_2, \ldots, x_{k-1})' \quad \text{as the} \quad (k-1) \times 1 \quad \text{vector of upper class boundaries,} \\ \boldsymbol{f} &= (f_1, f_2, \ldots, f_k)' \quad \text{as the} \quad k \times 1 \quad \text{frequency vector and} \\ \boldsymbol{p} &= \frac{\boldsymbol{f}}{n} \quad \text{as the} \quad k \times 1 \quad \text{relative frequency vector.} \end{aligned}$$

$\boldsymbol{f}$ is a discrete random vector with a multinomial$(n, \boldsymbol{\pi})$ distribution, where $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_k)'$ and π_i is the probability that an observed lifetime falls in the i^{th} lifetime interval.

The relative frequency vector $\boldsymbol{p}$ is an observed probability vector from a multinomial population with $n\boldsymbol{p} = \boldsymbol{f}$ being multinomial$(n, \boldsymbol{\pi})$ distributed.

$$E(\boldsymbol{p}) = \boldsymbol{\pi} \tag{3.3 .3}$$

$$Cov(\boldsymbol{p}) = \boldsymbol{V} = \frac{1}{n}\left[diag(\boldsymbol{\pi}) - \boldsymbol{\pi}\boldsymbol{\pi}'\right] \tag{3.3 .4}$$

Note that $\boldsymbol{p}$ is the MLE of $\boldsymbol{\pi}$ in the case of no constraints.

The MLE of $\boldsymbol{\pi}$ should be determined in terms of constraints imposed by the survival distribution to be fitted.

Note that $\boldsymbol{f}$ is a discrete random vector with a multinomial$(n, \boldsymbol{\pi})$ distribution and that the multinomial distribution is a member of the exponential family. Therefore Equation 3.3 .1 of Proposition 1 in [11] can be reformulated in terms of the survival analysis problem under consideration.

The MLE of $\boldsymbol{\pi}$ subject to the constraints $\boldsymbol{g}(\boldsymbol{\pi}) = \boldsymbol{0}$ is

$$\widehat{\boldsymbol{\pi}}_c = \boldsymbol{p} - (\boldsymbol{G}_\pi \boldsymbol{V}_\pi)' \left(\boldsymbol{G}_p \boldsymbol{V}_\pi \boldsymbol{G}_\pi'\right)^* \boldsymbol{g}(\boldsymbol{p}) \tag{3.3 .5}$$

with

$$\boldsymbol{G}_\pi = \frac{\partial \boldsymbol{g}(\boldsymbol{\pi})}{\partial \boldsymbol{\pi}} \quad \text{and} \quad \boldsymbol{G}_p = \left.\frac{\partial \boldsymbol{g}(\boldsymbol{\pi})}{\partial \boldsymbol{\pi}}\right|_{\boldsymbol{\pi} = \boldsymbol{p}} \tag{3.3 .6}$$

This result implies that the MLE of $\boldsymbol{\pi}$ must be obtained iteratively by means of double iterations. The variance-covariance matrix $\boldsymbol{V}_\pi$ to be used is the estimated variance-covariance matrix of the multinomial distribution as stated in Equation 3.3 .4.

A double iteration takes place over $\boldsymbol{p}$ and $\boldsymbol{\pi}$. For every value of $\boldsymbol{\pi}$ the iteration is performed over $\boldsymbol{p}$ to obtain a new estimate for $\boldsymbol{\pi}$.

The observed relative frequency vector $\boldsymbol{p} = \boldsymbol{p}_0$ is used as an initial estimate for $\boldsymbol{\pi}$ and $\boldsymbol{p}$. In the first iteration over $\boldsymbol{p}$, the $\boldsymbol{p}$ in Equation 3.3 .5 is replaced by this initial estimate, while the $\boldsymbol{V}_\pi$ in Equation 3.3 .5 is estimated by $\widehat{\boldsymbol{V}}_{p_0} = \frac{1}{n}\left[diag(\boldsymbol{p}_0) - \boldsymbol{p}_0\boldsymbol{p}_0'\right]$ and the $\boldsymbol{G}_\pi$ in Equation 3.3 .5 is replaced by $\boldsymbol{G}_{p_0} = \left.\frac{\partial \boldsymbol{g}(\boldsymbol{\pi})}{\partial \boldsymbol{\pi}}\right|_{\boldsymbol{\pi} = \boldsymbol{p}_0}$. This results in a new estimate for $\boldsymbol{\pi}$. In the second iteration over $\boldsymbol{p}$, only $\boldsymbol{p}$ in Equation 3.3 .5 is replaced to obtain the second estimate for $\boldsymbol{\pi}$, while $\boldsymbol{V}_\pi$ and $\boldsymbol{G}_\pi$ are kept constant at $\widehat{\boldsymbol{V}}_{p_0}$ and $\boldsymbol{G}_{p_0}$, since iteration at this stage is over $\boldsymbol{p}$. This is repeated until convergence is attained over $\boldsymbol{p}$. The final estimate for $\boldsymbol{\pi}$ at convergence during this first stage of iteration over $\boldsymbol{p}$ then becomes the second estimate for $\boldsymbol{\pi}$ in $\boldsymbol{G}_\pi$ and $\boldsymbol{V}_\pi$. Once again iteration takes place over $\boldsymbol{p}$, again starting with the observed relative frequency vector $\boldsymbol{p} = \boldsymbol{p}_0$ as estimate for $\boldsymbol{\pi}$ and keeping $\boldsymbol{V}_\pi$ and $\boldsymbol{G}_\pi$ constant at the estimated value at convergence. Iteration over $\boldsymbol{p}$ gives the third estimate for $\boldsymbol{\pi}$ in $\boldsymbol{G}_\pi$ and $\boldsymbol{V}_\pi$ and once again iteration takes place over $\boldsymbol{p}$, again starting with the initial $\boldsymbol{p}_0$ vector as estimator for $\boldsymbol{\pi}$ and keeping $\boldsymbol{V}_\pi$ and $\boldsymbol{G}_\pi$ constant at the new estimated value at convergence. This procedure continues and convergence will be attained when the final estimate for $\boldsymbol{\pi}$ in the iteration over $\boldsymbol{p}$ corresponds with the final estimate of $\boldsymbol{\pi}$ in the iteration over $\boldsymbol{\pi}$. This value then will be the MLE for $\boldsymbol{\pi}_c$.

Define $\boldsymbol{\pi}_S$ as the cumulative sum vector of the π_j's. Hence

$$\boldsymbol{\pi}_S = \boldsymbol{S} \times \boldsymbol{\pi} = \begin{pmatrix} \pi_1 \\ \pi_1 + \pi_2 \\ \cdots \\ \pi_1 + \pi_2 + \cdots + \pi_{k-1} \end{pmatrix}$$

where $\boldsymbol{S}$ is a $(k-1) \times k$ matrix of the form

$$\boldsymbol{S} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 1 & \cdots & 0 & 0 \\ \cdots & & \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & 1 & \cdots & 1 & 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\pi} = \begin{pmatrix} \pi_1 \\ \pi_2 \\ \cdots \\ \pi_k \end{pmatrix}$$

Let

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{k-1} \end{pmatrix}$$

then the cumulative distribution function

$$F(\boldsymbol{x}) = \boldsymbol{S} \times \boldsymbol{\pi} = \boldsymbol{\pi}_S \tag{3.3 .7}$$

It follows that

$$\boldsymbol{x} = F^{-1}(\boldsymbol{\pi}_S) \tag{3.3 .8}$$

specifies the constraints on the elements of $\boldsymbol{\pi}_S$ and hence on $\boldsymbol{\pi}$.

By using Equation 3.3 .2 of Proposition 2, the asymptotic variance-covariance matrix of $\widehat{\boldsymbol{\pi}}_c$ is

$$cov\,(\widehat{\boldsymbol{\pi}}_c) = \boldsymbol{V} - (\boldsymbol{G}_\pi \boldsymbol{V})' \,(\boldsymbol{G}_\pi \boldsymbol{V} \boldsymbol{G}_\pi')^* \,\boldsymbol{G}_\pi \boldsymbol{V} \tag{3.3 .9}$$

Next it will be shown how the estimation procedure may be utilized to fit continuous survival distributions, such as the Weibull, log-logistic and lognormal to **grouped survival data**. For the Weibull and log-logistic survival distributions some function of the survival function $S(\boldsymbol{x}) = \boldsymbol{1} - F(\boldsymbol{x}) = \boldsymbol{1} - \boldsymbol{\pi}_S$ may be written in terms of a **linear model**, from which the parameters of the survival distribution may be estimated. For the lognormal survival distribution some function of the cumulative distribution function of the standard normal distribution may be expressed in terms of a **linear model**, from which the parameters of the lognormal distribution may be estimated.

The procedure to find the ML estimates of these three survival distributions can be easily implemented using a matrix algebra package, for example the SAS/IML procedure of the SAS System.

3.3.3 Fitting of a Weibull distribution to grouped survival data

From Equation 2.4 .8 follows that

$$\ln(-\ln S(t)) = \ln\lambda + \alpha \ln t \qquad (3.3\ .10)$$

where t denotes the **continuous** survival time.

In the current notation for **grouped survival data** in terms of the vector of upper class boundaries, Equation 3.3 .10 becomes

$$\ln(-\ln S(\boldsymbol{x})) = \ln\lambda \cdot \mathbf{1} + \alpha \cdot \ln \boldsymbol{x} \qquad (3.3\ .11)$$

or from $S(\boldsymbol{x}) = \mathbf{1} - F(\boldsymbol{x})$

$$\ln\{-\ln(\mathbf{1} - F(\boldsymbol{x}))\} = \ln\lambda \cdot \mathbf{1} + \alpha \cdot \ln \boldsymbol{x} \qquad (3.3\ .12)$$

or from Equation 3.3 .7

$$\begin{aligned}
\ln\{-\ln(\mathbf{1} - \boldsymbol{\pi}_S)\} &= \ln\lambda \cdot \mathbf{1} + \alpha \cdot \ln \boldsymbol{x} \qquad (3.3\ .13)\\
&= \ln\lambda \cdot \begin{pmatrix} 1 \\ 1 \\ \cdots \\ 1 \end{pmatrix} + \alpha \cdot \begin{pmatrix} \ln x_1 \\ \ln x_2 \\ \cdots \\ \ln x_{k-1} \end{pmatrix}\\
&= \begin{pmatrix} 1 & \ln x_1 \\ 1 & \ln x_2 \\ \cdots & \cdots \\ 1 & \ln x_{k-1} \end{pmatrix} \cdot \begin{pmatrix} \ln\lambda \\ \alpha \end{pmatrix}\\
&= \underbrace{(\mathbf{1}\ , \ln \boldsymbol{x})} \cdot \begin{pmatrix} \ln\lambda \\ \alpha \end{pmatrix}\\
\Rightarrow \ln\{-\ln(\mathbf{1} - \boldsymbol{\pi}_S)\} &= \quad \boldsymbol{X}_1 \quad \cdot \begin{pmatrix} \ln\lambda \\ \alpha \end{pmatrix} \qquad (3.3\ .14)
\end{aligned}$$

Equation 3.3 .14 is a **linear model** in the parameters $\ln\lambda$ and α. According to the general result for a linear model $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}$ where $E(\boldsymbol{y}) = \boldsymbol{X\beta}$ is equivalent to $\boldsymbol{C} \cdot E(\boldsymbol{y}) = \boldsymbol{0}$ with $\boldsymbol{C} = \boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$, Equation 3.3 .14 is equivalent to

$$\begin{aligned}\underbrace{\left(\boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}'\right)}_{} \cdot \ln\left(-\ln\left[\boldsymbol{1} - \boldsymbol{\pi}_S\right]\right) &= 0 \\ \underbrace{\boldsymbol{C} \qquad \cdot \ln\left(-\ln\left[\boldsymbol{1} - \boldsymbol{\pi}_S\right]\right)}_{} &= 0 \\ \boldsymbol{g}(\boldsymbol{\pi}) &= 0\end{aligned}$$

$\boldsymbol{C}$ is the projection matrix orthogonal to the columns of the design matrix $\boldsymbol{X}_1$. Note that $\boldsymbol{C}\boldsymbol{X}_1 = \boldsymbol{0}$.

The function $g(\boldsymbol{\pi}) = \boldsymbol{0}$ satisfies the conditions of Proposition 1 and the estimation algorithm in [11] can be used to estimate the parameters λ and α of the Weibull distribution.

To summarize, the constraints imposed by the Weibull distribution are specified by

$$g(\boldsymbol{\pi}) = \boldsymbol{C}.\ln\{-\ln(\boldsymbol{1} - \boldsymbol{\pi}_S)\} = \boldsymbol{C}.\ln\{-\ln(\boldsymbol{1} - \boldsymbol{S}\cdot\boldsymbol{\pi})\} = \boldsymbol{0} \tag{3.3 .15}$$

with

$$\boldsymbol{C} = \boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}' \text{ and } \boldsymbol{X}_1 = (\boldsymbol{1}, \ln\boldsymbol{x}) \quad . \tag{3.3 .16}$$

The derivative of $g(\boldsymbol{\pi})$ with respect to $\boldsymbol{\pi}$ is given by

$$\begin{aligned}\boldsymbol{G}_\pi &= \frac{\partial \boldsymbol{g}(\boldsymbol{\pi})}{\partial \boldsymbol{\pi}} \\ &= \boldsymbol{C} \cdot \left[diag\left(\frac{1}{-\ln(\boldsymbol{1} - \boldsymbol{S}\cdot\boldsymbol{\pi})}\right)\right] \cdot diag\left(\frac{1}{-(\boldsymbol{1} - \boldsymbol{S}\cdot\boldsymbol{\pi})}\right) \cdot (-\boldsymbol{S}) \\ &= -\boldsymbol{C} \cdot diag\left(\frac{1}{\ln(\boldsymbol{1} - \boldsymbol{\pi}_S)}\right) \cdot diag\left(\frac{1}{\boldsymbol{1} - \boldsymbol{\pi}_S}\right) \cdot \boldsymbol{S} &\quad (3.3\ .17) \\ &= -\boldsymbol{C} \cdot \boldsymbol{D}_1{}^{-1} \cdot \boldsymbol{D}_2{}^{-1} \cdot \boldsymbol{S} &\quad (3.3\ .18)\end{aligned}$$

where $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ are diagonal matrices with the elements of $\ln(\boldsymbol{1} - \boldsymbol{\pi}_S)$ and $(\boldsymbol{1} - \boldsymbol{\pi}_S)$, respectively, on the main diagonal and

$$\boldsymbol{S} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 1 & \cdots & 0 & 0 \\ \cdots & & \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & 1 & \cdots & 1 & 0 \end{pmatrix}.$$

From Equation 3.3 .5 follows that the MLE of $\boldsymbol{\pi}$, the vector of probabilities, is

$$\widehat{\boldsymbol{\pi}}_c = \boldsymbol{p} - (\boldsymbol{G}_\pi \boldsymbol{V})'(\boldsymbol{G}_p \boldsymbol{V} \boldsymbol{G}_\pi')^* \cdot \boldsymbol{C}.\ln\{-\ln(\boldsymbol{1} - \boldsymbol{S}\cdot\boldsymbol{p})\} \tag{3.3 .19}$$

with $\boldsymbol{p} = \dfrac{\boldsymbol{f}}{n}$ where $\boldsymbol{f} = (f_1, f_2, \ldots, f_k)'$ is the frequency vector being multinomial$(n, \boldsymbol{\pi})$ distributed.

The variance-covariance matrix $\boldsymbol{V}$ to be used is the estimated variance-covariance matrix of the multinomial distribution, which follows from Equation 3.3 .4 as

$$\widehat{\boldsymbol{V}} = \frac{1}{n}\left[diag(\boldsymbol{p}) - \boldsymbol{p}\boldsymbol{p}'\right] \qquad (3.3\ .20)$$

Since Equation 3.3 .19 is still a function of the unknown parameter $\boldsymbol{\pi}$, the double iterative procedure in [11] must be implemented. Once the iterative procedure in Equation 3.3 .19 has converged, the estimated parameters of the Weibull distribution can be solved from

$$\begin{pmatrix} \widehat{\ln\lambda} \\ \widehat{\alpha} \end{pmatrix} = (\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}' \cdot \ln\left(-\ln(\mathbf{1} - \boldsymbol{S} \cdot \widehat{\boldsymbol{\pi}}_c)\right). \qquad (3.3\ .21)$$

The estimated lambda parameter of the Weibull distribution then is

$$\widehat{\lambda} = \exp(\widehat{\ln\lambda})$$

and the estimated alpha parameter $\widehat{\alpha}$.

The SAS/IML program to fit a Weibull distribution to grouped survival data with a fixed censoring time appears in Appendix A.

3.3.4 Fitting of a log-logistic distribution to grouped survival data

From Equation 2.4 .17 follows that

$$\ln\left(\frac{1 - S(t)}{S(t)}\right) = \ln\lambda + \alpha\ln t \qquad (3.3\ .22)$$

where t denotes the **continuous** survival time.

In the current notation for **grouped survival data** in terms of the vector of upper class boundaries, Equation 3.3 .22 becomes

$$\begin{aligned} \ln\left(\frac{\mathbf{1} - S(\boldsymbol{x})}{S(\boldsymbol{x})}\right) &= \ln\left(\frac{F(\boldsymbol{x})}{\mathbf{1} - F(\boldsymbol{x})}\right) \\ &= \ln\left(\frac{\boldsymbol{\pi}_S}{\mathbf{1} - \boldsymbol{\pi}_S}\right) \\ &= \ln(\boldsymbol{\pi}_S) - \ln(\mathbf{1} - \boldsymbol{\pi}_S) \\ &= \ln\lambda \cdot \mathbf{1} + \alpha \cdot \ln\boldsymbol{x} \end{aligned} \qquad (3.3\ .23)$$

$$\begin{aligned}
&= \ln\lambda \cdot \begin{pmatrix} 1 \\ 1 \\ \cdots \\ 1 \end{pmatrix} + \alpha \cdot \begin{pmatrix} \ln x_1 \\ \ln x_2 \\ \cdots \\ \ln x_{k-1} \end{pmatrix} \\
&= \begin{pmatrix} 1 & \ln x_1 \\ 1 & \ln x_2 \\ \cdots & \cdots \\ 1 & \ln x_{k-1} \end{pmatrix} \cdot \begin{pmatrix} \ln\lambda \\ \alpha \end{pmatrix} \\
&= \underbrace{(\mathbf{1}\,, \ln \boldsymbol{x})}\cdot \begin{pmatrix} \ln\lambda \\ \alpha \end{pmatrix} \\
\Rightarrow \ln\left(\frac{\boldsymbol{\pi}_S}{\mathbf{1}-\boldsymbol{\pi}_S}\right) = \ln(\boldsymbol{\pi}_S) - \ln(\mathbf{1}-\boldsymbol{\pi}_S) &= \boldsymbol{X}_1 \cdot \begin{pmatrix} \ln\lambda \\ \alpha \end{pmatrix} \qquad (3.3\ .24)
\end{aligned}$$

Equation 3.3 .24 is a **linear model** in the parameters $\ln\lambda$ and α.

Similar to the case of the Weibull distribution, Equation 3.3 .24 is equivalent to

$$\begin{aligned}
\underbrace{\left(\boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}'\right)} \cdot \ln\left(\frac{\boldsymbol{\pi}_S}{\mathbf{1}-\boldsymbol{\pi}_S}\right) &= \mathbf{0} \\
\underbrace{\boldsymbol{C} \cdot \ln\left(\frac{\boldsymbol{\pi}_S}{\mathbf{1}-\boldsymbol{\pi}_S}\right)} &= \mathbf{0} \\
\boldsymbol{g}(\boldsymbol{\pi}) &= \mathbf{0}
\end{aligned}$$

The function $g(\boldsymbol{\pi}) = \mathbf{0}$ satisfies the conditions of Proposition 1 and the estimation algorithm can be used to estimate the parameters λ and α of the log-logistic distribution.

To summarize, the constraints imposed by the log-logistic distribution are specified by

$$g(\boldsymbol{\pi}) = \boldsymbol{C}.\ln\left\{\frac{\boldsymbol{\pi}_S}{\mathbf{1}-\boldsymbol{\pi}_S}\right\} = \boldsymbol{C}.\ln\left[\frac{\boldsymbol{S}\cdot\boldsymbol{\pi}}{\mathbf{1}-\boldsymbol{S}\cdot\boldsymbol{\pi}}\right] = \boldsymbol{C}.\left[\ln(\boldsymbol{S}\cdot\boldsymbol{\pi}) - \ln(\mathbf{1}-\boldsymbol{S}\cdot\boldsymbol{\pi})\right] = \mathbf{0} \qquad (3.3\ .25)$$

with

$$\boldsymbol{C} = \boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}' \text{ and } \boldsymbol{X}_1 = (\mathbf{1}, \ln\boldsymbol{x}) \quad . \qquad (3.3\ .26)$$

The derivative of $g(\boldsymbol{\pi})$ with respect to $\boldsymbol{\pi}$ is given by

$$\begin{aligned}
\boldsymbol{G}_\pi &= \frac{\partial \boldsymbol{g}(\boldsymbol{\pi})}{\partial \boldsymbol{\pi}} \\
&= \boldsymbol{C}\cdot\left[diag\left(\frac{1}{\boldsymbol{S}\cdot\boldsymbol{\pi}}\right)\cdot\boldsymbol{S} - diag\left(\frac{1}{\mathbf{1}-\boldsymbol{S}\cdot\boldsymbol{\pi}}\right)\cdot(-\boldsymbol{S})\right]
\end{aligned}$$

$$= \boldsymbol{C} \cdot \left[diag\left(\frac{1}{\boldsymbol{\pi}_S}\right) + diag\left(\frac{1}{\boldsymbol{1}-\boldsymbol{\pi}_S}\right)\right] \cdot \boldsymbol{S} \quad (3.3\ .27)$$

$$= \boldsymbol{C} \cdot \left[\boldsymbol{D}_3{}^{-1} + \boldsymbol{D}_2{}^{-1}\right] \cdot \boldsymbol{S} \quad (3.3\ .28)$$

where $\boldsymbol{D}_3$ and $\boldsymbol{D}_2$ are diagonal matrices with the elements of $\boldsymbol{\pi}_S$ and $\boldsymbol{1}-\boldsymbol{\pi}_S$, respectively, on the main diagonal and

$$\boldsymbol{S} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 1 & \cdots & 0 & 0 \\ \cdots & & \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & 1 & \cdots & 1 & 0 \end{pmatrix}.$$

The MLE of $\boldsymbol{\pi}$, the vector of probabilities, is in this case

$$\widehat{\boldsymbol{\pi}}_c = \boldsymbol{p} - (\boldsymbol{G}_\pi \boldsymbol{V})' (\boldsymbol{G}_p \boldsymbol{V} \boldsymbol{G}_\pi')^* \cdot \boldsymbol{C}. \ln\left\{\frac{\boldsymbol{S}\cdot\boldsymbol{p}}{\boldsymbol{1}-\boldsymbol{S}\cdot\boldsymbol{p}}\right\} \quad (3.3\ .29)$$

with $\boldsymbol{p} = \frac{\boldsymbol{f}}{n}$ where $\boldsymbol{f} = (f_1, f_2, \ldots, f_k)'$ is the frequency vector being multinomial$(n, \boldsymbol{\pi})$ distributed. The variance-covariance matrix $\boldsymbol{V}$ to be used is again the estimated variance-covariance matrix of the multinomial distribution, namely

$$\widehat{\boldsymbol{V}} = \frac{1}{n}\left[diag(\boldsymbol{p}) - \boldsymbol{p}\boldsymbol{p}'\right] \quad (3.3\ .30)$$

Since Equation 3.3 .29 is still a function of the unknown parameter $\boldsymbol{\pi}$, the double iterative procedure must be implemented. Once the iterative procedure in Equation 3.3 .29 has converged, the estimated parameters of the log-logistic distribution can be solved from

$$\begin{pmatrix} \widehat{\ln\lambda} \\ \widehat{\alpha} \end{pmatrix} = (\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}' \cdot \ln\left\{\frac{\boldsymbol{S}\cdot\widehat{\boldsymbol{\pi}}_c}{\boldsymbol{1}-\boldsymbol{S}\cdot\widehat{\boldsymbol{\pi}}_c}\right\}. \quad (3.3\ .31)$$

The estimated lambda parameter of the log-logistic distribution then is

$$\widehat{\lambda} = \exp(\widehat{\ln\lambda})$$

and the estimated alpha parameter $\widehat{\alpha}$.

The SAS/IML program to fit a log-logistic distribution to grouped survival data with a fixed censoring time appears in Appendix A.

3.3.5 Fitting of a lognormal distribution to grouped survival data

From Equation 2.5 .10 follows that

$$T \sim \text{lognormal}(\mu, \sigma^2) \iff \ln(T) \sim \text{normal}(\mu, \sigma^2). \quad (3.3\ .32)$$

where T denotes the **continuous** survival time.

In the current notation for **grouped survival data** in terms of the vector of upper class boundaries, Equation 3.3 .32 becomes

$$\begin{aligned} \boldsymbol{x} \sim \text{lognormal}(\mu, \sigma^2) &\Longleftrightarrow \ln \boldsymbol{x} \sim \text{normal}(\mu, \sigma^2) \\ &\Longleftrightarrow \frac{\ln \boldsymbol{x} - \mu \cdot \mathbf{1}}{\sigma} \sim \text{normal}(0,1). \end{aligned} \qquad (3.3\ .33)$$

From Equation 3.3 .8 and Equation 3.3 .33 follow that

$$\frac{\ln \boldsymbol{x} - \mu.\mathbf{1}}{\sigma} = \Phi^{-1}(\boldsymbol{\pi}_S)$$

specifies the constraints on the elements of $\boldsymbol{\pi}_S = \boldsymbol{\pi}.\boldsymbol{S}$ and hence on $\boldsymbol{\pi}$ where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution.

$$\begin{aligned} \Rightarrow \Phi^{-1}(\boldsymbol{\pi}_S) &= -\frac{\mu}{\sigma} \cdot \mathbf{1} + \frac{1}{\sigma} \cdot \ln \boldsymbol{x} && (3.3\ .34) \\ &= (\mathbf{1}\ , \ln \boldsymbol{x}) \cdot \begin{pmatrix} -\frac{\mu}{\sigma} \\ \frac{1}{\sigma} \end{pmatrix} \\ \Rightarrow \Phi^{-1}(\boldsymbol{\pi}_S) &= \boldsymbol{X}_1 \cdot \begin{pmatrix} -\frac{\mu}{\sigma} \\ \frac{1}{\sigma} \end{pmatrix} && (3.3\ .35) \end{aligned}$$

Equation 3.3 .35 is a **linear model** in the parameters $-\frac{\mu}{\sigma}$ and $\frac{1}{\sigma}$.

Equation 3.3 .35 is equivalent to

$$\begin{aligned} \underbrace{\left(\boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\right)}_{} \cdot \Phi^{-1}(\boldsymbol{\pi}_S) &= \mathbf{0} \\ \underbrace{\boldsymbol{C} \qquad\qquad \cdot \Phi^{-1}(\boldsymbol{\pi}_S)}_{} &= \mathbf{0} \\ \boldsymbol{g}(\boldsymbol{\pi}) \qquad &= \mathbf{0} \end{aligned}$$

The function $\boldsymbol{g}(\boldsymbol{\pi}) = \mathbf{0}$ satisfies the conditions of Proposition 1 and the estimation algorithm can be used to estimate the parameters μ and σ^2 of the lognormal distribution.

To summarize, the constraints imposed by the lognormal distribution are specified by

$$\boldsymbol{g}(\boldsymbol{\pi}) = \boldsymbol{C}.\Phi^{-1}(\boldsymbol{\pi}_S) = \boldsymbol{C}.\Phi^{-1}(\boldsymbol{S} \cdot \boldsymbol{\pi}) = \mathbf{0} \qquad (3.3\ .36)$$

with

$$\boldsymbol{C} = \boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1' \text{ and } \boldsymbol{X}_1 = (\mathbf{1}, \ln \boldsymbol{x}) \quad . \tag{3.3 .37}$$

The derivative of $g(\boldsymbol{\pi})$ with respect to $\boldsymbol{\pi}$ is

$$\boldsymbol{G}_\pi = \frac{\partial g(\boldsymbol{\pi})}{\partial \boldsymbol{\pi}} = \boldsymbol{C} \cdot \frac{\partial}{\partial \boldsymbol{\pi}} \Phi^{-1}(\boldsymbol{S} \cdot \boldsymbol{\pi}) \cdot \boldsymbol{S} \tag{3.3 .38}$$

In order to find $\frac{\partial}{\partial \boldsymbol{\pi}} \Phi^{-1}(\boldsymbol{S} \cdot \boldsymbol{\pi})$ consider the scalar case. Let

$$\pi_i = \Phi\left(\frac{\ln(x_i) - \mu}{\sigma}\right) = \Phi(z_i) \ .$$

Then $\Phi^{-1}(\pi_i) = \dfrac{\ln(x_i) - \mu}{\sigma} = z_i$, so that

$$\frac{\partial}{\partial \pi_i} \Phi^{-1}(\pi_i) = \frac{\partial z_i}{\partial \pi_i} = \frac{1}{\partial \pi_i / \partial z_i} = \frac{1}{\phi(z_i)} \ ,$$

with $\phi(\cdot)$ the probability density function of the standard normal distribution.

Applying this result to the vector of derivatives, Equation 3.3 .38 becomes

$$\boldsymbol{G}_\pi = \boldsymbol{C} \cdot \left[\frac{1}{diag\left\{\phi\left(\frac{\ln \boldsymbol{x} - \mu \cdot \mathbf{1}}{\sigma}\right)\right\}} \right] \boldsymbol{S} \ . \tag{3.3 .39}$$

Since $\boldsymbol{G}_\pi$ depends on μ and σ in the iterative procedure, these parameters will be estimated within the iterative stages and the final estimates will be obtained on convergence.

The SAS/IML program to fit a lognormal distribution to grouped survival data with a fixed censoring time appears in Appendix A.

3.3.6 A measure to compare the fit of survival distributions

A simple **measure of discrepancy** for comparing the fit of the survival distributions, is the statistic

$$D_{\chi^2} = \frac{\chi^2_W}{n} \tag{3.3 .40}$$

where χ^2_W is the Wald goodness of fit statistic (refer to [1]).

The Wald statistic in the survival analysis context is defined as

$$\chi^2_W = g(\boldsymbol{p})' \cdot (\boldsymbol{G}_p \boldsymbol{V} \boldsymbol{G}_p')^* \cdot g(\boldsymbol{p}).$$

with $\boldsymbol{V} = \frac{1}{n}\left[diag(\boldsymbol{p}) - \boldsymbol{p}\boldsymbol{p}'\right]$ the estimated variance-covariance matrix of the multinomial distribution.

When fitting a Weibull distribution

$$g(\boldsymbol{p}) = \boldsymbol{C}.\ln\{-\ln(\mathbf{1} - \boldsymbol{S}\cdot\boldsymbol{p})\}$$

and

$$\boldsymbol{G}_p = -\boldsymbol{C}\cdot diag\left(\frac{1}{\ln(\mathbf{1} - \boldsymbol{S}\cdot\boldsymbol{p})}\right)\cdot diag\left(\frac{1}{\mathbf{1} - \boldsymbol{S}\cdot\boldsymbol{p}}\right)\cdot\boldsymbol{S}.$$

When fitting a log-logistic distribution

$$g(\boldsymbol{p}) = \boldsymbol{C}.\left[\ln(\boldsymbol{S}\cdot\boldsymbol{p}) - \ln(\mathbf{1} - \boldsymbol{S}\cdot\boldsymbol{p})\right]$$

and

$$\boldsymbol{G}_p = \boldsymbol{C}\cdot\left[diag\left(\frac{1}{\boldsymbol{S}\cdot\boldsymbol{p}}\right) + diag\left(\frac{1}{\mathbf{1} - \boldsymbol{S}\cdot\boldsymbol{p}}\right)\right]\cdot\boldsymbol{S}.$$

When fitting a lognormal distribution

$$g(\boldsymbol{p}) = \boldsymbol{C}.\Phi^{-1}(\boldsymbol{S}\cdot\boldsymbol{\pi})$$

and

$$\boldsymbol{G}_p = \boldsymbol{C}\cdot\left[\frac{1}{diag\left\{\phi\left(\frac{\ln\boldsymbol{x} - \mu_p\cdot\mathbf{1}}{\sigma_p}\right)\right\}}\right]\boldsymbol{S}$$

with μ_p and σ_p the estimated values of μ and σ at the first iteration.

The number of degrees of freedom equals the number of independent constraints imposed by the model. In general, a value of D_{χ^2} less than 0.05 may be regarded as a good fit.

The Pearson's χ^2 statistic and the maximum likelihood χ^2 statistic (refer to [38, page 16-18] are asymptotically equivalent to the Wald statistic.

The calculation of the Wald statistic and the associated discrepancy is shown in the SAS/IML programs in Appendix A.

3.4 MLE subject to Constraints - Staggered Entry

3.4.1 Introduction

Consider the following experimental design as illustrated in Figure 3.1. Policies enter the study at different times (**staggered entry**). The event to be occurred is a lapse. The lifetime of a policy is measured from inception date up to the lapsing date. If the lapsing date is prior to a fixed termination date (cutoff date) of the study, determined in advance, then the lifetime is observed (an uncensored observation). If a policy is still in force (alive) when the termination point is reached, the lifetime of this policy is **right-censored**. Random entries to the study are assumed. This type of censoring is known as **random right-censoring**. The censoring is **noninformative** in that the lapse and censoring times are independent.

3.4.2 Notation for staggered entry

C_j is the potential censoring time for the j^{th} policy, associated with lifetime X_j.
$C_1, C_2, ..., C_n$ are independent identically distributed random variables, each with distribution function G and density function g. A further assumption that X_i and C_i are independent is made.

The survival data, based on a sample of size n, can then be represented by pairs $(T_1, \delta_1), (T_2, \delta_2), ..., (T_n, \delta_n)$ where

$$\begin{aligned} T_j &= min(X_j, C_j) \text{ for the } j^{th} \text{ policy} \\ \delta_j &= \begin{cases} 1 & \text{if} \quad X_j \leq C_j \quad \text{, that is, } X_j \text{ is not censored} \\ 0 & \text{if} \quad X_j > C_j \quad \text{, that is, } X_j \text{ is censored} \end{cases} \end{aligned}$$

$T_1, T_2, ..., T_n$ are independent identically distributed random variables with distribution function F if $T_j = X_j$ and distribution function G if $T_j = C_j$.

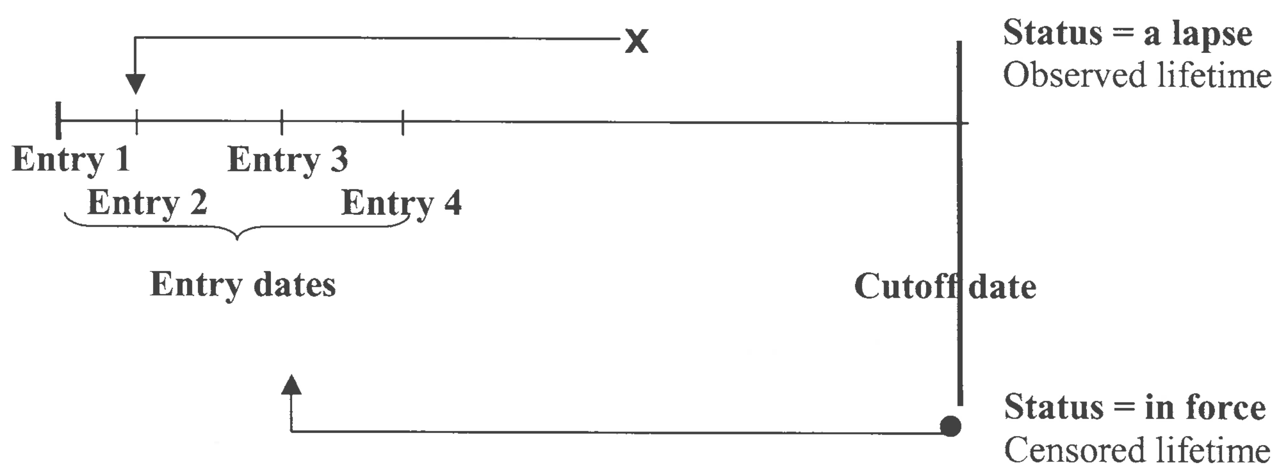

Figure 3.1: **Experimental design for staggered entry of policies**

To set notation for the staggered entry case, assume for illustration purposes **four different entry times** for the policies. The lifetimes of the n_1 policies that enter the study at the first entry time (called the first sample of size n_1) can be grouped into k adjacent, non-overlapping fixed intervals

$$I_j = [x_{j-1}; x_j) \quad j = 1, 2, ..., k$$

with $x_0 = 0$ and $x_k = \infty$. The last interval is an open interval containing all the censored lifetimes of the first sample.

The lifetimes of the n_2 policies that enter the study at the second entry time (called the second sample of size n_2) can be grouped into $(k-1)$ adjacent, non-overlapping fixed intervals

$$I_j = [x_{j-1}; x_j) \quad j = 1, 2, ..., k-1$$

with $x_0 = 0$ and $x_{k-1} = \infty$. The last interval is an open interval containing all the censored lifetimes of the second sample .

The lifetimes of the n_3 policies that enter the study at the third entry time (called the third sample of size n_3) can be grouped into $(k-2)$ adjacent, non-overlapping fixed intervals

$$I_j = [x_{j-1}; x_j) \quad j = 1, 2, ..., k-2$$

with $x_0 = 0$ and $x_{k-2} = \infty$. The last interval is an open interval containing all the censored lifetimes of the third sample.

The lifetimes of the n_4 policies that enter the study at the last entry time (called the fourth sample of size n_4) can be grouped into $(k-3)$ adjacent, non-overlapping fixed intervals

$$I_j = [x_{j-1}; x_j) \quad j = 1, 2, ..., k-3$$

with $x_0 = 0$ and $x_{k-3} = \infty$. The last interval is an open interval containing all the censored lifetimes of the fourth sample.

Four frequency distributions are formed when the observed and censored lifetimes of all the policies are grouped into the different lifetime intervals. The total number of observations in the data set is $n = n_1 + n_2 + n_3 + n_4$.

The four vectors of upper class boundaries are defined as follows:

$\boldsymbol{x}_1 = (x_1, x_2, ..., x_{k-1})'$ is a $(k-1) \times 1$ vector (sample 1)
$\boldsymbol{x}_2 = (x_1, x_2, ..., x_{k-2})'$ is a $(k-2) \times 1$ vector (sample 2)
$\boldsymbol{x}_3 = (x_1, x_2, ..., x_{k-3})'$ is a $(k-3) \times 1$ vector (sample 3)
$\boldsymbol{x}_4 = (x_1, x_2, ..., x_{k-4})'$ is a $(k-4) \times 1$ vector (sample 4)

The four relative frequency vectors are observed probability vectors from four independent multinomial populations. Let $\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3$ and $\boldsymbol{p}_4$ be the four relative frequency vectors.

$\boldsymbol{p}_1 = (p_{1,1}, p_{1,2}, p_{1,3}, ..., p_{1,k})'$ is an observed probability vector (sample 1)
$\boldsymbol{p}_2 = (p_{2,1}, p_{2,2}, p_{2,3}, ..., p_{2,k-1})'$ is an observed probability vector (sample 2)
$\boldsymbol{p}_3 = (p_{3,1}, p_{3,2}, p_{3,3}, ..., p_{3,k-2})'$ is an observed probability vector (sample 3)
$\boldsymbol{p}_4 = (p_{4,1}, p_{4,2}, p_{4,3}, ..., p_{4,k-3})'$ is an observed probability vector (sample 4)

Each sample is from a multinomial population $\quad i = 1, 2, 3, 4 \quad$ with

$$E(\boldsymbol{p}_i) = \boldsymbol{\pi}_i$$

$$Cov(\boldsymbol{p}_i) = \boldsymbol{V}_i = \frac{1}{n_i}\left[diag(\boldsymbol{\pi}_i) - \frac{1}{n_i}\boldsymbol{\pi}_i\boldsymbol{\pi}_i'\right]$$

where $\boldsymbol{\pi}_1 = (\pi_{1,1}, \pi_{1,2}, \pi_{1,3}, ..., \pi_{1,k})'$ is a $k \times 1$ probability vector
$\boldsymbol{\pi}_2 = (\pi_{2,1}, \pi_{2,2}, \pi_{2,3}, ..., \pi_{2,k-1})'$ is a $(k-1) \times 1$ probability vector
$\boldsymbol{\pi}_3 = (\pi_{3,1}, \pi_{3,2}, \pi_{3,3}, ..., \pi_{3,k-2})'$ is a $(k-2) \times 1$ probability vector
$\boldsymbol{\pi}_4 = (\pi_{4,1}, \pi_{4,2}, \pi_{4,3}, ..., \pi_{4,k-3})'$ is a $(k-3) \times 1$ probability vector

$\pi_{i,j}$ is the probability that an observation from sample i will fall in the j^{th} interval, that is the interval probability of the j^{th} interval from sample $i \quad i = 1, 2, 3, 4 \quad j = 1, 2, ..., k$.

$\Rightarrow \quad \boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3$ and $\boldsymbol{p}_4$ are four observed probability vectors corresponding to

$$n_i\boldsymbol{p}_i \quad \text{being} \quad \text{multinomial}(n_i; \boldsymbol{\pi}_i)$$

with n_i the number of observations in the i^{th} sample $\quad i = 1, 2, 3, 4$.

Table 3.2 gives the relative frequency distributions of the four samples.

The vectors $\boldsymbol{x}_i \quad i = 1, 2, 3, 4 \quad$ of upper class boundaries for the i^{th} sample (entry group) are

$$\boldsymbol{x}_1 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ x_{k-1} \end{pmatrix} \quad \boldsymbol{x}_2 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ \cdots \\ \cdots \\ x_{k-2} \end{pmatrix} \quad \boldsymbol{x}_3 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ \cdots \\ x_{k-3} \end{pmatrix} \quad \text{and} \quad \boldsymbol{x}_4 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ x_{k-4} \end{pmatrix} .$$

Define the combined vector of relative frequencies (combined observed probability vector) as $\boldsymbol{p}' = (\boldsymbol{p}'_1, \boldsymbol{p}'_2, \boldsymbol{p}'_3, \boldsymbol{p}'_4)$ and the combined probability vector as $\boldsymbol{\pi}' = (\boldsymbol{\pi}'_1, \boldsymbol{\pi}'_2, \boldsymbol{\pi}'_3, \boldsymbol{\pi}'_4)$.

Note that $\boldsymbol{p}$ is the MLE of $\boldsymbol{\pi}$ in the case of no constraints. $\boldsymbol{\pi}$ is to be estimated under certain constraints.

The MLE of $\boldsymbol{\pi}$ should be determined in terms of

- constraints imposed by the experimental design
- constraints imposed by the survival distribution to be fitted

Table 3.2: **Relative frequency distributions of survival data - staggered entry**

Interval	Lifetime Intervals				Observed Probability Vector				Probability Vector				Vector of Upper Boundarie			
number	Entry 1	Entry 2	Entry 3	Entry 4	p_1	p_2	p_3	p_4	π_1	π_2	π_3	π_4	x_1	x_2	x_3	x_4
first	$[0, x_1)$	$[0, x_1)$	$[0, x_1)$	$[0, x_1)$	$p_{1,1}$	$p_{2,1}$	$p_{3,1}$	$p_{4,1}$	$\pi_{1,1}$	$\pi_{2,1}$	$\pi_{3,1}$	$\pi_{4,1}$	x_1	x_1	x_1	x_1
second	$[x_1, x_2)$	$[x_1, x_2)$	$[x_1, x_2)$	$[x_1, x_2)$	$p_{1,2}$	$p_{2,2}$	$p_{3,2}$	$p_{4,2}$	$\pi_{1,2}$	$\pi_{2,2}$	$\pi_{3,2}$	$\pi_{4,2}$	x_2	x_2	x_2	x_2
third	$[x_2, x_3)$	$[x_2, x_3)$	$[x_2, x_3)$	$[x_2, x_3)$	$p_{1,3}$	$p_{2,3}$	$p_{3,3}$	$p_{4,3}$	$\pi_{1,3}$	$\pi_{2,3}$	$\pi_{3,3}$	$\pi_{4,3}$	x_3	x_3	x_3	x_3
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	x_{k-4}
$(k-3)^{th}$	$[x_{k-4}, x_{k-3})$	$[x_{k-4}, x_{k-3})$	$[x_{k-4}, x_{k-3})$	$[x_{k-4}, \infty)$	$p_{1,k-3}$	$p_{2,k-3}$	$p_{3,k-3}$	$p_{4,k-3}$	$\pi_{1,k-3}$	$\pi_{2,k-3}$	$\pi_{3,k-3}$	$\pi_{4,k-3}$	x_{k-3}	x_{k-3}	x_{k-3}	
$(k-2)^{th}$	$[x_{k-3}, x_{k-2})$	$[x_{k-3}, x_{k-2})$	$[x_{k-3}, \infty)$		$p_{1,k-2}$	$p_{2,k-2}$	$p_{3,k-2}$		$\pi_{1,k-2}$	$\pi_{2,k-2}$	$\pi_{3,k-2}$		x_{k-2}	x_{k-2}		
$(k-1)^{th}$	$[x_{k-2}, x_{k-1})$	$[x_{k-2}, \infty)$			$p_{1,k-1}$	$p_{2,k-1}$			$\pi_{1,k-1}$	$\pi_{2,k-1}$			x_{k-1}			
k^{th}	$[x_{k-1}, \infty)$				$p_{1,k}$				$\pi_{1,k}$							

3.4.3 Definition of Constraints

Constraints imposed by the survival distribution to be fitted

Constraints imposed by the Weibull distribution

A Weibull distribution with parameters λ and α subject to the constraints $\boldsymbol{\pi}_S$ can be written from Equation 3.3 .13 as

$$\boxed{\ln\left(-\ln(\mathbf{1}-\boldsymbol{\pi}_S)\right) \;=\; \ln\lambda\cdot\mathbf{1}+\alpha\cdot\ln\boldsymbol{x}}$$

Constraints imposed by the log-logistic distribution

A log-logistic distribution with parameters λ and α subject to the constraints $\boldsymbol{\pi}_S$ can be written from Equation 3.3 .23 as

$$\boxed{\ln\left(\boldsymbol{\pi}_S\right)-\ln\left(\mathbf{1}-\boldsymbol{\pi}_S\right) \;=\; \ln\lambda\cdot\mathbf{1}+\alpha\cdot\ln\boldsymbol{x}}$$

Constraints imposed by the lognormal distribution

A lognormal distribution with parameters μ and σ^2 subject to the constraints $\boldsymbol{\pi}_s$ can be written from Equation 3.3 .34 as

$$\boxed{\Phi^{-1}(\boldsymbol{\pi}_S) \;=\; -\frac{\mu}{\sigma}\cdot\mathbf{1}+\frac{1}{\sigma}\cdot\ln\boldsymbol{x}}$$

Constraints imposed by the experimental design

Consider Figure 3.2, illustrating the constraints imposed by the experimental design.

- $\pi_{1,j} = \pi_{2,j} = \pi_{3,j} = \pi_{4,j} \quad j = 1, 2, ..., k-4$

- $$\begin{aligned} \pi_{1,k} + \pi_{1,k-1} + \pi_{1,k-2} + \pi_{1,k-3} &= \pi_{2,k-1} + \pi_{2,k-2} + \pi_{2,k-3} \\ &= \pi_{3,k-2} + \pi_{3,k-3} \\ &= \pi_{4,k-3} \end{aligned}$$

- $$\begin{aligned} \pi_{1,k-2} &= \pi_{2,k-2} \\ \pi_{1,k-3} &= \pi_{2,k-3} \end{aligned}$$

where $\pi_{i,j}$ = probability of an observation from sample i will fall in the j^{th} interval
= interval probability of j^{th} interval from sample i $\quad i = 1, 2, 3, 4 \quad j = 1, 2, ..., k$

These constraints can be written as

- $1 \cdot \pi_{1,j} - 1 \cdot \pi_{2,j} = 0$
  $1 \cdot \pi_{1,j} - 1 \cdot \pi_{3,j} = 0$
  $1 \cdot \pi_{1,j} - 1 \cdot \pi_{4,j} = 0 \quad j = 1, 2, ..., k-4$

- $1 \cdot \pi_{1,k} + 1 \cdot \pi_{1,k-1} + 1 \cdot \pi_{1,k-2} + 1 \cdot \pi_{1,k-3} - 1 \cdot \pi_{2,k-1} - 1 \cdot \pi_{2,k-2} - 1 \cdot \pi_{2,k-3} = 0$
  $1 \cdot \pi_{1,k} + 1 \cdot \pi_{1,k-1} + 1 \cdot \pi_{1,k-2} + 1 \cdot \pi_{1,k-3} - 1 \cdot \pi_{3,k-2} - 1 \cdot \pi_{3,k-3} = 0$
  $1 \cdot \pi_{1,k} + 1 \cdot \pi_{1,k-1} + 1 \cdot \pi_{1,k-2} + 1 \cdot \pi_{1,k-3} - 1 \cdot \pi_{4,k-3} = 0$

- $1 \cdot \pi_{1,k-2} - 1 \cdot \pi_{2,k-2} = 0$
  $1 \cdot \pi_{1,k-3} - 1 \cdot \pi_{2,k-3} = 0$

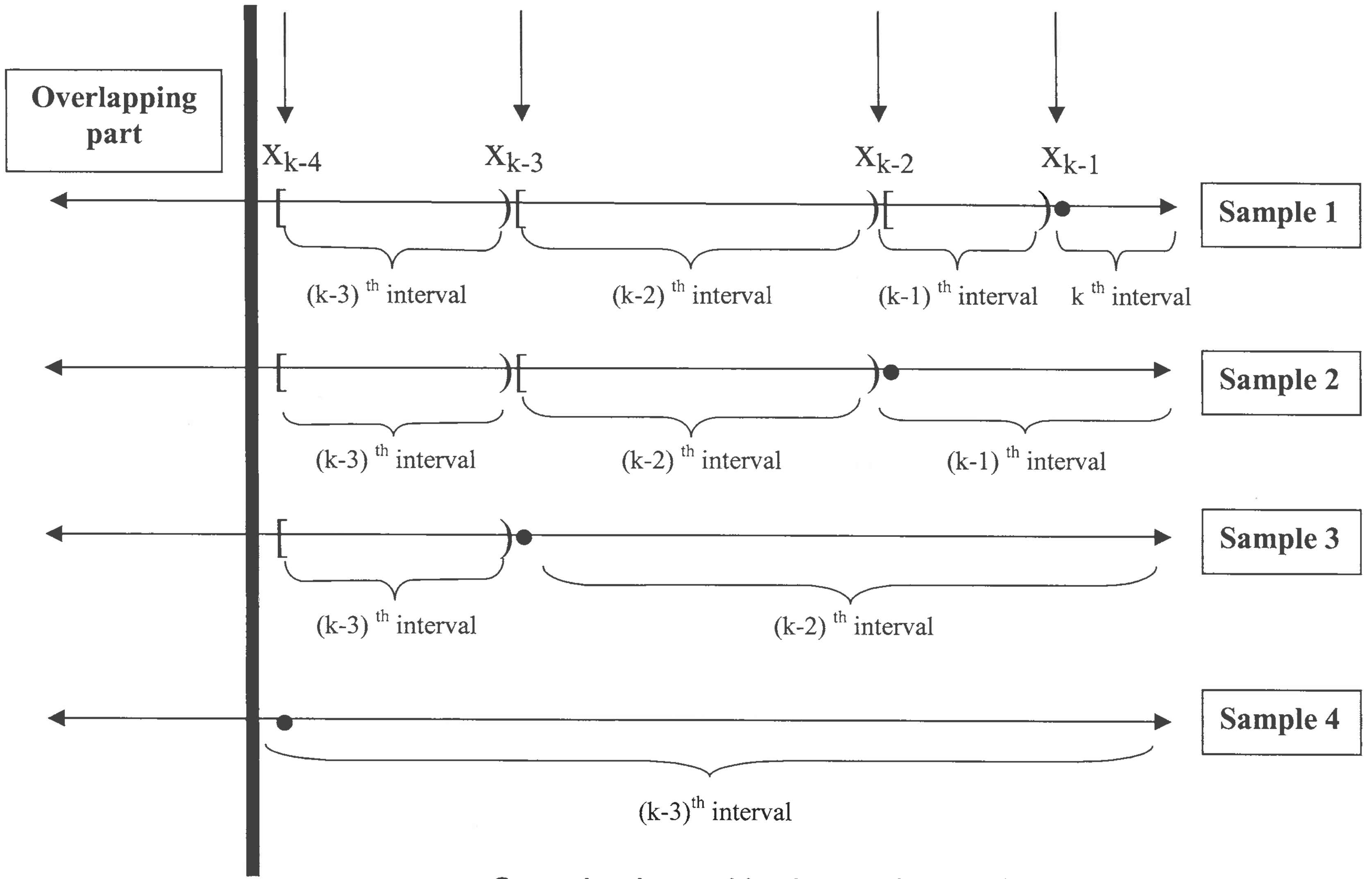

Figure 3.2: **Constraints imposed by the experimental design**

These constraints in matrix form are $\boldsymbol{G} \cdot \boldsymbol{\pi} = 0$ with $\boldsymbol{\pi}' = (\boldsymbol{\pi}_1', \boldsymbol{\pi}_2', \boldsymbol{\pi}_3', \boldsymbol{\pi}_4')$ where

$\boldsymbol{\pi}_1 = (\pi_{1,1}, \pi_{1,2}, \pi_{1,3}, ..., \pi_{1,k})'$ is a $k \times 1$ probability vector
$\boldsymbol{\pi}_2 = (\pi_{2,1}, \pi_{2,2}, \pi_{2,3}, ..., \pi_{2,k-1})'$ is a $(k-1) \times 1$ probability vector
$\boldsymbol{\pi}_3 = (\pi_{3,1}, \pi_{3,2}, \pi_{3,3}, ..., \pi_{3,k-2})'$ is a $(k-2) \times 1$ probability vector
$\boldsymbol{\pi}_4 = (\pi_{4,1}, \pi_{4,2}, \pi_{4,3}, ..., \pi_{4,k-3})'$ is a $(k-3) \times 1$ probability vector
and

$$\boldsymbol{G} = \begin{pmatrix} \boldsymbol{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -\boldsymbol{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -\boldsymbol{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -\boldsymbol{I} & \mathbf{0} \\ & & & & & & & & & & & & & \\ \mathbf{0}' & 1 & 1 & 1 & 1 & \mathbf{0}' & -1 & -1 & -1 & \mathbf{0}' & 0 & 0 & \mathbf{0}' & 0 \\ \mathbf{0}' & 1 & 1 & 1 & 1 & \mathbf{0}' & 0 & 0 & 0 & \mathbf{0}' & -1 & -1 & \mathbf{0}' & 0 \\ \mathbf{0}' & 1 & 1 & 1 & 1 & \mathbf{0}' & 0 & 0 & 0 & \mathbf{0}' & 0 & 0 & \mathbf{0}' & -1 \\ & & & & & & & & & & & & & \\ \mathbf{0}' & 0 & 1 & 0 & 0 & \mathbf{0}' & 0 & -1 & 0 & \mathbf{0}' & 0 & 0 & \mathbf{0}' & 0 \\ \mathbf{0}' & 1 & 0 & 0 & 0 & \mathbf{0}' & -1 & 0 & 0 & \mathbf{0}' & 0 & 0 & \mathbf{0}' & 0 \end{pmatrix}.$$

3.4.4 Method of maximum likelihood estimation subject to constraints: staggered entry

The technique of maximum likelihood estimation subject to constraints is implemented in the following way:

1. One survival model is fitted under constraints imposed by the Weibull/log-logistic/lognormal distribution over the four entry groups.
2. Four survival models (Weibull/log-logistic/lognormal models), one for each entry time, are fitted under constraints imposed by the Weibull/log-logistic/lognormal distribution and under **further constraints** that
   - λ_i's are equal and α_i's are equal when fitting a Weibull or log-logistic

     or

   - μ_i's are equal and σ_i's are equal when fitting a lognormal
3. A joint histogram is fitted to the four histograms of the four relative frequency distributions under constraints imposed by the experimental design.

3.4.5 Fitting of one survival distribution to the four histograms

Fitting of one Weibull distribution to the four histograms

Recall that a Weibull distribution with parameters λ and α under the constraints π_S can be written as

$$\ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_S)\} = \ln\lambda \cdot \mathbf{1} + \alpha \cdot \ln \boldsymbol{x} \tag{3.4 .1}$$

or

$$\begin{aligned}
\ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_S)\} &= \ln\lambda \cdot \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \end{pmatrix} + \alpha \cdot \begin{pmatrix} \ln \boldsymbol{x}_1 \\ \ln \boldsymbol{x}_2 \\ \ln \boldsymbol{x}_3 \\ \ln \boldsymbol{x}_4 \end{pmatrix} \\
&= \underbrace{\begin{pmatrix} \mathbf{1} & \ln \boldsymbol{x}_1 \\ \mathbf{1} & \ln \boldsymbol{x}_2 \\ \mathbf{1} & \ln \boldsymbol{x}_3 \\ \mathbf{1} & \ln \boldsymbol{x}_4 \end{pmatrix}} \cdot \begin{pmatrix} \ln\lambda \\ \alpha \end{pmatrix} \\
\Rightarrow \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_S)\} &= \qquad \boldsymbol{X}_1 \qquad \cdot \begin{pmatrix} \ln\lambda \\ \alpha \end{pmatrix}
\end{aligned} \tag{3.4 .2}$$

where

$$\boldsymbol{x}_1 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ x_{k-1} \end{pmatrix} \quad \boldsymbol{x}_2 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ \cdots \\ \cdots \\ x_{k-2} \end{pmatrix} \quad \boldsymbol{x}_3 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ \cdots \\ x_{k-3} \end{pmatrix} \quad \text{and} \quad \boldsymbol{x}_4 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ x_{k-4} \end{pmatrix}.$$

Equation 3.4 .2 is a **linear model** in the parameters $\ln\lambda$ and α. This model is equivalent to

$$\begin{aligned}
\underbrace{\left(\boldsymbol{I} - \mathbf{X}_1(\mathbf{X}_1{}'\mathbf{X}_1)^{-1}\mathbf{X}_1{}'\right)} \cdot \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_S)\} &= 0 \\
\underbrace{\boldsymbol{C} \qquad\qquad \cdot \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_S)\}} &= 0 \\
g(\boldsymbol{\pi}) &= 0
\end{aligned}$$

$\boldsymbol{C}$ is the projection matrix orthogonal to the columns of the design matrix $\boldsymbol{X}_1$. Note that $\boldsymbol{C}\boldsymbol{X}_1 = 0$.

The function $g(\boldsymbol{\pi}) = \mathbf{0}$ satisfies the conditions of Proposition 1 and the estimation algorithm can be used to estimate the λ and α of the Weibull distribution.

To summarize, the constraints imposed by the Weibull distribution are specified by

$$g(\boldsymbol{\pi}) = \boldsymbol{C}.\ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_S)\} = \boldsymbol{C}.\ln\{-\ln(\mathbf{1}-\boldsymbol{S}\cdot\boldsymbol{\pi})\} = \mathbf{0} \qquad (3.4\ .3)$$

with

$$\boldsymbol{C} = \boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}' \quad . \qquad (3.4\ .4)$$

The derivative of $g(\boldsymbol{\pi})$ with respect to $\boldsymbol{\pi}$ is

$$\begin{aligned} \boldsymbol{G}_\pi &= \frac{\partial g(\boldsymbol{\pi})}{\partial \boldsymbol{\pi}} \\ &= -\boldsymbol{C}\cdot diag\left(\frac{1}{\ln(\mathbf{1}-\boldsymbol{\pi}_S)}\right)\cdot diag\left(\frac{1}{\mathbf{1}-\boldsymbol{\pi}_S}\right)\cdot\boldsymbol{S} \qquad (3.4\ .5) \\ &= -\boldsymbol{C}\cdot\boldsymbol{D}_1{}^{-1}\cdot\boldsymbol{D}_2{}^{-1}\cdot\boldsymbol{S} \qquad (3.4\ .6) \end{aligned}$$

where

$\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ are diagonal matrices with the elements of $\ln(\mathbf{1}-\boldsymbol{\pi}_S)$ and $(\mathbf{1}-\boldsymbol{\pi}_S)$, respectively, on the main diagonal and $\boldsymbol{S}$ is a block-diagonal matrix created from four matrices $\boldsymbol{S}_1, \boldsymbol{S}_2, \boldsymbol{S}_3$ and $\boldsymbol{S}_4$ associated with the four entry periods.

The estimated vector of probabilities is in this case

$$\hat{\boldsymbol{\pi}}_c = \boldsymbol{p} - (\boldsymbol{G}_\pi\boldsymbol{V})'(\boldsymbol{G}_p\boldsymbol{V}\boldsymbol{G}_\pi')^*\cdot\boldsymbol{C}.\ln\{-\ln(1-\boldsymbol{S}\cdot\boldsymbol{p})\} \qquad (3.4\ .7)$$

with $\boldsymbol{p}' = (\boldsymbol{p}_1', \boldsymbol{p}_2', \boldsymbol{p}_3', \boldsymbol{p}_4')$ where $\boldsymbol{p}_1 = (p_{1,1}, p_{1,2}, p_{1,3}, ..., p_{1,k})'$ $\boldsymbol{p}_2 = (p_{2,1}, p_{2,2}, p_{2,3}, ..., p_{2,k-1})'$ $\boldsymbol{p}_3 = (p_{3,1}, p_{3,2}, p_{3,3}, ..., p_{3,k-2})'$ and $\boldsymbol{p}_4 = (p_{4,1}, p_{4,2}, p_{4,3}, ..., p_{4,k-3})'$ are four relative frequency vectors corresponding to $n_i\boldsymbol{p}_i$ being multinomial$(n_i; \boldsymbol{\pi}_i)$ $\quad i = 1, 2, 3, 4$ distributed.

The variance-covariance matrix $\boldsymbol{V}$ to be used is the estimated variance-covariance matrix of the multinomial distribution **for each entry period**.

$$\Longrightarrow \quad \widehat{\boldsymbol{V}} = \text{block}(\widehat{\boldsymbol{V}}_1, \widehat{\boldsymbol{V}}_2, \widehat{\boldsymbol{V}}_3, \widehat{\boldsymbol{V}}_4)$$

and

$$\widehat{\boldsymbol{V}}_i = \frac{1}{n_i}[diag(\boldsymbol{p}_i) - \boldsymbol{p}_i\boldsymbol{p}_i'] \quad i = 1, 2, 3, 4$$

In the notation for staggered entry of policies, as described in Table 3.2 with four entry periods and k, the number of class intervals, for illustration purposes equal to seven,

$\boldsymbol{p}_1 = (p_{1,1}, p_{1,2}, p_{1,3}, p_{1,4}, p_{1,5}, p_{1,6}, p_{1,7})'$ is a 7×1 relative frequency vector
$\boldsymbol{p}_2 = (p_{2,1}, p_{2,2}, p_{2,3}, p_{2,4}, p_{2,5}, p_{2,6})'$ is a 6×1 relative frequency vector
$\boldsymbol{p}_3 = (p_{3,1}, p_{3,2}, p_{3,3}, p_{3,4}, p_{3,5})'$ is a 5×1 relative frequency vector
$\boldsymbol{p}_4 = (p_{4,1}, p_{4,2}, p_{4,3}, p_{4,4})'$ is a 4×1 relative frequency vector

$\boldsymbol{\pi}_1 = (\pi_{1,1}, \pi_{1,2}, \pi_{1,3}, \pi_{1,4}, \pi_{1,5}, \pi_{1,6}, \pi_{1,7})'$ is a 7×1 probability vector
$\boldsymbol{\pi}_2 = (\pi_{2,1}, \pi_{2,2}, \pi_{2,3}, \pi_{2,4}, \pi_{2,5}, \pi_{2,6})'$ is a 6×1 probability vector
$\boldsymbol{\pi}_3 = (\pi_{3,1}, \pi_{3,2}, \pi_{3,3}, \pi_{3,4}, \pi_{3,5})'$ is a 5×1 probability vector
$\boldsymbol{\pi}_4 = (\pi_{4,1}, \pi_{4,2}, \pi_{4,3}, \pi_{4,4})'$ is a 4×1 probability vector.

$\boldsymbol{S}$ is a 18×22 block-diagonal matrix, that is

$$\boldsymbol{S} = \text{block}(\boldsymbol{S}_1, \boldsymbol{S}_2, \boldsymbol{S}_3, \boldsymbol{S}_4)$$

with

$$\boldsymbol{S}_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}$$

$$\boldsymbol{S}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}$$

$$\boldsymbol{S}_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix}$$

$$\boldsymbol{S}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

The matrix $\boldsymbol{S}$ is given below.

$$\boldsymbol{S} = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0
\end{pmatrix}.$$

Note that $\widehat{\boldsymbol{V}}_1$ is a 7×7, $\widehat{\boldsymbol{V}}_2$ is a 6×6, $\widehat{\boldsymbol{V}}_3$ is a 5×5 and $\widehat{\boldsymbol{V}}_4$ is a 4×4 matrix so that $\widehat{\boldsymbol{V}}$ is a 22×22 matrix.

Since Equation 3.4 .7 is still a function of the unknown parameter $\boldsymbol{\pi}$, the double iterative procedure must be implemented. Once the iterative procedure in Equation 3.4 .7 has converged, the estimated parameters of the Weibull distribution can be solved from

$$\begin{pmatrix} \widehat{\ln\lambda} \\ \widehat{\alpha} \end{pmatrix} = (\mathbf{X}_1{}'\mathbf{X}_1)^{-1}\mathbf{X}_1{}' \cdot \ln\left\{-\ln(\mathbf{1} - \boldsymbol{S} \cdot \widehat{\boldsymbol{\pi}}_c)\right\}. \tag{3.4 .8}$$

The estimated lambda parameter of the Weibull distribution then is

$$\widehat{\lambda} = \exp(\widehat{\ln\lambda})$$

and the estimated alpha parameter $\widehat{\alpha}$.

The SAS/IML program to fit a Weibull model to grouped survival data with staggered entry of policies appears in Appendix A.

Fitting of one log-logistic distribution to the four histograms

Recall that a log-logistic distribution with parameters λ and α under the constraints $\boldsymbol{\pi}_S$ can be written as

$$\ln(\boldsymbol{\pi}_S) - \ln(\mathbf{1} - \boldsymbol{\pi}_S) = \ln\lambda \cdot \mathbf{1} + \alpha \cdot \ln\boldsymbol{x} \tag{3.4 .9}$$

or

$$\begin{aligned}
\ln\left(\frac{\boldsymbol{\pi}_S}{\mathbf{1}-\boldsymbol{\pi}_S}\right) &= \ln(\boldsymbol{\pi}_S) - \ln(\mathbf{1}-\boldsymbol{\pi}_S) \\
&= \ln\lambda \cdot \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \end{pmatrix} + \alpha \cdot \begin{pmatrix} \ln \boldsymbol{x}_1 \\ \ln \boldsymbol{x}_2 \\ \ln \boldsymbol{x}_3 \\ \ln \boldsymbol{x}_4 \end{pmatrix} \\
&= \underbrace{\begin{pmatrix} \mathbf{1} & \ln \boldsymbol{x}_1 \\ \mathbf{1} & \ln \boldsymbol{x}_2 \\ \mathbf{1} & \ln \boldsymbol{x}_3 \\ \mathbf{1} & \ln \boldsymbol{x}_4 \end{pmatrix}} \cdot \begin{pmatrix} \ln\lambda \\ \alpha \end{pmatrix} \\
\Rightarrow \ln\left(\frac{\boldsymbol{\pi}_S}{\mathbf{1}-\boldsymbol{\pi}_S}\right) &= \quad \boldsymbol{X}_1 \quad \cdot \begin{pmatrix} \ln\lambda \\ \alpha \end{pmatrix} \qquad (3.4\ .10)
\end{aligned}$$

where

$$\boldsymbol{x}_1 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ x_{k-1} \end{pmatrix} \quad \boldsymbol{x}_2 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ \cdots \\ \cdots \\ x_{k-2} \end{pmatrix} \quad \boldsymbol{x}_3 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ \cdots \\ x_{k-3} \end{pmatrix} \quad \text{and} \quad \boldsymbol{x}_4 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ x_{k-4} \end{pmatrix}.$$

Equation 3.4 .10 is a **linear model** in the parameters $\ln\lambda$ and α. This model is equivalent to

$$\begin{aligned}
\underbrace{\left(\boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}'\right)} \cdot \ln\left(\frac{\boldsymbol{\pi}_S}{\mathbf{1}-\boldsymbol{\pi}_S}\right) &= \mathbf{0} \\
\underbrace{\boldsymbol{C} \qquad \cdot \ln\left(\frac{\boldsymbol{\pi}_S}{\mathbf{1}-\boldsymbol{\pi}_S}\right)} &= \mathbf{0} \\
\boldsymbol{g}(\boldsymbol{\pi}) &= \mathbf{0}
\end{aligned}$$

$\boldsymbol{C}$ is the projection matrix orthogonal to the columns of the design matrix $\boldsymbol{X}_1$. Note that $\boldsymbol{C}\boldsymbol{X}_1 = \mathbf{0}$.

The function $g(\pi) = 0$ satisfies the conditions of Proposition 1 and the estimation algorithm can be used to estimate the λ's and α's of the log-logistic distribution.

To summarize, the constraints imposed by the log-logistic distribution are specified by

$$g(\boldsymbol{\pi}) = \boldsymbol{C}.\ln\left\{\frac{\boldsymbol{\pi}_S}{\mathbf{1}-\boldsymbol{\pi}_S}\right\} = \boldsymbol{C}.\ln\left[\frac{\boldsymbol{S}\cdot\boldsymbol{\pi}}{\mathbf{1}-\boldsymbol{S}\cdot\boldsymbol{\pi}}\right] = \boldsymbol{C}.\left[\ln(\boldsymbol{S}\cdot\boldsymbol{\pi}) - \ln(\mathbf{1}-\boldsymbol{S}\cdot\boldsymbol{\pi})\right] = \mathbf{0} \tag{3.4 .11}$$

with

$$\boldsymbol{C} = \boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}' \quad . \tag{3.4 .12}$$

The derivative of $g(\boldsymbol{\pi})$ with respect to $\boldsymbol{\pi}$ is

$$\begin{aligned}\boldsymbol{G}_\pi &= \frac{\partial g(\boldsymbol{\pi})}{\partial\boldsymbol{\pi}} \\ &= \boldsymbol{C}\cdot\left[diag\left(\frac{1}{\boldsymbol{\pi}_S}\right) + diag\left(\frac{1}{\mathbf{1}-\boldsymbol{\pi}_S}\right)\right]\cdot\boldsymbol{S} \qquad (3.4\ .13) \\ &= \boldsymbol{C}\cdot\boldsymbol{D}_3{}^{-1} + \boldsymbol{D}_2{}^{-1}\cdot\boldsymbol{S} \qquad (3.4\ .14)\end{aligned}$$

where

$\boldsymbol{D}_3$ and $\boldsymbol{D}_2$ are diagonal matrices with the elements of $\boldsymbol{\pi}_S$ and $\mathbf{1}-\boldsymbol{\pi}_S$, respectively, on the main diagonal. The matrix $\boldsymbol{S}$ is the same $\boldsymbol{S}$ matrix that was used when fitting a Weibull model.

The estimated vector of probabilities in this case is

$$\widehat{\boldsymbol{\pi}}_c = \boldsymbol{p} - (\boldsymbol{G}_\pi\boldsymbol{V})'\,(\boldsymbol{G}_p\boldsymbol{V}\boldsymbol{G}_\pi')^*\cdot\boldsymbol{C}.\ln\left\{\frac{\boldsymbol{S}\cdot\boldsymbol{p}}{\mathbf{1}-\boldsymbol{S}\cdot\boldsymbol{p}}\right\} \tag{3.4 .15}$$

with $\boldsymbol{p}' = (\boldsymbol{p}_1',\boldsymbol{p}_2',\boldsymbol{p}_3',\boldsymbol{p}_4')$ where $\boldsymbol{p}_1 = (p_{1,1}, p_{1,2}, p_{1,3}, ..., p_{1,k})'$ $\boldsymbol{p}_2 = (p_{2,1}, p_{2,2}, p_{2,3}, ..., p_{2,k-1})'$ $\boldsymbol{p}_3 = (p_{3,1}, p_{3,2}, p_{3,3}, ..., p_{3,k-2})'$ and $\boldsymbol{p}_4 = (p_{4,1}, p_{4,2}, p_{4,3}, ..., p_{4,k-3})'$ are four relative frequency vectors corresponding to $n_i\boldsymbol{p}_i$ being multinomial$(n_i;\boldsymbol{\pi}_i)$ $\; i = 1,2,3,4$ distributed. The variance-covariance matrix $\boldsymbol{V}$ to be used is the estimated variance-covariance matrix of the multinomial distribution **for each entry period**.

Since Equation 3.4 .15 is still a function of the unknown parameter $\boldsymbol{\pi}$, the double iterative procedure must be implemented. Once the iterative procedure in Equation 3.4 .15 has converged, the estimated parameters of the log-logistic distribution can be solved from

$$\begin{pmatrix}\widehat{\ln\lambda} \\ \widehat{\alpha}\end{pmatrix} = (\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}'\cdot\ln\left\{\frac{\boldsymbol{S}\cdot\widehat{\boldsymbol{\pi}}_c}{\mathbf{1}-\boldsymbol{S}\cdot\widehat{\boldsymbol{\pi}}_c}\right\}. \tag{3.4 .16}$$

The estimated lambda parameter of the log-logistic distribution then is

$$\widehat{\lambda} = \exp(\widehat{\ln\lambda})$$

and the estimated alpha parameter $\widehat{\alpha}$.

The SAS/IML program to fit a log-logistic model to grouped survival data with staggered entry of policies appears in Appendix A.

Fitting of one lognormal distribution to the four histograms

Recall that a lognormal distribution with parameters μ and σ^2 under the constraints $\boldsymbol{\pi}_s$ can be written as

$$\Phi^{-1}(\boldsymbol{\pi}_S) = -\frac{\mu}{\sigma}\cdot\mathbf{1} + \frac{1}{\sigma}\cdot\ln\boldsymbol{x} \qquad (3.4\ .17)$$

$$= -\frac{\mu}{\sigma}\cdot\begin{pmatrix}\mathbf{1}\\\mathbf{1}\\\mathbf{1}\\\mathbf{1}\end{pmatrix} + \frac{1}{\sigma}\cdot\begin{pmatrix}\ln\boldsymbol{x}_1\\\ln\boldsymbol{x}_2\\\ln\boldsymbol{x}_3\\\ln\boldsymbol{x}_4\end{pmatrix}$$

$$= \underbrace{\begin{pmatrix}\mathbf{1} & \ln\boldsymbol{x}_1\\\mathbf{1} & \ln\boldsymbol{x}_2\\\mathbf{1} & \ln\boldsymbol{x}_3\\\mathbf{1} & \ln\boldsymbol{x}_4\end{pmatrix}}\cdot\begin{pmatrix}-\frac{\mu}{\sigma}\\\frac{1}{\sigma}\end{pmatrix}$$

$$\Rightarrow \Phi^{-1}(\boldsymbol{\pi}_S) = \boldsymbol{X}_1 \cdot \begin{pmatrix}-\frac{\mu}{\sigma}\\\frac{1}{\sigma}\end{pmatrix} \qquad (3.4\ .18)$$

where

$$\boldsymbol{x}_1 = \begin{pmatrix}x_1\\x_2\\x_3\\\cdots\\\cdots\\\cdots\\\cdots\\x_{k-1}\end{pmatrix} \quad \boldsymbol{x}_2 = \begin{pmatrix}x_1\\x_2\\x_3\\\cdots\\\cdots\\\cdots\\x_{k-2}\end{pmatrix} \quad \boldsymbol{x}_3 = \begin{pmatrix}x_1\\x_2\\x_3\\\cdots\\\cdots\\x_{k-3}\end{pmatrix} \quad \text{and} \quad \boldsymbol{x}_4 = \begin{pmatrix}x_1\\x_2\\x_3\\\cdots\\x_{k-4}\end{pmatrix}$$

Equation 3.4 .18 is a **linear model** in the parameters $-\frac{\mu}{\sigma}$ and $\frac{1}{\sigma}$.

Equation 3.3 .35 is equivalent to

$$\underbrace{\left(\boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}'\right)}\cdot\Phi^{-1}(\boldsymbol{\pi}_S) = 0$$

$$\underbrace{\boldsymbol{C} \qquad\qquad \cdot \Phi^{-1}(\boldsymbol{\pi}_S)}_{g(\boldsymbol{\pi})} = 0$$
$$g(\boldsymbol{\pi}) = 0$$

$\boldsymbol{C}$ is the projection matrix orthogonal to the columns of the design matrix $\boldsymbol{X}_1$. Note that $\boldsymbol{C}\boldsymbol{X}_1 = \boldsymbol{0}$.

The function $g(\boldsymbol{\pi}) = \boldsymbol{0}$ satisfies the conditions of Proposition 1 and the estimation algorithm can be used to estimate the parameters μ and σ^2 of the lognormal distribution.

To summarize, the constraints imposed by the lognormal distribution are specified by

$$g(\boldsymbol{\pi}) = \boldsymbol{C}.\Phi^{-1}(\boldsymbol{\pi}_S) = \boldsymbol{C}.\Phi^{-1}(\boldsymbol{S} \cdot \boldsymbol{\pi}) = \boldsymbol{0} \qquad (3.4\ .19)$$

with

$$\boldsymbol{C} = \boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}' \quad . \qquad (3.4\ .20)$$

The derivative of $g(\boldsymbol{\pi})$ with respect to $\boldsymbol{\pi}$ is

$$\boldsymbol{G}_\pi = \frac{\partial g(\boldsymbol{\pi})}{\partial \boldsymbol{\pi}} = \boldsymbol{C} \cdot \frac{\partial}{\partial \boldsymbol{\pi}} \Phi^{-1}(\boldsymbol{S} \cdot \boldsymbol{\pi}) \cdot \boldsymbol{S} \qquad (3.4\ .21)$$

that is equal to

$$\boldsymbol{G}_\pi = \boldsymbol{C} \cdot \left[\frac{1}{diag\left\{ \phi\left(\frac{\ln \boldsymbol{x} - \mu \cdot \boldsymbol{1}}{\sigma} \right) \right\}} \right] \boldsymbol{S} \ . \qquad (3.4\ .22)$$

The matrix $\boldsymbol{S}$ to be used is the same matrix as defined at the fitting of the Weibull or the log-logistic model.

Since $\boldsymbol{G}_\pi$ depends on μ and σ in the iterative procedure, these parameters will be estimated within the iterative stages and the final estimates will be obtained on convergence.

The SAS/IML program to fit a lognormal model to grouped survival data with staggered entry of policies appears in Appendix A.

3.4.6 Fitting of four survival distributions to the four histograms

Fitting of four Weibull distributions

Four Weibull distributions are to be fitted to the four histograms.

Consider four Weibull distributions with parameters (λ_1, α_1), (λ_2, α_2), (λ_3, α_3) and (λ_4, α_4) respectively.

Maximum likelihood estimation of the parameters is done subject to constraints imposed by the four Weibull distributions and further constraints that the λ_i's are equal and α_i's are equal.

The four Weibull models under constraints $\boldsymbol{\pi}_{S1}$, $\boldsymbol{\pi}_{S2}$, $\boldsymbol{\pi}_{S3}$, $\boldsymbol{\pi}_{S4}$, can be written from Equation 3.3 .14 as follows:

$$\begin{pmatrix} \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S1})\} \\ \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S2})\} \\ \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S3})\} \\ \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S4})\} \end{pmatrix} = \begin{pmatrix} \mathbf{1} & \ln\boldsymbol{x}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \ln\boldsymbol{x}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \ln\boldsymbol{x}_3 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \ln\boldsymbol{x}_4 \end{pmatrix} \cdot \begin{pmatrix} \ln\lambda_1 \\ \alpha_1 \\ \ln\lambda_2 \\ \alpha_2 \\ \ln\lambda_3 \\ \alpha_3 \\ \ln\lambda_4 \\ \alpha_4 \end{pmatrix} \tag{3.4 .23}$$

where

$$\boldsymbol{x}_1 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ x_{k-1} \end{pmatrix} \quad \boldsymbol{x}_2 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ \cdots \\ \cdots \\ x_{k-2} \end{pmatrix} \quad \boldsymbol{x}_3 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ \cdots \\ x_{k-3} \end{pmatrix} \quad \text{and} \quad \boldsymbol{x}_4 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ x_{k-4} \end{pmatrix}.$$

Equation 3.4 .23 is a **linear model** in the parameters $\ln\lambda_1, \alpha_1, \ln\lambda_2, \alpha_2, \ln\lambda_3, \alpha_3, \ln\lambda_4$ and α_4.

Maximum likelihood estimation of these parameters subject to further constraints that the λ_i's are equal and the α_i's are equal can be done similar to the fitting of one Weibull to the four histograms, when the following changes are made.

From Equation 3.4 .23 follows that the design matrix for the fitting of four Weibull distributions is

$$\boldsymbol{X}_1 = \begin{pmatrix} \mathbf{1} & \ln \boldsymbol{x}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \ln \boldsymbol{x}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \ln \boldsymbol{x}_3 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \ln \boldsymbol{x}_4 \end{pmatrix}.$$

The four Weibull models in Equation 3.4 .23 are equivalent to

$$\begin{aligned}
\underbrace{\left(\boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}'\right)}\cdot \begin{pmatrix} \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S1})\} \\ \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S2})\} \\ \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S3})\} \\ \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S4})\} \end{pmatrix} &= \mathbf{0} \\
\underbrace{\boldsymbol{C} \cdot \begin{pmatrix} \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S1})\} \\ \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S2})\} \\ \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S3})\} \\ \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S4})\} \end{pmatrix}} &= \mathbf{0} \\
\boldsymbol{g}(\boldsymbol{\pi}) &= \mathbf{0}
\end{aligned}$$

$\boldsymbol{C}$ is the projection matrix orthogonal to the columns of the design matrix $\boldsymbol{X}_1$. Note that $\boldsymbol{CX}_1 = \mathbf{0}$.

The constraints that the λ_i's are equal and the α_i's are equal are specified by

$$\boldsymbol{H}. \begin{pmatrix} \ln\lambda_1 \\ \alpha_1 \\ \ln\lambda_2 \\ \alpha_2 \\ \ln\lambda_3 \\ \alpha_3 \\ \ln\lambda_4 \\ \alpha_4 \end{pmatrix} = \mathbf{0} \qquad (3.4\ .24)$$

where

$$\boldsymbol{H} = \begin{pmatrix} 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}.$$

Equation 3.4 .24 is equivalent to

$$\begin{pmatrix} \ln\lambda_1 \\ \alpha_1 \\ \ln\lambda_2 \\ \alpha_2 \\ \ln\lambda_3 \\ \alpha_3 \\ \ln\lambda_4 \\ \alpha_4 \end{pmatrix} = (\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}' \cdot \begin{pmatrix} \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S1})\} \\ \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S2})\} \\ \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S3})\} \\ \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S4})\} \end{pmatrix} \tag{3.4 .25}$$

that is

$$\boldsymbol{H}.\begin{pmatrix} \ln\lambda_1 \\ \alpha_1 \\ \ln\lambda_2 \\ \alpha_2 \\ \ln\lambda_3 \\ \alpha_3 \\ \ln\lambda_4 \\ \alpha_4 \end{pmatrix} = \underbrace{\boldsymbol{H}.(\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}'} \cdot \begin{pmatrix} \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S1})\} \\ \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S2})\} \\ \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S3})\} \\ \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S4})\} \end{pmatrix}$$

or

$$\mathbf{0} = \boldsymbol{D} \cdot \begin{pmatrix} \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S1})\} \\ \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S2})\} \\ \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S3})\} \\ \ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_{S4})\} \end{pmatrix} \tag{3.4 .26}$$

Equation 3.4 .26 specifies the further constraints of equal parameters for the four histograms.

In the notation for staggered entry of policies, as described in Table 3.2 with four entry periods and k, the number of class intervals, equal to seven, matrix $\boldsymbol{D}$ is a 6×18 matrix.

A new matrix is formed that takes the further constraints into account. This matrix is created by concatenating the six rows of $\boldsymbol{D}$ to the 18 rows of $\boldsymbol{C}$. This new matrix is then used instead of the matrix $\boldsymbol{C}$ in further calculations.

Fitting of four log-logistic distributions

Four log-logistic distributions are to be fitted to the four histograms.

Consider four log-logistic distributions with parameters (λ_1, α_1), (λ_2, α_2), (λ_3, α_3) and (λ_4, α_4) respectively.

Maximum likelihood estimation of the parameters is done subject to constraints imposed by the four log-logistic distributions and further constraints that the λ_i's are equal and α_i's are equal.

The four log-logistic models under constraints $\boldsymbol{\pi}_{S1}$, $\boldsymbol{\pi}_{S2}$, $\boldsymbol{\pi}_{S3}$, $\boldsymbol{\pi}_{S4}$, can be written from Equation 3.3 .24 as follows:

$$\begin{pmatrix} \ln\{\boldsymbol{\pi}_{S1}\} - \ln\{\mathbf{1} - \boldsymbol{\pi}_{S1}\} \\ \ln\{\boldsymbol{\pi}_{S2}\} - \ln\{\mathbf{1} - \boldsymbol{\pi}_{S2}\} \\ \ln\{\boldsymbol{\pi}_{S3}\} - \ln\{\mathbf{1} - \boldsymbol{\pi}_{S3}\} \\ \ln\{\boldsymbol{\pi}_{S4}\} - \ln\{\mathbf{1} - \boldsymbol{\pi}_{S4}\} \end{pmatrix} = \begin{pmatrix} \mathbf{1} & \ln \boldsymbol{x}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \ln \boldsymbol{x}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \ln \boldsymbol{x}_3 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \ln \boldsymbol{x}_4 \end{pmatrix} \cdot \begin{pmatrix} \ln \lambda_1 \\ \alpha_1 \\ \ln \lambda_2 \\ \alpha_2 \\ \ln \lambda_3 \\ \alpha_3 \\ \ln \lambda_4 \\ \alpha_4 \end{pmatrix} \qquad (3.4\ .27)$$

where

$$\boldsymbol{x}_1 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ x_{k-1} \end{pmatrix} \quad \boldsymbol{x}_2 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ \cdots \\ \cdots \\ x_{k-2} \end{pmatrix} \quad \boldsymbol{x}_3 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ \cdots \\ x_{k-3} \end{pmatrix} \quad \text{and} \quad \boldsymbol{x}_4 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ x_{k-4} \end{pmatrix}.$$

Equation 3.4 .27 is a **linear model** in the parameters $\ln \lambda_1, \alpha_1, \ln \lambda_2, \alpha_2, \ln \lambda_3, \alpha_3, \ln \lambda_4$ and α_4.

Maximum likelihood estimation of these parameters subject to further constraints that the λ_i's are equal and the α_i's are equal can be done similar to the fitting of one log-logistic to the four histograms, when the following changes are made.

From Equation 3.4 .27 follows that the design matrix for the fitting of four log-logistic distributions is

$$\boldsymbol{X}_1 = \begin{pmatrix} \mathbf{1} & \ln \boldsymbol{x}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \ln \boldsymbol{x}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \ln \boldsymbol{x}_3 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \ln \boldsymbol{x}_4 \end{pmatrix}.$$

The four log-logistic models in Equation 3.4 .27 are equivalent to

$$\underbrace{\left(\boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}'\right)}\cdot \begin{pmatrix} \ln\{\boldsymbol{\pi}_{S1}\} - \ln\{\mathbf{1} - \boldsymbol{\pi}_{S1}\} \\ \ln\{\boldsymbol{\pi}_{S2}\} - \ln\{\mathbf{1} - \boldsymbol{\pi}_{S2}\} \\ \ln\{\boldsymbol{\pi}_{S3}\} - \ln\{\mathbf{1} - \boldsymbol{\pi}_{S3}\} \\ \ln\{\boldsymbol{\pi}_{S4}\} - \ln\{\mathbf{1} - \boldsymbol{\pi}_{S4}\} \end{pmatrix} = \mathbf{0}$$

$$\underbrace{\boldsymbol{C} \cdot \begin{pmatrix} \ln\{\boldsymbol{\pi}_{S1}\} - \ln\{\mathbf{1} - \boldsymbol{\pi}_{S1}\} \\ \ln\{\boldsymbol{\pi}_{S2}\} - \ln\{\mathbf{1} - \boldsymbol{\pi}_{S2}\} \\ \ln\{\boldsymbol{\pi}_{S3}\} - \ln\{\mathbf{1} - \boldsymbol{\pi}_{S3}\} \\ \ln\{\boldsymbol{\pi}_{S4}\} - \ln\{\mathbf{1} - \boldsymbol{\pi}_{S4}\} \end{pmatrix}}_{\boldsymbol{g}(\boldsymbol{\pi})} = \mathbf{0}$$

$$\boldsymbol{g}(\boldsymbol{\pi}) = \mathbf{0}$$

$\boldsymbol{C}$ is the projection matrix orthogonal to the columns of the design matrix $\boldsymbol{X}_1$. Note that $\boldsymbol{C}\boldsymbol{X}_1 = \mathbf{0}$.

The constraints that the λ_i's are equal and the α_i's are equal are specified by

$$\boldsymbol{H}. \begin{pmatrix} \ln \lambda_1 \\ \alpha_1 \\ \ln \lambda_2 \\ \alpha_2 \\ \ln \lambda_3 \\ \alpha_3 \\ \ln \lambda_4 \\ \alpha_4 \end{pmatrix} = \mathbf{0} \qquad (3.4\ .28)$$

where

$$\boldsymbol{H} = \begin{pmatrix} 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}.$$

Equation 3.4 .28 is equivalent to

$$\begin{pmatrix} \ln\lambda_1 \\ \alpha_1 \\ \ln\lambda_2 \\ \alpha_2 \\ \ln\lambda_3 \\ \alpha_3 \\ \ln\lambda_4 \\ \alpha_4 \end{pmatrix} = (\mathbf{X}_1{'}\mathbf{X}_1)^{-1}\mathbf{X}_1{'} \cdot \begin{pmatrix} \ln\{\boldsymbol{\pi}_{S1}\} - \ln\{\mathbf{1}-\boldsymbol{\pi}_{S1}\} \\ \ln\{\boldsymbol{\pi}_{S2}\} - \ln\{\mathbf{1}-\boldsymbol{\pi}_{S2}\} \\ \ln\{\boldsymbol{\pi}_{S3}\} - \ln\{\mathbf{1}-\boldsymbol{\pi}_{S3}\} \\ \ln\{\boldsymbol{\pi}_{S4}\} - \ln\{\mathbf{1}-\boldsymbol{\pi}_{S4}\} \end{pmatrix} \tag{3.4 .29}$$

that is

$$\boldsymbol{H}. \begin{pmatrix} \ln\lambda_1 \\ \alpha_1 \\ \ln\lambda_2 \\ \alpha_2 \\ \ln\lambda_3 \\ \alpha_3 \\ \ln\lambda_4 \\ \alpha_4 \end{pmatrix} = \underbrace{\boldsymbol{H}.(\mathbf{X}_1{'}\mathbf{X}_1)^{-1}\mathbf{X}_1{'}} \cdot \begin{pmatrix} \ln\{\boldsymbol{\pi}_{S1}\} - \ln\{\mathbf{1}-\boldsymbol{\pi}_{S1}\} \\ \ln\{\boldsymbol{\pi}_{S2}\} - \ln\{\mathbf{1}-\boldsymbol{\pi}_{S2}\} \\ \ln\{\boldsymbol{\pi}_{S3}\} - \ln\{\mathbf{1}-\boldsymbol{\pi}_{S3}\} \\ \ln\{\boldsymbol{\pi}_{S4}\} - \ln\{\mathbf{1}-\boldsymbol{\pi}_{S4}\} \end{pmatrix}$$

or

$$\mathbf{0} = \boldsymbol{D} \cdot \begin{pmatrix} \ln\{\boldsymbol{\pi}_{S1}\} - \ln\{\mathbf{1}-\boldsymbol{\pi}_{S1}\} \\ \ln\{\boldsymbol{\pi}_{S2}\} - \ln\{\mathbf{1}-\boldsymbol{\pi}_{S2}\} \\ \ln\{\boldsymbol{\pi}_{S3}\} - \ln\{\mathbf{1}-\boldsymbol{\pi}_{S3}\} \\ \ln\{\boldsymbol{\pi}_{S4}\} - \ln\{\mathbf{1}-\boldsymbol{\pi}_{S4}\} \end{pmatrix} \tag{3.4 .30}$$

Equation 3.4 .30 specifies the further constraints of equal parameters for the four histograms.

If k, the number of class intervals, is equal to seven, then matrix $\boldsymbol{D}$ is a 6×18 matrix.

A new matrix is formed that takes the further constraints into account. This matrix is created by concatenating the six rows of $\boldsymbol{D}$ to the 18 rows of $\boldsymbol{C}$. This new matrix is then used instead of the matrix $\boldsymbol{C}$ in further calculations.

Fitting of four lognormal distributions

Four lognormal distributions are to be fitted to the four histograms.

Consider four lognormal distributions with parameters $(\mu_1, {\sigma_1}^2)$, $(\mu_2, {\sigma_2}^2)$, $(\mu_3, {\sigma_3}^2)$ and $(\mu_4, {\sigma_4}^2)$ respectively.

Maximum likelihood estimation of the parameters is done subject to constraints imposed by the four lognormal distributions and further constraints that the μ_i's are equal and σ_i's are equal.

The four lognormal models under constraints $\boldsymbol{\pi}_{S1}$, $\boldsymbol{\pi}_{S2}$, $\boldsymbol{\pi}_{S3}$, $\boldsymbol{\pi}_{S4}$, can be written from Equation 3.3 .35 as follows:

$$\begin{pmatrix} \ln\{\Phi^{-1}(\boldsymbol{\pi}_{S1})\} \\ \ln\{\Phi^{-1}(\boldsymbol{\pi}_{S2})\} \\ \ln\{\Phi^{-1}(\boldsymbol{\pi}_{S3})\} \\ \ln\{\Phi^{-1}(\boldsymbol{\pi}_{S4})\} \end{pmatrix} = \begin{pmatrix} \mathbf{1} & \ln \boldsymbol{x}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \ln \boldsymbol{x}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \ln \boldsymbol{x}_3 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \ln \boldsymbol{x}_4 \end{pmatrix} \cdot \begin{pmatrix} -\frac{\mu_1}{\sigma_1} \\ \frac{1}{\sigma_1} \\ -\frac{\mu_2}{\sigma_2} \\ \frac{1}{\sigma_2} \\ -\frac{\mu_3}{\sigma_3} \\ \frac{1}{\sigma_3} \\ -\frac{\mu_4}{\sigma_4} \\ \frac{1}{\sigma_4} \end{pmatrix} \tag{3.4 .31}$$

where

$$\boldsymbol{x}_1 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ x_{k-1} \end{pmatrix} \quad \boldsymbol{x}_2 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ \cdots \\ \cdots \\ x_{k-2} \end{pmatrix} \quad \boldsymbol{x}_3 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ \cdots \\ x_{k-3} \end{pmatrix} \quad \text{and} \quad \boldsymbol{x}_4 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ x_{k-4} \end{pmatrix}.$$

Equation 3.4 .31 is a **linear model** in the parameters $\mu_1, {\sigma_1}^2, \mu_2, {\sigma_2}^2, \mu_3, {\sigma_3}^2, \mu_4$ and ${\sigma_4}^2$.

Maximum likelihood estimation of these parameters subject to further constraints that the μ_i's are equal and the σ_i's are equal can be done similar to the fitting of one lognormal to the four histograms, when the following changes are made.

From Equation 3.4 .31 follows that the design matrix for the fitting of four lognormal distributions is

$$\boldsymbol{X}_1 = \begin{pmatrix} 1 & \ln x_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \ln x_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \ln x_3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & \ln x_4 \end{pmatrix}.$$

The four lognormal models in Equation 3.4 .31 are equivalent to

$$\underbrace{\left(\boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\right)}\cdot \begin{pmatrix} \ln\{\Phi^{-1}(\pi_{S1})\} \\ \ln\{\Phi^{-1}(\pi_{S2})\} \\ \ln\{\Phi^{-1}(\pi_{S3})\} \\ \ln\{\Phi^{-1}(\pi_{S4})\} \end{pmatrix} = 0$$

$$\underbrace{\boldsymbol{C} \cdot \begin{pmatrix} \ln\{\Phi^{-1}(\pi_{S1})\} \\ \ln\{\Phi^{-1}(\pi_{S2})\} \\ \ln\{\Phi^{-1}(\pi_{S3})\} \\ \ln\{\Phi^{-1}(\pi_{S4})\} \end{pmatrix}}_{g(\boldsymbol{\pi})} = 0$$

$$g(\boldsymbol{\pi}) = 0$$

$\boldsymbol{C}$ is the projection matrix orthogonal to the columns of the design matrix $\boldsymbol{X}_1$. Note that $\boldsymbol{C}\boldsymbol{X}_1 = \boldsymbol{0}$.

The constraints that the μ_i's are equal and the σ_i's are equal are specified by

$$\boldsymbol{H}. \begin{pmatrix} -\frac{\mu_1}{\sigma_1} \\ \frac{1}{\sigma_1} \\ -\frac{\mu_2}{\sigma_2} \\ \frac{1}{\sigma_2} \\ -\frac{\mu_3}{\sigma_3} \\ \frac{1}{\sigma_3} \\ -\frac{\mu_4}{\sigma_4} \\ \frac{1}{\sigma_4} \end{pmatrix} = \boldsymbol{0} \qquad (3.4\ .32)$$

where

$$\boldsymbol{H} = \begin{pmatrix} 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}.$$

Equation 3.4 .32 is equivalent to

$$\begin{pmatrix} -\frac{\mu_1}{\sigma_1} \\ \frac{1}{\sigma_1} \\ -\frac{\mu_2}{\sigma_2} \\ \frac{1}{\sigma_2} \\ -\frac{\mu_3}{\sigma_3} \\ \frac{1}{\sigma_3} \\ -\frac{\mu_4}{\sigma_4} \\ \frac{1}{\sigma_4} \end{pmatrix} = (\boldsymbol{X_1}'\boldsymbol{X_1})^{-1}\boldsymbol{X_1}' \cdot \begin{pmatrix} \ln\{\Phi^{-1}(\pi_{S1})\} \\ \ln\{\Phi^{-1}(\pi_{S2})\} \\ \ln\{\Phi^{-1}(\pi_{S3})\} \\ \ln\{\Phi^{-1}(\pi_{S4})\} \end{pmatrix} \tag{3.4 .33}$$

that is

$$\boldsymbol{H}. \begin{pmatrix} -\frac{\mu_1}{\sigma_1} \\ \frac{1}{\sigma_1} \\ -\frac{\mu_2}{\sigma_2} \\ \frac{1}{\sigma_2} \\ -\frac{\mu_3}{\sigma_3} \\ \frac{1}{\sigma_3} \\ -\frac{\mu_4}{\sigma_4} \\ \frac{1}{\sigma_4} \end{pmatrix} = \underbrace{\boldsymbol{H}.(\boldsymbol{X_1}'\boldsymbol{X_1})^{-1}\boldsymbol{X_1}'} \cdot \begin{pmatrix} \ln\{\Phi^{-1}(\pi_{S1})\} \\ \ln\{\Phi^{-1}(\pi_{S2})\} \\ \ln\{\Phi^{-1}(\pi_{S3})\} \\ \ln\{\Phi^{-1}(\pi_{S4})\} \end{pmatrix}$$

or

$$\boldsymbol{0} = \boldsymbol{D} \cdot \begin{pmatrix} \ln\{\Phi^{-1}(\pi_{S1})\} \\ \ln\{\Phi^{-1}(\pi_{S2})\} \\ \ln\{\Phi^{-1}(\pi_{S3})\} \\ \ln\{\Phi^{-1}(\pi_{S4})\} \end{pmatrix} \tag{3.4 .34}$$

Equation 3.4 .34 specifies the further constraints of equal parameters for the four histograms.

If k, the number of class intervals, is equal to seven, then matrix $\boldsymbol{D}$ is a 6×18 matrix.

A new matrix is formed that takes the further constraints into account. This matrix is created by concatenating the six rows of $\boldsymbol{D}$ to the 18 rows of $\boldsymbol{C}$. This new matrix is then used instead of the matrix $\boldsymbol{C}$ in further calculations.

3.4.7 Fitting of a joint histogram to the four histograms

A joint histogram is to be fitted to the four histograms of the four relative frequency distributions under constraints imposed by the experimental design.

These constraints in matrix form are $\boldsymbol{G} \cdot \boldsymbol{\pi} = \mathbf{0}$ with $\boldsymbol{\pi}' = (\boldsymbol{\pi}_1',\boldsymbol{\pi}_2',\boldsymbol{\pi}_3',\boldsymbol{\pi}_4')$ where

$\boldsymbol{\pi}_1 = (\pi_{1,1}, \pi_{1,2}, \pi_{1,3}, ..., \pi_{1,k})'$ is a $k \times 1$ probability vector
$\boldsymbol{\pi}_2 = (\pi_{2,1}, \pi_{2,2}, \pi_{2,3}, ..., \pi_{2,k-1})'$ is a $(k-1) \times 1$ probability vector
$\boldsymbol{\pi}_3 = (\pi_{3,1}, \pi_{3,2}, \pi_{3,3}, ..., \pi_{3,k-2})'$ is a $(k-2) \times 1$ probability vector
$\boldsymbol{\pi}_4 = (\pi_{4,1}, \pi_{4,2}, \pi_{4,3}, ..., \pi_{4,k-3})'$ is a $(k-3) \times 1$ probability vector
and

$$\boldsymbol{G} = \begin{pmatrix} \boldsymbol{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -\boldsymbol{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -\boldsymbol{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -\boldsymbol{I} & \mathbf{0} \\ \\ \mathbf{0}' & 1 & 1 & 1 & 1 & \mathbf{0}' & -1 & -1 & -1 & \mathbf{0}' & 0 & 0 & \mathbf{0}' & 0 \\ \mathbf{0}' & 1 & 1 & 1 & 1 & \mathbf{0}' & 0 & 0 & 0 & \mathbf{0}' & -1 & -1 & \mathbf{0}' & 0 \\ \mathbf{0}' & 1 & 1 & 1 & 1 & \mathbf{0}' & 0 & 0 & 0 & \mathbf{0}' & 0 & 0 & \mathbf{0}' & -1 \\ \\ \mathbf{0}' & 0 & 1 & 0 & 0 & \mathbf{0}' & 0 & -1 & 0 & \mathbf{0}' & 0 & 0 & \mathbf{0}' & 0 \\ \mathbf{0}' & 1 & 0 & 0 & 0 & \mathbf{0}' & -1 & 0 & 0 & \mathbf{0}' & 0 & 0 & \mathbf{0}' & 0 \end{pmatrix}.$$

The function $g(\boldsymbol{\pi}) = \mathbf{0}$ satisfies the conditions of Proposition 1 and the estimation algorithm can be used to estimate the relative frequencies of the joint relative frequency distribution.

Note that the constraints $\boldsymbol{G} \cdot \boldsymbol{\pi} = g(\boldsymbol{\pi})$ is a **linear** function of $\boldsymbol{\pi}$. This implies that $\boldsymbol{G} = \dfrac{\partial g(\boldsymbol{\pi})}{\partial \boldsymbol{\pi}}$ and only a **single iteration** is needed in the iterative procedure to determine the MLE of $\boldsymbol{\pi}$ under the constraints $\boldsymbol{g}(\boldsymbol{\pi}) = \mathbf{0}$.

This MLE of $\boldsymbol{\pi}$ is

$$\widehat{\boldsymbol{\pi}}_c = \boldsymbol{p} - (\boldsymbol{GV})' (\boldsymbol{GVG}')^* \, \boldsymbol{g}(\boldsymbol{p})$$

with asymptotic variance-covariance matrix

$$cov\,(\widehat{\boldsymbol{\pi}}_c) = \boldsymbol{V} - (\boldsymbol{GV})' (\boldsymbol{GVG}')^* \, \boldsymbol{GV}.$$

The variance-covariance matrix $\boldsymbol{V}$ to be used is the estimated variance-covariance matrix of the multinomial distribution **for each entry period**.

$$\Longrightarrow \quad \widehat{\boldsymbol{V}} = \text{block}(\widehat{\boldsymbol{V}}_1, \widehat{\boldsymbol{V}}_2, \widehat{\boldsymbol{V}}_3, \widehat{\boldsymbol{V}}_4)$$

and

$$\widehat{\boldsymbol{V}}_i = \frac{1}{n_i} \left[diag(\boldsymbol{p}_i) - \boldsymbol{p}_i \boldsymbol{p}_i'\right] \quad i = 1, 2, 3, 4$$

Once the iterative procedure has converged, the estimated joint relative frequencies can be read off from the estimated vector of probabilities $\widehat{\pi}_c$. The histogram of the fitted joint relative frequency distribution is a representative image of the four histograms.

The SAS/IML program to fit a joint histogram to the four histograms of the entry groups appears in Appendix A.

3.4.8 Estimated survivor and hazard functions and percentiles

Once the parameters of the Weibull and log-logistic survival distributions have been estimated, estimated hazard and survivor functions and the odds of a lapse can be calculated for time t. Percentiles of these survival distributions can also be estimated.

Survival distribution

Weibull

Estimated hazard function

$$\widehat{h}(t) = \widehat{\lambda} \cdot \widehat{\alpha} \cdot t^{\widehat{\alpha}-1}$$

Estimated survivor function

$$\widehat{S}(t) = \exp(-\widehat{\lambda} \cdot t^{\widehat{\alpha}})$$

Estimated odds of a lapse

$$\widehat{odds}(t) = \frac{1-\widehat{S}(t)}{\widehat{S}(t)} = \exp(\widehat{\lambda} \cdot t^{\widehat{\alpha}-1})$$

Estimated percentiles

$$\widehat{t_p} = (\frac{1}{\widehat{\lambda}} \cdot \ln \frac{100}{100-p})^{\frac{1}{\widehat{\alpha}}}$$

Log-logistic

Estimated hazard function

$$\widehat{h}(t) = \frac{\widehat{\lambda} \cdot \widehat{\alpha} \cdot t^{\widehat{\alpha}-1}}{(1+\widehat{\lambda} \cdot t^{\widehat{\alpha}})}$$

Estimated survivor function

$$\widehat{S}(t) = \frac{1}{1+\widehat{\lambda} \cdot t^{\widehat{\alpha}}}$$

Estimated odds of a lapse

$$\widehat{odds}(t) = \frac{1-\widehat{S}(t)}{\widehat{S}(t)} = \widehat{\lambda} \cdot t^{\widehat{\alpha}}$$

Estimated percentiles

$$\widehat{t_p} = (\frac{1}{\widehat{\lambda}} \cdot \frac{p}{100-p})^{\frac{1}{\widehat{\alpha}}}$$

3.5 Simulation Studies

Simulations are used to compare the maximum likelihood estimation under constraints procedure with the standard maximum likelihood estimation procedure that is used by SAS.

A thousand sets of right-censored data with a fixed censoring time are simulated from each of three survival distributions, namely the Weibull, the log-logistic and the lognormal distributions. The percentage censoring is about 35% and censoring at continuous data as well as censoring at grouped data, the special case of interval-censored data, are considered.

For the **continuous case**, three groups of 20, 50, 100 and 200 observations are generated from each of three survival distributions, with different scale and shape parameters for the three groups. A scale parameter of λ=0.15 and a shape parameter of α=0.5 are selected for the first group generated from the Weibull(λ, α) and the log-logistic(λ, α) distributions, while a μ-value of 2 and a σ-value of 0.5 are used for the generation from the lognormal(μ, σ) distribution. The second and third groups use extreme λ-values of 10 and 30 and extreme α-values of 3 and 1.8 for generation from the Weibull and log-logistic distributions and μ-values of 5 and 3 and σ-values of 0.2 and 0.03 for generation from the lognormal distribution. In order to be able to apply the IML program for maximum likelihood estimation subject to constraints, the continuous data are grouped into intervals with boundaries the means of two adjacent observed survival times with frequency 1 in each interval. The frequency of the last open interval is equal to the number of censored survival times. The standard method of maximum likelihood estimation used by PROC LIFEREG of SAS is applied to the continuous data without grouping into such intervals.

For the **grouped data case**, three groups of
100 observations (grouped into five intervals),
200 observations (grouped into five intervals),
2000 observations (grouped into five intervals) and
2000 observations (grouped into ten intervals)
are generated from each of the three survival distributions, with parameters similar to the continuous case.

Programs to generate lifetime data and to run simulations with the technique of maximum likelihood estimation under constraints, appear in Appendix A. Maximum likelihood estimates by the standard method are also found using PROC LIFEREG of SAS. For comparison purposes, both estimation techniques are performed on the same set of simulated data in the same program. The means of the thousand simulated $\hat{\lambda}$- and $\hat{\alpha}$-values are computed as maximum likelihood estimators of the model parameters λ and α.

These estimators of the model parameters are considered to be significantly biased if the absolute difference between the model parameter and the estimator is greater than three standard errors of the mean. This criterium can be motivated as follows. For large sample sizes, the sampling distribution of the mean is approximately normally distributed and then the probability that the absolute difference between the model parameter and the estimator will lie within three standard errors of the mean, is 0.9999 . For small sample sizes (not assuming normality) the inequality of Chebyshev specifies that this probability is at least 0.8889 . Thus the confidence coefficient in this case will be between 0.8889 and 0.9999. A significantly biased estimator over-estimates the model parameter if its value is more than three standard errors of the mean to the right of the model parameter, and is an under-estimator if its value is more than three standard errors of the mean to the left of the model parameter.

Table 3.3 represents the simulation results for the continuous case for samples of various sizes when generating from a Weibull model.

Table 3.4 represents the simulation results for the grouped data case for samples of various sizes, classified in five or ten intervals, when generating from a Weibull model.

Table 3.5 represents the simulation results for the continuous case for samples of various sizes when generating from a log-logistic model.

Table 3.6 represents the simulation results for the grouped data case for samples of various sizes, classified in five or ten intervals, when generating from a log-logistic model.

Table 3.7 represents the simulation results for the continuous case for samples of various sizes when generating from a lognormal model.

Table 3.8 represents the simulation results for the grouped data case for samples of various sizes, classified in five or ten intervals, when generating from a lognormal model.

For generation from the Weibull and log-logistic distributions, the maximum likelihood estimates of the IML method compare very well to the maximum likelihood estimates of the SAS method. The biasness of these estimates is very small for values of λ and α that are usually used in practice, namely λ=0.15 and α=0.5 . For larger values of λ and α, the model parameters seem to be overestimated. The biasness of the estimates reduces when the sample size increases, as can be expected from maximum likelihood estimates which are asymptotically unbiased. The same conclusions can be made for generation from the lognormal distribution.

Table 3.3: **Weibull Simulation - Continuous Data**

Sample Size	Group	Model Parameter	Maximum Likelihood Estimator	Absolute difference |parameter-estimator|	3× Std Error	Significantly Biased
n=20	first	λ=0.15	$\hat{\lambda}_{SAS}$=0.14996619	0.00003381	0.0069663	No
			$\hat{\lambda}_{IML}$=0.14668896	0.00331104	0.0068990	No
		α=0.5	$\hat{\alpha}_{SAS}$ =0.54013638	0.04013638	0.0140021	Yes Over
			$\hat{\alpha}_{IML}$ =0.54593890	0.0459389	0.0140514	Yes Over
	second	λ=10	$\hat{\lambda}_{SAS}$=14.5457664	4.5457664	1.1514872	Yes Over
			$\hat{\lambda}_{IML}$=13.4422577	3.4422577	1.0386621	Yes Over
		α=3	$\hat{\alpha}_{SAS}$ =3.18959524	0.18959524	0.0717345	Yes Over
			$\hat{\alpha}_{IML}$ =3.09723538	0.09723538	0.0701570	Yes Over
	third	λ=30	$\hat{\lambda}_{SAS}$=120.665220	90.66522	48.796461	Yes Over
			$\hat{\lambda}_{IML}$=101.053564	71.053564	36.689586	Yes Over
		α=1.8	$\hat{\alpha}_{SAS}$ =1.95511189	0.15511189	0.0537395	Yes Over
			$\hat{\alpha}_{IML}$ =1.90197261	0.10197261	0.0527703	Yes Over
n=50	first	λ=0.15	$\hat{\lambda}_{SAS}$=0.14816260	0.0018374	0.0044372	No
			$\hat{\lambda}_{IML}$=0.14813664	0.00186336	0.0029639	No
		α=0.5	$\hat{\alpha}_{SAS}$ =0.51746332	0.01746332	0.0044459	Yes Over
			$\hat{\alpha}_{IML}$ =0.51743153	0.01743153	0.0079694	Yes Over
	second	λ=10	$\hat{\lambda}_{SAS}$=11.9257535	1.9257535	0.5200346	Yes Over
			$\hat{\lambda}_{IML}$=11.5073420	1.5073420	0.4952037	Yes Over
		α=3	$\hat{\alpha}_{SAS}$ =3.10165140	0.10165140	0.0447723	Yes Over
			$\hat{\alpha}_{IML}$ =3.05756831	0.05756831	0.0442101	Yes Over
	third	λ=30	$\hat{\lambda}_{SAS}$=41.4510357	11.4510357	3.4742741	Yes Over
			$\hat{\lambda}_{IML}$=39.5947272	9.5947272	3.3339734	Yes Over
		α=1.8	$\hat{\alpha}_{SAS}$ =1.85067820	0.0506782	0.0292124	Yes Over
			$\hat{\alpha}_{IML}$ =1.82857605	0.02857605	0.0289569	No

Table 3.3 (continued)

Sample Size	Group	Model Parameter	Maximum Likelihood Estimator	Absolute difference |parameter-estimator|	3× Std Error	Significantly Biased
n=100	first	λ=0.15	$\hat{\lambda}_{SAS}$=0.15177147	0.00177147	0.0032904	No
			$\hat{\lambda}_{IML}$=0.15250267	0.00250267	0.0021865	Yes Over
		α=0.5	$\hat{\alpha}_{SAS}$ =0.50761848	0.00761848	0.0032798	Yes Over
			$\hat{\alpha}_{IML}$ =0.50625014	0.00625014	0.0055814	Yes Over
	second	λ=10	$\hat{\lambda}_{SAS}$=10.5995925	0.5995925	0.3027174	Yes Over
			$\hat{\lambda}_{IML}$=10.3895455	0.3895455	0.2970639	Yes Over
		α=3	$\hat{\alpha}_{SAS}$ =3.02213801	0.02213801	0.0304752	No
			$\hat{\alpha}_{IML}$ =2.99591283	0.00408717	0.0305282	No
	third	λ=30	$\hat{\lambda}_{SAS}$=35.3120154	5.3120154	1.8132791	Yes Over
			$\hat{\lambda}_{IML}$=34.3043403	4.3043403	1.7604518	Yes Over
		α=1.8	$\hat{\alpha}_{SAS}$ =1.82845863	0.02845863	0.0201914	Yes Over
			$\hat{\alpha}_{IML}$ =1.81352486	0.01352486	0.0202521	No
n=200	first	λ=0.15	$\hat{\lambda}_{SAS}$=0.15180449	0.00180449	0.0023565	No
			$\hat{\lambda}_{IML}$=0.15250004	0.00250004	0.0002381	Yes Over
		α=0.5	$\hat{\alpha}_{SAS}$ =0.50087100	0.00087100	0.0038819	No
			$\hat{\alpha}_{IML}$ =0.49973979	0.00026021	0.0039027	No
	second	λ=10	$\hat{\lambda}_{SAS}$=10.1859060	0.185906	0.1980131	No
			$\hat{\lambda}_{IML}$=10.0588244	0.0588244	0.1955345	No
		α=3	$\hat{\alpha}_{SAS}$ =2.99899145	0.00100855	0.0215052	No
			$\hat{\alpha}_{IML}$ =2.98236429	0.01763571	0.0214703	No
	third	λ=30	$\hat{\lambda}_{SAS}$=32.2848790	2.284879	1.0252172	Yes Over
			$\hat{\lambda}_{IML}$=31.6653840	1.665384	1.0081139	Yes Under
		α=1.8	$\hat{\alpha}_{SAS}$ =1.85067820	0.0506782	0.0292124	Yes Over
			$\hat{\alpha}_{IML}$ =1.80067207	0.00067207	0.0141369	No

Table 3.4: **Weibull Simulation - Grouped Data**

Sample Size	Group	Model Parameter	Maximum Likelihood Estimator	Absolute difference |parameter-estimator|	3× Std Error	Significantly Biased
n=100 5 classes	first	λ=0.15	$\hat{\lambda}_{SAS}$=0.15535337	0.00535337	0.0049053	Yes Over
			$\hat{\lambda}_{IML}$=0.15535344	0.00535344	0.0049053	Yes Over
		α=0.5	$\hat{\alpha}_{SAS}$ =0.50497058	0.00497058	0.0083862	No
			$\hat{\alpha}_{IML}$ =0.50497044	0.00497044	0.0083862	No
	second	λ=10	$\hat{\lambda}_{SAS}$=10.8529348	0.8529348	0.3285815	Yes Over
			$\hat{\lambda}_{IML}$=10.8529223	0.8529223	0.3285812	Yes Over
		α=3	$\hat{\alpha}_{SAS}$ =3.04895140	0.04895140	0.0335870	Yes Over
			$\hat{\alpha}_{IML}$ =3.04895031	0.04895031	0.0335870	Yes Over
	third	λ=30	$\hat{\lambda}_{SAS}$=36.3075637	6.3075637	2.4547211	Yes Over
			$\hat{\lambda}_{IML}$=36.3075006	6.3075006	2.4547164	Yes Over
		α=1.8	$\hat{\alpha}_{SAS}$ =1.80779115	0.00779115	0.0257724	No
			$\hat{\alpha}_{IML}$ =1.80779049	0.00779049	0.0257723	No
n=200 5 classes	first	λ=0.15	$\hat{\lambda}_{SAS}$=0.15331594	0.00331594	0.0033041	Yes Over
			$\hat{\lambda}_{IML}$=0.15331603	0.00331603	0.0033041	Yes Over
		α=0.5	$\hat{\alpha}_{SAS}$ =0.50052889	0.00052889	0.0056775	No
			$\hat{\alpha}_{IML}$ =0.50052873	0.00052873	0.0056528	No
	second	λ=10	$\hat{\lambda}_{SAS}$=10.3602667	0.3602667	0.2201871	Yes Over
			$\hat{\lambda}_{IML}$=10.3602529	0.3602529	0.2201864	Yes Over
		α=3	$\hat{\alpha}_{SAS}$ =3.01886552	0.01886552	0.0237776	No
			$\hat{\alpha}_{IML}$ =3.01886427	0.01886427	0.0237774	No
	third	λ=30	$\hat{\lambda}_{SAS}$=33.6583074	3.6583074	1.4001587	Yes Over
			$\hat{\lambda}_{IML}$=33.6582478	3.6582478	1.4001549	Yes Over
		α=1.8	$\hat{\alpha}_{SAS}$ =1.81684031	0.01684031	0.0183728	No
			$\hat{\alpha}_{IML}$ =1.81683962	0.01683962	0.0183728	No

Table 3.4 (continued)

Sample Size	Group	Model Parameter	Maximum Likelihood Estimator	Absolute difference |parameter-estimator|	3× Std Error	Significantly Biased
n=2000 5 classes	first	λ=0.15	$\hat{\lambda}_{SAS}$=0.14992101	0.00007899	0.0010535	No
			$\hat{\lambda}_{IML}$=0.14992104	0.00007896	0.0010535	No
		α=0.5	$\hat{\alpha}_{SAS}$ =0.50057904	0.00057904	0.0018729	No
			$\hat{\alpha}_{IML}$ =0.50057897	0.00057897	0.0018729	No
	second	λ=10	$\hat{\lambda}_{SAS}$=10.0365052	0.0365052	0.0601692	No
			$\hat{\lambda}_{IML}$=10.0364959	0.0364959	0.0601688	No
		α=3	$\hat{\alpha}_{SAS}$ =3.00197758	0.00197758	0.0069176	No
			$\hat{\alpha}_{IML}$ =3.00197667	0.00197667	0.0069176	No
	third	λ=30	$\hat{\lambda}_{SAS}$=30.3212508	0.3212508	0.3504578	No
			$\hat{\lambda}_{IML}$=30.3212202	0.3212202	0.3504531	No
		α=1.8	$\hat{\alpha}_{SAS}$ =1.80186572	0.00186572	0.0055631	No
			$\hat{\alpha}_{IML}$ =1.80186532	0.00186532	0.0055631	No
n=2000 10 classes	first	λ=0.15	$\hat{\lambda}_{SAS}$=0.15028167	0.00028167	0.0008550	No
			$\hat{\lambda}_{IML}$=0.15028175	0.00028175	0.0008550	No
		α=0.5	$\hat{\alpha}_{SAS}$ =0.49991711	0.00008289	0.0014807	No
			$\hat{\alpha}_{IML}$ =0.49991693	0.00008307	0.0014807	No
	second	λ=10	$\hat{\lambda}_{SAS}$=10.0448501	0.0448501	0.0635568	No
			$\hat{\lambda}_{IML}$=10.0448461	0.0448461	0.0635568	No
		α=3	$\hat{\alpha}_{SAS}$ =3.00247119	0.00247119	0.0072035	No
			$\hat{\alpha}_{IML}$ =3.00247076	0.00247076	0.0072035	No
	third	λ=30	$\hat{\lambda}_{SAS}$=30.1459685	0.1459685	0.3150615	No
			$\hat{\lambda}_{IML}$=30.1459519	0.1459519	0.3150614	No
		α=1.8	$\hat{\alpha}_{SAS}$ =1.80002906	0.00002906	0.0049601	No
			$\hat{\alpha}_{IML}$ =1.80002879	0.00002879	0.0049601	No

Table 3.5: **Log-logistic Simulation - Continuous Data**

Sample Size	Group	Model Parameter	Maximum Likelihood Estimator	Absolute difference |parameter-estimator|	3× Std Error	Significantly Biased
n=20	first	λ=0.15	$\hat{\lambda}_{SAS}$=0.15145327	0.00145327	0.0079043	No
			$\hat{\lambda}_{IML}$=0.14796354	0.00203646	0.0075909	No
		α=0.5	$\hat{\alpha}_{SAS}$ =0.53723029	0.03723029	0.0123594	Yes Over
			$\hat{\alpha}_{IML}$ =0.53770964	0.03770964	0.0121689	Yes Over
	second	λ=10	$\hat{\lambda}_{SAS}$=20.1335459	10.1335459	3.2219037	Yes Over
			$\hat{\lambda}_{IML}$=18.2308334	8.2308334	3.1020546	Yes Over
		α=3	$\hat{\alpha}_{SAS}$ =3.28465392	0.28465392	0.0866151	Yes Over
			$\hat{\alpha}_{IML}$ =3.18022776	0.18022776	0.0841209	Yes Over
	third	λ=30	$\hat{\lambda}_{SAS}$=89.9665445	59.9665445	27.531053	Yes Over
			$\hat{\lambda}_{IML}$=86.0467171	56.0467171	34.960206	Yes Over
		α=1.8	$\hat{\alpha}_{SAS}$ =1.96003040	0.1600304	0.0475061	Yes Over
			$\hat{\alpha}_{IML}$ =1.91475105	0.11475105	0.0467700	Yes Over
n=50	first	λ=0.15	$\hat{\lambda}_{SAS}$=0.14899903	0.00100097	0.0050486	No
			$\hat{\lambda}_{IML}$=0.1485948	0.0014052	0.0050454	No
		α=0.5	$\hat{\alpha}_{SAS}$ =0.51827824	0.01827824	0.0074885	Yes Over
			$\hat{\alpha}_{IML}$ =0.51851833	0.01851833	0.0075098	Yes Over
	second	λ=10	$\hat{\lambda}_{SAS}$=12.8922606	2.8922606	0.7619717	Yes Over
			$\hat{\lambda}_{IML}$=12.3350982	2.3350982	0.7227566	Yes Over
		α=3	$\hat{\alpha}_{SAS}$ =3.13906180	0.13906180	0.0242147	Yes Over
			$\hat{\alpha}_{IML}$ =3.08961706	0.08961706	0.0497274	Yes Over
	third	λ=30	$\hat{\lambda}_{SAS}$=41.9117887	11.9117887	3.4527255	Yes Over
			$\hat{\lambda}_{IML}$=40.1049803	10.1049803	3.2656557	Yes Over
		α=1.8	$\hat{\alpha}_{SAS}$ =1.85576484	0.05576484	0.0278412	Yes Over
			$\hat{\alpha}_{IML}$ =1.83471490	0.03471490	0.0276355	Yes Over

Table 3.5 (continued)

Sample Size	Group	Model Parameter	Maximum Likelihood Estimator	Absolute difference |parameter-estimator|	3× Std Error	Significantly Biased
n=100	first	λ=0.15	$\hat{\lambda}_{SAS}$=0.14829547	0.00170453	0.0036117	No
			$\hat{\lambda}_{IML}$=0.14868802	0.00131198	0.0036063	No
		α=0.5	$\hat{\alpha}_{SAS}$ =0.50958388	0.00958388	0.0053112	Yes Over
			$\hat{\alpha}_{IML}$ =0.50898401	0.00898401	0.0052796	Yes Over
	second	λ=10	$\hat{\lambda}_{SAS}$=11.2448956	1.2448956	0.4224530	Yes Over
			$\hat{\lambda}_{IML}$=10.9567399	0.9567399	0.4093430	Yes Over
		α=3	$\hat{\alpha}_{SAS}$ =3.06515628	0.06515628	0.0349263	Yes Over
			$\hat{\alpha}_{IML}$ =3.03466202	0.03466202	0.0346883	No
	third	λ=30	$\hat{\lambda}_{SAS}$=34.5809933	4.5809933	1.5117978	Yes Over
			$\hat{\lambda}_{IML}$=33.6560830	3.6560830	1.4586887	Yes Over
		α=1.8	$\hat{\alpha}_{SAS}$ =1.82539203	0.02539203	0.0178437	Yes Over
			$\hat{\alpha}_{IML}$ =1.81188676	0.01188676	0.0176387	No
n=200	first	λ=0.15	$\hat{\lambda}_{SAS}$=0.15012895	0.00012895	0.0025574	No
			$\hat{\lambda}_{IML}$=0.15073248	0.00073248	0.0025749	No
		α=0.5	$\hat{\alpha}_{SAS}$ =0.50478209	0.00478209	0.0035717	Yes Over
			$\hat{\alpha}_{IML}$ =0.50392805	0.00392805	0.0035856	Yes Over
	second	λ=10	$\hat{\lambda}_{SAS}$=10.5491323	0.5491323	0.2693544	Yes Over
			$\hat{\lambda}_{IML}$=10.3753797	0.3753797	0.2650004	Yes Over
		α=3	$\hat{\alpha}_{SAS}$ =3.02213542	0.02213542	0.0236147	No
			$\hat{\alpha}_{IML}$ =3.00137177	0.00137177	0.0236348	No
	third	λ=30	$\hat{\lambda}_{SAS}$=32.4239353	2.4239353	0.9744693	Yes Over
			$\hat{\lambda}_{IML}$=31.8716829	1.8716829	0.9617250	Yes Over
		α=1.8	$\hat{\alpha}_{SAS}$ =1.81737307	0.01737307	0.0132893	Yes Over
			$\hat{\alpha}_{IML}$ =1.80795453	0.00795453	0.0133557	No

Table 3.6: **Log-logistic Simulation - Grouped Data**

Sample Size	Group	Model Parameter	Maximum Likelihood Estimator	Absolute difference |parameter-estimator|	3× Std Error	Significantly Biased
n=100 5 classes	first	λ=0.15	$\hat{\lambda}_{SAS}$=0.16521023	0.01521023	0.0094249	Yes Over
			$\hat{\lambda}_{IML}$=0.16521040	0.01521040	0.0094250	Yes Over
		α=0.5	$\hat{\alpha}_{SAS}$ =0.55605556	0.05605556	0.0124358	Yes Over
			$\hat{\alpha}_{IML}$ =0.55605532	0.05605532	0.0124358	Yes Over
	second	λ=10	$\hat{\lambda}_{SAS}$=11.4755280	1.475528	0.4744770	Yes Over
			$\hat{\lambda}_{IML}$=11.4755188	1.4755188	0.4744766	Yes Over
		α=3	$\hat{\alpha}_{SAS}$ =3.05054222	0.05054222	0.0386571	Yes Over
			$\hat{\alpha}_{IML}$ =3.05054135	0.05054135	0.0386571	Yes Over
	third	λ=30	$\hat{\lambda}_{SAS}$=35.5083302	5.5083302	1.8496092	Yes Over
			$\hat{\lambda}_{IML}$=35.5082841	5.5082841	1.8496053	Yes Over
		α=1.8	$\hat{\alpha}_{SAS}$ =1.82348176	0.02348176	0.0213981	Yes Over
			$\hat{\alpha}_{IML}$ =1.82348122	0.02348122	0.0213981	Yes Over
n=200 5 classes	first	λ=0.15	$\hat{\lambda}_{SAS}$=0.15312020	0.0031202	0.0061112	No
			$\hat{\lambda}_{IML}$=0.15312036	0.00312036	0.0061112	No
		α=0.5	$\hat{\alpha}_{SAS}$ =0.55197887	0.05197887	0.0089604	Yes Over
			$\hat{\alpha}_{IML}$ =0.55197863	0.05197863	0.0089604	Yes Over
	second	λ=10	$\hat{\lambda}_{SAS}$=10.4917160	0.4917160	0.2787344	Yes Over
			$\hat{\lambda}_{IML}$=10.4917107	0.4917107	0.2787339	Yes Over
		α=3	$\hat{\alpha}_{SAS}$ =3.01591703	0.01591703	0.0260622	No
			$\hat{\alpha}_{IML}$ =3.01591651	0.01591651	0.0260621	No
	third	λ=30	$\hat{\lambda}_{SAS}$=33.0284088	3.0284088	1.1623553	Yes Over
			$\hat{\lambda}_{IML}$=33.0283646	3.0283646	1.1623533	Yes Over
		α=1.8	$\hat{\alpha}_{SAS}$ =1.81587417	0.01587417	0.0157598	Yes Over
			$\hat{\alpha}_{IML}$ =1.81587357	0.01587357	0.0157598	Yes Over

Table 3.6 (continued)

Sample Size	Group	Model Parameter	Maximum Likelihood Estimator	Absolute difference |parameter-estimator|	3× Std Error	Significantly Biased
n=2000 5 classes	first	λ=0.15	$\hat{\lambda}_{SAS}$=0.14345146	0.00654854	0.0017930	Yes Under
			$\hat{\lambda}_{IML}$=0.14345176	0.00654824	0.0017930	Yes Under
		α=0.5	$\hat{\alpha}_{SAS}$ =0.54985916	0.04985916	0.0028193	Yes Over
			$\hat{\alpha}_{IML}$ =0.54985868	0.04985868	0.0028193	Yes Over
	second	λ=10	$\hat{\lambda}_{SAS}$=10.0348477	0.0348477	0.0871307	No
			$\hat{\lambda}_{IML}$=10.0348476	0.0348476	0.0871307	No
		α=3	$\hat{\alpha}_{SAS}$ =3.00041664	0.00041664	0.0089861	No
			$\hat{\alpha}_{IML}$ =3.00041662	0.00041662	0.0089861	No
	third	λ=30	$\hat{\lambda}_{SAS}$=30.1441862	0.1441862	0.3068751	No
			$\hat{\lambda}_{IML}$=30.1441400	0.1441400	0.3068718	No
		α=1.8	$\hat{\alpha}_{SAS}$ =1.80061919	0.00061919	0.0046998	No
			$\hat{\alpha}_{IML}$ =1.80061848	0.00061848	0.0046998	No
n=2000 10 classes	first	λ=0.15	$\hat{\lambda}_{SAS}$=0.14345099	0.00654901	0.0016850	Yes Under
			$\hat{\lambda}_{IML}$=0.14345135	0.00654865	0.0016850	Yes Under
		α=0.5	$\hat{\alpha}_{SAS}$ =0.54915511	0.04915511	0.0026267	Yes Over
			$\hat{\alpha}_{IML}$ =0.54915455	0.04915455	0.0026267	Yes Over
	second	λ=10	$\hat{\lambda}_{SAS}$=10.0650169	0.0650169	0.0885605	No
			$\hat{\lambda}_{IML}$=10.0650121	0.0650121	0.0885596	No
		α=3	$\hat{\alpha}_{SAS}$ =3.00168243	0.00168243	0.0084723	No
			$\hat{\alpha}_{IML}$ =3.00168201	0.00168201	0.0084722	No
	third	λ=30	$\hat{\lambda}_{SAS}$=30.3229680	0.3229680	0.2967881	Yes Over
			$\hat{\lambda}_{IML}$=30.3229672	0.3229672	0.2967879	Yes Over
		α=1.8	$\hat{\alpha}_{SAS}$ =1.80166754	0.00166754	0.0049601	No
			$\hat{\alpha}_{IML}$ =1.80166752	0.00166752	0.0049601	No

Table 3.7: **Lognormal Simulation - Continuous Data**

Sample Size	Group	Model Parameter	Maximum Likelihood Estimator	Absolute difference |parameter-estimator|	3× Std Error	Significantly Biased
n=20	first	μ=2	$\hat{\mu}_{SAS}$ =2.01370552	0.01370552	0.0132221	Yes Over
			$\hat{\mu}_{IML}$ =2.01849401	0.01849401	0.0138065	Yes Over
		σ=0.5	$\hat{\sigma}_{SAS}$ =0.48826184	0.01173816	0.0114902	Yes Over
			$\hat{\sigma}_{IML}$ =0.50773538	0.00773538	0.0124704	No
	second	μ=5	$\hat{\mu}_{SAS}$ =5.00535672	0.00535672	0.0053853	No
			$\hat{\mu}_{IML}$ =5.0081579	0.0081579	0.0057089	Yes Over
		σ=0.2	$\hat{\sigma}_{SAS}$ =0.19530672	0.00469328	0.0047150	No
			$\hat{\sigma}_{IML}$ =0.20488857	0.00488857	0.0051035	No
	third	μ=3	$\hat{\mu}_{SAS}$ =3.00032999	0.00032999	0.0009290	No
			$\hat{\mu}_{IML}$ =3.00101803	0.00101803	0.0009774	Yes Over
		σ=0.03	$\hat{\sigma}_{SAS}$ =0.02910627	0.00089373	0.0007827	Yes Under
			$\hat{\sigma}_{IML}$ =0.03096809	0.00096809	0.0008820	Yes Over
n=50	first	μ=2	$\hat{\mu}_{SAS}$ =1.99992428	0.00145327	0.0078302	No
			$\hat{\mu}_{IML}$ =2.0009279	0.0009279	0.0079442	No
		σ=0.5	$\hat{\sigma}_{SAS}$ =0.49328237	0.00671763	0.0070923	No
			$\hat{\sigma}_{IML}$ =0.50120403	0.00120403	0.0072662	No
	second	μ=5	$\hat{\mu}_{SAS}$ =5.00275189	0.00275189	0.0033918	No
			$\hat{\mu}_{IML}$ =5.00361253	0.00361253	0.0034631	Yes Over
		σ=0.2	$\hat{\sigma}_{SAS}$ =0.19899201	0.00100799	0.0031028	No
			$\hat{\sigma}_{IML}$ =0.20325814	0.00325814	0.0032217	Yes Over
	third	μ=3	$\hat{\mu}_{SAS}$ =3.00035126	0.00035126	0.0005712	No
			$\hat{\mu}_{IML}$ =3.00060422	0.00060422	0.0005883	Yes Over
		σ=0.03	$\hat{\sigma}_{SAS}$ =0.0295653	0.0004347	0.0005112	No
			$\hat{\sigma}_{IML}$ =0.03032833	0.00032833	0.0005343	No

Table 3.7 (continued)

Sample Size	Group	Model Parameter	Maximum Likelihood Estimator	Absolute difference |parameter-estimator|	3× Std Error	Significantly Biased
n=100	first	μ=2	$\hat{\mu}_{SAS}$ =1.99791343	0.00208657	0.0056862	No
			$\hat{\mu}_{IML}$ =1.99824256	0.00175744	0.0057267	No
		σ=0.5	$\hat{\sigma}_{SAS}$ =0.49740803	0.00259197	0.0050457	No
			$\hat{\sigma}_{IML}$ =0.50205305	0.00205305	0.0051075	No
	second	μ=5	$\hat{\mu}_{SAS}$ =5.00073008	0.00073008	0.0023390	No
			$\hat{\mu}_{IML}$ =5.00106256	0.00106256	0.0023610	No
		σ=0.2	$\hat{\sigma}_{SAS}$ =0.19929321	0.00070679	0.0020387	No
			$\hat{\sigma}_{IML}$ =0.20155982	0.00155982	0.0020783	No
	third	μ=3	$\hat{\mu}_{SAS}$ =3.00029441	0.00029441	0.0003854	No
			$\hat{\mu}_{IML}$ =3.00041118	0.00041118	0.0003908	Yes Over
		σ=0.03	$\hat{\sigma}_{SAS}$ =0.02996311	0.00003689	0.0003504	No
			$\hat{\sigma}_{IML}$ =0.03038029	0.00038029	0.0035913	No
n=200	first	μ=2	$\hat{\mu}_{SAS}$ =1.99963933	0.00036067	0.0038171	No
			$\hat{\mu}_{IML}$ =1.99974197	0.00025803	0.0038318	No
		σ=0.5	$\hat{\sigma}_{SAS}$ =0.49820589	0.00179411	0.0034412	No
			$\hat{\sigma}_{IML}$ =0.50103108	0.00103108	0.0034835	No
	second	μ=5	$\hat{\mu}_{SAS}$ =5.00034985	0.00034985	0.0016695	No
			$\hat{\mu}_{IML}$ =5.00050944	0.00050944	0.0016796	No
		σ=0.2	$\hat{\sigma}_{SAS}$ =0.19992307	0.00007693	0.0014759	No
			$\hat{\sigma}_{IML}$ =0.20115462	0.00115462	0.0014906	No
	third	μ=3	$\hat{\mu}_{SAS}$ =3.00001599	0.00001599	0.0002696	No
			$\hat{\mu}_{IML}$ =3.00007712	0.00007712	0.0002721	No
		σ=0.03	$\hat{\sigma}_{SAS}$ =0.02994986	0.00005014	0.0002402	No
			$\hat{\sigma}_{IML}$ =0.03019125	0.00019125	0.0002436	No

Table 3.8: **Lognormal Simulation - Grouped Data**

Sample Size	Group	Model Parameter	Maximum Likelihood Estimator	Absolute difference |parameter-estimator|	3× Std Error	Significantly Biased
n=100 5 classes	first	μ=2	$\hat{\mu}_{SAS}$ =2.00192369	0.00192369	0.0055787	No
			$\hat{\mu}_{IML}$ =2.00192375	0.00192375	0.0055787	No
		σ=0.5	$\hat{\sigma}_{SAS}$ =0.50117455	0.00117455	0.0060003	No
			$\hat{\sigma}_{IML}$ =0.5011747	0.0011747	0.0060003	No
	second	μ=5	$\hat{\mu}_{SAS}$ =5.00401992	0.00401992	0.0024615	Yes Over
			$\hat{\mu}_{IML}$ =5.00401995	0.00401995	0.0024615	Yes Over
		σ=0.2	$\hat{\sigma}_{SAS}$ =0.20294216	0.00294216	0.0028697	Yes Over
			$\hat{\sigma}_{IML}$ =0.20294222	0.00294222	0.0028697	Yes Over
	third	μ=3	$\hat{\mu}_{SAS}$ =3.00033514	0.00033514	0.0004122	No
			$\hat{\mu}_{IML}$ =3.00033514	0.00033514	0.0004122	No
		σ=0.03	$\hat{\sigma}_{SAS}$ =0.03061107	0.00061107	0.0005295	Yes Over
			$\hat{\sigma}_{IML}$ =0.03061108	0.00061108	0.0005295	Yes Over
n=200 5 classes	first	μ=2	$\hat{\mu}_{SAS}$ =2.00140175	0.00140175	0.0039281	No
			$\hat{\mu}_{IML}$ =2.00140179	0.00140179	0.0039281	No
		σ=0.5	$\hat{\sigma}_{SAS}$ =0.5032537	0.0032537	0.0042402	No
			$\hat{\sigma}_{IML}$ =0.50325377	0.00325377	0.0042402	No
	second	μ=5	$\hat{\mu}_{SAS}$ =4.99932522	0.00067478	0.0016349	No
			$\hat{\mu}_{IML}$ =4.99932524	0.00067476	0.0016349	No
		σ=0.2	$\hat{\sigma}_{SAS}$ =0.19964898	0.00035102	0.0018044	No
			$\hat{\sigma}_{IML}$ =0.19964903	0.00035097	0.0018044	No
	third	μ=3	$\hat{\mu}_{SAS}$ =3.00019211	0.00019211	0.0002874	No
			$\hat{\mu}_{IML}$ =3.00019211	0.00019211	0.0002874	No
		σ=0.03	$\hat{\sigma}_{SAS}$ =0.03033419	0.00033419	0.0003491	No
			$\hat{\sigma}_{IML}$ =0.0303342	0.0003342	0.0003491	No

Table 3.8 (continued)

Sample Size	Group	Model Parameter	Maximum Likelihood Estimator	Absolute difference |parameter-estimator|	3× Std Error	Significantly Biased
n=2000 5 classes	first	μ=2	$\hat{\mu}_{SAS}$ =1.99936468	0.00063532	0.0012531	No
			$\hat{\mu}_{IML}$ =1.99936468	0.00063532	0.0012531	No
		σ=0.5	$\hat{\sigma}_{SAS}$ =0.49931475	0.00068525	0.0013452	No
			$\hat{\sigma}_{IML}$ =0.49931475	0.00068525	0.0013452	No
	second	μ=5	$\hat{\mu}_{SAS}$ =4.99993747	0.00006253	0.0005217	No
			$\hat{\mu}_{IML}$ =4.99993747	0.00006253	0.0005217	No
		σ=0.2	$\hat{\sigma}_{SAS}$ =0.1999748	0.0000252	0.0005898	No
			$\hat{\sigma}_{IML}$ =0.1999748	0.0000252	0.0005898	No
	third	μ=3	$\hat{\mu}_{SAS}$ =3.00004037	0.00004037	0.0000872	No
			$\hat{\mu}_{IML}$ =3.00004038	0.00004038	0.0000872	No
		σ=0.03	$\hat{\sigma}_{SAS}$ =0.02997699	0.00002301	0.0001070	No
			$\hat{\sigma}_{IML}$ =0.02997701	0.00002299	0.0001070	No
n=2000 10 classes	first	μ=2	$\hat{\mu}_{SAS}$ =1.99950829	0.00049171	0.0012132	No
			$\hat{\mu}_{IML}$ =1.99950833	0.00049167	0.0012132	No
		σ=0.5	$\hat{\sigma}_{SAS}$ =0.50012979	0.00012979	0.0012959	No
			$\hat{\sigma}_{IML}$ =0.50012988	0.00012988	0.0012959	No
	second	μ=5	$\hat{\mu}_{SAS}$ =5.00003176	0.00003176	0.0000509	No
			$\hat{\mu}_{IML}$ =5.00003186	0.00003186	0.0000509	No
		σ=0.2	$\hat{\sigma}_{SAS}$ =0.20016008	0.00016008	0.0005625	No
			$\hat{\sigma}_{IML}$ =0.20016027	0.00016027	0.0005625	No
	third	μ=3	$\hat{\mu}_{SAS}$ =3.00003361	0.00003361	0.0000870	No
			$\hat{\mu}_{IML}$ =3.00003362	0.00003362	0.0000870	No
		σ=0.03	$\hat{\sigma}_{SAS}$ =0.03008051	0.00008051	0.0001053	No
			$\hat{\sigma}_{IML}$ =0.03008052	0.00008052	0.0001053	No