UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Fitting of survival functions for grouped data on insurance policies

by

Elizabeth Magrietha Louw

Submitted in fulfilment of part of the requirements
for the degree of Doctor of Philosophiae
in the Faculty of Natural and Agricultural Sciences
University of Pretoria

Supervisor: Professor N.A.S. Crowther

January 2002

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Preface

I am extremely grateful to my supervisor, Professor N.A.S. Crowther, for his careful supervision, positive guidance and constructive comments throughout.

I appreciated the enthusiastic encouragement of my family and friends. The patience and understanding and continuous support of my husband, Johann, and daughters, Hannelie and Annemarie, kept me going.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Contents

List of Tables

List of Figures

Chapter 1

INTRODUCTION

1.1 Background

Survival data consist of a response variable that measures the duration of time until a specified event occurs, and a set of covariates that can be either discrete or continuous. Some observations are right-censored, that is, the time until the event occurs, is not observed due to withdrawal or termination of the study. Thus survival data consist of a response variable (duration of time), as well as an additional variable indicating which responses are **observed** survival times and which are **censored** times and some covariates.

In this thesis, a structured framework for the analysis of policy survival times has been developed. Such a framework must be embedded in the experimental design. The lifetime of a policy is measured from the inception date up to the lapsing date or a cut-off point. If the lapsing date is prior to a cut-off point of the study, determined in advance, then the lifetime is **observed** (an uncensored observation). If a policy is still in force (alive) when the cut-off point is reached, the lifetime of this policy is said to be **right-censored**. This thesis focuses on **grouped data** for lifetimes. This scenario has become extremely important, not only for application in the actuarial context, but also in other fields.

In this thesis, the analysis takes account of the actual lifetime (duration) of the policy rather than just recording the fact that the policy lapsed or was still alive after (say) twelve months. In other words, the response variable, lifetime, is a continuous one, and the whole distribution of lifetimes can be used. This is in contrast to a categorical response where a loglinear logit model can be used to express the relationship among the categorical response and some covariates at a fixed time point of twelve months, conditional on a restricted experimental design where all the policies must have an exposure of at least twelve months when investigating the lapses of policies in the first year. There is no such restrictions in the more general experimental design used in this thesis where all the policies can be used in the analysis, even those policies with inception dates very close to the cut-off point.

All the policies that are written during a certain month and are recorded at the end of that month as that month's business, can be considered as a group of policies with a common starting point or entry time. In this way policies written in different months or time-periods have different entry times. This is called **staggered entry** of policies.

The policy lifetimes (observed and censored lifetimes) are typically of a discrete nature and principles of grouped data have to be taken into account when fitting lifetime distributions to such survival data. Statistical analysing techniques for this kind of grouped data are currently insufficient.

1.2 The Research Problem

The aim of the research is the statistical modelling of parametric survival distributions of **grouped survival data** of long- and shortterm policies in the insurance industry, by means of a method of maximum likelihood estimation **subject to constraints**. The statistical methodology is developed from previous research on maximum likelihood estimation subject to constraints (refer to [11, 30]) and special attention is given to the **staggered entry** of policies and the fitting of **parametric** regression models.

Much literature exist on the fitting of survival distributions, mainly in connection with non-parametric analyses or the semi-parametric proportional hazards regression model of Cox [6, 23, 41]. Theoretical results for continuous data of lifetimes are well developed. Corresponding results for grouped data of lifetimes or the so-called interval-censored data are available only in special cases, and then again mainly in the non-parametric and semi-parametric case (refer to [33, 15, 14]). Only in recent years interval censoring in parametric model fitting were addressed by [27, 22, 14, 36]. This was made possible by modern computing power and sophisticated statistical programming languages.

The standard method of estimation to be used in the literature is to maximize the partial likelihood (at Cox's model) or the full likelihood (at parametric regression models). The estimates of the parameters are then found by maximizing the likelihood function numerically. The methodology of maximum likelihood estimation subject to constraints, used in this thesis, leads to **explicit expressions** for the estimates of the parameters, as well as for approximated variances and covariances of the estimates, which gives **exact** maximum likelihood estimates of the parameters. This makes direct extension to more complex designs feasible.

Once the parameters of the survival distributions have been estimated, estimated hazard and survivor functions, odds of a lapse, **odds ratios** and **hazard ratios** at time t can be directly calculated, as well as estimated percentiles for the fitted survival distributions. These estimates form the statistical foundation for scientific decisionmaking with respect to actuarial design, maintenance and marketing of insurance policies.

The intrinsic value of the methodology described in this thesis is of fundamental importance to the actuarial science. The statistical modelling offers parametric models for survival distributions, in contrast with non-parametric models that are used commonly in the actuarial profession. This leads to more accurate estimation procedures. When the parametric models provide a good fit to data, they tend to give more precise estimates of the quantities of interest such as odds ratios, hazard ratios or median lifetimes. Estimates of these quantities will tend to have smaller standard errors than they would in the absence of distributional assumptions.

Data from the insurance industry are extensive data sets with very large sample sizes. The focus in this thesis will be on the estimation of lifetime distributions, based on a large sample of lifetimes of policies that are grouped into intervals of lifetimes. The variance- covariance matrix can be estimated as part of this maximum likelihood estimation procedure subject to constraints, but no meaningful interpretation can be given to the exceedance probability values that are associated with the very small standard errors of the parameter estimates due to the large sample size. Therefore discrepancy values will be used to evaluate the fitting of the lifetime distributions.

1.3 Generalization of the Statistical Procedures

One important aspect of the research field of survival analysis is the wide range of types of application. Although the methodology in this thesis is developed specifically for the insurance industry, it may be applied in the normal context of research and scientific decisionmaking, that includes for example survival distributions for the medical, biological, engineering, econometric and sociological sciences.

1.4 Future Research

The potential for extending the methodology to other realistic practical application is unlimited. This can be to the advance of the insurance industry in general. The models should reflect an interactive adaptability for direct application in practice by salesforce (the marketing people on ground level), as well as for actuarial planning.

1.5 Outlay of chapters

In chapter two, basic survival functions are defined and a description of the survival distributions that are used in this thesis, together with their properties, are given. The construction of different likelihood functions is also discussed.

In chapter three, parametric models are fitted to a single sample of survival data. The standard method of maximum likelihood estimation is reviewed. This is followed by the methodology of maximum likelihood estimation subject to constraints, applied in two different scenarios: a fixed censoring time as well as staggered entry of policies. The chapter is concluded with simulation studies to compare the estimates of the parameters obtained by the standard and the proposed estimating procedures.

In chapter four parametric regression models are fitted and important indicators of the effect of the covariates are defined such as risk scores (hazard ratios) and indices (odds ratios).

The developed theoretical results are applied in chapter five to a typical data set from an insurance company.

Chapter six concludes with a summary of the most important results in this thesis and focuses on application of these results in practice, especially in the actuarial industry.

All computer programs that were written in the SAS/IML language to perform the new methodology are given in appendix A.

The abstract appears at the end of the thesis, just after the references.

Chapter 2

SURVIVAL FUNCTIONS AND DISTRIBUTIONS

2.1 Introduction

The aim of this chapter is to introduce notation and survival concepts and to summarize properties of the survival distributions to be used in this thesis.

2.2 Notation

2.2.1 Right-censored continuous survival data

Let X be a nonnegative continuous random variable denoting the lifetime (or survival time) of a policy. 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 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 said to be **right-censored**.

A fixed censoring time

Consider the simplest scenario where all policies enter the study at the **same time**. Let C be the fixed termination date of the study, determined in advance. C is then the preassigned **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 $(\delta_j = 0)$. This type of censoring is known as **Type I right-censoring**.

Staggered entry of policies

Staggered entry of policies occur when policies enter the study at **different times**.

Define C_j as 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.

Only pairs $(T_1, \delta_1), (T_2, \delta_2), ..., (T_n, \delta_n)$ can be observed where

$$T_j = min(X_j, C_j) \text{ for the } j^{th} \text{ policy}$$

$$\delta_j = \begin{cases} 1 & \text{if} \quad X_j \leq C_j \quad \text{, that is, } X_j \text{ is not censored} \\ 0 & \text{if} \quad X_j > C_j \quad \text{, that is, } X_j \text{ is censored} \end{cases}$$

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

Random entries to the study are assumed. This type of censoring is known as **random right-censoring**. Type I right-censoring is a special case of the random censoring model by simply setting C_i=C. With random censoring the crucial **assumption** that X_i and C_i are independent, is made. This assumption means that censoring is not related to any factors associated with the actual survival time. This is called **non-informative censoring**. A graphical test to examine the assumption of non-informative censoring is given by [5, page

274].

2.2.2 Discrete data

Sometimes the lifetime T is a discrete random variable. Discrete lifetimes arise due to rounding off time measurements. Lifetimes of policies are usually measured on a discrete time-scale (to the nearest month).

If the lifetimes of the policies are distinct and $t_1 < t_2 < ... < t_n$ denotes the ordered lifetimes, then $R(t_i)$ is the risk set at time t_i, that is the set of policies that are still alive at a time **just prior** to t_i $\quad i = 1, 2, ..., n$.

To allow for possible ties in the data, suppose that the lapses occur at D distinct times $t_1 < t_2 < ... < t_D$. Define

$$
\begin{aligned}
d_i &= \text{number of lapses at time } t_i \\
Y_i &= \text{number of policies that are at risk at an instant just prior to time } t_i \\
&= \text{number of policies that are still alive at time } t_i \text{ or lapse at time } t_i \\
&= \text{number of policies in the risk set } R(t_i)
\end{aligned}
$$

2.2.3 Interval-censored data

Introduction

In the case where a large group of policies are followed from a common starting point (the inception date) over certain periods of time, the data consist of only the **number** of policies that lapse or are censored within various time-intervals. Interval censoring is used to describe this situation where a policy's lifetime is known only to lie between two values (interval boundaries). Interval-censored data can involve left and right censoring, as being outlayed in the scheme of [4, page 144-145].

Notation

Assume that the lifetime for the j^{th} policy is bounded between two known values, denoted $b_j \leq T < c_j$ and it is known whether the policy lapsed. The intervals do not necessarily coincide and may be overlapping. Left- and right-censoring are special cases of interval-censored data. Denote the left-censoring time by C_l and the right-censoring time by C_r. Observations that are left-censored have $b_j = 0$, $c_j = C_l$ and $\delta_j = 1$. Observations that are right-censored have $b_j = C_r$, $c_j = \infty$ and $\delta_j = 0$.

Grouped data

Grouped data arise due to the grouping of the continuous lifetimes of policies into k adjacent, non-overlapping fixed intervals

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

with $a_0 = 0$ and $a_k = \infty$.

Thus grouped data consist of the **numbers** of observed and censored lifetimes falling into each of the k intervals. In the case of the last interval $I_k = [a_k; \infty)$ the policies will lapse some time in this open interval and all the lifetimes in this interval can be considered as observed.

Define

d_j = number of policies that lapsed in I_j

Y'_j = number of policies entering interval I_j that have not lapsed

W_j = number of policies with censored lifetimes in I_j

assuming that censored lifetimes are independent of the time those policies would have had they been observed until the lapse

$$Y_j = Y'_j - \frac{W_j}{2}$$

= number of policies at risk of lapsing in I_j (policies that are still alive at a_{j-1})

assuming that censored and observed lifetimes are uniformly distributed over the interva

2.3 Basic Survival Functions

2.3.1 Introduction

Four functions characterize the distribution of T, namely the **survivor function**, which is the probability of a policy surviving beyond time t, the **hazard rate function** which is the chance a policy, surviving up to time t, will lapse in the next instant, the **probability density (or mass) function** which is the unconditional probability of a lapse at time t, and the **mean residual life** at time t, which is the mean time to a lapse, given no lapse at t. If any one of these four functions is known, the other three then can be uniquely determined.

2.3.2 Probability distribution and distribution function

For a **continuous** random variable T the probability density function of T is

$$f(t) = \lim_{dt \to 0} \frac{P(t < T < t + dt)}{dt} = \lim_{dt \to 0} \frac{F(t + dt) - F(t)}{dt} \qquad (2.3\ .1)$$

with cumulative distribution function

$$F(t) = P(T \leq t) = \int_0^t f(x)\, dx \qquad (2.3\ .2)$$

$f(t)$ is the unconditional probability of a lapse at time t (probability per unit of time over a small interval of time) and is called the unconditional lapse rate.

For a **discrete** random variable T the probability mass function of T at time t_i is

$$p(t_i) = \begin{cases} P(T = t_i) & i = 1, 2, \ldots \\ 0 & \text{elsewhere} \end{cases} \qquad (2.3\ .3)$$

The distribution of T is characterized by either one of two equivalent functions, namely the survivor function and the hazard rate function.

2.3.3 Survivor function

The survivor function at time t is the probability of a policy surviving beyond time t, also called the survival rate at time t.

The survivor function for **continuous** T is defined as

$$S(t) = P(T > t) = \int_t^\infty f(x)\,dx = 1 - F(t) \quad \text{for} \quad t > 0 \tag{2.3 .4}$$

- The graph of $S(t)$ versus t is called the survival curve
- $S(t)$ is a continuous, strictly decreasing function with $S(0) = 1$ and $S(\infty) = 0$
- $f(t) = -\frac{d}{dt}S(t)$

The survivor function for **discrete** T is defined as

$$S(t) = P(T > t) = \sum_{t_j > t} p(t_j) \quad \text{for} \quad t > 0 \tag{2.3 .5}$$

- The graph of $S(t)$ versus t is for discrete T a nonincreasing step function with jumps downward at $t_1, t_2, \ldots$

2.3.4 Hazard rate function

The hazard rate function at time t is the **conditional** probability of a lapse in the next instant just beyond time t **given that the policy has survived up to time** t. This is also called the instantaneous lapse rate assuming that the policy was alive up to time t.

For a **continuous** random variable T the hazard rate function at time t is defined as

$$\begin{aligned} h(t) &= \lim_{dt \to 0} \frac{P(t < T < t + dt \mid T > t)}{dt} \\ &= \lim_{dt \to 0} \frac{P[(t < T < t + dt) \cap (T > t)]}{P(T > t)dt} \\ &= \lim_{dt \to 0} \frac{P(t < T < t + dt)}{dt} \cdot \frac{1}{P(T > t)} \\ &= f(t) \cdot \frac{1}{S(t)} \end{aligned}$$

$$\Rightarrow h(t) = \frac{f(t)}{S(t)} = \frac{-\frac{d}{dt}S(t)}{S(t)} \tag{2.3 .6}$$

For a **discrete** random variable T the hazard rate function at time t_j is defined as

$$\begin{aligned} h(t_j) &= P(T = t_j \mid T \geq t_j) &&= \frac{P[(T = t_j) \cap (T \geq t_j)]}{P(T \geq t_j)} \\ &= \frac{P(T = t_j)}{P(T > t_{j-1})} &&= \frac{p(t_j)}{S(t_{j-1})} \\ &= \frac{S(t_{j-1}) - S(t_j)}{S(t_{j-1})} &&= 1 - \frac{S(t_j)}{S(t_{j-1})} \end{aligned}$$

$$\Rightarrow \frac{S(t_j)}{S(t_{j-1})} = 1 - h(t_j)$$

Note that $S(t) = \frac{S(t_1)}{S(t_0)} \cdot \frac{S(t_2)}{S(t_1)} \cdot \ldots \cdot \frac{S(t)}{S(t_{j-1})}$

$$\Rightarrow S(t) = \prod_{t_j \leq t} \frac{S(t_j)}{S(t_{j-1})} = \prod_{t_j \leq t} [1 - h(t_j)] \tag{2.3 .7}$$

The hazard rate is zero for a discrete random variable T except at points where a lapse could occur.

2.3.5 Cumulative hazard function

For **continuous** T the cumulative hazard function is defined as

$$\begin{aligned} H(t) &= \int_0^t h(x)\,dx \\ &= \int_0^t \frac{-\frac{d}{dx}S(x)}{S(x)}\,dx \\ &= -[\ln S(t) - \ln S(0)] \\ &= -\ln S(t) \end{aligned}$$

$$\Rightarrow h(t) = \frac{d}{dt}H(t) = -\frac{d}{dt}\ln S(t) \quad (2.3\ .8)$$

For **discrete** T the cumulative hazard function is defined as

$$H(t) = \sum_{t_j \le t} h(t_j)$$

2.3.6 Median lifetime, mean lifetime and mean residual lifetime

The survivor function is used to determine the median lifetime (median time to a lapse) and other percentiles.

$$\text{median} \quad t_{50} = \text{time point at which} \quad S(t_{50}) is equal to 0.5 \quad (2.3\ .9)$$

$$p^{th}\text{percentile} = t_p \quad \text{so that} \quad S(t_p) = \frac{100-p}{100} \quad (2.3\ .10)$$

The mean lifetime is represented as the area under the survival curve.

$$\mu = E(T) = \int_0^\infty S(t)\,dt \quad (2.3\ .11)$$

The variance of T is related to the survivor function by

$$var(T) = 2\int_0^\infty tS(t)\,dt - \left[\int_0^\infty S(t)\,dt\right]^2 \quad (2.3\ .12)$$

The mean residual lifetime (mean amount of lifetime remaining after a particular time t) is defined as

$$mrl(t) = E(T - t \mid T > t) = \frac{\int_t^\infty (x-t) f(x)\, dx}{P(T > t)} = \frac{\int_t^\infty S(x)\, dx}{S(t)} \qquad (2.3\ .13)$$

2.4 Survival Models

2.4.1 Introduction

In this section analytical forms of survival distributions that are often used in fitting survival data are given. Properties of the Weibull survival model are discussed by [2, 25, 26], while [22] describes the log-logistic survival model and [34, 26] the lognormal survival model.

2.4.2 Weibull distribution

Consider $T \sim \text{Weib}(\lambda, \alpha)$ $\quad \lambda$ = scale parameter $\lambda > 0$
α = shape parameter $\alpha > 0$

density function	$f(t)$	=	$\lambda \alpha t^{\alpha-1} \exp\{-\lambda t^\alpha\}$	$t > 0$	(2.4 .1)
expected value	$E(T)$	=	$\Gamma(1 + \frac{1}{\alpha}) \cdot \lambda^{\frac{-1}{\alpha}}$		(2.4 .2)
variance	$var(T)$	=	$\left[\Gamma(1 + \frac{2}{\alpha}) - \Gamma^2(1 + \frac{1}{\alpha})\right] \cdot \lambda^{\frac{-2}{\alpha}}$		(2.4 .3)
survivor function	$S(t)$	=	$\exp(-\lambda t^\alpha)$		(2.4 .4)
hazard function	$h(t)$	=	$\lambda \alpha t^{\alpha-1}$		(2.4 .5)
p^{th} percentile	t_p	=	$\left[\frac{1}{\lambda} \ln\left(\frac{100}{100-p}\right)\right]^{\frac{1}{\alpha}}$	$p = 1, 2, ..., 99$	(2.4 .6)

$$\Rightarrow \quad \text{median lifetime} \quad t_{50} = \left[\frac{1}{\lambda} \ln\left(\frac{100}{100-50}\right)\right]^{\frac{1}{\alpha}}$$

$$= \left[\frac{1}{\lambda} \ln(2)\right]^{\frac{1}{\alpha}} \qquad (2.4\ .7)$$

Note that

$$-\ln S(t) = \lambda t^{\alpha}$$

$$\Rightarrow \ln(-\ln S(t)) = \ln\lambda + \alpha \ln t \tag{2.4 .8}$$

2.4.3 The log-logistic distribution

Consider $T \sim$ log-logistic(λ, α) $\quad \lambda$ = scale parameter $\lambda > 0$
α = shape parameter $\alpha > 0$

$$\text{density function} \quad f(t) = \frac{\lambda \alpha t^{\alpha-1}}{(1+\lambda t^{\alpha})^2} \quad \text{with} \quad t > 0 \tag{2.4 .9}$$

$$\text{expected value} \quad E(T) = \frac{\pi csc\frac{\pi}{\alpha}}{\alpha\lambda^{\frac{1}{\alpha}}} \quad \text{if} \quad \alpha > 1 \tag{2.4 .10}$$

$$\text{variance} \quad var(T) = \frac{2\pi csc\frac{2\pi}{\alpha}}{\alpha\lambda^{\frac{2}{\alpha}}} - E(T^2) \quad \text{if} \quad \alpha > 2 \tag{2.4 .11}$$

$$\text{survivor function} \quad S(t) = (1+\lambda t^{\alpha})^{-1} \tag{2.4 .12}$$

$$\text{hazard function} \quad h(t) = \frac{\lambda \alpha t^{\alpha-1}}{(1+\lambda t^{\alpha})} \tag{2.4 .13}$$

$$\text{odds of a lapse at time t} \quad \frac{1-S(t)}{S(t)} = \lambda t^{\alpha} \tag{2.4 .14}$$

$$p^{th} \quad \text{percentile} \quad t_p = \left[\frac{1}{\lambda} \cdot \frac{p}{100-p}\right]^{\frac{1}{\alpha}} \tag{2.4 .15}$$

$$\Rightarrow \quad \text{median lifetime} \quad t_{50} = \left(\frac{1}{\lambda}\right)^{\frac{1}{\alpha}} \tag{2.4 .16}$$

Note that

$$\ln\left(\frac{1-S(t)}{S(t)}\right) = \ln\lambda + \alpha \ln t \tag{2.4 .17}$$

2.4.4 The lognormal distribution

Consider $T \sim$ lognormal(μ, σ^2).

$$\text{density function} \quad f(t) = \frac{\exp\left\{-\frac{1}{2}\left(\frac{ln(t)-\mu}{\sigma}\right)^2\right\}}{t\sigma(2\pi)^{\frac{1}{2}}} \quad \text{with} \quad t > 0 \tag{2.4 .18}$$

expected value $E(T) = \exp(\mu + 0.5\sigma^2)$ (2.4 .19)

variance $var(T) = \exp(2\mu + \sigma^2) \cdot [\exp(\sigma^2) - 1]$ (2.4 .20)

survivor function $S(t) = 1 - \Phi[\frac{ln(t) - \mu}{\sigma}]$ (2.4 .21)

with Φ the standard normal distribution function

hazard function $h(t) = f(t)/S(t)$

p^{th} percentile $t_p = \exp(\mu + \sigma z_p)$ (2.4 .22)

with z_p the p^{th} percentile of the n(0;1) distribution.

2.4.5 Location-scale parameter survival models

Introduction

A univariate location-scale parameter distribution is described by [13] as a distribution with a probability density function of the form

$$f(y; \mu, \sigma) = \frac{1}{\sigma} \cdot g(\frac{y - \mu}{\sigma}) \quad -\infty < y < \infty \tag{2.4 .23}$$

where

μ = location parameter $-\infty < \mu < \infty$
σ = scale parameter $\sigma > 0$
g = a fully specified probability density function defined on $(-\infty, \infty)$

The survivor function corresponding to Equation 2.4 .23 is

$$G[\frac{y - \mu}{\sigma}] \quad \text{where} \quad G(x) = \int_x^\infty g(z)\, dz \tag{2.4 .24}$$

The extreme value, logistic and normal distributions are examples of location-scale parameter distributions that are used in fitting survival models to survival data.

The extreme value distribution

Consider $Y =\sim$ EV(σ, μ) μ = location parameter
σ = scale parameter $\sigma > 0$

$$\text{density function} \qquad f(y) = \frac{1}{\sigma} \cdot \exp\left\{\left(\frac{y-\mu}{\sigma}\right) - \exp\left(\frac{y-\mu}{\sigma}\right)\right\} \qquad -\infty < y < \infty \tag{2.4 .25}$$

$$\text{expected value} \qquad E(Y) = \mu - 0.5772\,\sigma \qquad \text{Euler's constant} = 0.5772 \tag{2.4 .26}$$

$$\text{variance} \qquad var(Y) = \frac{\pi^2\sigma^2}{6} \tag{2.4 .27}$$

The extreme value distribution with $\mu = 0$ and $\sigma = 1$ is termed the standard extreme value distribution. A discussion of the exreme value distribution is given by [34].

The logistic distribution

Consider $Y \sim$ logistic(μ, σ) $\quad \mu =$ location parameter

$\sigma =$ scale parameter $\quad \sigma > 0$

$$\text{density function} \qquad f(y) = \frac{1}{\sigma} \cdot \frac{\exp\left(\frac{y-\mu}{\sigma}\right)}{\left\{1 + \exp\left(\frac{y-\mu}{\sigma}\right)\right\}^2} \qquad -\infty < y < \infty \tag{2.4 .28}$$

$$\text{expected value} \qquad E(Y) = \mu \tag{2.4 .29}$$

$$\text{variance} \qquad var(Y) = \frac{\pi^2\sigma^2}{3} \tag{2.4 .30}$$

The logistic distribution is a symmetrical distribution whose probability density function is very similar to that of the normal distribution.

The logistic distribution with $\mu = 0$ and $\sigma = 1$ is termed the standard logistic distribution.

The properties of the logistic distribution are discussed by [2].

The normal distribution

Consider $Y \sim$ normal(μ, σ) $\quad \mu =$ location parameter

$\sigma =$ scale parameter $\quad \sigma > 0$

$$\text{density function} \quad f(y) = \frac{1}{\sigma(2\pi)^{\frac{1}{2}}} \cdot \exp\left\{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right\} \quad -\infty < y < \infty \tag{2.4 .31}$$

$$\text{expected value} \quad E(Y) = \mu \tag{2.4 .32}$$

$$\text{variance} \quad var(Y) = \sigma^2 \tag{2.4 .33}$$

The normal distribution with $\mu = 0$ and $\sigma = 1$ is termed the standard normal distribution.

2.5 Relationships between Survival Distributions

2.5.1 Introduction

In analyzing survival data it is often convenient to work with $Y = \ln(T)$, the logarithm of the lifetimes.

2.5.2 Relationship between the Weibull and extreme value distributions

The extreme value distribution arises when lifetimes are taken to be Weibull distributed. The following relationship between the Weibull and extreme value distributions exists:

$$T \sim \text{Weib}(\lambda, \alpha) \iff Y = \ln(T) \sim \text{EV}\left(\frac{1}{\alpha}, \frac{-\ln\lambda}{\alpha}\right) \tag{2.5 .1}$$

This result can be proved in the following way. If $T \sim \text{Weib}(\lambda, \alpha)$ with density function

$$f_T(t) = \lambda\alpha t^{\alpha-1} \exp\{-\lambda t^\alpha\} \quad t > 0$$

and survivor function

$$S_T(t) = \exp\{-\lambda t^\alpha\}$$

then, by using the transformation technique, $Y = ln(T)$ has density function

$$\begin{aligned} f_Y(y) &= f_T(t) \cdot \left| \frac{dt}{dy} \right| \\ &= \lambda\alpha(e^y)^{\alpha-1} \exp\{-\lambda(e^y)^\alpha\} \cdot | e^y | \end{aligned}$$

$$= \lambda\alpha e^{\alpha y} \exp\{-\lambda e^{\alpha y}\} \quad -\infty < y < \infty \tag{2.5 .2}$$

$$= \exp\{-\lambda e^{\alpha y}\} \cdot (-\lambda e^{\alpha y}) \cdot \alpha \cdot (-1). \tag{2.5 .3}$$

From Equation 2.5 .3 and the relationship $f_Y(y) = -\frac{d}{dy}S_Y(y)$ follow that the survivor function of Y is

$$S_Y(y) = \exp\{-\lambda e^{\alpha y}\}. \tag{2.5 .4}$$

Equation 2.5 .2 implies that

$$\begin{aligned} f_Y(y) &= \lambda\alpha \exp\{\alpha y - \lambda e^{\alpha y}\} \\ &= \exp\{ln(\lambda)\} \cdot \alpha \cdot \exp\{\alpha y - \exp\{ln(\lambda)\} \exp\{\alpha y\}\} \\ &= \alpha \exp\{\alpha y + ln(\lambda) - \exp\{\alpha y + ln(\lambda)\}\} \\ &= \alpha \exp\left\{\alpha\left[y - \frac{-ln(\lambda)}{\alpha}\right] - \exp\left\{\alpha\left[y - \frac{-ln(\lambda)}{\alpha}\right]\right\}\right\}. \end{aligned} \tag{2.5 .5}$$

Let $\mu = \frac{-ln(\lambda)}{\alpha}$ and $\sigma = \frac{1}{\alpha}$ in Equation 2.5 .5.

$$\Rightarrow f_Y(y) = \frac{1}{\sigma}\exp\left\{\frac{y-\mu}{\sigma} - \exp\left\{\frac{y-\mu}{\sigma}\right\}\right\}. \tag{2.5 .6}$$

Equation 2.5 .6 is the density of the **extreme value distribution** with parameters σ and μ

$$\Rightarrow Y \sim \text{EV}(\sigma, \mu) \equiv \text{EV}\left(\frac{1}{\alpha}, \frac{-ln(\lambda)}{\alpha}\right). \tag{2.5 .7}$$

This leads to the relationship between the Weibull distribution and the extreme value distribution:

$$T \sim \text{Weib}(\lambda, \alpha) \Longleftrightarrow Y = ln(T) \sim \text{EV}\left(\frac{1}{\alpha}, \frac{-\ln\lambda}{\alpha}\right) \tag{2.5 .8}$$

2.5.3 Relationship between the log-logistic and logistic distributions

The log-logistic distribution is related to the logistic distribution by the relationship

$$T \sim \quad \text{log-logistic}(\lambda, \alpha) \quad \Longleftrightarrow \quad Y = \ln(T) \sim \text{logistic}(\ln\lambda, \frac{1}{\alpha}). \tag{2.5 .9}$$

2.5.4 Relationship between the lognormal and normal distributions

The lognormal distribution is related to the normal distribution by the relationship

$$T \sim \quad \text{lognormal}(\mu, \sigma^2) \qquad \Longleftrightarrow \qquad Y = \ln(T) \sim \text{normal}(\mu, \sigma^2). \qquad (2.5\ .10)$$

In view of the similarity of the normal and logistic distributions, the lognormal model will tend to be very similar to the loglogistic model.

2.5.5 A linear model representation in log-time

Consider the following log-linear model that describes the basic underlying distribution of lifetimes:

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

where W is the error distribution, μ is the location parameter and σ is the scale parameter. A variety of distributions for W can be assumed, for example the standard extreme value distribution, the standard logistic distribution or the standard normal distribution.

- If this linear model format $Y = \ln T = \mu + \sigma W$ is used where W has the standard extreme value distribution, that is $W \sim EV(1, 0)$, with density

  $$f(w) = \exp\{w - \exp(w)\} \qquad -\infty < w < \infty$$

  then $Y = ln(T)$ has an extreme value EV $\left(\frac{1}{\alpha}, \frac{-\ln\lambda}{\alpha}\right)$ distribution and T has an underlying Weibull(λ, α) distribution with parameters

  $$\lambda = \exp\left\{\frac{-\mu}{\sigma}\right\} \quad \text{and} \quad \alpha = \frac{1}{\sigma}. \qquad (2.5\ .11)$$

- If this linear model format $Y = \ln T = \mu + \sigma W$ is used where W has the standard logistic distribution with density

  $$f(w) = \frac{\exp(w)}{(1 + \exp(w))^2} \qquad -\infty < w < \infty$$

  then $Y = ln(T)$ has a logistic$\left(\frac{1}{\alpha}, \frac{-\ln\lambda}{\alpha}\right)$ distribution and T has an underlying log-logistic(λ, α) distribution also with parameters

  $$\lambda = \exp\left\{\frac{-\mu}{\sigma}\right\} \quad \text{and} \quad \alpha = \frac{1}{\sigma}. \qquad (2.5\ .12)$$

- If this linear model format $Y = \ln T = \mu + \sigma W$ is used where W has the standard normal distribution with density

$$f(w) = \frac{\exp(w^2)}{(2\pi)^{\frac{1}{2}}} \qquad -\infty < w < \infty$$

  then $Y = ln(T)$ has a normal(μ, σ^2) distribution and T has an underlying lognormal(μ, σ^2) distribution.

The first of the three results is now proven. The other two results follow in a similar way.

If the error distribution $W \sim$ EV$(1,0)$, with density

$$f(w) = \exp\{w - \exp(w)\} \qquad -\infty < w < \infty$$

then, by using the transformation technique and the fact that $W = \dfrac{ln(T) - \mu}{\sigma}$, the density function of T is

$$\begin{aligned} f_T(t) &= f_W(w) \cdot \mid \frac{dw}{dt} \mid \\ &= \exp\left\{\frac{\ln(t) - \mu}{\sigma} - \exp\left\{\frac{\ln(t) - \mu}{\sigma}\right\}\right\} \cdot \mid \frac{1}{\sigma t} \mid \\ &= \exp\left\{\frac{1}{\sigma} ln(t) - \frac{\mu}{\sigma} - \exp\left\{\frac{ln(t)}{\sigma} - \frac{\mu}{\sigma}\right\}\right\} \cdot \frac{1}{\sigma t} \\ &= t^{\frac{1}{\sigma}} \cdot t^{-1} \cdot \frac{1}{\sigma} \cdot \exp\left(\frac{-\mu}{\sigma}\right) \cdot \exp\left(-t^{\frac{1}{\sigma}} \cdot \exp\left\{\frac{-\mu}{\sigma}\right\}\right) \\ &= \exp(\frac{-\mu}{\sigma}) \cdot \frac{1}{\sigma} \cdot t^{\frac{1}{\sigma}-1} \cdot \exp\left(-\exp\left\{\frac{-\mu}{\sigma}\right\} \cdot t^{\frac{1}{\sigma}}\right) \end{aligned} \qquad (2.5\ .13)$$

Equation 2.5 .13 is the density of the **Weibull distribution** with parameters

$$\lambda = \exp\left\{\frac{-\mu}{\sigma}\right\} \quad \text{and} \quad \alpha = \frac{1}{\sigma}$$

$$\Rightarrow T \sim \text{Weibull}(\lambda, \alpha)$$

In a similar way, if $W \sim$ EV$(1,0)$, with density

$$f(w) = \exp\{w - \exp w\} \qquad -\infty < w < \infty$$

then, by using the transformation technique and the fact that $W = \dfrac{Y - \mu}{\sigma}$, the density function of Y is

$$\begin{aligned} f_Y(y) &= f_W(w) \cdot \mid \frac{dw}{dy} \mid \\ &= \exp\left\{\frac{y - \mu}{\sigma} - \exp\left\{\frac{y - \mu}{\sigma}\right\}\right\} \cdot \mid \frac{1}{\sigma} \mid \\ &= \frac{1}{\sigma} \cdot \exp\left\{\frac{y - \mu}{\sigma} - \exp\left\{\frac{y - \mu}{\sigma}\right\}\right\} \\ & -\infty < y < \infty \end{aligned} \qquad (2.5\ .14)$$

Equation 2.5 .14 is the density of the **extreme value distribution** with parameters σ and μ

$$\Rightarrow Y \sim \text{EV}(\sigma, \mu) \equiv \text{EV}\left(\frac{1}{\alpha}, \frac{-ln(\lambda)}{\alpha}\right)$$

To summarize, if the general linear model format $Y = \ln T = \mu + \sigma W$ is used, where the error distribution W is the standard extreme value $EV(1, 0)$ distribution, then $Y = ln(T)$ has an extreme value EV $\left(\frac{1}{\alpha}, \frac{-\ln\lambda}{\alpha}\right)$ distribution and T has an underlying Weibull(λ, α) distribution with parameters

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

2.6 Construction of Likelihood Functions

2.6.1 Introduction

The standard method of fitting survival models, specified in section 2.4, to survival data is the method of Maximum Likelihood Estimation (MLE). The first step is to create the specific likelihood function to be maximized. Likelihoods for different scenarios are now given.

2.6.2 Likelihood function for random right-censored continuous data

Assume the random right-censoring model 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 $(\delta_j = 0)$.

Consider the pair (T_i, δ_i) for the i^{th} policy. The likelihood function is constructed by considering the contribution to the likelihood of the pairs $(t_i, \delta_i = 1)$ and $(t_i, \delta_i = 0)$ separately (refer to [25]).

- The contribution to the likelihood of the pair $(t_i, \delta_i = 1)$ is the probability that the i^{th} policy lapses at time $t_i = x_i$. This probability is

$$\begin{aligned} P(t_i \in (t_i, t_i + dt_i), \delta_i = 1) &= P(T_i \in (t_i, t_i + dt_i), C_i > t_i) \\ &= P(t_i < T_i < t_i + dt_i) \cdot P(C_i > t_i) \quad (2.6\ .1) \end{aligned}$$

$$= f(t_i)dt_i \cdot [1 - G(t_i)] \qquad (2.6\ .2)$$

Equation 2.6 .1 follows from the fact that the observed survival times are independent of the censoring times.

- The contribution to the likelihood of the pair $(t_i, \delta_i = 0)$ is the probability that the i^{th} policy survives at least time $t_i = C_i$. This probability is

$$\begin{aligned} P(C_i \in (t_i, t_i + dt_i), \delta_i = 0) &= P(C_i \in (t_i, t_i + dt_i), \boldsymbol{X_i} > t_i) \\ &= P(t_i < C_i < t_i + dt_i) \cdot P(X_i > t_i) \\ &= g(t_i)dt_i \cdot S(t_i) \end{aligned} \qquad (2.6\ .3)$$

The complete likelihood for the i^{th} policy under random censoring is

$$L(t_i, \delta_i) = \{f(t_i)dt_i[1 - G(t_i)]\}^{\delta_i} \cdot \{g(t_i)dt_iS(t_i)\}^{1-\delta_i} \qquad (2.6\ .4)$$

Under the assumption of n independent censored and observed survival times, the full likelihood function is obtained by multiplying the respective contributions of the n pairs $(t_i, \delta_i) \quad i = 1, 2, ..., n$ in the data set. This likelihood function of the full sample is

$$\begin{aligned} L(t_1, t_2, ..., t_n; \delta_1, \delta_2, ..., \delta_n) &= \prod_{i=1}^{n} L(t_i, \delta_i) \\ &= \prod_{i=1}^{n} [f(t_i)dt_i[1 - G(t_i)]]^{\delta_i} \cdot [g(t_i)dt_iS(t_i)]^{1-\delta_i} \\ &= \prod_{i=1}^{n} \left\{[f(t_i)dt_i]^{\delta_i} [S(t_i)]^{1-\delta_i}\right\} \cdot \left\{[1 - G(t_i)]^{\delta_i} [g(t_i)dt_i]^{1-\delta_i}\right\} \end{aligned} \qquad (2.6\ .5)$$

Let $\boldsymbol{\theta}$ be the vector of parameters of the survival model. The survival model specifies the distribution of T_i, independent from the distribution of C_i. Only the first term of the product in Equation 2.6 .5 involves the unknown lifetime parameters $\boldsymbol{\theta}$, so that the last term of the product, namely

$$\left\{[1 - G(t_i)]^{\delta_i} [g(t_i)dt_i]^{1-\delta_i}\right\}$$

can be treated like constants when maximizing $L(t_1, t_2, ..., t_n; \delta_1, \delta_2, ..., \delta_n)$. Thus the likelihood function, up to a multiple constant, is

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} [f(t_i)]^{\delta_i} \cdot [S(t_i)]^{1-\delta_i} \qquad (2.6\ .6)$$

The log-likelihood function

$$\ln L(\boldsymbol{\theta}) = \sum_{i=1}^{n} \delta_i \cdot \ln[f(t_i)] + \sum_{i=1}^{n} (1 - \delta_i) \cdot \ln[S(t_i)] \qquad (2.6\ .7)$$

is maximized to obtain the maximum likelihood estimators of the unknown parameters $\boldsymbol{\theta}$. The procedure to obtain the values of the MLE involves taking derivatives of $lnL(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$, setting these equations equal to zero, and solving for $\boldsymbol{\theta}$.

2.6.3 Likelihood function for interval-censored data

Consider the lifetime of the j^{th} policy that is bounded between two known values, that is $b_j \leq T_j < c_j$. It is also known whether this policy lapsed. Observations that are left-censored have $b_j = 0$, $c_j = C_l$ and $\delta_j = 1$, while observations that are right-censored have $b_j = C_r$, $c_j = \infty$ and $\delta_j = 0$, with C_l and C_r the left- and right-censoring times respectively.

In constructing a likelihood function for interval-censored data, the information each observation provides, needs to be considered as the contribution of that observation to the likelihood (refer to [4]).

- An observation corresponding to an **exact lifetime** t_i provides information on the probability that the lapse occurs at this time t_i, which is approximately equal to the density function of T at this time. This probability is $f(t_i)$.
- For a **right-censored observation**, it is known that the lifetime is larger than C_r with C_r the right-censoring time. Thus the information provided is the survival function evaluated at C_r, that is $S(C_r)$.
- For a **left-censored observation**, it is known that the lapse has already occurred, so that the contribution to the likelihood is the cumulative distribution function evaluated at C_l, that is $F(C_l) = 1 - S(C_l)$ with C_l the left-censoring time.
- For **interval-censored data**, it is known that the lapse occurred within the interval, so that the information is the probability that the lifetime is in this interval. This probability is $S(b_i) - S(c_i)$.

The likelihood function may be constructed by putting together the above-mentioned components.

$$L(\boldsymbol{\theta}) \quad \propto \quad \prod_{i \in D} f(t_i) \cdot \prod_{i \in RC} S(C_r) \cdot \prod_{i \in LC} [1 - S(C_l)] \cdot \prod_{i \in I} [S(b_i) - S(c_i)] \qquad (2.6\ .8)$$

where

D is the set of observed lifetimes (lapses)
RC is the set of right-censored observations
LC is the set of left-censored observations
I is the set of interval-censored observations

Consider the n pairs $(t_i, \delta_i) \quad i = 1, 2, ..., n$ in the data set. Some of the responses are observed, while other responses are left, right or interval-censored. By making use of the principles of construction of the likelihood for interval-censored data in Equation 2.6 .8 it follows that

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} [f(t_i)]^{\delta_i} \cdot [S(t_i)]^{1-\delta_i} \cdot [1 - S(t_i)]^{\delta_i} \cdot [S(b_i) - S(t_i)]^{1-\delta_i} \qquad (2.6\ .9)$$

with b_i the lower end of a censoring interval.

2.6.4 Likelihood function for right-censored grouped data

Consider the grouped data case as a special case of interval-censored data where the n lifetimes of policies are grouped into k adjacent, non-overlapping fixed intervals

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

with $a_0 = 0$ and $a_k = \infty$.

The likelihood function for right-censored grouped data is stated by [24].

For **complete data** with no censored lifetimes, the n observed lifetimes are grouped into k intervals so that
$n = d_1 + d_2 + ... + d_k$ with d_j=number of lapses in I_j.

The unconditional probability of a lapse in I_j is

$$\pi_j = S(a_{j-1}) - S(a_j) \quad j = 1, 2, ..., k$$

. Then $(d_1, d_2, ..., d_k)$ has a multinomial probability function

$$\frac{n!}{d_1!d_2!...d_k!}{\pi_1}^{d_1}{\pi_2}^{d_2}...{\pi_k}^{d_k}$$

. The likelihood function can thus be taken as

$$L(\boldsymbol{\theta}) = n! \prod_{j=1}^{k} \left\{ \frac{[S(a_{j-1}) - S(a_j)]^{d_j}}{d_j!} \right\} \qquad (2.6\ .10)$$

For **incomplete data**, where the n censored and observed lifetimes are grouped into k intervals, it is **further assumed** that the W_j censored lifetimes in I_j occur at the midpoint of the interval $a_j^\star = a_{j-1} + \frac{1}{2}h_j$ with $h_j = a_j - a_{j-1}$ the length of interval I_j.

For interval $I_j = [a_{j-1}; a_j)$, **conditional** on surviving till a_{j-1},

- the probability of a lapse is
$$q_j = \frac{S(a_{j-1}) - S(a_j)}{S(a_{j-1})}$$
- the probability of surviving until $a_j^\star$ is
$$p_j^\star = \frac{S(a_{j-1}) - S(a_j^\star)}{S(a_{j-1})}$$
- the probability of surviving the full interval I_j is
$$\begin{aligned} p_j &= 1 - q_j \\ &= 1 - \frac{S(a_{j-1}) - S(a_j)}{S(a_{j-1})} \\ &= \frac{S(a_j)}{S(a_{j-1})} \end{aligned}$$

The **conditional** likelihood for interval I_j is

$$L_j(\boldsymbol{\theta}) \quad \propto \quad [q_j]^{d_j} \cdot [p_j^\star]^{W_j} \cdot [p_j]^{Y_j - d_j - W_j} \qquad (2.6\ .11)$$

where Y_j is the number of policies at risk of lapsing in I_j, that is policies that are still alive at a_{j-1}.

The overall likelihood function is

$$L(\boldsymbol{\theta}) = \prod_{j=1}^{k} L_j(\boldsymbol{\theta}) \qquad (2.6\ .12)$$

If class intervals are narrow, another possibility is to treat the data as continuous and assume that all lifetimes in interval I_j occur at the interval midpoint.

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 $(y_i, \delta_i) \quad i = 1, 2, ..., n$ 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

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

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 $\widehat{\boldsymbol{V}}_{p0} = \frac{1}{n}\left[diag(\boldsymbol{p}_0) - \boldsymbol{p}_0\boldsymbol{p}_0'\right]$ and the $\boldsymbol{G}_\pi$ in Equation 3.3 .5 is replaced by $\boldsymbol{G}_{p0} = \left.\frac{\partial \boldsymbol{g}(\boldsymbol{\pi})}{\partial \boldsymbol{\pi}}\right|_{\boldsymbol{\pi} = \boldsymbol{p}_0}$. 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 $\widehat{\boldsymbol{V}}_{p0}$ and $\boldsymbol{G}_{p0}$, 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

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

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 $S(\boldsymbol{x}) = \mathbf{1} - F(\boldsymbol{x}) = \mathbf{1} - \boldsymbol{\pi}_S$ 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

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

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

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

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

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

or from Equation 3.3 .7

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

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

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

$\boldsymbol{C}$ 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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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.

$$\Rightarrow \Phi^{-1}(\boldsymbol{\pi}_S) = -\frac{\mu}{\sigma} \cdot \mathbf{1} + \frac{1}{\sigma} \cdot \ln \boldsymbol{x} \tag{3.3 .34}$$

$$= (\mathbf{1}\ , \ln \boldsymbol{x}) \cdot \begin{pmatrix} -\frac{\mu}{\sigma} \\ \frac{1}{\sigma} \end{pmatrix}$$

$$\Rightarrow \Phi^{-1}(\boldsymbol{\pi}_S) = \boldsymbol{X}_1 \cdot \begin{pmatrix} -\frac{\mu}{\sigma} \\ \frac{1}{\sigma} \end{pmatrix} \tag{3.3 .35}$$

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

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

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

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

with

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

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

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

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

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

and

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

When fitting a log-logistic distribution

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

and

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

When fitting a lognormal distribution

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

and

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

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

$$T_j = min(X_j, C_j) \text{ for the } j^{th} \text{ policy}$$

$$\delta_j = \begin{cases} 1 & \text{if } \; X_j \leq C_j \quad \text{, that is, } X_j \text{ is not censored} \\ 0 & \text{if } \; X_j > C_j \quad \text{, that is, } X_j \text{ is censored} \end{cases}$$

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

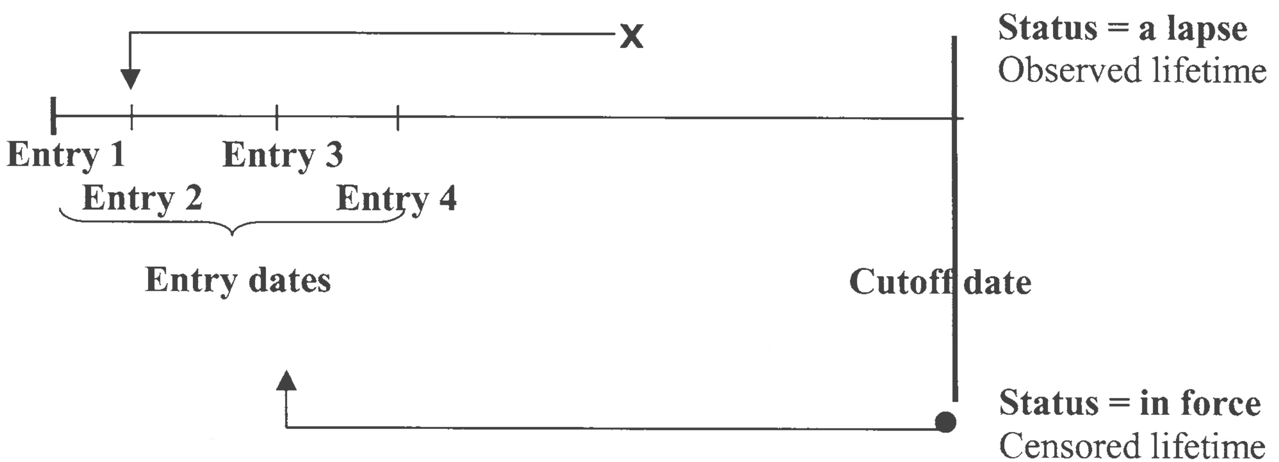

Figure 3.1: **Experimental design for staggered entry of policies**

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 number	Lifetime Intervals				Observed Probability Vector				Probability Vector				Vector of Upper Boundarie			
	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

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

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

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

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

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

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

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

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

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

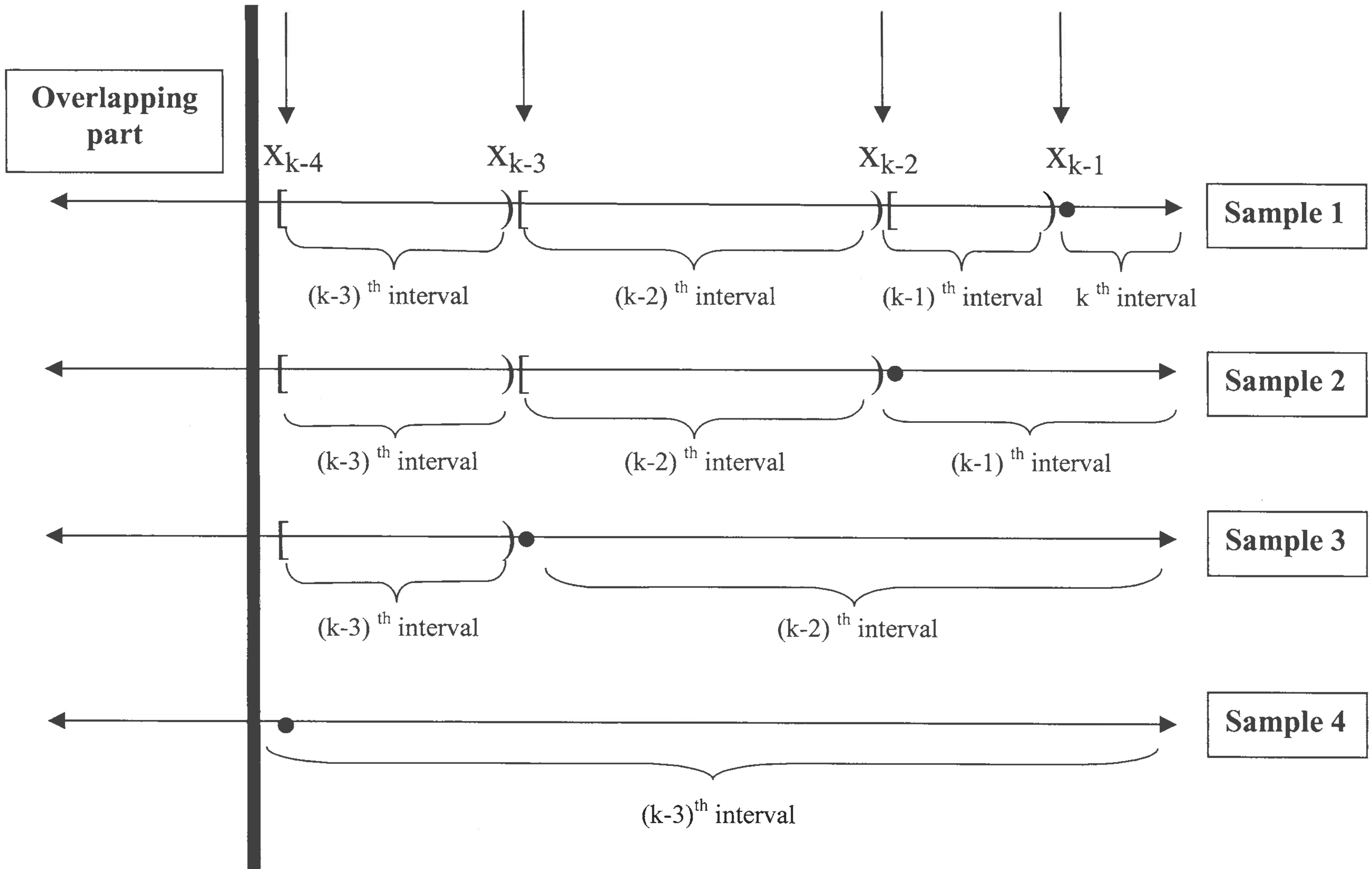

Figure 3.2: **Constraints imposed by the experimental design**

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 $\boldsymbol{\pi}_S$ 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

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

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

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

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

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

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

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

or

$$\begin{aligned}
\ln\left(\frac{\pi_S}{1-\pi_S}\right) &= \ln(\pi_S) - \ln(1-\pi_S) \\
&= \ln\lambda \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \alpha \cdot \begin{pmatrix} \ln x_1 \\ \ln x_2 \\ \ln x_3 \\ \ln x_4 \end{pmatrix} \\
&= \underbrace{\begin{pmatrix} 1 & \ln x_1 \\ 1 & \ln x_2 \\ 1 & \ln x_3 \\ 1 & \ln x_4 \end{pmatrix}} \cdot \begin{pmatrix} \ln\lambda \\ \alpha \end{pmatrix} \\
\Rightarrow \ln\left(\frac{\pi_S}{1-\pi_S}\right) &= X_1 \cdot \begin{pmatrix} \ln\lambda \\ \alpha \end{pmatrix} \qquad (3.4\ .10)
\end{aligned}$$

where

$$x_1 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ x_{k-1} \end{pmatrix} \quad x_2 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ \cdots \\ \cdots \\ x_{k-2} \end{pmatrix} \quad x_3 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ \cdots \\ x_{k-3} \end{pmatrix} \quad \text{and} \quad x_4 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ x_{k-4} \end{pmatrix}.$$

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

$$\begin{aligned}
\underbrace{\left(I - X_1({X_1}'X_1)^{-1}{X_1}'\right)} \cdot \ln\left(\frac{\pi_S}{1-\pi_S}\right) &= 0 \\
\underbrace{C \qquad \cdot \ln\left(\frac{\pi_S}{1-\pi_S}\right)} &= 0 \\
g(\pi) &= 0
\end{aligned}$$

C is the projection matrix orthogonal to the columns of the design matrix X_1. Note that $CX_1 = 0$.

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

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

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

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

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$(n_i;\boldsymbol{\pi}_i)$ $\quad i = 1,2,3,4$ distributed. The variance-covariance matrix $\boldsymbol{V}$ to be used is the estimated variance-covariance matrix of the multinomial distribution **for each entry period**.

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

$$\begin{aligned}
\Phi^{-1}(\boldsymbol{\pi}_S) &= -\frac{\mu}{\sigma}\cdot\mathbf{1}+\frac{1}{\sigma}\cdot\ln\boldsymbol{x} && (3.4\ .17)\\
&= -\frac{\mu}{\sigma}\cdot\begin{pmatrix}\mathbf{1}\\ \mathbf{1}\\ \mathbf{1}\\ \mathbf{1}\end{pmatrix}+\frac{1}{\sigma}\cdot\begin{pmatrix}\ln\boldsymbol{x}_1\\ \ln\boldsymbol{x}_2\\ \ln\boldsymbol{x}_3\\ \ln\boldsymbol{x}_4\end{pmatrix}\\
&= \underbrace{\begin{pmatrix}\mathbf{1} & \ln\boldsymbol{x}_1\\ \mathbf{1} & \ln\boldsymbol{x}_2\\ \mathbf{1} & \ln\boldsymbol{x}_3\\ \mathbf{1} & \ln\boldsymbol{x}_4\end{pmatrix}}\cdot\begin{pmatrix}-\frac{\mu}{\sigma}\\ \frac{1}{\sigma}\end{pmatrix}\\
\Rightarrow \Phi^{-1}(\boldsymbol{\pi}_S) &= \boldsymbol{X}_1\cdot\begin{pmatrix}-\frac{\mu}{\sigma}\\ \frac{1}{\sigma}\end{pmatrix} && (3.4\ .18)
\end{aligned}$$

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

Equation 3.3 .35 is equivalent to

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

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

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

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

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

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

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:

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

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

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

The four 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 $\boldsymbol{C}\boldsymbol{X}_1 = \mathbf{0}$.

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

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

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:

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

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

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

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

$\boldsymbol{C}$ 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

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

that is

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

or

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

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

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

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

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

$\boldsymbol{C}$ 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 $\boldsymbol{G} \cdot \boldsymbol{\pi} = \boldsymbol{0}$ 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

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

The function $g(\boldsymbol{\pi}) = \boldsymbol{0}$ 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 $\boldsymbol{G} = \frac{\partial g(\boldsymbol{\pi})}{\partial \boldsymbol{\pi}}$ and only a **single iteration** is needed in the iterative procedure to determine the MLE of $\boldsymbol{\pi}$ under the constraints $\boldsymbol{g}(\boldsymbol{\pi}) = \boldsymbol{0}$.

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

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

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=200 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	first	λ=0.15	$\hat{\lambda}_{SAS}$=0.16521023	0.01521023	0.0094249	Yes Over
5 classes			$\hat{\lambda}_{IML}$=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	first	λ=0.15	$\hat{\lambda}_{SAS}$=0.15312020	0.0031202	0.0061112	No
5 classes			$\hat{\lambda}_{IML}$=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

Chapter 4

PARAMETRIC REGRESSION MODELS FOR SURVIVAL DATA WITH COVARIATES

4.1 Notation

Suppose that the distribution of T depends on a vector of fixed-time explanatory variables (covariates) $\boldsymbol{Z} = (Z_1, Z_2, ..., Z_p)'$. The right-censored survival data set then consists of triples $(T_j, \delta_j, \boldsymbol{Z}_j)$ $\quad j = 1, 2, ..., n$ $\quad$ where

T_j = lifetime for the j^{th} policy

δ_j = lapse indicator for the j^{th} policy $= \begin{cases} 1 & \text{if policy has lapsed} \\ 0 & \text{if lifetime is right-censored} \end{cases}$

$\boldsymbol{Z}_j$ = $(Z_{j1}, Z_{j2}, ..., Z_{jp})'$ is the vector of explanatory variables for the j^{th} policy at a fixed time

4.2 Three Approaches to Regression Modelling

The effect of covariates (risk factors on the lapse of policies) must be modelled in order to predict lifetimes of policies.

Either the conditional survivor function or the conditional hazard function can be modelled as a function of p **fixed** covariates or risk factors $\boldsymbol{Z} = (Z_1, Z_2, ..., Z_p)'$.

Three general approaches to regression modelling of survival data will be discussed.

- Accelerated Failure Time Model (AFM)

  This model states that the **survivor function** at time t of a policy with covariate $\boldsymbol{Z}$ is the same as the survivor function of a policy with a baseline survivor function at a time $t \cdot \exp(\boldsymbol{\theta}'\boldsymbol{Z})$ where $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_p)'$ is a vector of regression coefficients.

- Proportional Hazards Model (PHM)

  This model states that the **relative hazard rate** or **hazard rate** of a lapse at time t of a policy, with risk vector $\boldsymbol{Z}$, compared to a policy with the baseline characteristics (that means $\boldsymbol{Z} = \mathbf{0}$), is a constant $e^{\boldsymbol{\beta}'\boldsymbol{Z}}$ where $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_p)'$ is a vector of regression coefficients.

- Proportional Odds Model (POM)

  This model states that the **relative odds** or **odds ratio** of a lapse at time t of a policy, with risk vector $\boldsymbol{Z}$, compared to a policy with the baseline characteristics (that means $\boldsymbol{Z} = \mathbf{0}$), is a constant $e^{\boldsymbol{\beta}'\boldsymbol{Z}}$ where $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_p)'$ is a vector of regression coefficients.

The effect of the covariates in all three models is to **alter the scale parameter**, while the **shape parameter remains constant**.

Parametric regression models are discussed by [9, 7, 13, 16, 24, 26, 35]. A comparison between the AFM and the PHM is done by [21, 18], while [18] also compare the AFM and POM. The Weibull AFM and Weibull PHM are compared by [5], and [5] also compares the log-logistic AFM and the log-logistic POM.

The properties of the different models are compared in the following tables.

PARAMETRIC REGRESSION MODELS

ACCELERATED FAILURE TIME MODEL

models the **survivor function** of a policy with risk vector $\boldsymbol{Z} = (Z_1, Z_2, ..., Z_p)'$ and regression coefficients $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_p)'$

$$\boxed{S(t|\boldsymbol{Z}) = S_0(e^{\boldsymbol{\theta}'\boldsymbol{Z}}t)}$$

$S(t|\boldsymbol{Z})$ is survivor function at time t for a policy with risk vector $\boldsymbol{Z}$

$S_0(t)$ is **baseline survivor function** at time t with a specified parametric form
(survivor function for a policy whose risk factors all take the value zero)
$\Rightarrow S_0(t) = S(t \mid \boldsymbol{Z} = \boldsymbol{0})$

- effect of covariates is multiplicative on **survival time**
- $e^{\boldsymbol{\theta}'\boldsymbol{Z}} = \exp\left\{\sum_{k=1}^{p} \theta_k Z_k\right\}$ is **acceleration factor**
- $e^{\theta_k Z_k}$ indicates how risk factor Z_k "speeds up" or "slows down" the lifetime of a policy
- median lifetime for a given $\boldsymbol{Z}$ is equal to baseline median lifetime $\times$ acceleration factor

$$\boxed{h(t|\boldsymbol{Z}) = e^{\boldsymbol{\theta}'\boldsymbol{Z}} h_0(e^{\boldsymbol{\theta}'\boldsymbol{Z}}t)}$$

$$\boxed{S(t|\boldsymbol{Z}) = S_0(e^{\boldsymbol{\theta}'\boldsymbol{Z}}t)}$$

PROPORTIONAL HAZARDS MODEL

models the **hazard rate** of a lapse for a policy with risk vector $\boldsymbol{Z} = (Z_1, Z_2, ..., Z_p)'$ and regression coefficients $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_p)'$

$$\boxed{h(t|\boldsymbol{Z}) = h_0(t)e^{\boldsymbol{\beta}'\boldsymbol{Z}}}$$

$h(t|\boldsymbol{Z})$ is hazard rate of a lapse at time t for a policy with risk vector $\boldsymbol{Z}$

$h_0(t)$ is **baseline hazard rate** for a policy at time t with a **specified parametric form**
(hazard rate for a policy whose risk factors all take the value zero)
$\Rightarrow h_0(t) = h(t \mid \boldsymbol{Z} = \boldsymbol{0})$

- effect of covariates is multiplicative on **hazard rate**
- $e^{\boldsymbol{\beta}'\boldsymbol{Z}} = \exp\left\{\sum_{k=1}^{p} \beta_k Z_k\right\}$ is **link function**
- $e^{\beta_k Z_k}$ is **risk score** for risk factor Z_k $k = 1, 2, ..., p$
- If $S_0(t)$ is unspecified, the PHM is the famous **semiparametric Cox's PHM**: [23]

$$\boxed{h(t|\boldsymbol{Z}) = e^{\boldsymbol{\beta}'\boldsymbol{Z}} h_0(t)}$$

$$\boxed{S(t|\boldsymbol{Z}) = [S_0(t)]^{e^{\boldsymbol{\beta}'\boldsymbol{Z}}}}$$

Special case: lifetimes are assumed to have a Weibull(λ, α)

Weibull AFM

$h_0(t) = \lambda\alpha t^{\alpha-1}$

From AFM: The hazard function of the j^{th} policy with risk vector $\boldsymbol{Z}$ is

$$\begin{aligned} h_j(t) &= e^{\boldsymbol{\theta}'\boldsymbol{Z}}\lambda\alpha\left[e^{\boldsymbol{\theta}'\boldsymbol{Z}}t\right]^{\alpha-1} \\ &= [(e^{\boldsymbol{\theta}'\boldsymbol{Z}})^{\alpha}\lambda]\alpha t^{\alpha-1} \\ &= (e^{\boldsymbol{\theta}'\boldsymbol{Z}})^{\alpha}h_0(t) \end{aligned}$$

Also $S_0(t) = \exp\{-\lambda t^{\alpha}\}$

From AFM: The survivor function of the j^{th} policy with risk vector $\boldsymbol{Z}$ is

$$\begin{aligned} S_j(t) &= S_0(e^{\boldsymbol{\theta}'\boldsymbol{Z}}t) \\ &= \exp\left\{-\lambda[e^{\boldsymbol{\theta}'\boldsymbol{Z}}t]^{\alpha}\right\} \\ &= \exp\left\{-[\lambda(e^{\boldsymbol{\theta}'\boldsymbol{Z}})^{\alpha}]\cdot t^{\alpha}\right\} \end{aligned}$$

$\Rightarrow$ The lifetime of the j^{th} policy has a Weibull$(\lambda(e^{\boldsymbol{\theta}'\boldsymbol{Z}})^{\alpha}, \alpha)$ distribution

Say that Weibull possesses the **accelerated failure time property**

Weibull PHM

$h_0(t) = \lambda\alpha t^{\alpha-1}$

From PHM: The hazard function of the j^{th} policy with risk vector $\boldsymbol{Z}$ is

$$\begin{aligned} h_j(t) &= e^{\boldsymbol{\beta}'\boldsymbol{Z}}\lambda\alpha t^{\alpha-1} \\ &= [e^{\boldsymbol{\beta}'\boldsymbol{Z}}\lambda]\alpha t^{\alpha-1} \\ &= e^{\boldsymbol{\beta}'\boldsymbol{Z}}h_0(t) \end{aligned}$$

Also $S_0(t) = \exp\{-\lambda t^{\alpha}\}$

From PHM: The survivor function of the j^{th} policy with risk vector $\boldsymbol{Z}$ is

$$\begin{aligned} S_j(t) &= \exp\{-\lambda t^{\alpha}\}^{e^{\boldsymbol{\beta}'\boldsymbol{Z}}} \\ &= \exp\left\{e^{\boldsymbol{\beta}'\boldsymbol{Z}}\cdot(-\lambda t^{\alpha})\right\} \\ &= \exp\left\{-[\lambda e^{\boldsymbol{\beta}'\boldsymbol{Z}}]\cdot t^{\alpha}\right\} \end{aligned}$$

$\Rightarrow$ The lifetime of the j^{th} policy has a Weibull$(\lambda e^{\boldsymbol{\beta}'\boldsymbol{Z}}, \alpha)$ distribution

Say that Weibull possesses the **proportional hazards property**

PARAMETRIC REGRESSION MODELS (continued)

ACCELERATED FAILURE TIME MODEL

models the **survivor function** of a policy with risk vector $\boldsymbol{Z} = (Z_1, Z_2, ..., Z_p)'$ and regression coefficients $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_p)'$

$$S(t|\boldsymbol{Z}) = S_0(e^{\boldsymbol{\theta}'\boldsymbol{Z}}t)$$

- effect of covariates is multiplicative on **survival time**
- median time to a lapse (with given $\boldsymbol{Z}$) $=$ [baseline median time to a lapse] $\cdot e^{\boldsymbol{\theta}'\boldsymbol{Z}}$

PROPORTIONAL ODDS MODEL

models the **odds of a lapse** at time t for a policy with risk vector $\boldsymbol{Z} = (Z_1, Z_2, ..., Z_p)'$ and regression coefficients $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_p)'$

$$\frac{1 - S(t|\boldsymbol{Z})}{S(t|\boldsymbol{Z})} = e^{\boldsymbol{\beta}'\boldsymbol{Z}} \cdot \frac{1 - S_0(t)}{S_0(t)}$$

- effect of covariates is multiplicative on the **odds of a lapse** at time t
- $e^{\beta_k Z_k}$ is **index** for covariate Z_k
  $k = 1, 2, ..., p$
- If $S_0(t)$ is unspecified, the POM is the **semiparametric Cox's regression model** that includes a **time-dependent** covariate to produce **non-proportional hazards**
  Refer to [25, 3].

Special case: lifetimes are assumed to have a log-logistic(λ, α)

log-logistic AFM

From AFM: $S_0(t) = \dfrac{1}{1+\lambda t^\alpha}$

From AFM: The survivor function of the j^{th} policy with risk vector $\boldsymbol{Z}$ is

$$\begin{aligned} S_j(t) &= S_0(e^{\boldsymbol{\theta}'\boldsymbol{Z}}t) \\ &= \frac{1}{1+\lambda(e^{\boldsymbol{\theta}'\boldsymbol{Z}}t)^\alpha} \\ &= \frac{1}{1+[\lambda(e^{\boldsymbol{\theta}'\boldsymbol{Z}})^\alpha]\cdot t^\alpha} \end{aligned}$$

$\Rightarrow$ The lifetime of the j^{th} policy has a log-logistic$(\lambda(e^{\boldsymbol{\theta}'\boldsymbol{Z}})^\alpha, \alpha)$ distribution

Say that log-logistic possesses the **accelerated failure time property**

log-logistic POM

From $S_0(t) = \dfrac{1}{1+\lambda t^\alpha}$ follows that the baseline odds of a lapse at time t is $\dfrac{1-S_0(t)}{S_0(t)} = \lambda t^\alpha$

From POM: The odds of a lapse of the j^{th} policy at time t with risk vector $\boldsymbol{Z}$ is

$$\begin{aligned} \frac{1-S_j(t)}{S_j(t)} &= e^{\boldsymbol{\beta}'\boldsymbol{Z}}\cdot\lambda t^\alpha \\ &= [\lambda e^{\boldsymbol{\beta}'\boldsymbol{Z}}]\cdot t^\alpha \end{aligned}$$

$\Rightarrow$ The lifetime of the j^{th} policy has a log-logistic$(\lambda e^{\boldsymbol{\beta}'\boldsymbol{Z}}, \alpha)$ distribution

Say that log-logistic possesses the **proportional odds property**

4.3 Comparison between the AFM and the PHM

- The Weibull is the only continuous distribution which has the property of being both an AFM and a PHM.
  The lifetimes under AFM $\sim$ Weibull$(\lambda(e^{\boldsymbol{\theta}'\boldsymbol{Z}})^{\alpha}, \alpha)$.
  The lifetimes under PHM $\sim$ Weibull$(\lambda e^{\boldsymbol{\beta}'\boldsymbol{Z}}, \alpha)$.

- There is a direct correspondence between the parameters of the Weibull under these two models. It follows that when the β_k's in the linear component of the PHM are divided by α, the corresponding θ_k's of the AFM are determined

$$\Rightarrow \theta_k = \frac{\beta_k}{\alpha} \quad \text{or} \quad \boldsymbol{\theta}' = \frac{\boldsymbol{\beta}'}{\alpha}$$

- The acceleration factor $e^{\boldsymbol{\theta}'\boldsymbol{Z}}$ at the Weibull AFM indicates how a change in covariate values changes the time scale from the baseline time scale.
  The factor $e^{\boldsymbol{\beta}'\boldsymbol{Z}}$ at the Weibull PHM indicates how much the baseline hazard rate of a lapse at any time changes when a policy has covariate vector $\boldsymbol{Z}$.
  Note that $e^{\boldsymbol{\beta}'\boldsymbol{Z}}$ is the **relative hazard rate of a lapse** for a policy with covariate $\boldsymbol{Z}$ compared to a policy with the baseline characteristics. This relative hazard rate is called the **hazard ratio**. This hazard ratio is constant over time (or the hazard rates are proportional).

- The PHM has the **property of proportional hazard rates for fixed covariates**. The hazard ratio (relative risk) of a lapse at time t for a policy with risk factor $\boldsymbol{Z}$, as compared to a policy with risk factor $\boldsymbol{Z}^*$, is

$$\frac{h(t \mid \boldsymbol{Z})}{h(t \mid \boldsymbol{Z}^*)} = \frac{h_0(t)\exp\left\{(\sum_{k=1}^{p} \beta_k Z_k)\right\}}{h_0(t)\exp\left\{(\sum_{k=1}^{p} \beta_k Z_k^*)\right\}} = \exp\left\{\sum_{k=1}^{p} \beta_k (Z_k - Z_k^*\right\}$$

  which is a constant. So the hazard rates are proportional (or the hazard ratio is constant).

- Estimates of the β_k's can be used to provide estimates of **hazard ratios**. For a **constant shape parameter** in the Weibull distributions, the hazard ratios may be estimated from the exponent of the $\widehat{\beta}$-values in the Weibull regression model. These estimated hazard ratios are called **risk scores**.

4.4 Comparison between the AFM and the POM

- The log-logistic is the only continuous distribution which has the property of being both an AFM and a POM.
  The lifetimes under AFM $\sim$ log-logistic$(\lambda(e^{\boldsymbol{\theta}'\boldsymbol{Z}})^{\alpha}, \alpha)$.
  The lifetimes under POM $\sim$ log-logistic$(\lambda e^{\boldsymbol{\beta}'\boldsymbol{Z}}, \alpha)$.

- There is a direct correspondence between the parameters of the log-logistic under these two models. It follows that when the β_k's in the linear component of the POM are divided by α, the corresponding θ_k's of the AFM are determined

$$\Rightarrow \theta_k = \frac{\beta_k}{\alpha} \quad \text{or} \quad \boldsymbol{\theta}' = \frac{\boldsymbol{\beta}'}{\alpha}$$

- The acceleration factor $e^{\boldsymbol{\theta}'\boldsymbol{Z}}$ at the log-logistic AFM indicates how a change in covariate values changes the time scale from the baseline time scale.
  The factor $e^{\boldsymbol{\beta}'\boldsymbol{Z}}$ at the log-logistic POM indicates how much the baseline odds of a lapse at any time changes when a policy has covariate vector $\boldsymbol{Z}$.
  Note that $e^{\boldsymbol{\beta}'\boldsymbol{Z}}$ is the **relative odds of a lapse** for a policy with covariate $\boldsymbol{Z}$ compared to a policy with the baseline characteristics. This relative odds is called the **odds ratio**. This odds ratio is constant over time (or the odds are proportional).

- The POM has the **property of convergent hazard rates** or the **property of proportional odds for time-dependent covariates** or **non-proportional hazard rates for time-dependent covariates**.
  The ratio of the hazard rate for the j^{th} policy to the baseline hazard rate, namely $\frac{h_j(t)}{h_0(t)}$, converges from the value $\exp(-\sum_{k=1}^{p} \beta_k Z_k)$ at time t=0 to the value 1 at time $t = \infty$.

- Estimates of the β_k's can be used to provide estimates of **odds ratios**. For a **constant shape parameter** in the log-logistic distributions, the odds ratios may be estimated from the exponent of the $\widehat{\beta}$-values in the log-logistic regression model. These estimated odds ratios are called **indices**.

4.5 Log-linear Presentation of Models for Survival Data

4.5.1 A linear regression model for the log of the hazard ratio

In the AFM the hazard rate of a lapse for a policy with risk vector $\boldsymbol{Z} = (Z_1, Z_2, ..., Z_p)'$ and regression coefficients $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_p)'$ is modelled as $h(t|\boldsymbol{Z}) = h_0(t)e^{\boldsymbol{\theta}'\boldsymbol{Z}}$ while in the PHM the hazard rate of a lapse for a policy with risk vector $\boldsymbol{Z} = (Z_1, Z_2, ..., Z_p)'$ and regression coefficients $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_p)'$ is modelled as $h(t|\boldsymbol{Z}) = h_0(t)e^{\boldsymbol{\beta}'\boldsymbol{Z}}$.

The relative hazard rate for a policy with covariate Z compared to a policy with the baseline characteristics is termed the hazard ratio (relative risk or risk score)

$\Rightarrow$ hazard ratio $= \dfrac{h(t|\boldsymbol{Z})}{h_0(t)} = e^{\boldsymbol{\theta}'\boldsymbol{Z}} = e^{\boldsymbol{\beta}'\boldsymbol{Z}}$ is constant over time

$\Rightarrow$ log(hazard ratio) is modelled as

$$\boxed{\ln \frac{h(t|\boldsymbol{Z})}{h_0(t)} = \boldsymbol{\theta}'\boldsymbol{Z} = \boldsymbol{\beta}'\boldsymbol{Z} = \sum_{k=1}^{p} \beta_k Z_k}$$

4.5.2 A linear regression model for the log of the odds ratio

In the AFM the odds of a lapse for a policy with risk vector $\boldsymbol{Z} = (Z_1, Z_2, ..., Z_p)'$ and regression coefficients $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_p)'$ is modelled as $\dfrac{1 - S(t|\boldsymbol{Z})}{S(t|\boldsymbol{Z})} = e^{\boldsymbol{\theta}'\boldsymbol{Z}} \cdot \dfrac{1 - S_0(t)}{S_0(t)}$ while in the POM the odds of a lapse for a policy with risk vector $\boldsymbol{Z} = (Z_1, Z_2, ..., Z_p)'$ and regression coefficients $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_p)'$ is modelled as $\dfrac{1 - S(t|\boldsymbol{Z})}{S(t|\boldsymbol{Z})} = e^{\boldsymbol{\beta}'\boldsymbol{Z}} \cdot \dfrac{1 - S_0(t)}{S_0(t)}$.

The relative odds of a lapse for a policy with covariate Z compared to a policy with the baseline characteristics is termed the **odds ratio** (relative odds or index)

$\Rightarrow$ odds ratio $= \dfrac{\dfrac{1 - S(t|\boldsymbol{Z})}{S(t|\boldsymbol{Z})}}{\dfrac{1 - S_0(t)}{S_0(t)}} = e^{\boldsymbol{\theta}'\boldsymbol{Z}} = e^{\boldsymbol{\beta}'\boldsymbol{Z}}$ is constant over time

$\Rightarrow$ log(odds ratio) is modelled as

$$\ln\left\{\frac{\frac{1-S(t|\boldsymbol{Z})}{S(t|\boldsymbol{Z})}}{\frac{1-S_0(t)}{S_0(t)}}\right\} = \boldsymbol{\beta}'\boldsymbol{Z} = \sum_{k=1}^{p}\beta_k Z_k$$

4.5.3 A linear regression model for log-time

Consider the following linear log-time regression model that describes the linear relationship between log-time and the covariate values.

$$Y = \ln T = \mu + \boldsymbol{\gamma}'\boldsymbol{Z} + \sigma W$$

W is the error distribution, μ is the intercept, σ is the scale parameter and $\boldsymbol{\gamma}' = (\gamma_1, \gamma_2, ..., \gamma_p)$ is a vector of regression coefficients that are interpretated similar to those in standard normal theory regression.

A variety of models is discussed by [4] that can be used for W, or equivalently for T or S_0. Note that $S_0(t)$ denotes the survivor function of $T = e^Y$ when $\boldsymbol{Z} = \boldsymbol{0}$, that is $S_0(t)$ is the survivor function of $e^{\mu + \sigma W}$. Then the linear log-time regression model is equivalent to the AFM with $\boldsymbol{\theta} = -\boldsymbol{\gamma}$.

- If W has the standard extreme value distribution, that is $W \sim EV(1,0)$, with density

  $$f(w) = \exp\{(w - \exp(w))\} \qquad -\infty < w < \infty$$

  then T has an underlying Weibull(λ, α) distribution. This model leads to

  1. an AFM for T with a Weibull baseline survivor function with parameters

     $$\lambda = \exp\left\{\frac{-\mu}{\sigma}\right\}, \alpha = \frac{1}{\sigma} \quad \text{and} \quad \theta_k = -\gamma_k \quad k = 1, 2, ..., p$$

  2. a PHM for T with a Weibull baseline hazard function with parameters

     $$\lambda = \exp\left\{\frac{-\mu}{\sigma}\right\}, \alpha = \frac{1}{\sigma} \quad \text{and} \quad \beta_k = \theta_k\alpha = \frac{-\gamma_k}{\sigma} \quad k = 1, 2, ..., p$$

- If W has the standard logistic distribution with density

  $$f(w) = \frac{\exp(w)}{(1 + \exp(w))^2} \qquad -\infty < w < \infty$$

  then T has an underlying log-logistic(λ, α) distribution. This model leads to

1. an AFM for T with a log-logistic baseline survivor function with parameters

$$\lambda = \exp\left\{\frac{-\mu}{\sigma}\right\}, \alpha = \frac{1}{\sigma} \quad \text{and} \quad \theta_k = -\gamma_k \quad k = 1, 2, ..., p$$

2. a POM for T with a log-logistic baseline survivor function with parameters

$$\lambda = \exp\left\{\frac{-\mu}{\sigma}\right\}, \alpha = \frac{1}{\sigma} \quad \text{and} \quad \beta_k = \theta_k \alpha = \frac{-\gamma_k}{\sigma} \quad k = 1, 2, ..., p$$

4.6 Maximum Likelihood Estimation

4.6.1 Introduction

The standard way of fitting parametric regression models to an observed set of survival data with covariates is to use the method of maximum likelihood. Maximum likelihood estimation for the Weibull and log-logistic regression models has been discussed in [18, 22]. The construction of the likelihood functions for continuous and grouped survival data with covariates is now discussed.

4.6.2 Likelihood function for random right-censored continuous data

The right-censored survival data set consists of triplets $(T_j, \delta_j, \boldsymbol{Z}_j) \quad j = 1, 2, ..., n$ where

$$\begin{aligned}
T_j &= \text{lifetime for the } j^{th} \text{ policy} \\
\delta_j &= \text{lapse indicator for the } j^{th} \text{ policy} = \begin{cases} 1 & \text{if policy has lapsed} \\ 0 & \text{if lifetime is right-censored} \end{cases} \\
\boldsymbol{Z}_j &= (Z_{j1}, Z_{j2}, ..., Z_{jp})' \text{ is the vector of explanatory variables for the } j^{th} \text{ policy at a fixed time.}
\end{aligned}$$

The likelihood function is constructed by considering the contribution to the likelihood of the triplets $(T_j, \delta_j = 1, \boldsymbol{Z}_j)$ and $(T_j, \delta_j = 0, \boldsymbol{Z}_j)$ separately $\quad j = 1, 2, ..., n$.

- For a specific triplet $(t, \delta_i = 1, \boldsymbol{Z})$ the observed survival time is t. Thus the contribution to the likelihood of this triplet is the probability that a policy with covariate vector $\boldsymbol{Z}$ lapse at time t. This probability is given by the density function $f(t|\boldsymbol{Z})$.
- For a specific triplet $(t, \delta_i = 0, \boldsymbol{Z})$ the survival time is at least t. Thus the contribution to the likelihood of this triplet is the probability that a policy with covariate vector $\boldsymbol{Z}$ survives at least time t. This probability is given by the survivor function $S(t|\boldsymbol{Z})$.

The complete likelihood for the i^{th} policy under random censoring is

$$[f(t_j|\boldsymbol{Z}_j)]^{\delta_j} \times [S(t_j|\boldsymbol{Z}_j)]^{1-\delta_j} \quad j = 1, 2, ..., n \tag{4.6 .1}$$

Under the assumption of n independent censored and observed survival times, the full likelihood function is obtained by multiplying the respective contributions of the n triplets. This gives the likelihood function

$$L(\boldsymbol{\eta}) = \prod_{i=1}^{n} [f(t_i|\boldsymbol{Z}_i)]^{\delta_i} \cdot [S(t_i|\boldsymbol{Z}_i)]^{1-\delta_i} \tag{4.6 .2}$$

where $\boldsymbol{\eta}$ is the vector of parameters of the survival model. The log-likelihood function

$$\ln L(\boldsymbol{\eta}) = \sum_{i=1}^{n} \{\delta_i \cdot \ln[f(t_i|\boldsymbol{Z}_i)] + (1 - \delta_i) \cdot \ln[S(t_i|\boldsymbol{Z}_i)]\} \tag{4.6 .3}$$

is maximized to obtain the maximum likelihood estimators of the unknown parameters $\boldsymbol{\eta}$. The procedure to obtain the values of the MLE involves taking derivatives of $\ln L(\boldsymbol{\eta})$ with respect to $\boldsymbol{\eta}$, setting these equations equal to zero, and solving for $\boldsymbol{\eta}$.

4.6.3 Likelihood function for right-censored grouped data

Consider the grouped data case, as in [24], where the n lifetimes of policies are grouped into k adjacent, non-overlapping fixed intervals

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

with $a_0 = 0$ and $a_k = \infty$.

For **complete data**, the n observed lifetimes are grouped into k intervals so that $n = d_1 + d_2 + ... + d_k$ with d_j=number of lapses in I_j.

The unconditional probability of a lapse in I_j is

$$\pi_j(\boldsymbol{\eta}) = S(a_{j-1}, \boldsymbol{\eta}) - S(a_j, \boldsymbol{\eta}) \quad j = 1, 2, ..., k.$$

Then $(d_1, d_2, ..., d_k)$ has a multinomial probability function

$$\frac{n!}{d_1!d_2!...d_k!} \pi_1(\boldsymbol{\eta})^{d_1} \pi_2(\boldsymbol{\eta})^{d_2} ... \pi_k(\boldsymbol{\eta})^{d_k}.$$

The likelihood function can thus be taken as

$$L(\boldsymbol{\eta}) = n! \prod_{j=1}^{k} \left\{ \frac{[S(a_{j-1}, \boldsymbol{\eta}) - S(a_j, \boldsymbol{\eta})]^{d_j}}{d_j!} \right\} \tag{4.6 .4}$$

For **incomplete data**, where the n censored and observed lifetimes are grouped into k intervals, it is further assumed that the W_j censored lifetimes in I_j occur at the midpoint of the interval $a_j^\star = a_{j-1} + \frac{1}{2}h_j$ with $h_j = a_j - a_{j-1}$ the length of interval I_j.

For interval $I_j = [a_{j-1}; a_j)$, **conditional** on surviving till a_{j-1},

- the probability of a lapse is
$$q_j(\boldsymbol{\eta}) = \frac{S(a_{j-1},\boldsymbol{\eta}) - S(a_j,\boldsymbol{\eta})}{S(a_{j-1},\boldsymbol{\eta})}$$
- the probability of surviving until $a_j^\star$ is
$$p_j^\star(\boldsymbol{\eta}) = \frac{S(a_{j-1},\boldsymbol{\eta}) - S(a_j^\star,\boldsymbol{\eta})}{S(a_{j-1},\boldsymbol{\eta})}$$
- the probability of surviving the full interval I_j is
$$\begin{aligned} p_j(\boldsymbol{\eta}) &= 1 - q_j(\boldsymbol{\eta}) \\ &= 1 - \frac{S(a_{j-1},\boldsymbol{\eta}) - S(a_j,\boldsymbol{\eta})}{S(a_{j-1},\boldsymbol{\eta})} \\ &= \frac{S(a_j,\boldsymbol{\eta})}{S(a_{j-1},\boldsymbol{\eta})} \end{aligned}$$

The **conditional** likelihood for interval I_j is

$$L_j(\boldsymbol{\eta}) \quad \propto \quad [q_j(\boldsymbol{\eta})]^{d_j} \cdot [p_j^\star(\boldsymbol{\eta})]^{W_j} \cdot [p_j(\boldsymbol{\eta})]^{Y_j - d_j - W_j} \tag{4.6 .5}$$

where Y_j is the number of policies at risk of lapsing in I_j, that is still alive at a_{j-1}.

The overall likelihood function is

$$L(\boldsymbol{\eta}) = \prod_{j=1}^{k} L_j(\boldsymbol{\eta}) \tag{4.6 .6}$$

If class intervals are narrow, another possibility is to treat the data as continuous and assume that all lifetimes in interval I_j occur at the interval midpoint.

4.6.4 Likelihood function for the linear model in log-time

A log-linear regression model (a linear regression model in log-time) is discussed by [22] and could be fitted to a survival data set of the form

$$\boxed{Y = \ln T = \mu + \boldsymbol{\gamma}'\mathbf{Z} + \sigma W}$$

where W is the error distribution, μ is the intercept, σ is the scale parameter and $\gamma' = (\gamma_1, \gamma_2, ..., \gamma_p)$ is a vector of regression coefficients.

Consider the n triplets $(y_j, \delta_j, \boldsymbol{Z}_j) \quad j = 1, 2, ..., n$ in the data set with $y_j = \ln(t_j)$.

The basic form of the likelihood function for **random right-censored continuous data** is, from Equation 4.6 .2, equal to

$$\begin{aligned} L(\mu, \gamma, \sigma) &= \prod_{i=1}^{n} [f_Y(y_i)]^{\delta_i} \cdot [S_Y(y_i)]^{1-\delta_i} \\ &= \prod_{i=1}^{n} \left[f_W(\frac{y_i - \mu - \gamma' \boldsymbol{Z}}{\sigma}) \cdot \frac{1}{\sigma} \right]^{\delta_i} \cdot \left[S_W(\frac{y_i - \mu - \gamma' \boldsymbol{Z}}{\sigma}) \right]^{1-\delta_i} \end{aligned} \quad (4.6\ .7)$$

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

$$\ln L(\mu, \gamma, \sigma) = \sum \delta_i \cdot \ln \left[f_W(\frac{y_i - \mu - \gamma' \boldsymbol{Z}}{\sigma}) \cdot \frac{1}{\sigma} \right] + \sum (1 - \delta_i) \cdot \ln \left[S_W(\frac{y_i - \mu - \gamma' \boldsymbol{Z}}{\sigma}) \right] \quad (4.6\ .8)$$

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** is

$$\begin{aligned} L(\mu, \gamma, \sigma) &= \prod_{i=1}^{n} [f_Y(y_i)]^{\delta_i} \cdot [S_Y(y_i)]^{1-\delta_i} \cdot [1 - S_Y(y_i)]^{\delta_i} \cdot [S_Y(b_i) - S_Y(y_i)]^{1-\delta_i} \\ &= \prod_{i=1}^{n} \left[f_W(\frac{y_i - \mu - \gamma' \boldsymbol{Z}}{\sigma}) \cdot \frac{1}{\sigma} \right]^{\delta_i} \cdot \left[S_W(\frac{y_i - \mu - \gamma' \boldsymbol{Z}}{\sigma}) \right]^{1-\delta_i} \cdot \\ &\quad \left[1 - S_W(\frac{y_i - \mu - \gamma' \boldsymbol{Z}}{\sigma}) \right]^{\delta_i} \cdot \left[S_W(\frac{b_i - \mu}{\sigma}) - S_W(\frac{y_i - \mu - \gamma' \boldsymbol{Z}}{\sigma}) \right]^{1-\delta_i} \end{aligned} \quad (4.6\ .9)$$

with b_i the lower end of a censoring interval.

The log-likelihood function for **interval-censored data** is

$$\begin{aligned} \ln L(\mu, \gamma, \sigma) &= \sum \delta_i \cdot \ln \left[f_W(\frac{y_i - \mu - \gamma' \boldsymbol{Z}}{\sigma}) \cdot \frac{1}{\sigma} \right] + \sum (1 - \delta_i) \cdot \ln \left[S_W(\frac{y_i - \mu - \gamma' \boldsymbol{Z}}{\sigma}) \right] \\ &\quad + \sum (\delta_i) \cdot \ln \left[1 - S_W(\frac{y_i - \mu - \gamma' \boldsymbol{Z}}{\sigma}) \right] \\ &\quad + \sum (1 - \delta_i) \cdot \ln \left[S_W(\frac{b_i - \mu}{\sigma}) - S_W(\frac{y_i - \mu - \gamma' \boldsymbol{Z}}{\sigma}) \right] \end{aligned}$$

(4.6 .10)

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.

4.6.5 Maximum likelihood estimators

Maximum likelihood estimates for the Weibull and log-logistic regression models must be found. Most computer software packages for survival analysis, including SAS, use the linear log-time regression model version. Refer to [4]. Maximum likelihood estimators of the log-linear parameters μ, σ and $\gamma' = (\gamma_1, \gamma_2, ..., \gamma_p)$ are found numerically, and routines to do so are available in the SAS statistical package. SAS allows for right-, left- and interval-censored data. The SAS program appear in Appendix A.

The invariance property of the SAS maximum likelihood estimators of μ, σ and $\gamma_k \quad k = 1, 2, ..., p$ in the log-linear model provides estimates of parameters of the three models (AFM, PHM or POM). Then the parameters of the three models are **functions** of these estimates and can be computed in the following way:

- If W has the standard extreme value distribution then T has an underlying Weibull(λ, α) distribution. The linear model for log-time then leads to

  1. an AFM for T with a Weibull baseline survivor function with estimated parameters

  $$\hat{\lambda} = \exp\left\{\frac{-\hat{\mu}}{\hat{\sigma}}\right\}, \hat{\alpha} = \frac{1}{\hat{\sigma}} \quad \text{and} \quad \hat{\theta}_k = -\hat{\gamma}_k \quad k = 1, 2, ..., p \qquad (4.6\ .11)$$

  2. a PHM for T with a Weibull baseline hazard function with parameters

  $$\hat{\lambda} = \exp\left\{\frac{-\hat{\mu}}{\hat{\sigma}}\right\}, \hat{\alpha} = \frac{1}{\hat{\sigma}} \quad \text{and} \quad \hat{\beta}_k = \hat{\theta}_k \hat{\alpha} = \frac{-\hat{\gamma}_k}{\hat{\sigma}} \quad k = 1, 2, ..., p \qquad (4.6\ .12)$$

- If W has the standard logistic distribution then T has an underlying log-logistic(λ, α) distribution. The linear model for log-time then leads to

  1. an AFM for T with a log-logistic baseline survivor function with parameters

  $$\hat{\lambda} = \exp\left\{\frac{-\hat{\mu}}{\hat{\sigma}}\right\}, \hat{\alpha} = \frac{1}{\hat{\sigma}} \quad \text{and} \quad \hat{\theta}_k = -\hat{\gamma}_k \quad k = 1, 2, ..., p \qquad (4.6\ .13)$$

2. a POM for T with a log-logistic baseline survivor function with parameters

$$\hat{\lambda} = \exp\left\{\frac{-\hat{\mu}}{\hat{\sigma}}\right\}, \hat{\alpha} = \frac{1}{\hat{\sigma}} \quad \text{and} \quad \hat{\beta}_k = \hat{\theta}_k\hat{\alpha} = \frac{-\hat{\gamma}_k}{\hat{\sigma}} \quad k = 1, 2, ..., p \qquad (4.6\ .14)$$

The variance-covariance matrix of the log-linear parameters μ, γ and σ, obtained from the observed information matrix, are also available in this package.

Using the delta method, the approximate variance-covariance matrix for these estimates, based on the estimates and their covariances in the log-linear model, is

$$cov(\hat{\beta}_j, \hat{\beta}_k) = \frac{cov(\hat{\gamma}_j, \hat{\gamma}_k)}{\hat{\sigma}^2} - \hat{\gamma}_j\frac{cov(\hat{\gamma}_j, \hat{\sigma})}{\hat{\sigma}^3} - \hat{\gamma}_k\frac{cov(\hat{\gamma}_k, \hat{\sigma})}{\hat{\sigma}^3} + \hat{\gamma}_j\hat{\gamma}_k\frac{var(\hat{\sigma})}{\hat{\sigma}^4} \qquad (4.6\ .15)$$

$$var(\hat{\lambda}) = \exp\left(\frac{-2\hat{\mu}}{\hat{\sigma}}\right) \cdot \left[\frac{var(\hat{\mu})}{\hat{\sigma}^2} + \hat{\mu}^2\frac{var(\hat{\sigma})}{\hat{\sigma}^4} - 2\hat{\mu}\frac{cov(\hat{\mu}, \hat{\sigma})}{\hat{\sigma}^3}\right] \qquad (4.6\ .16)$$

$$var(\hat{\alpha}) = \frac{var(\hat{\sigma})}{\hat{\sigma}^4} \qquad (4.6\ .17)$$

$$cov(\hat{\beta}_j, \hat{\lambda}) = \exp\left(\frac{-\hat{\mu}}{\hat{\sigma}}\right) \cdot \left[\frac{cov(\hat{\gamma}_j, \hat{\mu})}{\hat{\sigma}^2} - \hat{\gamma}_j\frac{cov(\hat{\gamma}_j, \hat{\sigma})}{\hat{\sigma}^3} - \hat{\mu}\frac{cov(\hat{\mu}, \hat{\sigma})}{\hat{\sigma}^3} + \hat{\gamma}_j\hat{\mu}\frac{var(\hat{\sigma})}{\hat{\sigma}^4}\right] \qquad (4.6\ .18)$$

$$cov(\hat{\beta}_j, \hat{\alpha}) = \frac{cov(\hat{\gamma}_j, \hat{\sigma})}{\hat{\sigma}^3} - \hat{\gamma}_j\frac{var(\hat{\sigma})}{\hat{\sigma}^4} \qquad (4.6\ .19)$$

$$cov(\hat{\lambda}, \hat{\alpha}) = \exp\left(\frac{-\hat{\mu}}{\hat{\sigma}}\right) \cdot \left[\frac{cov(\hat{\mu}, \hat{\sigma})}{\hat{\sigma}^3} - \hat{\mu}\frac{var(\hat{\sigma})}{\hat{\sigma}^4}\right] \qquad (4.6\ .20)$$

Once maximum likelihood estimates of the parameters 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) and log-logistic (or logistic).

In the above regression models, the effect of the covariates is to alter the scale parameter, while the shape parameter remains constant. The article [28] discusses how to extend the semi-parametric Cox's PHM to alter both the scale and the shape parameters. The standard parametric regression model fitting that is performed by PROC LIFEREG of SAS can not alter both parameters, but the method of maximum likelihood estimation subject to constraints in the next section can do it.

Graphical checks to determine whether or not a certain parametric model is reasonable, is given by [4]. These tests are based on the linear relationship between some function of the survivor function and some function of time.

4.7 Maximum Likelihood Estimation subject to Constraints

4.7.1 Introduction

The parametric regression model must describe the basic underlying distribution of survival time, but it must also characterize how that distribution changes as a function of the covariates. The effect of the covariates is

- to alter the scale parameter, while the shape parameter remains constant.
- to alter both the scale and the shape parameters.

In the case of grouped survival data, a survival distribution is fitted for each level of a risk factor or combination of levels of risk factors by using maximum likelihood estimation subject to constraints for estimating the parameters of the regression model. A detailed description of the development of this theory is given for **one categorical risk factor** on three levels, where the effect of the risk factor is either to keep the shape parameter α constant or to alter it. Then it is shown how to deal with a **continuous risk factor** when fitting the regression model when the shape parameter remains constant. In the last part of this chapter, the theory is extended to a regression model with two categorical risk factors.

The fitting of a log-logistic regression model and a Weibull regression model will be discussed for **staggered entry of policies**. These two distributions are used, because the log-logistic is the only continuous distribution which has the property of being both an AFM and a POM and the Weibull is the only continuous distribution which has the property of being both an AFM and a PHM.

4.7.2 Notation for a regression model with one risk factor at staggered entry

Consider a categorical risk factor A on three levels A_1, A_2 and A_3.

The notation for staggered entry of policies with four different entry periods, as described in chapter three, can be extended in the following way when fitting a regression model with one risk factor. For simplicity, assume that k, the number of class intervals for the first entry, is equal to seven,

The combined relative frequency vector $\boldsymbol{p}'$ is defined as

$\boldsymbol{p}' = (\boldsymbol{p}'_{11}, \boldsymbol{p}'_{21}, \boldsymbol{p}'_{31}, \boldsymbol{p}'_{41}, \boldsymbol{p}'_{12}, \boldsymbol{p}'_{22}, \boldsymbol{p}'_{32}, \boldsymbol{p}'_{42}, \boldsymbol{p}'_{13}, \boldsymbol{p}'_{23}, \boldsymbol{p}'_{33}, \boldsymbol{p}'_{43})$

$\boldsymbol{p}_{il}$ is the relative frequency vector for the i^{th} entry group and the l^{th} level of risk factor A, corresponding to $n_{il}\boldsymbol{p}_{il}$ being multinomial $(n_{il}; \boldsymbol{\pi}_{il})$ distributed

$i = 1, 2, 3, 4$ and $l = 1, 2, 3.$

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

and

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

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

$$\boldsymbol{x}_1 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} \quad \boldsymbol{x}_2 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} \quad \boldsymbol{x}_3 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \quad \text{and} \quad \boldsymbol{x}_4 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

The number of entries per cell in the cross tabulation of entry period and risk factor A can be summarized in table 4.1.

Table 4.1: **Number of entries per cell in cross table of entry period and risk factor A**

Entry Period	Risk Factor A Level A_1	Level A_2	Level A_3	Total
1	7	7	7	21
2	6	6	6	18
3	5	5	5	15
4	4	4	4	12
Total	22	22	22	66

Define matrix $\boldsymbol{S}$ as

$$\boldsymbol{S} = \begin{pmatrix} \boldsymbol{S}_{A_1} \\ \boldsymbol{S}_{A_2} \\ \boldsymbol{S}_{A_3} \end{pmatrix}$$

where $\boldsymbol{S}_{A_l} = \text{block}(\boldsymbol{S}_1, \boldsymbol{S}_2, \boldsymbol{S}_3, \boldsymbol{S}_4) \quad l = 1, 2, 3$ with

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

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

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

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

$\boldsymbol{S}$ is a 54×66 matrix. $\boldsymbol{S}_{A_l}$ is a 18×22 matrix l=1,2,3 and has the following form.

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

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

$$\Longrightarrow \quad \widehat{\boldsymbol{V}} = \text{block}(\widehat{\boldsymbol{V}}_{A_1}, \widehat{\boldsymbol{V}}_{A_2}, \widehat{\boldsymbol{V}}_{A_3})$$

where

$$\widehat{\boldsymbol{V}}_{A_l} = \text{block}(\widehat{\boldsymbol{V}}_{1,A_l}, \widehat{\boldsymbol{V}}_{2,A_l}, \widehat{\boldsymbol{V}}_{3,A_l}, \widehat{\boldsymbol{V}}_{4,A_l}) \quad l = 1,2,3$$

and

$$\widehat{\boldsymbol{V}}_{i,A_l} = \frac{1}{n_{il}} [diag(\boldsymbol{p}_{il}) - \boldsymbol{p}_{il}\boldsymbol{p}_{il}'] \quad i = 1,2,3,4 \quad \text{and} \quad l = 1,2,3$$

Note that $\widehat{\boldsymbol{V}}_{1,A_l}$ is a 7×7, $\widehat{\boldsymbol{V}}_{2,A_l}$ is a 6×6, $\widehat{\boldsymbol{V}}_{3,A_l}$ is a 5×5 and $\widehat{\boldsymbol{V}}_{4,A_l}$ is a 4×4 matrix so that $\widehat{\boldsymbol{V}}_{A_l}$ is a 22×22 matrix $l = 1,2,3$ and $\widehat{\boldsymbol{V}}$ is a 66×66 matrix.

4.7.3 The log-logistic regression model: staggered entry, shape parameter remains constant

In this model the effect of the risk factor is to alter the scale parameter λ, while the shape parameter α remains constant.

Equation 3.3 .23 can be extended to take covariates into account as follows:

$$\ln\left(\frac{\mathbf{1} - S(\boldsymbol{x})}{S(\boldsymbol{x})}\right) = \ln\lambda \cdot \mathbf{1} + \boldsymbol{\beta}'\boldsymbol{Z} + \alpha \cdot \ln \boldsymbol{x} \tag{4.7 .1}$$

Consider a risk factor A on three levels A_1, A_2 and A_3. Two dummy variables Z_{A_1} and Z_{A_2} are defined for levels A_1 and A_2 in such a way that the regression coefficient β_{A_3} of level A_3 is equal to $-(\beta_{A_1} + \beta_{A_2})$, that means $\left\{\sum_{k=1}^{3} \beta_{A_k}\right\} = 0$.

Then Equation 4.7 .1 becomes

$$\ln\left(\frac{\mathbf{1}-S(\boldsymbol{x})}{S(\boldsymbol{x})}\right) = \ln\lambda\cdot\mathbf{1} + (\beta_{A_1}Z_{A_1} + \beta_{A_2}Z_{A_2}) + \alpha\cdot\ln\boldsymbol{x} \tag{4.7 .2}$$

or

$$\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\left(\boldsymbol{\pi}_S\right) - \ln\left(\mathbf{1}-\boldsymbol{\pi}_S\right) \\
&= \ln\lambda\cdot\begin{pmatrix} \mathbf{1}\\ \mathbf{1}\\ \mathbf{1}\\ \mathbf{1}\\ \\ \mathbf{1}\\ \mathbf{1}\\ \mathbf{1}\\ \mathbf{1}\\ \\ \mathbf{1}\\ \mathbf{1}\\ \mathbf{1}\\ \mathbf{1} \end{pmatrix} + \beta_{A_1}\cdot\begin{pmatrix} \mathbf{1}\\ \mathbf{1}\\ \mathbf{1}\\ \mathbf{1}\\ \\ \mathbf{0}\\ \mathbf{0}\\ \mathbf{0}\\ \mathbf{0}\\ \\ -\mathbf{1}\\ -\mathbf{1}\\ -\mathbf{1}\\ -\mathbf{1} \end{pmatrix} + \beta_{A_2}\cdot\begin{pmatrix} \mathbf{0}\\ \mathbf{0}\\ \mathbf{0}\\ \mathbf{0}\\ \\ \mathbf{1}\\ \mathbf{1}\\ \mathbf{1}\\ \mathbf{1}\\ \\ -\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\\ \\ \ln\boldsymbol{x}_1\\ \ln\boldsymbol{x}_2\\ \ln\boldsymbol{x}_3\\ \ln\boldsymbol{x}_4\\ \\ \ln\boldsymbol{x}_1\\ \ln\boldsymbol{x}_2\\ \ln\boldsymbol{x}_3\\ \ln\boldsymbol{x}_4 \end{pmatrix} \\
&= \underbrace{\begin{pmatrix}
\mathbf{1} & \mathbf{1} & \mathbf{0} & \ln\boldsymbol{x}_1\\
\mathbf{1} & \mathbf{1} & \mathbf{0} & \ln\boldsymbol{x}_2\\
\mathbf{1} & \mathbf{1} & \mathbf{0} & \ln\boldsymbol{x}_3\\
\mathbf{1} & \mathbf{1} & \mathbf{0} & \ln\boldsymbol{x}_4\\
 & & & \\
\mathbf{1} & \mathbf{0} & \mathbf{1} & \ln\boldsymbol{x}_1\\
\mathbf{1} & \mathbf{0} & \mathbf{1} & \ln\boldsymbol{x}_2\\
\mathbf{1} & \mathbf{0} & \mathbf{1} & \ln\boldsymbol{x}_3\\
\mathbf{1} & \mathbf{0} & \mathbf{1} & \ln\boldsymbol{x}_4\\
 & & & \\
\mathbf{1} & -\mathbf{1} & -\mathbf{1} & \ln\boldsymbol{x}_1\\
\mathbf{1} & -\mathbf{1} & -\mathbf{1} & \ln\boldsymbol{x}_2\\
\mathbf{1} & -\mathbf{1} & -\mathbf{1} & \ln\boldsymbol{x}_3\\
\mathbf{1} & -\mathbf{1} & -\mathbf{1} & \ln\boldsymbol{x}_4
\end{pmatrix}} \cdot \begin{pmatrix} \ln\lambda\\ \beta_{A_1}\\ \beta_{A_2}\\ \alpha \end{pmatrix}
\end{aligned}$$

$$\Rightarrow \ln\left(\frac{\boldsymbol{\pi}_S}{\mathbf{1}-\boldsymbol{\pi}_S}\right) = \boldsymbol{X}_1 \cdot \begin{pmatrix} \ln\lambda \\ \beta_{A_1} \\ \beta_{A_2} \\ \alpha \end{pmatrix} \qquad (4.7\ .3)$$

Equation 4.7 .3 is a linear model in the parameters $\ln\lambda$, β_{A_1}, β_{A_2} 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 $g(\boldsymbol{\pi}) = \mathbf{0}$ 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 distributions for the three levels of the risk factor A.

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

$$g(\boldsymbol{\pi}) = \boldsymbol{C}.\ln\left\{\frac{\boldsymbol{\pi}_S}{\mathbf{1}-\boldsymbol{\pi}_S}\right\} = \boldsymbol{C}.\ln\left[\frac{\boldsymbol{S}\cdot\boldsymbol{\pi}}{\mathbf{1}-\boldsymbol{S}\cdot\boldsymbol{\pi}}\right] = \boldsymbol{C}.\left[\ln(\boldsymbol{S}\cdot\boldsymbol{\pi}) - \ln(\mathbf{1}-\boldsymbol{S}\cdot\boldsymbol{\pi})\right] = \mathbf{0} \qquad (4.7\ .4)$$

with

$$\boldsymbol{C} = \boldsymbol{I}-\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1' \quad . \qquad (4.7\ .5)$$

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

$$\begin{aligned} \boldsymbol{G}_\pi &= \frac{\partial \boldsymbol{g}(\boldsymbol{\pi})}{\partial \boldsymbol{\pi}} \\ &= \boldsymbol{C}\cdot\left[diag\left(\frac{1}{\boldsymbol{\pi}_S}\right) + diag\left(\frac{1}{\mathbf{1}-\boldsymbol{\pi}_S}\right)\right]\cdot\boldsymbol{S} \qquad (4.7\ .6) \\ &= \boldsymbol{C}\cdot\left[\boldsymbol{D}_3^{-1}+\boldsymbol{D}_2^{-1}\right]\cdot\boldsymbol{S} \qquad (4.7\ .7) \end{aligned}$$

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 estimated vector of probabilities in this case is

$$\widehat{\boldsymbol{\pi}}_c = \boldsymbol{p} - (\boldsymbol{G}_\pi \boldsymbol{V})' (\boldsymbol{G}_p \boldsymbol{V} \boldsymbol{G}_\pi')^* \cdot \boldsymbol{C}.\ln\left\{\frac{\boldsymbol{S}\cdot\boldsymbol{p}}{\boldsymbol{1}-\boldsymbol{S}\cdot\boldsymbol{p}}\right\} \quad (4.7\ .8)$$

with $\boldsymbol{S}$ and $\widehat{\boldsymbol{V}}$, the estimated variance-covariance matrix, defined in section 4.7.2 .

Since Equation 4.7 .8 is still a function of the unknown parameter $\boldsymbol{\pi}$, the double iterative procedure must be implemented. Once the iterative procedure in Equation 4.7 .8 has converged, the estimated parameters of the three log-logistic distributions can be solved from

$$\begin{pmatrix} \widehat{\ln\lambda} \\ \widehat{\beta}_{A_1} \\ \widehat{\beta}_{A_2} \\ \widehat{\alpha} \end{pmatrix} = (\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}' \cdot \ln\left\{\frac{\boldsymbol{S}\cdot\widehat{\boldsymbol{\pi}}_c}{\boldsymbol{1}-\boldsymbol{S}\cdot\widehat{\boldsymbol{\pi}}_c}\right\} \quad (4.7\ .9)$$

and $\widehat{\beta}_{A_3} = -\left(\widehat{\beta}_{A_1} + \widehat{\beta}_{A_2}\right)$.

The estimated lambda parameters of the three log-logistic distributions for the three risk factor levels then are

$$\begin{aligned} \widehat{\lambda}_{A_1} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_1}) \\ \widehat{\lambda}_{A_2} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_2}) \\ \widehat{\lambda}_{A_3} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_3}). \end{aligned}$$

Consider Equation 4.7 .2 where $\ln(odds)$ is modelled in terms of dummy variables Z_{A_1} and Z_{A_2}

$$\Rightarrow \ln(odds) = \ln\lambda\cdot\boldsymbol{1} + (\beta_{A_1}Z_{A_1} + \beta_{A_2}Z_{A_2}) + \alpha\cdot\ln\boldsymbol{x}$$

Take the summation over the risk factor levels, that gives

$$\sum\{\ln(odds)\} = \sum\{\ln\lambda\cdot\boldsymbol{1}\} + \sum\{(\beta_{A_1}\cdot Z_{A_1} + \beta_{A_2}\cdot Z_{A_2})\} + \sum\{\alpha\cdot\ln\boldsymbol{x}\},$$

but

$$\begin{aligned} \sum\{(\beta_{A_1}\cdot Z_{A_1} + \beta_{A_2}\cdot Z_{A_2})\} &= (\beta_{A_1}.1 + \beta_{A_2}.0) + (\beta_{A_1}.0 + \beta_{A_2}.1) + (\beta_{A_1}.(-1) + \beta_{A_2}.(-1)) \\ &= \beta_{A_1} + \beta_{A_2} + \beta_{A_3} \\ &= 0, \end{aligned}$$

therefore

$$\sum\{\ln(odds)\} = \sum\{\ln\lambda\cdot\boldsymbol{1}\} + \sum\{\alpha\cdot\ln\boldsymbol{x}\}.$$

$\Rightarrow$ average$\{\ln(odds)\} = \ln\lambda\cdot\boldsymbol{1} + \alpha\cdot\ln\boldsymbol{x}$.

The **baseline** log-logistic model in this context is defined as the log-logistic model at this average value of $\ln(odds)$.

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

$$\widehat{\lambda}_0 = \exp(\widehat{\ln\lambda})$$

so that

$$\begin{aligned}\widehat{\lambda}_{A_1} &= \widehat{\lambda}_0 \times \exp(\widehat{\beta}_{A_1})\\ \widehat{\lambda}_{A_2} &= \widehat{\lambda}_0 \times \exp(\widehat{\beta}_{A_2})\\ \widehat{\lambda}_{A_3} &= \widehat{\lambda}_0 \times \exp(\widehat{\beta}_{A_3}).\end{aligned}$$

These four log-logistic distributions all have the same estimated alpha parameter $\widehat{\alpha}$.

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

4.7.4 The log-logistic regression model: staggered entry, shape parameter alters

In this model the effect of the risk factor is to alter both the scale parameter λ and the shape parameter α.

Then Equation 4.7 .1 becomes

$$\ln\left(\frac{\mathbf{1}-S(\boldsymbol{x})}{S(\boldsymbol{x})}\right)=\ln\left(\frac{F(\boldsymbol{x})}{\mathbf{1}-F(\boldsymbol{x})}\right)=\ln\left(\frac{\boldsymbol{\pi}_S}{\mathbf{1}-\boldsymbol{\pi}_S}\right)=\ln\left(\boldsymbol{\pi}_S\right)-\ln\left(\mathbf{1}-\boldsymbol{\pi}_S\right)$$

$$=\ln\lambda\cdot\begin{pmatrix}1\\1\\1\\1\\ \\1\\1\\1\\1\\ \\1\\1\\1\\1\end{pmatrix}+\beta_{A_1}\cdot\begin{pmatrix}1\\1\\1\\1\\ \\0\\0\\0\\0\\ \\-1\\-1\\-1\\-1\end{pmatrix}+\beta_{A_2}\cdot\begin{pmatrix}0\\0\\0\\0\\ \\1\\1\\1\\1\\ \\-1\\-1\\-1\\-1\end{pmatrix}+\alpha_{A_1}\cdot\begin{pmatrix}\ln x_1\\\ln x_2\\\ln x_3\\\ln x_4\\ \\0\\0\\0\\0\\ \\0\\0\\0\\0\end{pmatrix}+\alpha_{A_2}\cdot\begin{pmatrix}0\\0\\0\\0\\ \\\ln x_1\\\ln x_2\\\ln x_3\\\ln x_4\\ \\0\\0\\0\\0\end{pmatrix}+\alpha_{A_3}\cdot\begin{pmatrix}0\\0\\0\\0\\ \\0\\0\\0\\0\\ \\\ln x_1\\\ln x_2\\\ln x_3\\\ln x_4\end{pmatrix}$$

$$= \underbrace{\begin{pmatrix} 1 & 1 & 0 & \ln x_1 & 0 & 0 \\ 1 & 1 & 0 & \ln x_2 & 0 & 0 \\ 1 & 1 & 0 & \ln x_3 & 0 & 0 \\ 1 & 1 & 0 & \ln x_4 & 0 & 0 \\ & & & & & \\ 1 & 0 & 1 & 0 & \ln x_1 & 0 \\ 1 & 0 & 1 & 0 & \ln x_2 & 0 \\ 1 & 0 & 1 & 0 & \ln x_3 & 0 \\ 1 & 0 & 1 & 0 & \ln x_4 & 0 \\ & & & & & \\ 1 & -1 & -1 & 0 & 0 & \ln x_1 \\ 1 & -1 & -1 & 0 & 0 & \ln x_2 \\ 1 & -1 & -1 & 0 & 0 & \ln x_3 \\ 1 & -1 & -1 & 0 & 0 & \ln x_4 \end{pmatrix}} \cdot \begin{pmatrix} \ln\lambda \\ \beta_{A_1} \\ \beta_{A_2} \\ \alpha_{A_1} \\ \alpha_{A_2} \\ \alpha_{A_3} \end{pmatrix}$$

$$\Rightarrow \ln\left(\frac{\pi_S}{1-\pi_S}\right) = \boldsymbol{X}_1 \cdot \begin{pmatrix} \ln\lambda \\ \beta_{A_1} \\ \beta_{A_2} \\ \alpha_{A_1} \\ \alpha_{A_2} \\ \alpha_{A_3} \end{pmatrix}$$

This is a linear model in the parameters $\ln\lambda$, β_{A_1}, β_{A_2}, α_{A_1}, α_{A_2} and α_{A_3}. This model is equivalent to

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

$\boldsymbol{C}$ 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 $g(\pi) = 0$ 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 distributions for the three levels of the risk factor A.

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

$$g(\boldsymbol{\pi}) = \boldsymbol{C}.\ln\left\{\frac{\boldsymbol{\pi}_S}{\mathbf{1}-\boldsymbol{\pi}_S}\right\} = \boldsymbol{C}.\ln\left[\frac{\boldsymbol{S}\cdot\boldsymbol{\pi}}{\mathbf{1}-\boldsymbol{S}\cdot\boldsymbol{\pi}}\right] = \boldsymbol{C}.\left[\ln(\boldsymbol{S}\cdot\boldsymbol{\pi}) - \ln(\mathbf{1}-\boldsymbol{S}\cdot\boldsymbol{\pi})\right] = \mathbf{0} \tag{4.7 .10}$$

with

$$\boldsymbol{C} = \boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}' \quad . \tag{4.7 .11}$$

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

$$\begin{aligned} \boldsymbol{G}_\pi &= \frac{\partial g(\boldsymbol{\pi})}{\partial \boldsymbol{\pi}} \\ &= \boldsymbol{C}\cdot\left[diag\left(\frac{1}{\boldsymbol{\pi}_S}\right) + diag\left(\frac{1}{\mathbf{1}-\boldsymbol{\pi}_S}\right)\right]\cdot\boldsymbol{S} && (4.7\ .12) \\ &= \boldsymbol{C}\cdot\left[\boldsymbol{D}_3{}^{-1} + \boldsymbol{D}_2{}^{-1}\right]\cdot\boldsymbol{S} && (4.7\ .13) \end{aligned}$$

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. $\boldsymbol{S}$ is a matrix composed from three matrices associated with the three levels of risk factor A.

The estimated vector of probabilities in this case is

$$\widehat{\boldsymbol{\pi}}_c = \boldsymbol{p} - (\boldsymbol{G}_\pi\boldsymbol{V})'\left(\boldsymbol{G}_p\boldsymbol{V}\boldsymbol{G}_\pi'\right)^{-1}\cdot\boldsymbol{C}.\ln\left\{\frac{\boldsymbol{S}\cdot\boldsymbol{p}}{\mathbf{1}-\boldsymbol{S}\cdot\boldsymbol{p}}\right\} \tag{4.7 .14}$$

with $\boldsymbol{S}$ and $\widehat{\boldsymbol{V}}$, the estimated variance-covariance matrix, defined in section 4.7.2 .

Since Equation 4.7 .14 is still a function of the unknown parameter $\boldsymbol{\pi}$, the double iterative procedure must be implemented. Once the iterative procedure in Equation 4.7 .14 has converged, the estimated parameters of the three log-logistic distributions can be solved from

$$\begin{pmatrix} \widehat{\ln\lambda} \\ \widehat{\beta}_{A_1} \\ \widehat{\beta}_{A_2} \\ \widehat{\alpha}_{A_1} \\ \widehat{\alpha}_{A_2} \\ \widehat{\alpha}_{A_3} \end{pmatrix} = (\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}'\cdot\ln\left\{\frac{\boldsymbol{S}\cdot\widehat{\boldsymbol{\pi}}_c}{\mathbf{1}-\boldsymbol{S}\cdot\widehat{\boldsymbol{\pi}}_c}\right\} \tag{4.7 .15}$$

and $\widehat{\beta}_{A_3} = -\left(\widehat{\beta}_{A_1} + \widehat{\beta}_{A_2}\right)$.

The estimated lambda parameters of the three log-logistic distributions for the three risk factor levels then are

$$\begin{aligned}\widehat{\lambda}_{A_1} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_1}) \\ \widehat{\lambda}_{A_2} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_2}) \\ \widehat{\lambda}_{A_3} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_3}).\end{aligned}$$

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

$$\widehat{\lambda}_0 = \exp(\widehat{\ln\lambda})$$

so that

$$\begin{aligned}\widehat{\lambda}_{A_1} &= \widehat{\lambda}_0 \times \exp(\widehat{\beta}_{A_1}) \\ \widehat{\lambda}_{A_2} &= \widehat{\lambda}_0 \times \exp(\widehat{\beta}_{A_2}) \\ \widehat{\lambda}_{A_3} &= \widehat{\lambda}_0 \times \exp(\widehat{\beta}_{A_3}).\end{aligned}$$

The estimated shape parameters of the baseline and the risk factor level log-logistic distributions are

$$\begin{aligned}&\widehat{\alpha}_{A_1} \\ &\widehat{\alpha}_{A_2} \\ &\widehat{\alpha}_{A_3}.\end{aligned}$$

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

4.7.5 Deriving of indices and risk scores from the log-logistic regression model

Once the parameters of the log-logistic baseline distribution and log-logistic risk factor level distributions have been estimated, estimated hazard and survivor functions, odds of a lapse, odds ratios and hazard ratios at time t can be calculated.

The odds ratio for risk factor level A_1 is the relative odds of a lapse at time t of a policy, with level A_1 characteristics, compared to a policy with the baseline characteristics. The odds ratios for the three risk factor levels result in a set of indices, showing the effect of each risk factor level on the baseline odds of a lapse at time t.

The hazard ratio for risk factor level A_1 is the relative hazard rate of a lapse at time t of a policy, with level A_1 characteristics, compared to a policy with the baseline characteristics. The hazard ratios for the three risk factor levels result in a set of risk scores, showing the effect of each risk factor level on the baseline hazard rate of a lapse at time t.

Percentiles of the four log-logistic survival distributions can also be estimated.

Log-logistic regression model

Shape remains constant

Estimated hazard function

$$\widehat{h}_0(t) = \frac{\widehat{\lambda}_0 \cdot \widehat{\alpha} \cdot t^{\widehat{\alpha}-1}}{(1+\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}})}$$

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

Estimated survivor function

$$\widehat{S}_0(t) = \frac{1}{1+\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}}}$$

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

Estimated odds of a lapse

$$\widehat{odds}_0(t) = \frac{1-\widehat{S}_0(t)}{\widehat{S}_0(t)} = \widehat{\lambda}_0 \cdot t^{\widehat{\alpha}}$$

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

Estimated odds ratio or index

$$\widehat{oddsratio}_{A_i}(t) = \frac{\widehat{odds}_{A_i}(t)}{\widehat{odds}_0(t)} = \frac{\widehat{\lambda}_{A_i} \cdot t^{\widehat{\alpha}}}{\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}}}$$

Estimated hazard ratio or risk score

$$\widehat{hazardratio}_{A_i}(t) = \frac{\widehat{h}_{A_i}(t)}{\widehat{h}_0(t)} = \frac{\widehat{\lambda}_{A_i}}{\widehat{\lambda}_0} \cdot \frac{(1+\widehat{\lambda}_0 t^{\widehat{\alpha}})}{(1+\widehat{\lambda}_{A_i} t^{\widehat{\alpha}})}$$

Estimated percentiles of lifetime distributions

$$\text{baseline } \widehat{t_p} = \left(\frac{1}{\widehat{\lambda}_0} \cdot \frac{p}{(100-p)}\right)^{\frac{1}{\widehat{\alpha}}}$$

$$\text{pred. level } \widehat{t_p} = \left(\frac{1}{\widehat{\lambda}_{A_i}} \cdot \frac{p}{(100-p)}\right)^{\frac{1}{\widehat{\alpha}}} = \text{baseline } \widehat{t_p} \cdot (\text{index})^{-\frac{1}{\widehat{\alpha}}}$$

Shape parameter alters

Estimated hazard function

$$\widehat{h}_0(t) = \frac{\widehat{\lambda}_0 \cdot \widehat{\alpha}_0 \cdot t^{\widehat{\alpha}_0-1}}{(1+\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}_0})}$$

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

Estimated survivor function

$$\widehat{S}_0(t) = \frac{1}{1+\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}_0}}$$

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

Estimated odds of a lapse

$$\widehat{odds}_0(t) = \frac{1-\widehat{S}_0(t)}{\widehat{S}_0(t)} = \widehat{\lambda}_0 \cdot t^{\widehat{\alpha}_0}$$

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

Estimated odds ratio or index

$$\widehat{oddsratio}_{A_i}(t) = \frac{\widehat{odds}_{A_i}(t)}{\widehat{odds}_0(t)} = \frac{\widehat{\lambda}_{A_i} \cdot t^{\widehat{\alpha}_{A_i}}}{\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}_0}}$$

Estimated hazard ratio or risk score

$$\widehat{hazardratio}_{A_i}(t) = \frac{\widehat{h}_{A_i}(t)}{\widehat{h}_0(t)} = \frac{\widehat{\lambda}_{A_i}}{\widehat{\lambda}_0} \cdot \frac{(1+\widehat{\lambda}_0 t^{\widehat{\alpha}_0})}{(1+\widehat{\lambda}_{A_i} t^{\widehat{\alpha}_{A_i}})}$$

Estimated percentiles of lifetime distributions

$$\text{baseline } \widehat{t_p} = \left(\frac{1}{\widehat{\lambda}_0} \cdot \frac{p}{(100-p)}\right)^{\frac{1}{\widehat{\alpha}_0}}$$

$$\text{pred. level } \widehat{t_p} = \left(\frac{1}{\widehat{\lambda}_{A_i}} \cdot \frac{p}{(100-p)}\right)^{\frac{1}{\widehat{\alpha}_{A_i}}}$$

$i = 1, 2, 3$

The estimated odds ratios are called **indices**. The index of a risk factor level shows the effect of this level on the baseline odds of a lapse. This effect is multiplicative on the baseline odds of a lapse and increases the baseline odds of a lapse (if the index > 1) or decreases the baseline odds of a lapse (if the index < 1).

The estimated hazard ratios are called **risk scores**. The risk score of a risk factor level shows the effect of this level on the baseline hazard rate of a lapse. This effect is multiplicative on the baseline hazard rate of a lapse and increases the baseline hazard rate of a lapse (if the risk score > 1) or decreases the baseline hazard rate of a lapse (if the risk score < 1).

Consider the risk factor A on three levels A_1, A_2 and A_3. Recall that the proportional odds model (POM) models the **odds of a lapse** at time t for a policy with risk vector $\boldsymbol{Z} = (Z_1, Z_2, ..., Z_p)'$ and regression coefficients $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_p)'$

$$\Rightarrow \boxed{odds_{A_l}(t|\boldsymbol{Z}) = e^{\boldsymbol{\beta}'\boldsymbol{Z}} \cdot odds_0(t) \quad l = 1, 2, 3.}$$

This property of constant odds ratios over time only holds when the shape parameter of the log-logistic distributions of the baseline and risk factor levels remains constant.

Two dummy variables Z_{A_1} and Z_{A_2} are defined for levels A_1 and A_2 in such a way that the regression coefficient β_{A_3} of level A_3 is equal to $-(\beta_{A_1} + \beta_{A_2})$. From the POM follows that

$$odds_{A_1}(t|Z_{A_1} = 1, Z_{A_2} = 0) = e^{(\beta_{A_1}.1+\beta_{A_2}.0)}.odds_0(t) \Rightarrow \frac{odds_{A_1}(t|Z_{A_1} = 1, Z_{A_2} = 0)}{odds_0(t)} = e^{\beta_{A_1}}$$

$$\Rightarrow \boxed{\widehat{oddsratio}_{A_1} = e^{\widehat{\beta}_{A_1}} = \frac{\widehat{\lambda}_{A_1} \cdot t^{\widehat{\alpha}}}{\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}}} = \frac{\widehat{\lambda}_{A_1}}{\widehat{\lambda}_0}}$$

This means that for a **constant shape parameter** in the log-logistic distributions, the indices (estimated odds ratios) may be obtained also from the exponent of the estimated β-values in the log-logistic regression model.

4.7.6 The Weibull regression model: staggered entry, the shape parameter remains constant

In this model the effect of the risk factor is to alter the scale parameter λ, while the shape parameter α remains constant.

Equation 3.3 .11 can be extended to take covariates into account as follows:

$$\ln(-\ln S(\boldsymbol{x})) = \ln \lambda \cdot \boldsymbol{1} + \boldsymbol{\beta}'\boldsymbol{Z} + \alpha \cdot \ln \boldsymbol{x} \qquad (4.7\ .16)$$

Consider again a risk factor A on three levels A_1, A_2 and A_3 for which two dummy variables Z_{A_1} and Z_{A_2} are defined. The staggered entry of policies occurs during four entry periods and k, the number of class intervals for the first entry group, equals seven.

Then Equation 4.7 .16 becomes

$$\ln(-\ln S(\boldsymbol{x})) = \ln\lambda \cdot \mathbf{1} + (\beta_{A_1} Z_{A_1} + \beta_{A_2} Z_{A_2}) + \alpha \cdot \ln \boldsymbol{x} \qquad (4.7\ .17)$$

or

$$\begin{aligned}
\ln(-\ln S(\boldsymbol{x})) &= \ln(-\ln(\mathbf{1} - F(\boldsymbol{x}))) = \ln(-\ln(\mathbf{1} - \boldsymbol{\pi}_S)) \\
&= \ln\lambda \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ \\ 1 \\ 1 \\ 1 \\ 1 \\ \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \beta_{A_1} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ \\ 0 \\ 0 \\ 0 \\ 0 \\ \\ -1 \\ -1 \\ -1 \\ -1 \end{pmatrix} + \beta_{A_2} \cdot \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \\ 1 \\ 1 \\ 1 \\ 1 \\ \\ -1 \\ -1 \\ -1 \\ -1 \end{pmatrix} + \alpha \cdot \begin{pmatrix} \ln x_1 \\ \ln x_2 \\ \ln x_3 \\ \ln x_4 \\ \\ \ln x_1 \\ \ln x_2 \\ \ln x_3 \\ \ln x_4 \\ \\ \ln x_1 \\ \ln x_2 \\ \ln x_3 \\ \ln x_4 \end{pmatrix} \\
&= \underbrace{\begin{pmatrix} 1 & 1 & 0 & \ln x_1 \\ 1 & 1 & 0 & \ln x_2 \\ 1 & 1 & 0 & \ln x_3 \\ 1 & 1 & 0 & \ln x_4 \\ & & & \\ 1 & 0 & 1 & \ln x_1 \\ 1 & 0 & 1 & \ln x_2 \\ 1 & 0 & 1 & \ln x_3 \\ 1 & 0 & 1 & \ln x_4 \\ & & & \\ 1 & -1 & -1 & \ln x_1 \\ 1 & -1 & -1 & \ln x_2 \\ 1 & -1 & -1 & \ln x_3 \\ 1 & -1 & -1 & \ln x_4 \end{pmatrix}} \cdot \begin{pmatrix} \ln\lambda \\ \beta_{A_1} \\ \beta_{A_2} \\ \alpha \end{pmatrix}
\end{aligned}$$

$$\Rightarrow \ln\left(-\ln(\mathbf{1}-\boldsymbol{\pi}_S)\right) = \boldsymbol{X}_1 \cdot \begin{pmatrix} \ln\lambda \\ \beta_{A_1} \\ \beta_{A_2} \\ \alpha \end{pmatrix} \tag{4.7 .18}$$

Equation 4.7 .18 is a linear model in the parameters $\ln\lambda$, β_{A_1}, β_{A_2} and α. This model is equivalent to

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

$\boldsymbol{C}$ 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 $g(\boldsymbol{\pi}) = \mathbf{0}$ satisfies the conditions of Proposition 1 and the estimation algorithm can be used to estimate the λ's and α's of the Weibull distributions for the three levels of the risk factor A.

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

$$g(\boldsymbol{\pi}) = \boldsymbol{C}.\ln\left\{-\ln(\mathbf{1}-\boldsymbol{\pi}_S)\right\} = \boldsymbol{C}.\ln\left\{-\ln(\mathbf{1}-\boldsymbol{S}\cdot\boldsymbol{\pi})\right\} = \mathbf{0} \tag{4.7 .19}$$

with

$$\boldsymbol{C} = \boldsymbol{I}-\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1' \quad . \tag{4.7 .20}$$

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

$$\begin{aligned} \boldsymbol{G}_\pi &= \frac{\partial \boldsymbol{g}(\boldsymbol{\pi})}{\partial \boldsymbol{\pi}} \\ &= -\boldsymbol{C}\cdot diag\left(\frac{1}{\ln(\mathbf{1}-\boldsymbol{\pi}_S)}\right)\cdot diag\left(\frac{1}{\mathbf{1}-\boldsymbol{\pi}_S}\right)\cdot \boldsymbol{S} && (4.7\ .21) \\ &= -\boldsymbol{C}\cdot \boldsymbol{D}_1^{-1}\cdot \boldsymbol{D}_2^{-1}\cdot \boldsymbol{S} && (4.7\ .22) \end{aligned}$$

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.

The estimated vector of probabilities is in this case

$$\widehat{\boldsymbol{\pi}}_c = \boldsymbol{p} - (\boldsymbol{G}_\pi \boldsymbol{V})' (\boldsymbol{G}_p \boldsymbol{V} \boldsymbol{G}_\pi')^* \cdot \boldsymbol{C}. \ln \left\{ \frac{\boldsymbol{S} \cdot \boldsymbol{p}}{\boldsymbol{1} - \boldsymbol{S} \cdot \boldsymbol{p}} \right\}. \qquad (4.7\ .23)$$

with $\boldsymbol{S}$ and $\widehat{\boldsymbol{V}}$, the estimated variance-covariance matrix, defined in section 4.7.2 .

Since Equation 4.7 .23 is still a function of the unknown parameter $\boldsymbol{\pi}$, the double iterative procedure must be implemented. Once the iterative procedure in Equation 4.7 .23 has converged, the estimated parameters of the three Weibull distributions can be solved from

$$\begin{pmatrix} \widehat{\ln \lambda} \\ \widehat{\beta}_{A_1} \\ \widehat{\beta}_{A_2} \\ \widehat{\alpha} \end{pmatrix} = (\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}' \cdot \ln\left(-\ln(\boldsymbol{1} - \boldsymbol{S} \cdot \widehat{\boldsymbol{\pi}}_c)\right) \qquad (4.7\ .24)$$

and $\widehat{\beta}_{A_3} = -\left(\widehat{\beta}_{A_1} + \widehat{\beta}_{A_2}\right)$.

The estimated lambda parameters of the three Weibull distributions for the three risk factor levels then are

$$\begin{aligned} \widehat{\lambda}_{A_1} &= \exp(\widehat{\ln \lambda} + \widehat{\beta}_{A_1}) \\ \widehat{\lambda}_{A_2} &= \exp(\widehat{\ln \lambda} + \widehat{\beta}_{A_2}) \\ \widehat{\lambda}_{A_3} &= \exp(\widehat{\ln \lambda} + \widehat{\beta}_{A_3}). \end{aligned}$$

The estimated lambda parameter of the baseline Weibull distribution then is

$$\widehat{\lambda}_0 = \exp(\widehat{\ln \lambda})$$

so that

$$\begin{aligned} \widehat{\lambda}_{A_1} &= \widehat{\lambda}_0 \times \exp(\widehat{\beta}_{A_1}) \\ \widehat{\lambda}_{A_2} &= \widehat{\lambda}_0 \times \exp(\widehat{\beta}_{A_2}) \\ \widehat{\lambda}_{A_3} &= \widehat{\lambda}_0 \times \exp(\widehat{\beta}_{A_3}). \end{aligned}$$

These four Weibull distributions all have the same estimated alpha parameter $\widehat{\alpha}$.

The SAS/IML program to fit a Weibull regression model (constant shape parameter) to grouped survival data with staggered entry of policies appears in Appendix A.

4.7.7 The Weibull regression model: staggered entry, the shape parameter alters

In this model the effect of the risk factor is to alter both the scale parameter λ and the shape parameter α.

Then Equation 4.7 .16 becomes

$$\ln(-\ln S(\boldsymbol{x})) = \ln(-\ln(1-F(\boldsymbol{x}))) = \ln(-\ln(\mathbf{1}-\boldsymbol{\pi}_S))$$

$$= \ln\lambda \cdot \begin{pmatrix} 1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1 \end{pmatrix} + \beta_{A_1} \cdot \begin{pmatrix} 1\\1\\1\\1\\0\\0\\0\\0\\-1\\-1\\-1\\-1 \end{pmatrix} + \beta_{A_2} \cdot \begin{pmatrix} 0\\0\\0\\0\\1\\1\\1\\1\\-1\\-1\\-1\\-1 \end{pmatrix} + \alpha_{A_1} \cdot \begin{pmatrix} \ln x_1\\\ln x_2\\\ln x_3\\\ln x_4\\0\\0\\0\\0\\0\\0\\0\\0 \end{pmatrix} + \alpha_{A_2} \cdot \begin{pmatrix} 0\\0\\0\\0\\\ln x_1\\\ln x_2\\\ln x_3\\\ln x_4\\0\\0\\0\\0 \end{pmatrix} + \alpha_{A_3} \cdot \begin{pmatrix} 0\\0\\0\\0\\0\\0\\0\\0\\\ln x_1\\\ln x_2\\\ln x_3\\\ln x_4 \end{pmatrix}$$

$$= \underbrace{\begin{pmatrix} 1 & 1 & 0 & \ln x_1 & 0 & 0\\ 1 & 1 & 0 & \ln x_2 & 0 & 0\\ 1 & 1 & 0 & \ln x_3 & 0 & 0\\ 1 & 1 & 0 & \ln x_4 & 0 & 0\\ 1 & 0 & 1 & 0 & \ln x_1 & 0\\ 1 & 0 & 1 & 0 & \ln x_2 & 0\\ 1 & 0 & 1 & 0 & \ln x_3 & 0\\ 1 & 0 & 1 & 0 & \ln x_4 & 0\\ 1 & -1 & -1 & 0 & 0 & \ln x_1\\ 1 & -1 & -1 & 0 & 0 & \ln x_2\\ 1 & -1 & -1 & 0 & 0 & \ln x_3\\ 1 & -1 & -1 & 0 & 0 & \ln x_4 \end{pmatrix}} \cdot \begin{pmatrix} \ln\lambda\\ \beta_{A_1}\\ \beta_{A_2}\\ \alpha_{A_1}\\ \alpha_{A_2}\\ \alpha_{A_3} \end{pmatrix}$$

$$\Rightarrow \ln(-\ln(\mathbf{1}-\boldsymbol{\pi}_S)) = \quad \boldsymbol{X}_1 \quad \cdot \begin{pmatrix} \ln\lambda\\ \beta_{A_1}\\ \beta_{A_2}\\ \alpha_{A_1}\\ \alpha_{A_2}\\ \alpha_{A_3} \end{pmatrix}$$

This is a linear model in the parameters $\ln\lambda$, β_{A_1}, β_{A_2}, α_{A_1}, α_{A_2} and α_{A_3}. This model is equivalent to

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

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

The function $g(\boldsymbol{\pi}) = \mathbf{0}$ satisfies the conditions of Proposition 1 and the estimation algorithm can be used to estimate the λ's and α's of the Weibull distributions for the three levels of the risk factor A.

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

$$g(\boldsymbol{\pi}) = \boldsymbol{C}.\ln\{-\ln(\mathbf{1}-\boldsymbol{\pi}_S)\} = \boldsymbol{C}.\ln\{-\ln(\mathbf{1}-\boldsymbol{S}\cdot\boldsymbol{\pi})\} = \mathbf{0} \qquad (4.7\ .25)$$

with

$$\boldsymbol{C} = \boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}' \quad . \qquad (4.7\ .26)$$

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

$$\begin{aligned}
\boldsymbol{G}_\pi &= \frac{\partial g(\boldsymbol{\pi})}{\partial \boldsymbol{\pi}} \\
&= -\boldsymbol{C}\cdot diag\left(\frac{1}{\ln(\mathbf{1}-\boldsymbol{\pi}_S)}\right)\cdot diag\left(\frac{1}{\mathbf{1}-\boldsymbol{\pi}_S}\right)\cdot\boldsymbol{S} \qquad (4.7\ .27) \\
&= -\boldsymbol{C}\cdot\boldsymbol{D}_1{}^{-1}\cdot\boldsymbol{D}_2{}^{-1}\cdot\boldsymbol{S} \qquad (4.7\ .28)
\end{aligned}$$

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.

The estimated vector of probabilities is in this case

$$\widehat{\boldsymbol{\pi}}_c = \boldsymbol{p} - (\boldsymbol{G}_\pi\boldsymbol{V})'\,(\boldsymbol{G}_p\boldsymbol{V}\boldsymbol{G}_\pi')^*\cdot\boldsymbol{C}.\ln\left\{\frac{\boldsymbol{S}\cdot\boldsymbol{p}}{\mathbf{1}-\boldsymbol{S}\cdot\boldsymbol{p}}\right\}. \qquad (4.7\ .29)$$

with $\boldsymbol{S}$ and $\widehat{\boldsymbol{V}}$, the estimated variance-covariance matrix, defined in section 4.7.2 .

Since Equation 4.7 .29 is still a function of the unknown parameter $\boldsymbol{\pi}$, the double iterative procedure must be implemented. Once the iterative procedure in Equation 4.7 .29 has converged, the estimated parameters of the three Weibull distributions can be solved from

$$\begin{pmatrix} \widehat{\ln\lambda} \\ \widehat{\beta}_{A_1} \\ \widehat{\beta}_{A_2} \\ \widehat{\alpha}_{A_1} \\ \widehat{\alpha}_{A_2} \\ \widehat{\alpha}_{A_3} \end{pmatrix} = (\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}' \cdot \ln\left\{\frac{\boldsymbol{S}\cdot\widehat{\boldsymbol{\pi}}_c}{\boldsymbol{1}-\boldsymbol{S}\cdot\widehat{\boldsymbol{\pi}}_c}\right\} \qquad (4.7\ .30)$$

and $\widehat{\beta}_{A_3} = -\left(\widehat{\beta}_{A_1} + \widehat{\beta}_{A_2}\right)$.

The estimated lambda parameters of the three Weibull distributions for the three risk factor levels then are

$$\begin{aligned} \widehat{\lambda}_{A_1} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_1}) \\ \widehat{\lambda}_{A_2} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_2}) \\ \widehat{\lambda}_{A_3} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_3}). \end{aligned}$$

The estimated lambda parameter of the baseline Weibull distribution then is

$$\widehat{\lambda}_0 = \exp(\widehat{\ln\lambda})$$

so that

$$\begin{aligned} \widehat{\lambda}_{A_1} &= \widehat{\lambda}_0 \times \exp(\widehat{\beta}_{A_1}) \\ \widehat{\lambda}_{A_2} &= \widehat{\lambda}_0 \times \exp(\widehat{\beta}_{A_2}) \\ \widehat{\lambda}_{A_3} &= \widehat{\lambda}_0 \times \exp(\widehat{\beta}_{A_3}). \end{aligned}$$

The estimated shape parameters of the baseline and the risk factor level Weibull distributions are

$$\begin{gathered} \widehat{\alpha}_{A_1} \\ \widehat{\alpha}_{A_2} \\ \widehat{\alpha}_{A_3}. \end{gathered}$$

The SAS/IML program to fit a Weibull regression model (shape parameter alters) to grouped survival data with staggered entry of policies appears in Appendix A.

4.7.8 Deriving of indices and risk scores from the Weibull regression model

Once the parameters of the Weibull baseline distribution and Weibull risk factor level distributions have been estimated, estimated hazard and survivor functions, odds of a lapse, odds ratios and hazard ratios at time t can be calculated.

The odds ratio for risk factor level A_1 is the relative odds of a lapse at time t of a policy, with level A_1 characteristics, compared to a policy with the baseline characteristics. The odds ratios for the three risk factor levels result in a set of indices, showing the effect of each risk factor level on the baseline odds of a lapse at time t.

The hazard ratio for risk factor level A_1 is the relative hazard rate of a lapse at time t of a policy, with level A_1 characteristics, compared to a policy with the baseline characteristics. The hazard ratios for the three risk factor levels result in a set of risk scores, showing the effect of each risk factor level on the baseline hazard rate of a lapse at time t.

Percentiles of the four Weibull survival distributions can also be estimated.

Weibull regression model

Shape remains constant

Estimated hazard function

$$\widehat{h}_0(t) = \widehat{\lambda}_0 \cdot \widehat{\alpha} \cdot t^{\widehat{\alpha}-1}$$

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

Estimated survivor function

$$\widehat{S}_0(t) = \exp(-\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}})$$

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

Estimated odds of a lapse

$$\widehat{odds}_0(t) = \frac{1-\widehat{S}_0(t)}{\widehat{S}_0(t)} = \exp(\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}-1})$$

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

Estimated odds ratio or index

$$\widehat{oddsratio}_{A_i}(t) = \frac{\widehat{odds}_{A_i}(t)}{\widehat{odds}_0(t)} = \frac{\exp(\widehat{\lambda}_{A_i} \cdot t^{\widehat{\alpha}-1})}{\exp(\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}-1})}$$

Estimated hazard ratio or risk score

$$\widehat{hazardratio}_{A_i}(t) = \frac{\widehat{h}_{A_i}(t)}{\widehat{h}_0(t)} = \frac{\widehat{\lambda}_{A_i} \cdot \widehat{\alpha} \cdot t^{\widehat{\alpha}-1}}{\widehat{\lambda}_0 \cdot \widehat{\alpha} \cdot t^{\widehat{\alpha}-1}}$$

Estimated percentiles of lifetime distributions

$$\text{baseline } \widehat{t_p} = \left(\frac{1}{\widehat{\lambda}_0} \cdot \ln \frac{100}{(100-p)}\right)^{\frac{1}{\widehat{\alpha}}}$$

$$\begin{aligned} \text{pred.level} \widehat{t_p} &= \left(\frac{1}{\widehat{\lambda}_{A_i}} \cdot \ln \frac{100}{(100-p)}\right)^{\frac{1}{\widehat{\alpha}}} \\ &= \text{baseline } \ \widehat{t_p} \cdot \ (\text{risk score})^{-\frac{1}{\widehat{\alpha}}} \end{aligned}$$

Shape parameter alters

Estimated hazard function

$$\widehat{h}_0(t) = \widehat{\lambda}_0 \cdot \widehat{\alpha}_0 \cdot t^{\widehat{\alpha}_0-1}$$

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

Estimated survivor function

$$\widehat{S}_0(t) = \exp(-\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}_0})$$

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

Estimated odds of a lapse

$$\widehat{odds}_0(t) = \frac{1-\widehat{S}_0(t)}{\widehat{S}_0(t)} = \exp(\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}_0-1})$$

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

Estimated odds ratio or index

$$\widehat{oddsratio}_{A_i}(t) = \frac{\widehat{odds}_{A_i}(t)}{\widehat{odds}_0(t)} = \frac{\exp(\widehat{\lambda}_{A_i} \cdot t^{\widehat{\alpha}_{A_i}-1})}{\exp(\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}_0-1})}$$

Estimated hazard ratio or risk score

$$\widehat{hazardratio}_{A_i}(t) = \frac{\widehat{h}_{A_i}(t)}{\widehat{h}_0(t)} = \frac{\widehat{\lambda}_{A_i} \cdot \widehat{\alpha}_{A_i} \cdot t^{\widehat{\alpha}_{A_i}-1}}{\widehat{\lambda}_0 \cdot \widehat{\alpha}_0 \cdot t^{\widehat{\alpha}_0-1}}$$

Estimated percentiles of lifetime distributions

$$\text{baseline } \widehat{t_p} = \left(\frac{\;}{\widehat{\lambda}_0} \cdot \ln \frac{100}{(100-p)}\right)^{\frac{1}{\widehat{\alpha}_0}}$$

$$\text{pred. level } \widehat{t_p} = \left(\frac{1}{\widehat{\lambda}_{A_i}} \cdot \ln \frac{100}{(100-p)}\right)^{\frac{1}{\widehat{\alpha}_{A_i}}}$$

$i = 1, 2, 3$

The estimated hazard ratios are called **risk scores**. The risk score of a risk factor level shows the effect of this level on the baseline hazard rate of a lapse. This effect is multiplicative on the baseline hazard rate of a lapse and increases the baseline hazard rate of a lapse (if the risk score > 1) or decreases the baseline hazard rate of a lapse (if the risk score < 1).

The estimated odds ratios are called **indices**. The index of a risk factor level shows the effect of this level on the baseline odds of a lapse. This effect is multiplicative on the baseline odds of a lapse and increases the baseline odds of a lapse (if the index > 1) or decreases the baseline odds of a lapse (if the index < 1).

Consider the risk factor A on three levels A_1, A_2 and A_3. Recall that the proportional hazards model (PHM) models $h(t|\boldsymbol{Z})$, the **hazard rate of a lapse** at time t for a policy with risk vector $\boldsymbol{Z} = (Z_1, Z_2, ..., Z_p)'$ and regression coefficients $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_p)'$

$$\Rightarrow \boxed{h_{A_l}(t|\boldsymbol{Z}) = e^{\boldsymbol{\beta}'\boldsymbol{Z}} \cdot h_0(t) \quad l = 1, 2, 3}$$

This property of constant hazard ratios over time only holds when the shape parameter of the Weibull distributions of the baseline and risk factor levels remains constant.

Two dummy variables Z_{A_1} and Z_{A_2} are defined for levels A_1 and A_2 in such a way that the regression coefficient β_{A_3} of level A_3 is equal to $-(\beta_{A_1} + \beta_{A_2})$. From the PHM follows that

$$h_{A_1}(t|Z_{A_1} = 1, Z_{A_2} = 0) = e^{(\beta_{A_1}.1 + \beta_{A_2}.0)} \cdot h_0(t) \Rightarrow \frac{h_{A_1}(t|Z_{A_1} = 1, Z_{A_2} = 0)}{h_0(t)} = e^{\beta_{A_1}}$$

$$\Rightarrow \boxed{\widehat{hazardratio}_{A_1} = e^{\widehat{\beta}_{A_1}} = \frac{\widehat{\lambda}_{A_i} \cdot \widehat{\alpha} \cdot t^{\widehat{\alpha}-1}}{\widehat{\lambda}_0 \cdot \widehat{\alpha} \cdot t^{\widehat{\alpha}-1}} = \frac{\widehat{\lambda}_{A_1}}{\widehat{\lambda}_0}}$$

This means that for a **constant shape parameter** in the Weibull distributions, the risk scores (estimated hazard ratios) may be also obtained from the exponent of the estimated β-values in the Weibull regression model.

4.7.9 The fitting of a regression model with a continuous risk factor

Consider a continuous risk factor that can be categorized into three groups. Define the ordinal covariate Z that takes on the values z=1 for the first group, z=2 for the second group and z=3 for the third group. Denote a vector of two's by $\mathbf{2}$ and a vector of three's by $\mathbf{3}$.

The log-logistic regression model with a continuous risk factor

The log-logistic regression model that models $\ln(odds)$ is then

$$\ln\left(\frac{\mathbf{1}-S(\boldsymbol{x})}{S(\boldsymbol{x})}\right)=\ln\lambda\cdot\mathbf{1}+\beta\cdot\boldsymbol{z}+\alpha\cdot\ln\boldsymbol{x} \qquad (4.7\ .31)$$

or

$$\begin{aligned}
\ln\left(\frac{\mathbf{1}-S(\boldsymbol{x})}{S(\boldsymbol{x})}\right) &= \ln\left(\frac{F(\boldsymbol{x})}{\mathbf{1}-F(\boldsymbol{x})}\right)=\ln\left(\frac{\boldsymbol{\pi}_S}{\mathbf{1}-\boldsymbol{\pi}_S}\right)=\ln\left(\boldsymbol{\pi}_S\right)-\ln\left(\mathbf{1}-\boldsymbol{\pi}_S\right)\\
&= \ln\lambda\cdot\begin{pmatrix}\mathbf{1}\\\mathbf{1}\\\mathbf{1}\\\mathbf{1}\\ \\\mathbf{1}\\\mathbf{1}\\\mathbf{1}\\\mathbf{1}\\ \\\mathbf{1}\\\mathbf{1}\\\mathbf{1}\\\mathbf{1}\end{pmatrix}+\alpha\cdot\begin{pmatrix}\ln\boldsymbol{x}_1\\\ln\boldsymbol{x}_2\\\ln\boldsymbol{x}_3\\\ln\boldsymbol{x}_4\\ \\\ln\boldsymbol{x}_1\\\ln\boldsymbol{x}_2\\\ln\boldsymbol{x}_3\\\ln\boldsymbol{x}_4\\ \\\ln\boldsymbol{x}_1\\\ln\boldsymbol{x}_2\\\ln\boldsymbol{x}_3\\\ln\boldsymbol{x}_4\end{pmatrix}+\beta\cdot\begin{pmatrix}\mathbf{1}\\\mathbf{1}\\\mathbf{1}\\\mathbf{1}\\ \\\mathbf{2}\\\mathbf{2}\\\mathbf{2}\\\mathbf{2}\\ \\\mathbf{3}\\\mathbf{3}\\\mathbf{3}\\\mathbf{3}\end{pmatrix}\\
&= \underbrace{\begin{pmatrix}\mathbf{1}&\ln\boldsymbol{x}_1&\mathbf{1}\\\mathbf{1}&\ln\boldsymbol{x}_2&\mathbf{1}\\\mathbf{1}&\ln\boldsymbol{x}_3&\mathbf{1}\\\mathbf{1}&\ln\boldsymbol{x}_4&\mathbf{1}\\ &&\\\mathbf{1}&\ln\boldsymbol{x}_1&\mathbf{2}\\\mathbf{1}&\ln\boldsymbol{x}_2&\mathbf{2}\\\mathbf{1}&\ln\boldsymbol{x}_3&\mathbf{2}\\\mathbf{1}&\ln\boldsymbol{x}_4&\mathbf{2}\\ &&\\\mathbf{1}&\ln\boldsymbol{x}_1&\mathbf{3}\\\mathbf{1}&\ln\boldsymbol{x}_2&\mathbf{3}\\\mathbf{1}&\ln\boldsymbol{x}_3&\mathbf{3}\\\mathbf{1}&\ln\boldsymbol{x}_4&\mathbf{3}\end{pmatrix}}\cdot\begin{pmatrix}\ln\lambda\\\alpha\\\beta\end{pmatrix}
\end{aligned}$$

$$\Rightarrow \ln\left(\frac{\pi_S}{\mathbf{1}-\pi_S}\right) = \boldsymbol{X}_1 \cdot \begin{pmatrix} \ln\lambda \\ \alpha \\ \beta \end{pmatrix} \qquad (4.7\ .32)$$

Equation 4.7 .32 is a linear model in the parameters $\ln\lambda$, α and β.

By proceeding in a similar way as in section 4.7.3, the estimated parameters of the log-logistic distributions for the three risk factor groups can be solved from

$$\begin{pmatrix} \widehat{\ln\lambda} \\ \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix} = (\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}' \cdot \ln\left\{\frac{\boldsymbol{S}\cdot\widehat{\boldsymbol{\pi}}_c}{\mathbf{1}-\boldsymbol{S}\cdot\widehat{\boldsymbol{\pi}}_c}\right\} \qquad (4.7\ .33)$$

with S and $\widehat{\boldsymbol{\pi}}_c$ defined in section 4.7.3 .

The estimated lambda parameters of these three log-logistic distributions then are

$$\begin{aligned} \widehat{\lambda}_{Z=1} &= \exp(\widehat{\ln\lambda} + \widehat{\beta} * 1) \\ \widehat{\lambda}_{Z=2} &= \exp(\widehat{\ln\lambda} + \widehat{\beta} * 2) \\ \widehat{\lambda}_{Z=3} &= \exp(\widehat{\ln\lambda} + \widehat{\beta} * 3). \end{aligned}$$

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

$$\widehat{\lambda}_0 = \exp(\widehat{\ln\lambda})$$

so that

$$\begin{aligned} \widehat{\lambda}_{Z=1} &= \widehat{\lambda}_0 \times \exp(\widehat{\beta} * 1) \\ \widehat{\lambda}_{Z=2} &= \widehat{\lambda}_0 \times \exp(\widehat{\beta} * 2) \\ \widehat{\lambda}_{Z=3} &= \widehat{\lambda}_0 \times \exp(\widehat{\beta} * 3). \end{aligned}$$

These four log-logistic distributions all have the same estimated alpha parameter $\widehat{\alpha}$.

The above procedure can also be applied when other continuous values of z, instead of the values 1,2,3, are used, for example the midpoints of the risk factor groupings. An application will be discussed in chapter 5.

The SAS/IML program to fit a log-logistic regression model with one continuous risk factor (constant shape parameter) to grouped survival data with staggered entry of policies appears in Appendix A.

The Weibull regression model with a continuous risk factor

The Weibull regression model is

$$\ln(-\ln S(\boldsymbol{x})) = \ln\lambda \cdot \mathbf{1} + \beta \cdot \boldsymbol{z} + \alpha \cdot \ln \boldsymbol{x} \qquad (4.7\ .34)$$

or

$$\begin{aligned}
\ln(-\ln S(\boldsymbol{x})) &= \ln(-\ln(\mathbf{1} - F(\boldsymbol{x}))) = \ln(-\ln(\mathbf{1} - \boldsymbol{\pi}_S)) \\
&= \ln\lambda \cdot \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \end{pmatrix} + \alpha \cdot \begin{pmatrix} \ln \boldsymbol{x}_1 \\ \ln \boldsymbol{x}_2 \\ \ln \boldsymbol{x}_3 \\ \ln \boldsymbol{x}_4 \\ \\ \ln \boldsymbol{x}_1 \\ \ln \boldsymbol{x}_2 \\ \ln \boldsymbol{x}_3 \\ \ln \boldsymbol{x}_4 \\ \\ \ln \boldsymbol{x}_1 \\ \ln \boldsymbol{x}_2 \\ \ln \boldsymbol{x}_3 \\ \ln \boldsymbol{x}_4 \end{pmatrix} + \beta \cdot \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \\ \mathbf{2} \\ \mathbf{2} \\ \mathbf{2} \\ \mathbf{2} \\ \\ \mathbf{3} \\ \mathbf{3} \\ \mathbf{3} \\ \mathbf{3} \end{pmatrix} \\
&= \underbrace{\begin{pmatrix} \mathbf{1} & \ln \boldsymbol{x}_1 & \mathbf{1} \\ \mathbf{1} & \ln \boldsymbol{x}_2 & \mathbf{1} \\ \mathbf{1} & \ln \boldsymbol{x}_3 & \mathbf{1} \\ \mathbf{1} & \ln \boldsymbol{x}_4 & \mathbf{1} \\ \\ \mathbf{1} & \ln \boldsymbol{x}_1 & \mathbf{2} \\ \mathbf{1} & \ln \boldsymbol{x}_2 & \mathbf{2} \\ \mathbf{1} & \ln \boldsymbol{x}_3 & \mathbf{2} \\ \mathbf{1} & \ln \boldsymbol{x}_4 & \mathbf{2} \\ \\ \mathbf{1} & \ln \boldsymbol{x}_1 & \mathbf{3} \\ \mathbf{1} & \ln \boldsymbol{x}_2 & \mathbf{3} \\ \mathbf{1} & \ln \boldsymbol{x}_3 & \mathbf{3} \\ \mathbf{1} & \ln \boldsymbol{x}_4 & \mathbf{3} \end{pmatrix}} \cdot \begin{pmatrix} \ln\lambda \\ \alpha \\ \beta \end{pmatrix}
\end{aligned}$$

$$\Rightarrow \ln(-\ln(\mathbf{1}-\boldsymbol{\pi}_S)) = \boldsymbol{X}_1 \cdot \begin{pmatrix} \ln\lambda \\ \alpha \\ \beta \end{pmatrix} \qquad (4.7\ .35)$$

Equation 4.7 .35 is a linear model in the parameters $\ln\lambda$, α and β.

By proceeding in a similar way as in section 4.7.6, the estimated parameters of the Weibull distributions for the three risk factor groups can be solved from

$$\begin{pmatrix} \widehat{\ln\lambda} \\ \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix} = (\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1' \cdot \ln\left\{\frac{\boldsymbol{S}\cdot\widehat{\boldsymbol{\pi}}_c}{\mathbf{1}-\boldsymbol{S}\cdot\widehat{\boldsymbol{\pi}}_c}\right\} \qquad (4.7\ .36)$$

with S and $\widehat{\boldsymbol{\pi}}_c$ defined in section 4.7.6.

The estimated lambda parameters of these three Weibull distributions then are

$$\begin{aligned} \widehat{\lambda}_{Z=1} &= \exp(\widehat{\ln\lambda} + \widehat{\beta} * 1) \\ \widehat{\lambda}_{Z=2} &= \exp(\widehat{\ln\lambda} + \widehat{\beta} * 2) \\ \widehat{\lambda}_{Z=3} &= \exp(\widehat{\ln\lambda} + \widehat{\beta} * 3). \end{aligned}$$

The estimated lambda parameter of the baseline Weibull distribution then is

$$\widehat{\lambda}_0 = \exp(\widehat{\ln\lambda})$$

so that

$$\begin{aligned} \widehat{\lambda}_{Z=1} &= \widehat{\lambda}_0 \times \exp(\widehat{\beta} * 1) \\ \widehat{\lambda}_{Z=2} &= \widehat{\lambda}_0 \times \exp(\widehat{\beta} * 2) \\ \widehat{\lambda}_{Z=3} &= \widehat{\lambda}_0 \times \exp(\widehat{\beta} * 3). \end{aligned}$$

These four Weibull distributions all have the same estimated alpha parameter $\widehat{\alpha}$.

The above procedure can also be applied when other continuous values of z, instead of the values 1,2,3, are used, for example the midpoints of the risk factor groupings. An application will be discussed in chapter 5.

The SAS/IML program to fit a Weibull regression model with one continuous risk factor (constant shape parameter) to grouped survival data with staggered entry of policies appears in Appendix A.

4.7.10 Notation for a regression model with two risk factors at staggered entry

Consider two risk factors A and B, each on three levels A_1, A_2 and A_3 and B_1, B_2 and B_3 respectively. Staggered entry of policies occur during four different entry periods and k, the number of class intervals for the first entry group, is equal to seven.

The combined relative frequency vector is $\boldsymbol{p}' = (\boldsymbol{p}'_{111}, \boldsymbol{p}'_{211}, \boldsymbol{p}'_{311}, \boldsymbol{p}'_{411}, \boldsymbol{p}'_{112}, \boldsymbol{p}'_{212}, \boldsymbol{p}'_{312}, \boldsymbol{p}'_{412}, \boldsymbol{p}'_{113}, \boldsymbol{p}'_{213}, \boldsymbol{p}'_{313}, \boldsymbol{p}'_{413}, \boldsymbol{p}'_{121}, \boldsymbol{p}'_{221}, \boldsymbol{p}'_{321}, \boldsymbol{p}'_{421}, \boldsymbol{p}'_{122}, \boldsymbol{p}'_{222}, \boldsymbol{p}'_{322}, \boldsymbol{p}'_{422}, \boldsymbol{p}'_{123}, \boldsymbol{p}'_{223}, \boldsymbol{p}'_{323}, \boldsymbol{p}'_{423}, \boldsymbol{p}'_{131}, \boldsymbol{p}'_{231}, \boldsymbol{p}'_{331}, \boldsymbol{p}'_{431}, \boldsymbol{p}'_{132}, \boldsymbol{p}'_{232}, \boldsymbol{p}'_{332}, \boldsymbol{p}'_{432}, \boldsymbol{p}'_{133}, \boldsymbol{p}'_{233}, \boldsymbol{p}'_{333}, \boldsymbol{p}'_{433})$.
$\boldsymbol{p}_{ilm}$ is the relative frequency vector for the i^{th} entry group, the l^{th} risk factor A level and the m^{th} risk factor B level corresponding to $n_{ilm}\boldsymbol{p}_{ilm}$ being multinomial $(n_{ilm}; \boldsymbol{\pi}_{ilm})$ distributed $i = 1, 2, 3, 4$ and $l = 1, 2, 3$ and $m = 1, 2, 3$.

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

and

$\boldsymbol{\pi}_{1lm} = (\pi_{1lm,1}, \pi_{1lm,2}, \pi_{1lm,3}, \pi_{1lm,4}, \pi_{1lm,5}, \pi_{1lm,6}, \pi_{1ml,7})'$ is a 7×1 probability vector
$\boldsymbol{\pi}_{2lm} = (\pi_{2lm,1}, \pi_{2lm,2}, \pi_{2lm,3}, \pi_{2lm,4}, \pi_{2lm,5}, \pi_{2lm,6})'$ is a 6×1 probability vector
$\boldsymbol{\pi}_{3lm} = (\pi_{3lm,1}, \pi_{3lm,2}, \pi_{3lm,3}, \pi_{3lm,4}, \pi_{3lm,5})'$ is a 5×1 probability vector
$\boldsymbol{\pi}_{4lm} = (\pi_{4lm,1}, \pi_{4lm,2}, \pi_{4lm,3}, \pi_{4lm,4})'$ is a 4×1 probability vector $\quad l =$ 1,2,3.

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

$$\boldsymbol{x}_1 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} \quad \boldsymbol{x}_2 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} \quad \boldsymbol{x}_3 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \quad \text{and} \quad \boldsymbol{x}_4 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

The number of entries per cell in the cross tabulation of entry period and risk factors A and B can be summarized in table 4.2.

Table 4.2: **Number of entries per cell in cross table of entry period and risk factors** A **and** B

Entry Period	Risk Factor Level									Total
	A_1			A_2			A_3			
	B_1	B_2	B_3	B_1	B_2	B_3	B_1	B_2	B_3	
1	7	7	7	7	7	7	7	7	7	63
2	6	6	6	6	6	6	6	6	6	54
3	5	5	5	5	5	5	5	5	5	45
4	4	4	4	4	4	4	4	4	4	12
Total	22	22	22	22	22	22	22	22	22	198

Matrix $\boldsymbol{S}$ is now a $(54 \times 3) \times (66 \times 3) = 162 \times 198$ matrix and is defined as

$$\boldsymbol{S} = \begin{pmatrix} \boldsymbol{S}_{11} \\ \boldsymbol{S}_{12} \\ \boldsymbol{S}_{13} \\ \boldsymbol{S}_{21} \\ \boldsymbol{S}_{22} \\ \boldsymbol{S}_{23} \\ \boldsymbol{S}_{31} \\ \boldsymbol{S}_{32} \\ \boldsymbol{S}_{33} \end{pmatrix}$$

where $\boldsymbol{S}_{lm} = \text{block}(\boldsymbol{S}_1, \boldsymbol{S}_2, \boldsymbol{S}_3, \boldsymbol{S}_4)$ is a 18×22 matrix of the form

$$\boldsymbol{S}_{lm} = \left(\begin{array}{ccccccc|cccccc|ccccc|cccc}
1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
1&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
1&1&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
1&1&1&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
1&1&1&1&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&1&1&1&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&1&1&1&1&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&1&1&1&1&1&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&1&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&0
\end{array}\right).$$

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

level and risk factor B level.

$$\Longrightarrow \quad \widehat{\boldsymbol{V}} = \text{block}(\widehat{\boldsymbol{V}}_{11}, \widehat{\boldsymbol{V}}_{12}, \widehat{\boldsymbol{V}}_{13}, \widehat{\boldsymbol{V}}_{21}, \widehat{\boldsymbol{V}}_{22}, \widehat{\boldsymbol{V}}_{23}, \widehat{\boldsymbol{V}}_{31}, \widehat{\boldsymbol{V}}_{32}, \widehat{\boldsymbol{V}}_{33})$$

where

$$\widehat{\boldsymbol{V}}_{lm} = \text{block}(\widehat{\boldsymbol{V}}_{1,lm}, \widehat{\boldsymbol{V}}_{2,lm}, \widehat{\boldsymbol{V}}_{3,lm}, \widehat{\boldsymbol{V}}_{4,lm}) \quad l = 1,2,3 \quad \text{and} \quad m = 1,2,3$$

and

$$\widehat{\boldsymbol{V}}_{i,lm} = \frac{1}{n_{ilm}}\left[diag(\boldsymbol{p}_{ilm}) - \boldsymbol{p}_{ilm}\boldsymbol{p}'_{ilm}\right] \quad i = 1,2,3,4 \quad \text{and} \quad l = 1,2,3 \quad \text{and} \quad m = 1,2,3.$$

Note that $\widehat{\boldsymbol{V}}_{1,lm}$ is a 21×21, $\widehat{\boldsymbol{V}}_{2,lm}$ is a 18×18, $\widehat{\boldsymbol{V}}_{3,lm}$ is a 15×15 and $\widehat{\boldsymbol{V}}_{4,lm}$ is a 12×12 matrix so that $\widehat{\boldsymbol{V}}_{lm}$ is a 66×66 matrix and $\widehat{\boldsymbol{V}}$ is a 198×198 matrix.

4.7.11 The log-logistic regression model with two risk factors at staggered entry

In this model the effect of the risk factors is to alter the scale parameter λ, while the shape parameter α remains constant.

Consider a risk factor A on three levels A_1, A_2 and A_3 and another risk factor B on three levels B_1, B_2 and B_3. Two dummy variables Z_{A_1} and Z_{A_2} are defined for levels A_1 and A_2 in such a way that the regression coefficient β_{A_3} of level A_3 is equal to $-(\beta_{A_1} + \beta_{A_2})$, that means $\left\{\sum_{k=1}^{3} \beta_{A_k}\right\} = 0$. Similarly two dummy variables Z_{B_1} and Z_{B_2} are defined for levels B_1 and B_2 in such a way that the regression coefficient β_{B_3} of level B_3 is equal to $-(\beta_{B_1} + \beta_{B_2})$, that means $\left\{\sum_{k=1}^{3} \beta_{B_k}\right\} = 0$.

The log-logistic regression model that models $\ln(odds)$ then is

$$\ln\left(\frac{1 - S(\boldsymbol{x})}{S(\boldsymbol{x})}\right) = \ln\lambda \cdot \mathbf{1} + (\beta_{A_1} Z_{A_1} + \beta_{A_2} Z_{A_2}) + (\beta_{B_1} Z_{B_1} + \beta_{B_2} Z_{B_2}) + \alpha \cdot \ln \boldsymbol{x}$$

or

$$\ln\left(\frac{\mathbf{1} - S(\boldsymbol{x})}{S(\boldsymbol{x})}\right) = \ln\left(\frac{F(\boldsymbol{x})}{\mathbf{1} - F(\boldsymbol{x})}\right) = \ln\left(\frac{\boldsymbol{\pi}_S}{\mathbf{1} - \boldsymbol{\pi}_S}\right) = \ln(\boldsymbol{\pi}_S) - \ln(\mathbf{1} - \boldsymbol{\pi}_S)$$

$$
= \ln\lambda \cdot \begin{pmatrix} 1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\ \\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\ \\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1 \end{pmatrix}
+ \beta_{A_1} \cdot \begin{pmatrix} 1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\ \\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\ \\-1\\-1\\-1\\-1\\-1\\-1\\-1\\-1\\-1\\-1\\-1\\-1 \end{pmatrix}
+ \beta_{A_2} \cdot \begin{pmatrix} 0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\ \\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\ \\-1\\-1\\-1\\-1\\-1\\-1\\-1\\-1\\-1\\-1\\-1\\-1 \end{pmatrix}
+ \beta_{B_1} \cdot \begin{pmatrix} 1\\1\\1\\1\\0\\0\\0\\0\\-1\\-1\\-1\\-1\\ \\1\\1\\1\\1\\0\\0\\0\\0\\-1\\-1\\-1\\-1\\ \\1\\1\\1\\1\\0\\0\\0\\0\\-1\\-1\\-1\\-1 \end{pmatrix}
+ \beta_{B_2} \cdot \begin{pmatrix} 0\\0\\0\\0\\1\\1\\1\\1\\-1\\-1\\-1\\-1\\ \\0\\0\\0\\0\\1\\1\\1\\1\\-1\\-1\\-1\\-1\\ \\0\\0\\0\\0\\1\\1\\1\\1\\-1\\-1\\-1\\-1 \end{pmatrix}
+ \alpha \cdot \begin{pmatrix} \ln\boldsymbol{x}_1\\\ln\boldsymbol{x}_2\\\ln\boldsymbol{x}_3\\\ln\boldsymbol{x}_4\\\ln\boldsymbol{x}_1\\\ln\boldsymbol{x}_2\\\ln\boldsymbol{x}_3\\\ln\boldsymbol{x}_4\\\ln\boldsymbol{x}_1\\\ln\boldsymbol{x}_2\\\ln\boldsymbol{x}_3\\\ln\boldsymbol{x}_4\\ \\\ln\boldsymbol{x}_1\\\ln\boldsymbol{x}_2\\\ln\boldsymbol{x}_3\\\ln\boldsymbol{x}_4\\\ln\boldsymbol{x}_1\\\ln\boldsymbol{x}_2\\\ln\boldsymbol{x}_3\\\ln\boldsymbol{x}_4\\\ln\boldsymbol{x}_1\\\ln\boldsymbol{x}_2\\\ln\boldsymbol{x}_3\\\ln\boldsymbol{x}_4\\ \\\ln\boldsymbol{x}_1\\\ln\boldsymbol{x}_2\\\ln\boldsymbol{x}_3\\\ln\boldsymbol{x}_4\\\ln\boldsymbol{x}_1\\\ln\boldsymbol{x}_2\\\ln\boldsymbol{x}_3\\\ln\boldsymbol{x}_4\\\ln\boldsymbol{x}_1\\\ln\boldsymbol{x}_2\\\ln\boldsymbol{x}_3\\\ln\boldsymbol{x}_4 \end{pmatrix}
$$

$$\Rightarrow \ln\left(\frac{\pi_S}{1-\pi_S}\right) = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & \ln x_1 \\ 1 & 1 & 0 & 1 & 0 & \ln x_2 \\ 1 & 1 & 0 & 1 & 0 & \ln x_3 \\ 1 & 1 & 0 & 1 & 0 & \ln x_4 \\ 1 & 1 & 0 & 0 & 1 & \ln x_1 \\ 1 & 1 & 0 & 0 & 1 & \ln x_2 \\ 1 & 1 & 0 & 0 & 1 & \ln x_3 \\ 1 & 1 & 0 & 0 & 1 & \ln x_4 \\ 1 & 1 & 0 & -1 & -1 & \ln x_1 \\ 1 & 1 & 0 & -1 & -1 & \ln x_2 \\ 1 & 1 & 0 & -1 & -1 & \ln x_3 \\ 1 & 1 & 0 & -1 & -1 & \ln x_4 \\ \\ 1 & 0 & 1 & 1 & 0 & \ln x_1 \\ 1 & 0 & 1 & 1 & 0 & \ln x_2 \\ 1 & 0 & 1 & 1 & 0 & \ln x_3 \\ 1 & 0 & 1 & 1 & 0 & \ln x_4 \\ 1 & 0 & 1 & 0 & 1 & \ln x_1 \\ 1 & 0 & 1 & 0 & 1 & \ln x_2 \\ 1 & 0 & 1 & 0 & 1 & \ln x_3 \\ 1 & 0 & 1 & 0 & 1 & \ln x_4 \\ 1 & 0 & 1 & -1 & -1 & \ln x_1 \\ 1 & 0 & 1 & -1 & -1 & \ln x_2 \\ 1 & 0 & 1 & -1 & -1 & \ln x_3 \\ 1 & 0 & 1 & -1 & -1 & \ln x_4 \\ \\ 1 & -1 & -1 & 1 & 0 & \ln x_1 \\ 1 & -1 & -1 & 1 & 0 & \ln x_2 \\ 1 & -1 & -1 & 1 & 0 & \ln x_3 \\ 1 & -1 & -1 & 1 & 0 & \ln x_4 \\ 1 & -1 & -1 & 0 & 1 & \ln x_1 \\ 1 & -1 & -1 & 0 & 1 & \ln x_2 \\ 1 & -1 & -1 & 0 & 1 & \ln x_3 \\ 1 & -1 & -1 & 0 & 1 & \ln x_4 \\ 1 & -1 & -1 & -1 & -1 & \ln x_1 \\ 1 & -1 & -1 & -1 & -1 & \ln x_2 \\ 1 & -1 & -1 & -1 & -1 & \ln x_3 \\ 1 & -1 & -1 & -1 & -1 & \ln x_4 \end{pmatrix} \cdot \begin{pmatrix} \ln\lambda \\ \beta_{A_1} \\ \beta_{A_2} \\ \beta_{B_1} \\ \beta_{B_2} \\ \alpha \end{pmatrix}$$

$$\Rightarrow \ln\left(\frac{\pi_S}{1-\pi_S}\right) = \boldsymbol{X}_1 \cdot \begin{pmatrix} \ln\lambda \\ \beta_{A_1} \\ \beta_{A_2} \\ \beta_{B_1} \\ \beta_{B_2} \\ \alpha \end{pmatrix}$$

This model is a linear model in the parameters $\ln\lambda$, β_{A_1}, β_{A_2}, β_{B_1}, β_{B_2} and α.

By proceeding in a similar way as in section 4.7.3, the estimated parameters of the log-logistic distributions for the nine combinations of risk factor A levels and risk factor B levels can be solved from

$$\begin{pmatrix} \widehat{\ln\lambda} \\ \widehat{\beta}_{A_1} \\ \widehat{\beta}_{A_2} \\ \widehat{\beta}_{B_1} \\ \widehat{\beta}_{B_2} \\ \widehat{\alpha} \end{pmatrix} = (\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}' \cdot \ln\left\{\frac{\boldsymbol{S}\cdot\widehat{\boldsymbol{\pi}}_c}{\boldsymbol{1}-\boldsymbol{S}\cdot\widehat{\boldsymbol{\pi}}_c}\right\} \qquad (4.7\ .37)$$

with $\widehat{\boldsymbol{\pi}}_c = \boldsymbol{p} - (\boldsymbol{G}_\pi\boldsymbol{V})'(\boldsymbol{G}_p\boldsymbol{V}\boldsymbol{G}_\pi')^* \cdot \boldsymbol{C}.\ln\left\{\frac{\boldsymbol{S}\cdot\boldsymbol{p}}{\boldsymbol{1}-\boldsymbol{S}\cdot\boldsymbol{p}}\right\}$ where $\boldsymbol{S}$, $\boldsymbol{V}$ and $\boldsymbol{p}$ are defined in section 4.7.10 .

Note that $\widehat{\beta}_{A_3} = -\left(\widehat{\beta}_{A_1} + \widehat{\beta}_{A_2}\right)$ and $\widehat{\beta}_{B_3} = -\left(\widehat{\beta}_{B_1} + \widehat{\beta}_{B_2}\right)$.

The estimated lambda parameters of the nine log-logistic distributions for the nine combinations of risk factor A levels and risk factor B levels then are

$$\begin{aligned}
\widehat{\lambda}_{A_1B_1} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_1} + \widehat{\beta}_{B_1}) \\
\widehat{\lambda}_{A_1B_2} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_1} + \widehat{\beta}_{B_2}) \\
\widehat{\lambda}_{A_1B_3} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_1} + \widehat{\beta}_{B_3}) \\
\widehat{\lambda}_{A_2B_1} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_2} + \widehat{\beta}_{B_1}) \\
\widehat{\lambda}_{A_2B_2} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_2} + \widehat{\beta}_{B_2}) \\
\widehat{\lambda}_{A_2B_3} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_2} + \widehat{\beta}_{B_3}) \\
\widehat{\lambda}_{A_3B_1} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_3} + \widehat{\beta}_{B_1}) \\
\widehat{\lambda}_{A_3B_2} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_3} + \widehat{\beta}_{B_2}) \\
\widehat{\lambda}_{A_3B_3} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_3} + \widehat{\beta}_{B_3}).
\end{aligned}$$

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

$$\widehat{\lambda}_0 = \exp(\widehat{\ln\lambda})$$

so that

$$\begin{aligned}
\widehat{\lambda}_{A_1B_1} &= \widehat{\lambda}_0 \times index_{A_1} \times index_{B_1}) \\
\widehat{\lambda}_{A_1B_2} &= \widehat{\lambda}_0 \times index_{A_1} \times index_{B_2}) \\
\widehat{\lambda}_{A_1B_3} &= \widehat{\lambda}_0 \times index_{A_1} \times index_{B_3}) \\
\widehat{\lambda}_{A_2B_1} &= \widehat{\lambda}_0 \times index_{A_2} \times index_{B_1}) \\
\widehat{\lambda}_{A_2B_2} &= \widehat{\lambda}_0 \times index_{A_2} \times index_{B_2}) \\
\widehat{\lambda}_{A_2B_3} &= \widehat{\lambda}_0 \times index_{A_2} \times index_{B_3}) \\
\widehat{\lambda}_{A_3B_1} &= \widehat{\lambda}_0 \times index_{A_3} \times index_{B_1}) \\
\widehat{\lambda}_{A_3B_2} &= \widehat{\lambda}_0 \times index_{A_3} \times index_{B_2}) \\
\widehat{\lambda}_{A_3B_3} &= \widehat{\lambda}_0 \times index_{A_3} \times index_{B_3}).
\end{aligned}$$

These ten log-logistic distributions all have the same estimated alpha parameter $\widehat{\alpha}$.

The SAS/IML program to fit a log-logistic regression model with two categorical risk factors (constant shape parameter) to grouped survival data with staggered entry of policies appears in Appendix A.

4.7.12 The Weibull regression model with two risk factors at staggered entry

In this model the effect of the risk factors is to alter the scale parameter λ, while the shape parameter α remains constant.

Consider again a risk factor A on three levels A_1, A_2 and A_3 and another risk factor B on three levels B_1, B_2 and B_3. Two dummy variables Z_{A_1} and Z_{A_2} are defined for levels A_1 and A_2 in such a way that the regression coefficient β_{A_3} of level A_3 is equal to $-(\beta_{A_1} + \beta_{A_2})$, that means $\left\{\sum_{k=1}^{3} \beta_{A_k}\right\} = 0$. Similarly two dummy variables Z_{B_1} and Z_{B_2} are defined for levels B_1 and B_2 in such a way that the regression coefficient β_{B_3} of level B_3 is equal to $-(\beta_{B_1} + \beta_{B_2})$, that means $\left\{\sum_{k=1}^{3} \beta_{B_k}\right\} = 0$.

The Weibull regression model then is

$$\ln(-\ln S(\boldsymbol{x})) = \ln \lambda \cdot \mathbf{1} + (\beta_{A_1} Z_{A_1} + \beta_{A_2} Z_{A_2}) + (\beta_{B_1} Z_{B_1} + \beta_{B_2} Z_{B_2}) + \alpha \cdot \ln \boldsymbol{x}$$

or

$$\ln(-\ln S(\boldsymbol{x})) = \ln\left(\frac{F(\boldsymbol{x})}{\mathbf{1} - F(\boldsymbol{x})}\right) = \ln\left(\frac{\boldsymbol{\pi}_S}{\mathbf{1} - \boldsymbol{\pi}_S}\right) = \ln(\boldsymbol{\pi}_S) - \ln(\mathbf{1} - \boldsymbol{\pi}_S)$$

$$
= \ln\lambda \cdot \begin{pmatrix} 1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\ \\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\ \\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1 \end{pmatrix}
+ \beta_{A_1} \cdot \begin{pmatrix} 1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\ \\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\ \\-1\\-1\\-1\\-1\\-1\\-1\\-1\\-1\\-1\\-1\\-1\\-1 \end{pmatrix}
+ \beta_{A_2} \cdot \begin{pmatrix} 0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\ \\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\ \\-1\\-1\\-1\\-1\\-1\\-1\\-1\\-1\\-1\\-1\\-1\\-1 \end{pmatrix}
+ \beta_{B_1} \cdot \begin{pmatrix} 1\\1\\1\\1\\0\\0\\0\\0\\-1\\-1\\-1\\-1\\ \\1\\1\\1\\1\\0\\0\\0\\0\\-1\\-1\\-1\\-1\\ \\1\\1\\1\\1\\0\\0\\0\\0\\-1\\-1\\-1\\-1 \end{pmatrix}
+ \beta_{B_2} \cdot \begin{pmatrix} 0\\0\\0\\0\\1\\1\\1\\1\\-1\\-1\\-1\\-1\\ \\0\\0\\0\\0\\1\\1\\1\\1\\-1\\-1\\-1\\-1\\ \\0\\0\\0\\0\\1\\1\\1\\1\\-1\\-1\\-1\\-1 \end{pmatrix}
+ \alpha \cdot \begin{pmatrix} \ln x_1\\\ln x_2\\\ln x_3\\\ln x_4\\\ln x_1\\\ln x_2\\\ln x_3\\\ln x_4\\\ln x_1\\\ln x_2\\\ln x_3\\\ln x_4\\ \\\ln x_1\\\ln x_2\\\ln x_3\\\ln x_4\\\ln x_1\\\ln x_2\\\ln x_3\\\ln x_4\\\ln x_1\\\ln x_2\\\ln x_3\\\ln x_4\\ \\\ln x_1\\\ln x_2\\\ln x_3\\\ln x_4\\\ln x_1\\\ln x_2\\\ln x_3\\\ln x_4\\\ln x_1\\\ln x_2\\\ln x_3\\\ln x_4 \end{pmatrix}
$$

$$
\ln(-\ln S(\boldsymbol{x})) = \begin{pmatrix}
1 & 1 & 0 & 1 & 0 & \ln x_1 \\
1 & 1 & 0 & 1 & 0 & \ln x_2 \\
1 & 1 & 0 & 1 & 0 & \ln x_3 \\
1 & 1 & 0 & 1 & 0 & \ln x_4 \\
1 & 1 & 0 & 0 & 1 & \ln x_1 \\
1 & 1 & 0 & 0 & 1 & \ln x_2 \\
1 & 1 & 0 & 0 & 1 & \ln x_3 \\
1 & 1 & 0 & 0 & 1 & \ln x_4 \\
1 & 1 & 0 & -1 & -1 & \ln x_1 \\
1 & 1 & 0 & -1 & -1 & \ln x_2 \\
1 & 1 & 0 & -1 & -1 & \ln x_3 \\
1 & 1 & 0 & -1 & -1 & \ln x_4 \\
\\
1 & 0 & 1 & 1 & 0 & \ln x_1 \\
1 & 0 & 1 & 1 & 0 & \ln x_2 \\
1 & 0 & 1 & 1 & 0 & \ln x_3 \\
1 & 0 & 1 & 1 & 0 & \ln x_4 \\
1 & 0 & 1 & 0 & 1 & \ln x_1 \\
1 & 0 & 1 & 0 & 1 & \ln x_2 \\
1 & 0 & 1 & 0 & 1 & \ln x_3 \\
1 & 0 & 1 & 0 & 1 & \ln x_4 \\
1 & 0 & 1 & -1 & -1 & \ln x_1 \\
1 & 0 & 1 & -1 & -1 & \ln x_2 \\
1 & 0 & 1 & -1 & -1 & \ln x_3 \\
1 & 0 & 1 & -1 & -1 & \ln x_4 \\
\\
1 & -1 & -1 & 1 & 0 & \ln x_1 \\
1 & -1 & -1 & 1 & 0 & \ln x_2 \\
1 & -1 & -1 & 1 & 0 & \ln x_3 \\
1 & -1 & -1 & 1 & 0 & \ln x_4 \\
1 & -1 & -1 & 0 & 1 & \ln x_1 \\
1 & -1 & -1 & 0 & 1 & \ln x_2 \\
1 & -1 & -1 & 0 & 1 & \ln x_3 \\
1 & -1 & -1 & 0 & 1 & \ln x_4 \\
1 & -1 & -1 & -1 & -1 & \ln x_1 \\
1 & -1 & -1 & -1 & -1 & \ln x_2 \\
1 & -1 & -1 & -1 & -1 & \ln x_3 \\
1 & -1 & -1 & -1 & -1 & \ln x_4
\end{pmatrix} \cdot \begin{pmatrix} \ln\lambda \\ \beta_{A_1} \\ \beta_{A_2} \\ \beta_{B_1} \\ \beta_{B_2} \\ \alpha \end{pmatrix}
$$

$$
\Rightarrow \ln(-\ln S(\boldsymbol{x})) = \mathbf{X}_1 \cdot \begin{pmatrix} \ln\lambda \\ \beta_{A_1} \\ \beta_{A_2} \\ \beta_{B_1} \\ \beta_{B_2} \\ \alpha \end{pmatrix}
$$

This model is a linear model in the parameters $\ln\lambda$, β_{A_1}, β_{A_2}, β_{B_1}, β_{B_2} and α.

By proceeding in a similar way as in section 4.7.6, the estimated parameters of the Weibull distributions for the nine combinations of risk factor A levels and risk factor B levels can be solved from

$$\begin{pmatrix} \widehat{\ln\lambda} \\ \widehat{\beta}_{A_1} \\ \widehat{\beta}_{A_2} \\ \widehat{\beta}_{B_1} \\ \widehat{\beta}_{B_2} \\ \widehat{\alpha} \end{pmatrix} = (\boldsymbol{X}_1{}'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1{}' \cdot \ln\left\{\frac{\boldsymbol{S}\cdot\widehat{\boldsymbol{\pi}}_c}{\boldsymbol{1}-\boldsymbol{S}\cdot\widehat{\boldsymbol{\pi}}_c}\right\} \tag{4.7 .38}$$

with $\widehat{\boldsymbol{\pi}}_c = \boldsymbol{p} - (\boldsymbol{G}_\pi \boldsymbol{V})'(\boldsymbol{G}_p \boldsymbol{V}\boldsymbol{G}_\pi')^* \cdot \boldsymbol{C}.\ln\left\{\frac{\boldsymbol{S}\cdot\boldsymbol{p}}{\boldsymbol{1}-\boldsymbol{S}\cdot\boldsymbol{p}}\right\}$ where $\boldsymbol{S}$, $\boldsymbol{V}$ and $\boldsymbol{p}$ are defined in section 4.7.10 .

Note that $\widehat{\beta}_{A_3} = -\left(\widehat{\beta}_{A_1} + \widehat{\beta}_{A_2}\right)$ and $\widehat{\beta}_{B_3} = -\left(\widehat{\beta}_{B_1} + \widehat{\beta}_{B_2}\right)$.

The estimated lambda parameters of the nine Weibull distributions for the nine combinations of risk factor A levels and risk factor B levels then are

$$\begin{aligned}
\widehat{\lambda}_{A_1B_1} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_1} + \widehat{\beta}_{B_1}) \\
\widehat{\lambda}_{A_1B_2} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_1} + \widehat{\beta}_{B_2}) \\
\widehat{\lambda}_{A_1B_3} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_1} + \widehat{\beta}_{B_3}) \\
\widehat{\lambda}_{A_2B_1} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_2} + \widehat{\beta}_{B_1}) \\
\widehat{\lambda}_{A_2B_2} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_2} + \widehat{\beta}_{B_2}) \\
\widehat{\lambda}_{A_2B_3} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_2} + \widehat{\beta}_{B_3}) \\
\widehat{\lambda}_{A_3B_1} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_3} + \widehat{\beta}_{B_1}) \\
\widehat{\lambda}_{A_3B_2} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_3} + \widehat{\beta}_{B_2}) \\
\widehat{\lambda}_{A_3B_3} &= \exp(\widehat{\ln\lambda} + \widehat{\beta}_{A_3} + \widehat{\beta}_{B_3}).
\end{aligned}$$

The estimated lambda parameter of the baseline Weibull distribution then is

$$\widehat{\lambda}_0 = \exp(\widehat{\ln\lambda})$$

so that

$$\begin{aligned}
\widehat{\lambda}_{A_1B_1} &= \widehat{\lambda}_0 \times index_{A_1} \times index_{B_1}) \\
\widehat{\lambda}_{A_1B_2} &= \widehat{\lambda}_0 \times index_{A_1} \times index_{B_2}) \\
\widehat{\lambda}_{A_1B_3} &= \widehat{\lambda}_0 \times index_{A_1} \times index_{B_3}) \\
\widehat{\lambda}_{A_2B_1} &= \widehat{\lambda}_0 \times index_{A_2} \times index_{B_1}) \\
\widehat{\lambda}_{A_2B_2} &= \widehat{\lambda}_0 \times index_{A_2} \times index_{B_2}) \\
\widehat{\lambda}_{A_2B_3} &= \widehat{\lambda}_0 \times index_{A_2} \times index_{B_3}) \\
\widehat{\lambda}_{A_3B_1} &= \widehat{\lambda}_0 \times index_{A_3} \times index_{B_1}) \\
\widehat{\lambda}_{A_3B_2} &= \widehat{\lambda}_0 \times index_{A_3} \times index_{B_2}) \\
\widehat{\lambda}_{A_3B_3} &= \widehat{\lambda}_0 \times index_{A_3} \times index_{B_3}).
\end{aligned}$$

These ten Weibull distributions all have the same estimated alpha parameter $\widehat{\alpha}$.

The SAS/IML program to fit a Weibull regression model with two categorical risk factors (constant shape parameter) to grouped survival data with staggered entry of policies appears in Appendix A.

Generalization to a regression model with more than two risk factors is obvious.

Chapter 5

APPLICATION TO DATA FROM THE INSURANCE INDUSTRY

5.1 Description of the Data Set

5.1.1 Introduction

An extensive data set from an insurance company, containing information on policies written over the last few years, is available and permission has been given by this company to use this data set to illustrate the theoretical principles developed in the previous two chapters.

5.1.2 The raw data set of policies

A subgroup of policies is formed by selecting only mortgage protection policies written during four selected months, namely March 1998, June 1998, November 1998 and March 1999. This subgroup or smaller data set consists of the lifetimes of 10077 policies, together with some concomitant information on other variables such as age of the policyholder, credit turnover of his bankaccount and a score value, determined by the company.

Consider the following experimental design as illustrated in Figure 5.1. The 10077 policies enter the study at four 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 the pre-determined cut-off date of 15 April 2001, 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**.

From the 2586 policies with entry date March 1998 (inception dates between 1 March 1998 and 31 March 1998), a total of 1666 policies have lifetimes 37 months and more and thus were right-censored. From the 2809 policies with entry date June 1998 (inception dates between 1 June 1998 and 30 June 1998), a total of 1924 policies have lifetimes 34 months and more and were censored. From the 2286 policies with entry date November 1998 (inception dates between 1 November 1998 and 30 November 1998), a total of 1674 policies have lifetimes 28 months and more and were censored. From the 2396 policies with entry date March 1999 (inception dates between 1 March 1999 and 31 March 1999), a total of 1848 policies have lifetimes 24 months and more and were censored.

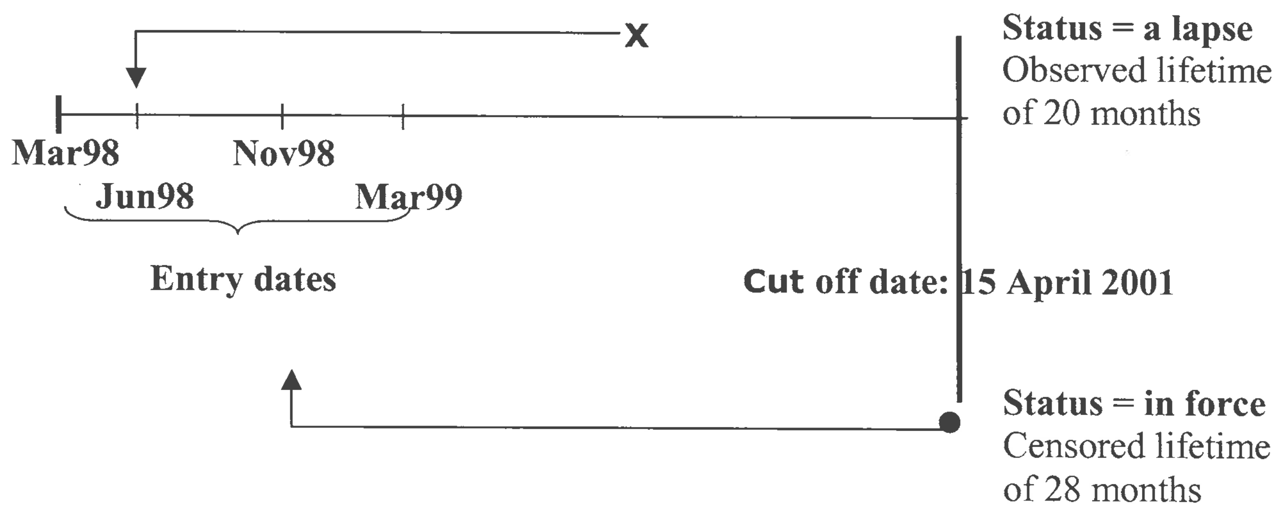

Figure 5.1: **Experimental design for illustrative data set**

5.1.3 The grouped data set of policies

The lifetimes of the policies that enter the study at March 1998 (called the first sample of size 2586) can be grouped into seven adjacent, non-overlapping fixed intervals

$$[0;12), [12;17), [17;24), [24;28), [28;34), [34;37) \quad \text{and} \quad [37;\infty).$$

The lifetimes of the policies that enter the study at June 1998 (called the second sample of size 2809) can be grouped into six adjacent, non-overlapping fixed intervals

$$[0;12), [12;17), [17;24), [24;28), [28;34) \quad \text{and} \quad [34;\infty).$$

The lifetimes of the policies that enter the study at November 1998 (called the third sample of size 2286) can be grouped into five adjacent, non-overlapping fixed intervals

$$[0;12), [12;17), [17;24), [24;28) \quad \text{and} \quad [28;\infty).$$

The lifetimes of the policies that enter the study at March 1999 (called the fourth sample of size 2396) can be grouped into four adjacent, non-overlapping fixed intervals

$$[0;12), [12;17), [17;24) \quad \text{and} \quad [24;\infty).$$

The four samples are assumed to be independent samples from multinomial populations. Four frequency distributions are formed when the observed and censored lifetimes of all the policies are grouped into the different class intervals and are shown in Table 5.1.

Table 5.1: **Frequency distributions of the four samples**

Interval	Class Intervals				Frequency Vector				Vector of Upper Bounds			
number	March 98	June 98	Nov 98	March 99	f_1	f_2	f_3	f_4	x_1	x_2	x_3	x_4
first	$[0,12)$	$[0,12)$	$[0,12)$	$[0,12)$	66	118	154	175	12	12	12	12
second	$[12,17)$	$[12,17)$	$[12,17)$	$[12,17)$	158	166	99	166	17	17	17	17
third	$[17,24)$	$[17,24)$	$[17,24)$	$[17,24)$	254	229	242	207	24	24	24	24
fourth	$[24,28)$	$[24,28)$	$[24,28)$	$[24,\infty)$	157	200	117	1848	28	28	28	
fifth	$[28,34)$	$[28,34)$	$[28,\infty)$		250	172	1674		34	34		
sixth	$[34,37)$	$[34,\infty)$			35	1924			37			
seventh	$[37,\infty)$				1666							
Total					2586	2809	2286	2396				

Figure 5.2 shows the histograms of the four relative frequency distributions.

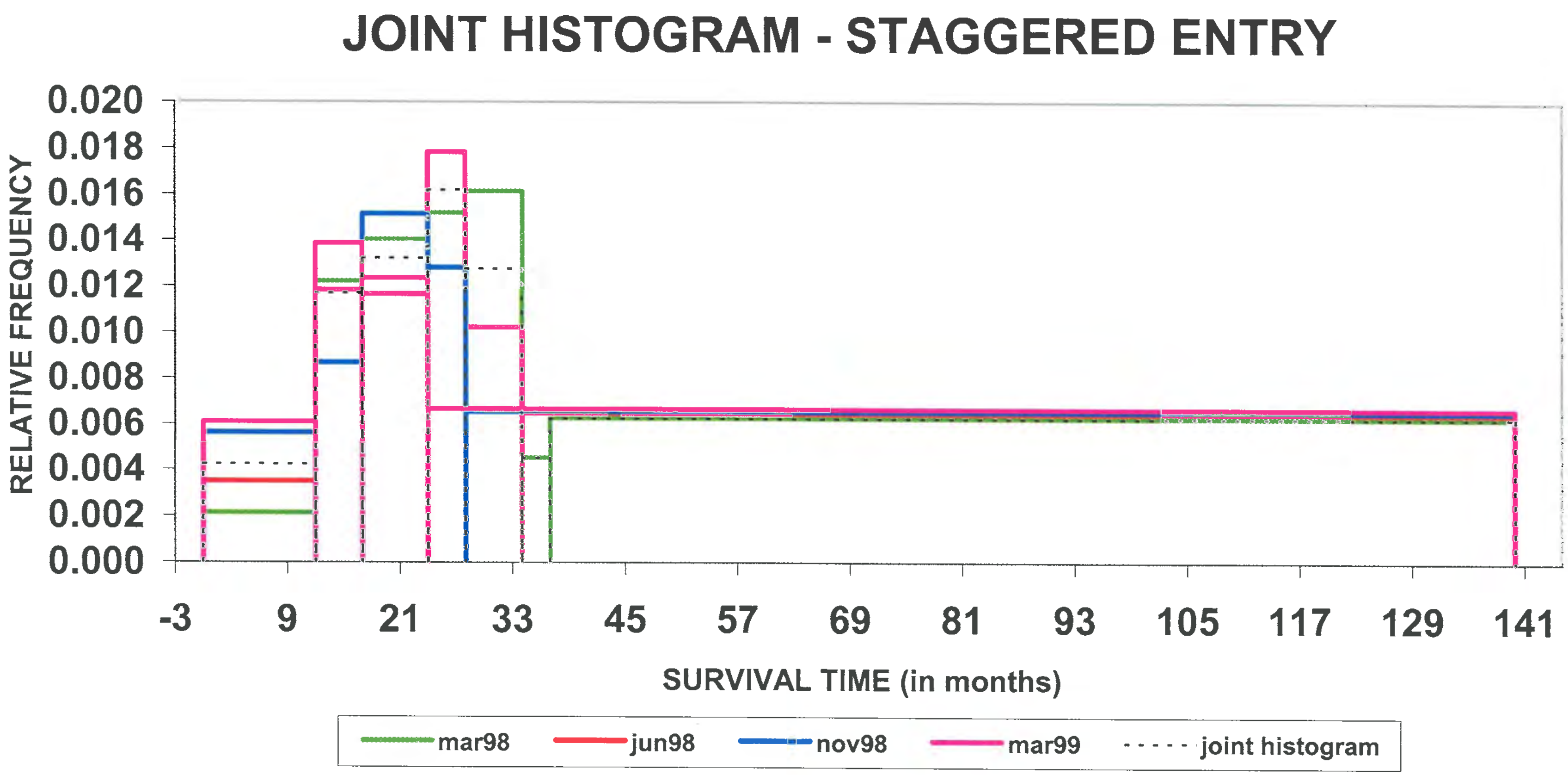

Figure 5.2: **Histograms of the four relative frequency distributions**

5.2 Fitting of a Lifetime Distribution

5.2.1 A fixed censoring time

Consider the June 1998 group of policies as an example of a data set with a pre-assigned fixed censoring time at 15 April 2001. All these 2809 policies start at the same time. The censored lifetimes are grouped into the open interval $[34;\infty)$, while the observed lifetimes are grouped into intervals $[0;12), [12;17), [17;24), [24;28)$, and $[28;34)$. Note that
$\boldsymbol{x} = (12, 17, 24, 28, 34)'$ is a 5×1 vector of upper class boundaries,
$\boldsymbol{f} = (118, 166, 229, 200, 172, 1924)'$ is a 6×1 vector of frequencies and
$\boldsymbol{p} = (0.0420078, 0.0590958, 0.0815237, 0.0711997, 0.0612318, 0.6849413)'$ is a 6×1 vector of relative frequencies. This vector is used as starting value for the iterative process.

One survival model is fitted to this grouped data set under constraints imposed by the Weibull/log-logistic/lognormal distribution. The double iterative procedure, as described in section 3.3.2 of chapter three, is illustrated by the following computer output when a Weibull model is fitted to the grouped data.

```
I J         P         M | I J         P         M | I J         P         M | I J         P         M | I J         P         M
                        |                         |                         |                         |
1 1 0.0550538 0.0420078 | 2 1 0.0504223 0.0562540 | 3 1 0.0505730 0.0507667 | 4 1 0.0505685 0.0509455 | 5 1 0.0505686 0.0509401
    0.0489497 0.0590958 |     0.0465755 0.0466379 |     0.0466658 0.0456550 |     0.0466631 0.0457035 |     0.0466632 0.0457021
    0.0801283 0.0815237 |     0.0815332 0.0803011 |     0.0815189 0.0815789 |     0.0815195 0.0815707 |     0.0815195 0.0815712
    0.0504270 0.0711997 |     0.0524268 0.0513481 |     0.0523893 0.0534019 |     0.0523906 0.0533565 |     0.0523905 0.0533581
    0.0811157 0.0612318 |     0.0861566 0.0812418 |     0.0860254 0.0857363 |     0.0860296 0.0856221 |     0.0860295 0.0856258
    0.6843255 0.6849413 |     0.6828856 0.6842171 |     0.6828276 0.6828612 |     0.6828286 0.6828017 |     0.6828286 0.6828027
                        |                         |                         |                         |
1 2 0.0562478 0.0420078 | 2 2 0.0507666 0.0562540 | 3 2 0.0509453 0.0507667 | 4 2 0.0509400 0.0509455 | 5 2 0.0509401 0.0509401
    0.0466462 0.0590958 |     0.0456548 0.0466379 |     0.0457034 0.0456550 |     0.0457021 0.0457035 |     0.0457021 0.0457021
    0.0803032 0.0815237 |     0.0815797 0.0803011 |     0.0815715 0.0815789 |     0.0815720 0.0815707 |     0.0815719 0.0815712
    0.0513481 0.0711997 |     0.0534018 0.0513481 |     0.0533565 0.0534019 |     0.0533580 0.0533565 |     0.0533580 0.0533581
    0.0812371 0.0612318 |     0.0857358 0.0812418 |     0.0856215 0.0857363 |     0.0856252 0.0856221 |     0.0856251 0.0856258
    0.6842177 0.6849413 |     0.6828613 0.6842171 |     0.6828017 0.6828612 |     0.6828028 0.6828017 |     0.6828027 0.6828027
                        |                         |                         |                         |
1 3 0.0562540 0.0420078 | 2 3 0.0507667 0.0562540 | 3 3 0.0509455 0.0507667 | 4 3 0.0509401 0.0509455 | 5 3 0.0509403 0.0509401
    0.0466379 0.0590958 |     0.0456550 0.0466379 |     0.0457035 0.0456550 |     0.0457021 0.0457035 |     0.0457022 0.0457021
    0.0803011 0.0815237 |     0.0815789 0.0803011 |     0.0815707 0.0815789 |     0.0815712 0.0815707 |     0.0815712 0.0815712
    0.0513481 0.0711997 |     0.0534019 0.0513481 |     0.0533565 0.0534019 |     0.0533581 0.0533565 |     0.0533580 0.0533581
    0.0812418 0.0612318 |     0.0857363 0.0812418 |     0.0856221 0.0857363 |     0.0856258 0.0856221 |     0.0856257 0.0856258
    0.6842171 0.6849413 |     0.6828612 0.6842171 |     0.6828017 0.6828612 |     0.6828027 0.6828017 |     0.6828027 0.6828027
                        |                                                                               |
1 4 0.0562540 0.0420078 |                                                                               |
    0.0466379 0.0590958 |                                                                               |        MLE for M_c
    0.0803011 0.0815237 |                                                                               |
    0.0513481 0.0711997 |                                                                               |         0.0509403
    0.0812418 0.0612318 |                                                                               |         0.0457022
    0.6842171 0.6849413 |                                                                               |         0.0815712
                                                                                                        |         0.053358
                                                                                                        |         0.0856257
                                                                                                        |         0.6828027
```

The estimated parameters of the fitted survival models and the Wald test with the discrepancy values are reported in Table 5.2.

Table 5.2: **Maximum likelihood estimation subject to constraints: a fixed censoring time**

	Survival model		
	Weibull	Log-logistic	Lognormal
Maximum likelihood estimates	$\ln\widehat{\lambda}$=-7.693382 $\widehat{\alpha}$=1.9084456	$\ln\widehat{\lambda}$=-8.243037 $\widehat{\alpha}$=1.9084456	$\widehat{\mu}$ =3.8910773 $\widehat{\sigma}$ =0.8319341
Wald test	51.5	39.8	25.0
Discrepancy	0.0183	0.0142	0.0089

Figure 5.3 shows the histogram of the relative frequency distribution and the fitted survival distributions. It is clear that the lognormal and log-logistic models fit very well.

5.2.2 Staggered entry

Introduction

Consider Figure 5.2, the four histograms of the relative frequency distributions. Maximum likelihood estimates are to be found in the following ways.

1. One survival model is fitted to the four histograms under constraints imposed by the Weibull/log-logistic/lognormal distribution.
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
   - 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.

The constraints imposed by the experimental design

Consider Figure 5.4, illustrating the constraints imposed by the experimental design.

- $\pi_{1,j} = \pi_{2,j} = \pi_{3,j} = \pi_{4,j} \quad j = 1, 2, ..., 3$
- $$\begin{aligned} \pi_{1,7} + \pi_{1,6} + \pi_{1,5} + \pi_{1,4} &= \pi_{2,6} + \pi_{2,5} + \pi_{2,4} \\ &= \pi_{3,5} + \pi_{3,4} \\ &= \pi_{4,4} \end{aligned}$$
- $$\begin{aligned} \pi_{1,5} &= \pi_{2,5} \\ \pi_{1,4} &= \pi_{2,4} \end{aligned}$$

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 $i \quad i = 1, 2, 3, 4 \quad j = 1, 2, ..., 7$

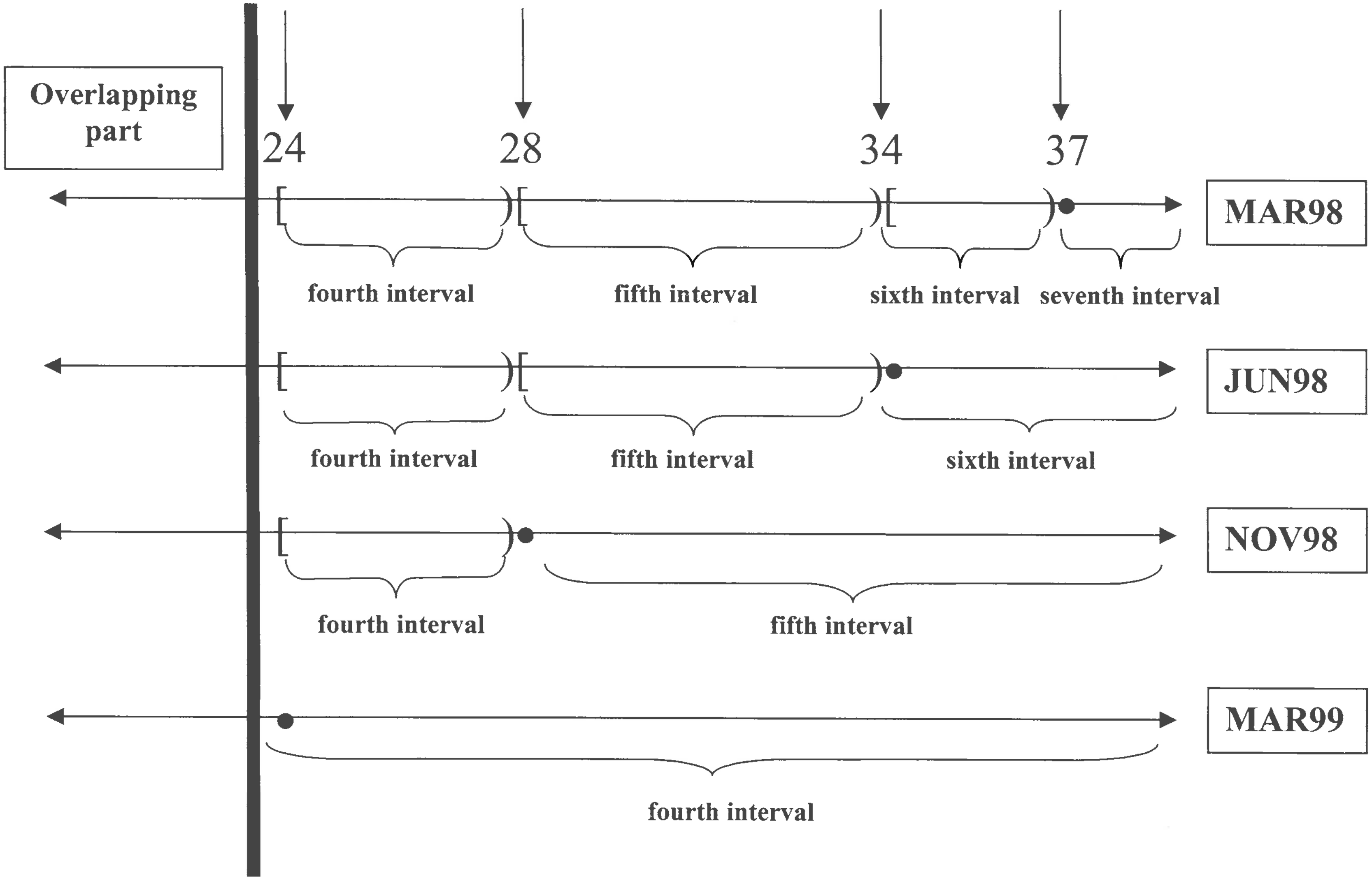

Figure 5.4: **Constraints imposed by the experimental design**

Fitting of one survival model to the four histograms

One survival model is fitted to the four histograms under constraints imposed by the Weibull/log-logistic/lognormal distribution. The estimated parameters and the Wald test with the discrepancy values are reported in Table 5.3.

Table 5.3: **Fitting of one survival model to the four histograms**

	Survival model		
	Weibull	Log-logistic	Lognormal
Maximum likelihood estimates	$\ln\widehat{\lambda}$=-7.39252	$\ln\widehat{\lambda}$=-7.959399	$\widehat{\mu}$ =3.8727296
	$\widehat{\alpha}$=1.8434286	$\widehat{\alpha}$=2.0647366	$\widehat{\sigma}$ =0.8636028
Wald test	302.5	253.6	254.9
Discrepancy	0.0300	0.0252	0.0253

The invariance property of the maximum likelihood estimator provides that the MLE of $\widehat{\ln\lambda}$ can be written as $\ln\widehat{\lambda}$.

Fitting of four survival models

Four survival models are fitted, one for each entry time, under constraints imposed by the Weibull/log-logistic/lognormal distribution and under further constraints that the parameters are equal.

The estimated parameters and the Wald test with the discrepancy values are reported in Table 5.4.

Table 5.4: **Fitting of one survival model to the four histograms**

Maximum likelihood estimates	Survival model: Weibull	Log-logistic	Lognormal
March 1998	$\ln\widehat{\lambda}$=-8.230773	$\ln\widehat{\lambda}$=-8.960949	$\widehat{\mu}$ =3.8234358
	$\widehat{\alpha}$=2.0570424	$\widehat{\alpha}$=2.3273887	$\widehat{\sigma}$ =0.7219206
June 1998	$\ln\widehat{\lambda}$=-7.693383	$\ln\widehat{\lambda}$=-8.243037	$\widehat{\mu}$ =3.8910773
	$\widehat{\alpha}$=1.9084457	$\widehat{\alpha}$=2.1214022	$\widehat{\sigma}$ =0.8319341
Nov 1998	$\ln\widehat{\lambda}$=-7.172834	$\ln\widehat{\lambda}$=-7.582113	$\widehat{\mu}$ =3.9182342
	$\widehat{\alpha}$=1.8026532	$\widehat{\alpha}$=1.9727851	$\widehat{\sigma}$ =0.9624843
March 1999	$\ln\widehat{\lambda}$=-6.781666	$\ln\widehat{\lambda}$=-7.113033	$\widehat{\mu}$ =3.9521501
	$\widehat{\alpha}$=1.7103598	$\widehat{\alpha}$=1.8569722	$\widehat{\sigma}$ =1.0417936
Over all four entry times	$\ln\widehat{\lambda}$=-7.39252	$\ln\widehat{\lambda}$=-7.959399	$\widehat{\mu}$ =3.8727296
	$\widehat{\alpha}$=1.8434286	$\widehat{\alpha}$=2.0647366	$\widehat{\sigma}$ =0.8636028
Wald test	302.5	253.6	254.9
Discrepancy	0.0300	0.0252	0.0253

A joint histogram to the four histograms is needed to make a graphical representation of the fitted models.

Fitting of a joint histogram to the four histograms

A joint histogram is fitted to the four histograms under constraints imposed by the experimental design. Table 5.5 gives the fitted joint relative frequencies to the four sets of relative frequencies of the samples. A graphical representation of the joint histogram over the four histograms appears in Figure 5.5.

Figure 5.6 shows the fitted joint histogram and the fitted survival distributions. The log-normal and log-logistic models again fit the data very well.

Table 5.5: **Fitted joint relative frequency distribution to the four samples**

Interval number	Lifetime Intervals				Relative Frequency Vector				Fitted Joint Relative Frequencies
	March 98	June 98	Nov 98	March 99	p_1	p_2	p_3	p_4	
first	$[0, 12)$	$[0, 12)$	$[0, 12)$	$[0, 12)$	0.025522	0.042008	0.067367	0.073038	0.050908
second	$[12, 17)$	$[12, 17)$	$[12, 17)$	$[12, 17)$	0.061098	0.059096	0.043307	0.069282	0.058450
third	$[17, 24)$	$[17, 24)$	$[17, 24)$	$[17, 24)$	0.098221	0.081524	0.105862	0.086394	0.092488
fourth	$[24, 28)$	$[24, 28)$	$[24, 28)$	$[24, \infty)$	0.060712	0.071200	0.051181	0.771286	0.064701
fifth	$[28, 34)$	$[28, 34)$	$[28, \infty)$		0.096674	0.061232	0.732284		0.076481
sixth	$[34, 37)$	$[34, \infty)$			0.013534	0.684941			0.013518
seventh	$[37, \infty)$				0.644238				0.643455

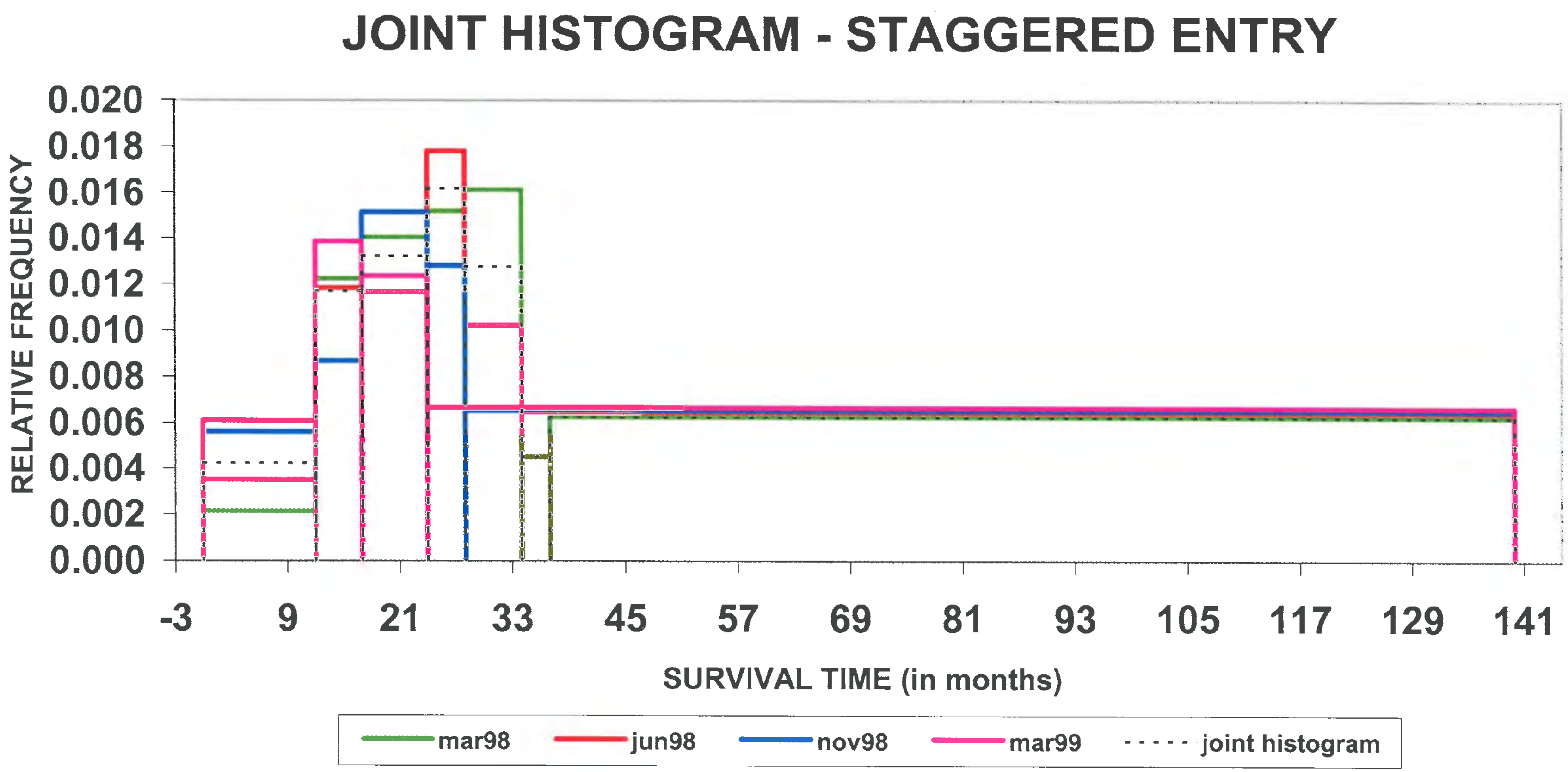

Figure 5.5: **Joint histogram over the four histograms**

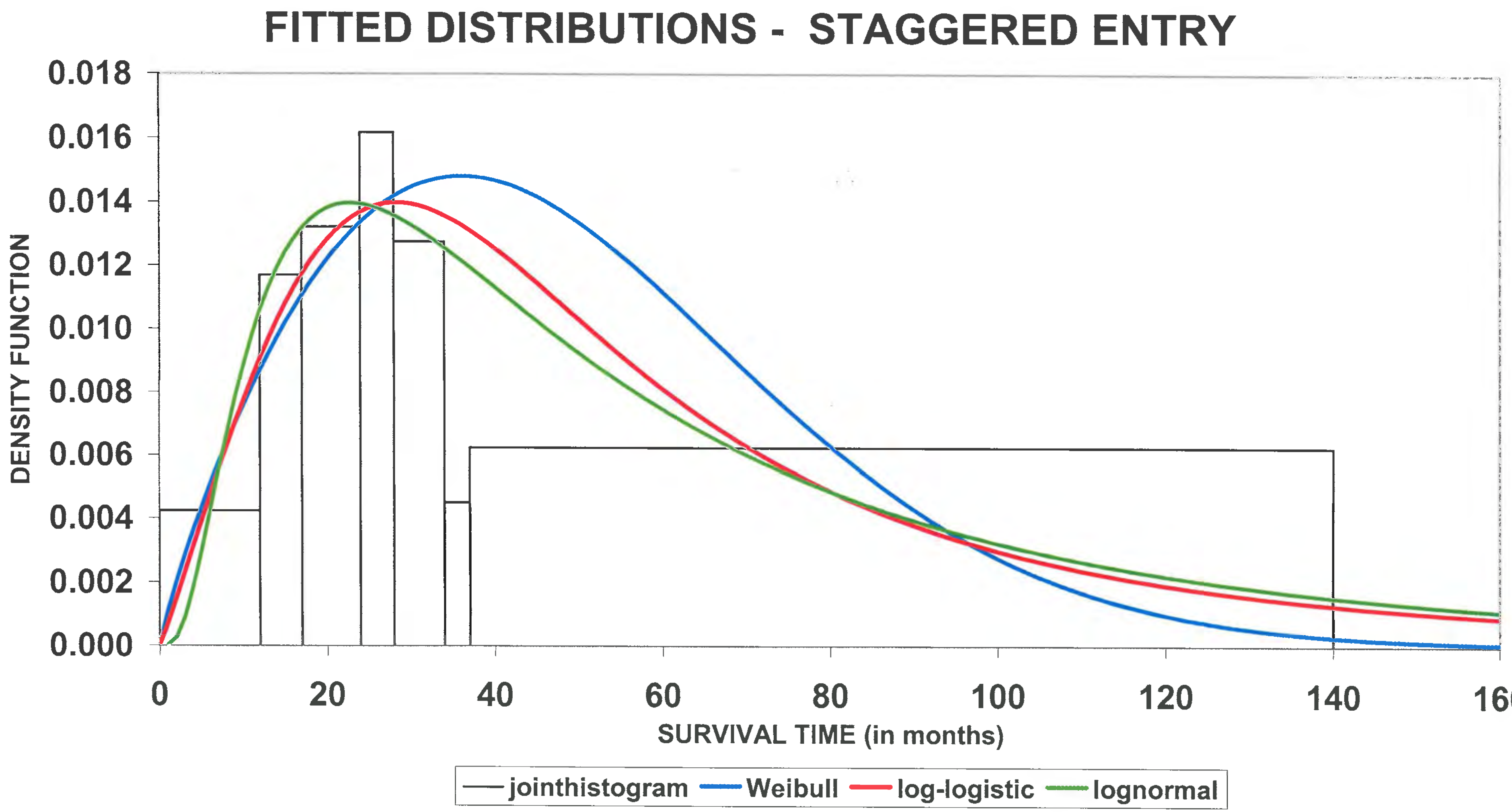

Figure 5.6: **Joint histogram and fitted survival distributions**

Estimated survivor and hazard functions and percentiles

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

The formulaes and examples of calculations of the estimated hazard rates, survivor functions, odds of a lapse and percentiles of the Weibull and log-logistic survival distributions are given on the next page.

The survival curves and the graphs of the hazard rates of the fitted Weibull and log-logistic models are shown respectively in Figure 5.7 and Figure 5.8. From Figure 5.7 is it clear that the two survivor functions are equal for $t-$values up to 40 months, and then the probability for a policy to survive longer than time t with $t >$40 becomes larger for the log-logistic fitting than for the Weibull fitting. Note in Figure 5.8 the increasing trend of the Weibull hazard rates as t increases.

Survival Model

WEIBULL

Estimated hazard function

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

$$\widehat{h}(t) = e^{-7.39252} \cdot 1.8434286 \cdot t^{1.8434286-1}$$

$$\widehat{h}(12) = 0.0092323$$
$$\widehat{h}(24) = 0.0165655$$

Estimated survivor function

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

$$\widehat{S}(t) = \exp(-e^{-7.39252} \cdot t^{1.8434286})$$

$$\widehat{S}(12) = 0.9416719$$
$$\widehat{S}(24) = 0.806001$$

Estimated odds of a lapse

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

$$\widehat{odds}(t) = \frac{1-\widehat{S}(t)}{\widehat{S}(t)} = \exp(e^{-7.39252} \cdot t^{1.8434286-1})$$

$$\widehat{odds}(12) = 0.061941$$
$$\widehat{odds}(24) = 0.240693$$

Estimated percentiles

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

$$\widehat{t_{50}} = \left(\frac{1}{e^{-7.39252}} \cdot \ln \frac{100}{100-50}\right)^{\frac{1}{1.8434286}} = 45.21$$

LOG-LOGISTIC

Estimated hazard function

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

$$\widehat{h}(t) = \frac{e^{-7.959399} \cdot 2.0647366 \cdot t^{2.0647366-1}}{(1+e^{-7.959399} \cdot t^{2.0647366})}$$

$$\widehat{h}(12) = 0.0095996$$
$$\widehat{h}(24) = 0.0170517$$

Estimated survivor function

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

$$\widehat{S}(t) = \frac{1}{1+e^{-7.959399} \cdot t^{2.0647366}}$$

$$\widehat{S}(12) = 0.9442083$$
$$\widehat{S}(24) = 0.8017956$$

Estimated odds of a lapse

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

$$\widehat{odds}(t) = e^{-7.959399} \cdot t^{2.0647366}$$

$$\widehat{odds_0}(12) = 0.0590884$$
$$\widehat{odds_0}(24) = 0.2472006$$

Estimated percentiles

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

$$\widehat{t_{50}} = \left(\frac{1}{e^{-7.959399}} \cdot \frac{50}{100-50}\right)^{\frac{1}{2.0647366}} = 47.22$$

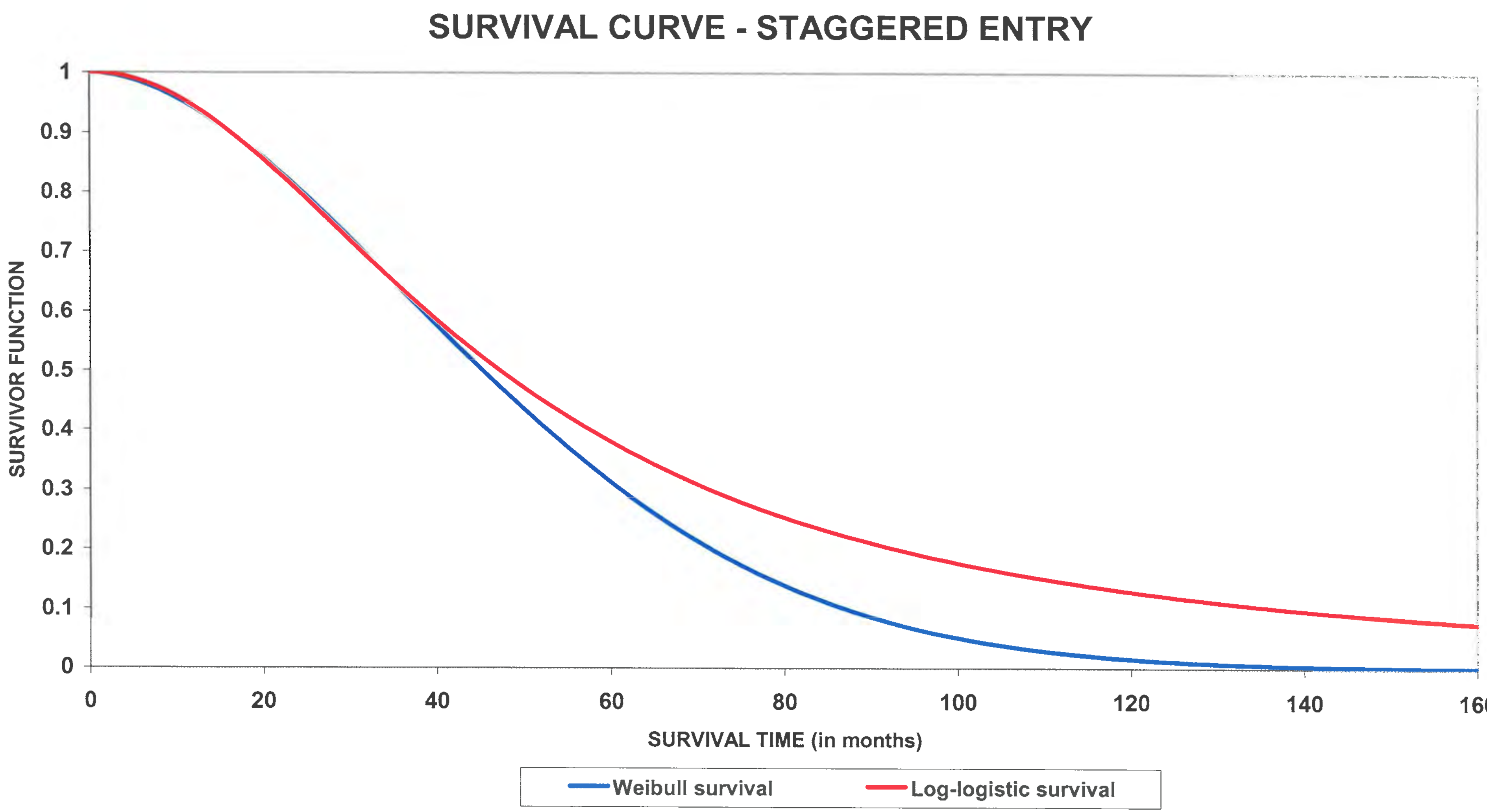

Figure 5.7: **Survival curves of fitted Weibull and log-logistic models**

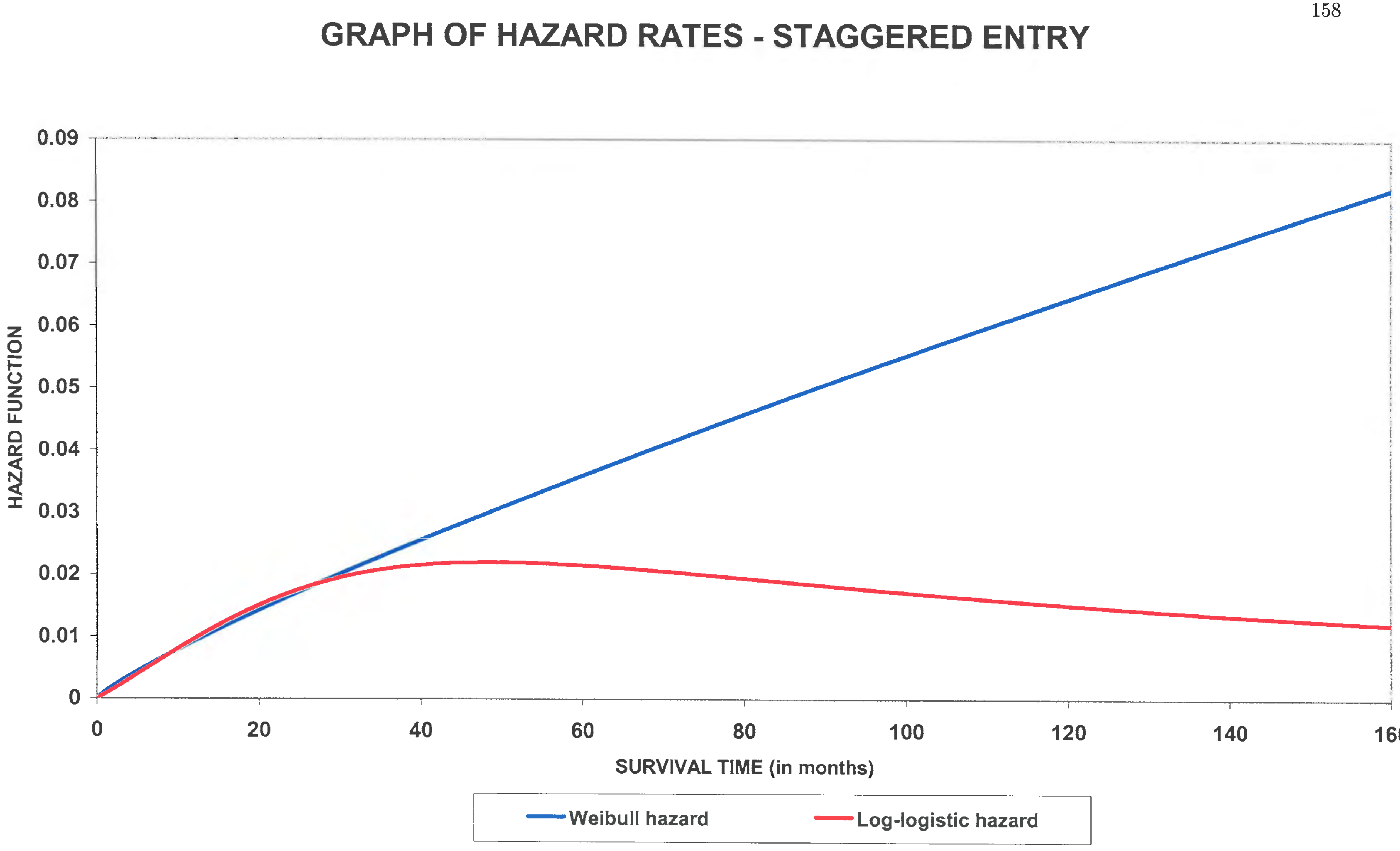

Figure 5.8: **Graphs of hazard rates of fitted Weibull and log-logistic models**

The estimated percentiles of the Weibull and log-logistic survival distributions are reported in Table 5.6.

Table 5.6: **Percentiles estimated from Weibull and log-logistic regression models**

	Survival model	
Percentile	Weibull	Log-logistic
P5	11.01	11.34
P10	16.27	16.29
P20	24.45	24.13
P25	28.06	27.74
P30	31.53	31.33
P40	38.31	38.80
P50	45.21	47.22
P60	52.60	57.47
P70	61.00	71.18
P75	65.85	80.40
P80	71.40	92.42
P90	86.71	136.88
P95	100.02	196.56

At the Weibull model, the median time to a lapse of a policy is estimated as 45.21 months and the odds of a lapse at 45.21 months is 1, that means $P(T >$45.21 months)$=P(T <$45.21 months). At the log-logistic model, the median time to a lapse of a policy is estimated as 47.22 months and the odds of a lapse at 47.22 months is 1, that means $P(T >$47.22 months)$=P(T <$47.22 months).

It is evident from the estimates of the percentile lifetimes that 20% of the policies will not lapse within 71 months under a Weibull model (see Weibull's P80), while 30% of the policies will not lapse within 71 months under a log-logistic model (see log-logistic's P70).

Note again the equal percentile estimates for the two distributions up to 40 months, confirming the pattern that was detected in the survival curves in Figure 5.7.

5.3 Fitting of Parametric Regression Models

5.3.1 Introduction

A survival model is fitted for **each level of a risk factor** or **combination of levels of risk factors** by using maximum likelihood estimation of parameters subject to constraints.

The fitting of regression models is illustrated where the effect of the risk factors (covariates) is to alter the scale parameter λ, while the shape parameter α remains constant. Applications are also done where both parameters alter.

The fitting of log-logistic regression models and Weibull regression models will be discussed only for **staggered entry of policies**.

5.3.2 A survival model for each level of a risk factor

Consider one risk factor AGE on three levels [18;35), [35;45) and [45+) years. The 10077 observations are distributed in the three age groups as follows: 3644 in age group [18;35), 3425 in age group [35;45) and 3008 in age group [45+). A regression model is fitted to the grouped survival data where each policy has information on the entry period as well as the age level. The grouped lifetimes of the policies with staggered entry as well as the concomitant information on AGE are given in Table 5.7.

The combined frequency vector $\boldsymbol{f}$ is defined as
$\boldsymbol{f}' = (\boldsymbol{f}'_{11}, \boldsymbol{f}'_{21}, \boldsymbol{f}'_{31}, \boldsymbol{f}'_{41}, \boldsymbol{f}'_{12}, \boldsymbol{f}'_{22}, \boldsymbol{f}'_{32}, \boldsymbol{f}'_{42}, \boldsymbol{f}'_{13}, \boldsymbol{f}'_{23}, \boldsymbol{f}'_{33}, \boldsymbol{f}'_{43})$
$\boldsymbol{f}_{il}$ is the frequency vector for the i^{th} entry group and the l^{th} AGE level,
$i = 1, 2, 3, 4$ and $l = 1, 2, 3$.

$\boldsymbol{f}_{11} = (29, 59, 95, 73, 108, 15, 642)'$
$\boldsymbol{f}_{12} = (21, 50, 91, 45, 75, 13, 553)'$
$\boldsymbol{f}_{13} = (16, 49, 68, 39, 67, 7, 471)'$

$\boldsymbol{f}_{21} = (41, 75, 103, 92, 83, 628)'$
$\boldsymbol{f}_{22} = (49, 62, 61, 66, 54, 753)'$

Table 5.7: **Multi-dimensional frequency table of grouped data set with one risk factor**

Entry	Age	Lifetime intervals						
March 98		$[0,12)$	$[12,17)$	$[17,24)$	$[24,28)$	$[28,34)$	$[34,37)$	$[37,\infty)$
	[18;35)	29	59	95	73	108	15	642
	[35;45)	21	50	91	45	75	13	553
	[45+)	16	49	68	39	67	7	471
June 98		$[0,12)$	$[12,17)$	$[17,24)$	$[24,28)$	$[28,34)$	$[34,\infty)$	
	[18;35)	41	75	103	92	83	628	
	[35;45)	49	62	61	66	54	753	
	[45+)	28	29	65	42	35	543	
Nov 98		$[0,12)$	$[12,17)$	$[17,24)$	$[24,28)$	$[28,\infty)$		
	[18;35)	68	34	99	57	570		
	[35;45)	40	44	83	33	533		
	[45+)	46	21	60	27	571		
March 99		$[0,12)$	$[12,17)$	$[17,24)$	$[24,\infty)$			
	[18;35)	71	60	69	573			
	[35;45)	54	61	68	616			
	[45+)	50	45	70	659			

$$\boldsymbol{f}_{23} = (28, 29, 65, 42, 35, 543)'$$

$$\boldsymbol{f}_{31} = (68, 34, 99, 57, 570)'$$
$$\boldsymbol{f}_{32} = (40, 44, 83, 33, 533)'$$
$$\boldsymbol{f}_{33} = (46, 21, 60, 27, 571)'$$

$$\boldsymbol{f}_{41} = (71, 60, 69, 573)'$$
$$\boldsymbol{f}_{42} = (54, 61, 68, 616)'$$
$$\boldsymbol{f}_{43} = (50, 45, 70, 659)'$$

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

$$\boldsymbol{x}_1 = \begin{pmatrix} 12 \\ 17 \\ 24 \\ 28 \\ 34 \\ 37 \end{pmatrix} \quad \boldsymbol{x}_2 = \begin{pmatrix} 12 \\ 17 \\ 24 \\ 28 \\ 34 \end{pmatrix} \quad \boldsymbol{x}_3 = \begin{pmatrix} 12 \\ 17 \\ 24 \\ 28 \end{pmatrix} \quad \text{and} \quad \boldsymbol{x}_4 = \begin{pmatrix} 12 \\ 17 \\ 24 \end{pmatrix}.$$

From the estimated regression parameters, survival model parameters can be found for each level of this risk factor as well as for the baseline distribution.

The shape parameter remains constant

The estimated regression coefficients of the regression model where the effect of the risk factor AGE is to alter the scale parameter λ, while the shape parameter α remains constant, are reported in Table 5.8.

Table 5.8: **Fitting a regression model (constant shape) to grouped data with one risk factor**

Effect	Maximum likelihood estimates	**Regression model** Log-logistic	Weibull
Baseline mean	$\widehat{\ln \lambda_0} = \ln \widehat{\lambda}_0$	-7.981750	-7.404312
Age [18;35)	$\widehat{\beta}_{A_1}$	0.180958	0.159090
Age [35;45)	$\widehat{\beta}_{A_2}$	-0.034975	-0.033957
Age [45+)	$\widehat{\beta}_{A_3}$	-0.145983	-0.125133
Constant shape	$\widehat{\alpha}$	2.066384	1.8423341

The estimated lambda parameters of the three survival distributions for the three AGE levels then are

$$\begin{aligned}\widehat{\lambda}_{A_1} &= \exp(\ln \widehat{\lambda}_0 + \widehat{\beta}_{A_1}) \\ \widehat{\lambda}_{A_2} &= \exp(\ln \widehat{\lambda}_0 + \widehat{\beta}_{A_2}) \\ \widehat{\lambda}_{A_3} &= \exp(\ln \widehat{\lambda}_0 + \widehat{\beta}_{A_3}).\end{aligned}$$

with the same estimated alpha parameter $\widehat{\alpha}$. These parameters are summarized for each AGE level in Table 5.9.

Table 5.9: **Parameters of a survival model (constant shape) for each level of risk factor AGE**

AGE level	Maximum likelihood estimates	**Survival model** Log-logistic	Weibull
Age [18;35)	$\ln \widehat{\lambda}_{A_1}$	-7.800792	-7.245223
	$\widehat{\alpha}$	2.066384	1.842334
Age [35;45)	$\ln \widehat{\lambda}_{A_2}$	-8.016725	-7.438269
	$\widehat{\alpha}$	2.066384	1.842334
Age [45+)	$\ln \widehat{\lambda}_{A_3}$	-8.127733	-7.529445
	$\widehat{\alpha}$	2.066384	1.842334
Baseline	$\ln \widehat{\lambda}_0$	-7.981750	-7.404312
	$\widehat{\alpha}$	2.066384	1.842334

The shape parameter alters

The fitting of regression models is illustrated where the effect of the risk factors (covariates) is to alter both the scale parameter λ and the shape parameter α. The estimated regression coefficients of this regression model are reported in Table 5.10.

Table 5.10: **Fitting a regression model (shape alters) to grouped data with one risk factor**

Effect	Maximum likelihood estimates	**Regression model** Log-logistic	Weibull
Baseline mean	$\widehat{\ln \lambda_0} = \ln \widehat{\lambda}_0$	-7.943357	-7.381423
Age [18;35)	$\widehat{\beta}_{A_1}$	-0.196012	-0.075175
	$\widehat{\alpha}_{A_1}$	2.168064	1.904217
Age [35;45)	$\widehat{\beta}_{A_2}$	0.156976	0.119892
	$\widehat{\alpha}_{A_2}$	1.9974967	1.790610
Age [45+)	$\widehat{\beta}_{A_3}$	0.039035	-0.044717
	$\widehat{\alpha}_{A_3}$	1.9995073	1.811986

The weighted mean of the $\widehat{\alpha}_{A_i}$'s $\quad i = 1, 2, 3$ is used as an estimate for the shape parameter of the baseline distribution.

The estimated lambda parameters of the three survival distributions for the three AGE levels are calculated from Table 5.10 as

$$\begin{aligned}
\widehat{\lambda}_{A_1} &= \exp(\ln \widehat{\lambda}_0 + \widehat{\beta}_{A_1}) \\
\widehat{\lambda}_{A_2} &= \exp(\ln \widehat{\lambda}_0 + \widehat{\beta}_{A_2}) \\
\widehat{\lambda}_{A_3} &= \exp(\ln \widehat{\lambda}_0 + \widehat{\beta}_{A_3}).
\end{aligned}$$

Each age group survival distribution has its own estimated alpha parameter.

These parameters for each AGE level are summarized in Table 5.11, together with the Wald test and discrepancy value to compare the fitted survival distributions at each AGE level.

Table 5.11: **Parameters of a survival model (shape alters) for each level of risk factor AGE**

AGE level	Maximum likelihood estimates	**Survival model** Log-logistic	Weibull
Age [18;35)	$\ln\widehat{\lambda}_{A_1}$	-8.139369	-7.456598
	$\widehat{\alpha}_{A_1}$	2.168064	1.904217
Wald test		128.5	144.2
Discrepancy		0.0353	0.0396
Age [35;45)	$\ln\widehat{\lambda}_{A_2}$	-7.786381	-7.261531
	$\widehat{\alpha}_{A_2}$	1.997497	1.790610
Wald test		93.1	108.3
Discrepancy		0.0271	0.0316
Age [45+)	$\ln\widehat{\lambda}_{A_3}$	-7.904321	-7.426139
	$\widehat{\alpha}_{A_3}$	1.999507	1.811986
Wald test		95.5	109.5
Discrepancy		0.0317	0.0364
Baseline	$\ln\widehat{\lambda}_0$	-7.943357	-7.381423
	$\widehat{\alpha}_0$	2.0597767	1.8380729

A joint histogram to the data of each AGE level **over the four entry groups** is needed to make a graphical representation of the fitted models for each AGE level. Table 5.12 gives the three sets of fitted joint frequencies for the three AGE levels. This fitting was done by maximum likelihood estimation subject to constraints imposed by the experimental design. The Wald test and discrepancy value measure the goodness-of-fit.

Table 5.12: **Fitted joint frequency distributions for the three AGE levels**

Interval number	Interval of survival times	Fitted Joint Frequencies		
		Age [18;35) years	Age [35;45) years	Age [45+) years
first	$[0, 12)$	209	164	140
second	$[12, 17)$	228	217	144
third	$[17, 24)$	366	303	263
fourth	$[24, 28)$	285.65814	195.15779	165.56561
fifth	$[28, 34)$	330.67093	226.805	208.49003
sixth	$[34, 37)$	50.791574	53.264105	30.561947
seventh	$[37, \infty)$	2173.8794	2265.7731	2056.3824
Wald		70.63	53.92	46.62
Discrepancy		0.0194	0.0157	0.0155

Figure 5.9 shows the fitted joint histogram and the fitted survival distributions for age group [18;35).

Figure 5.10 shows the fitted joint histogram and the fitted survival distributions for age group [35;45).

Figure 5.11 shows the fitted joint histogram and the fitted survival distributions for age group [45+).

In all the cases the survival models fit very well, with the log-logistic model slightly better than the Weibull model, as indicated by the discrepancy values in Table 5.11.

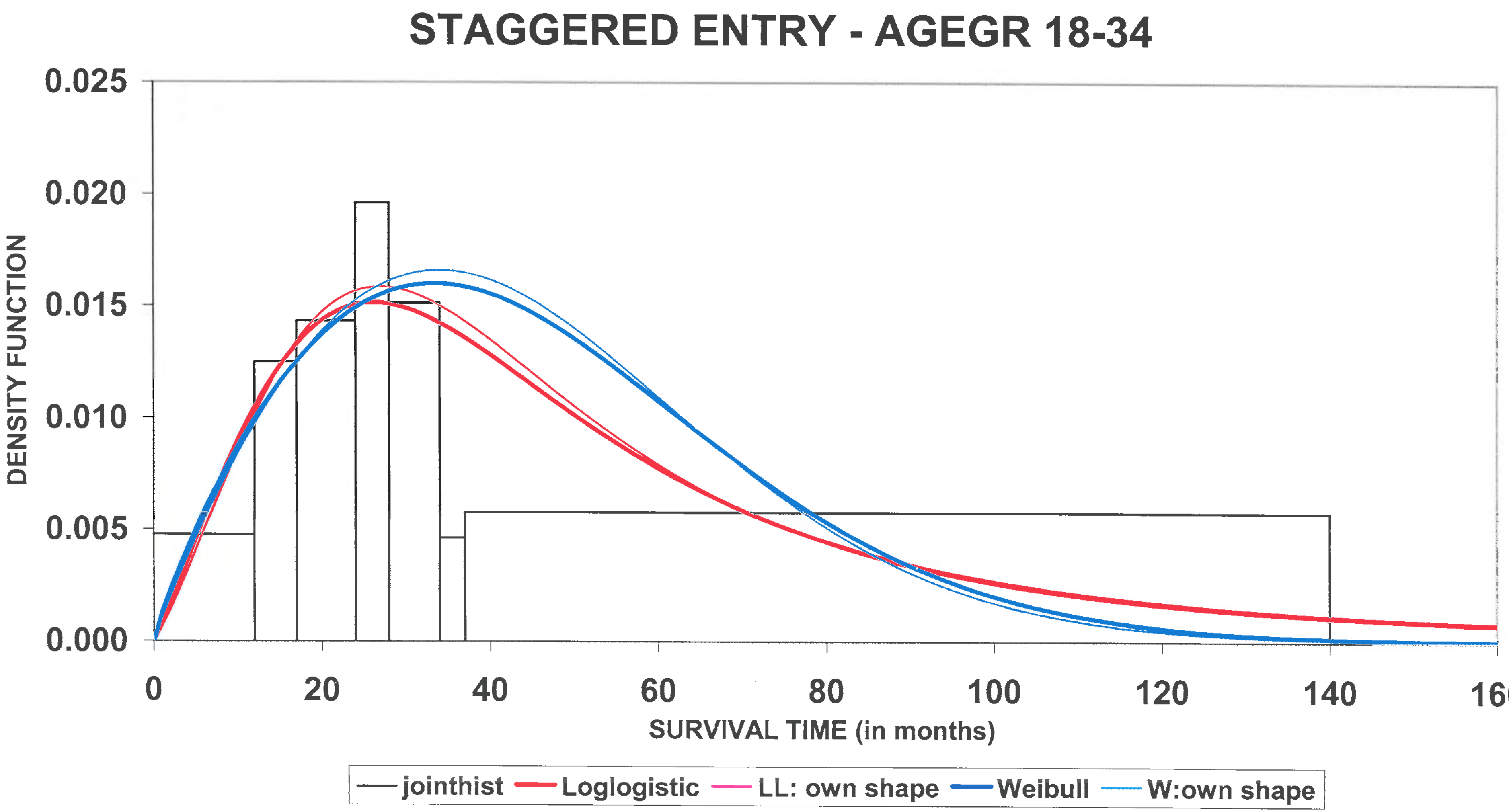

Figure 5.9: **Joint histogram and fitted survival distributions for age group [18;35)**

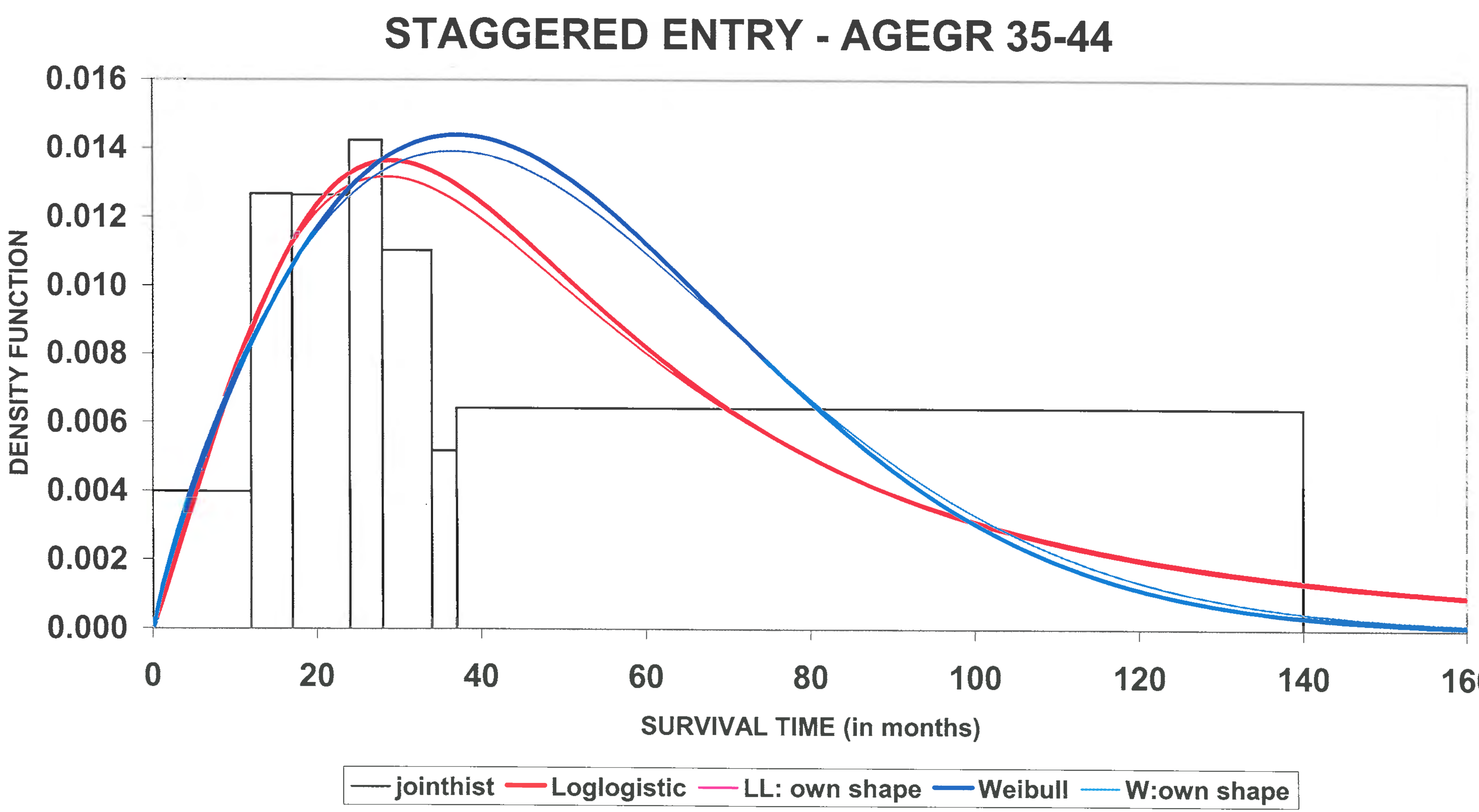

Figure 5.10: **Joint histogram and fitted survival distributions for age group [35;45)**

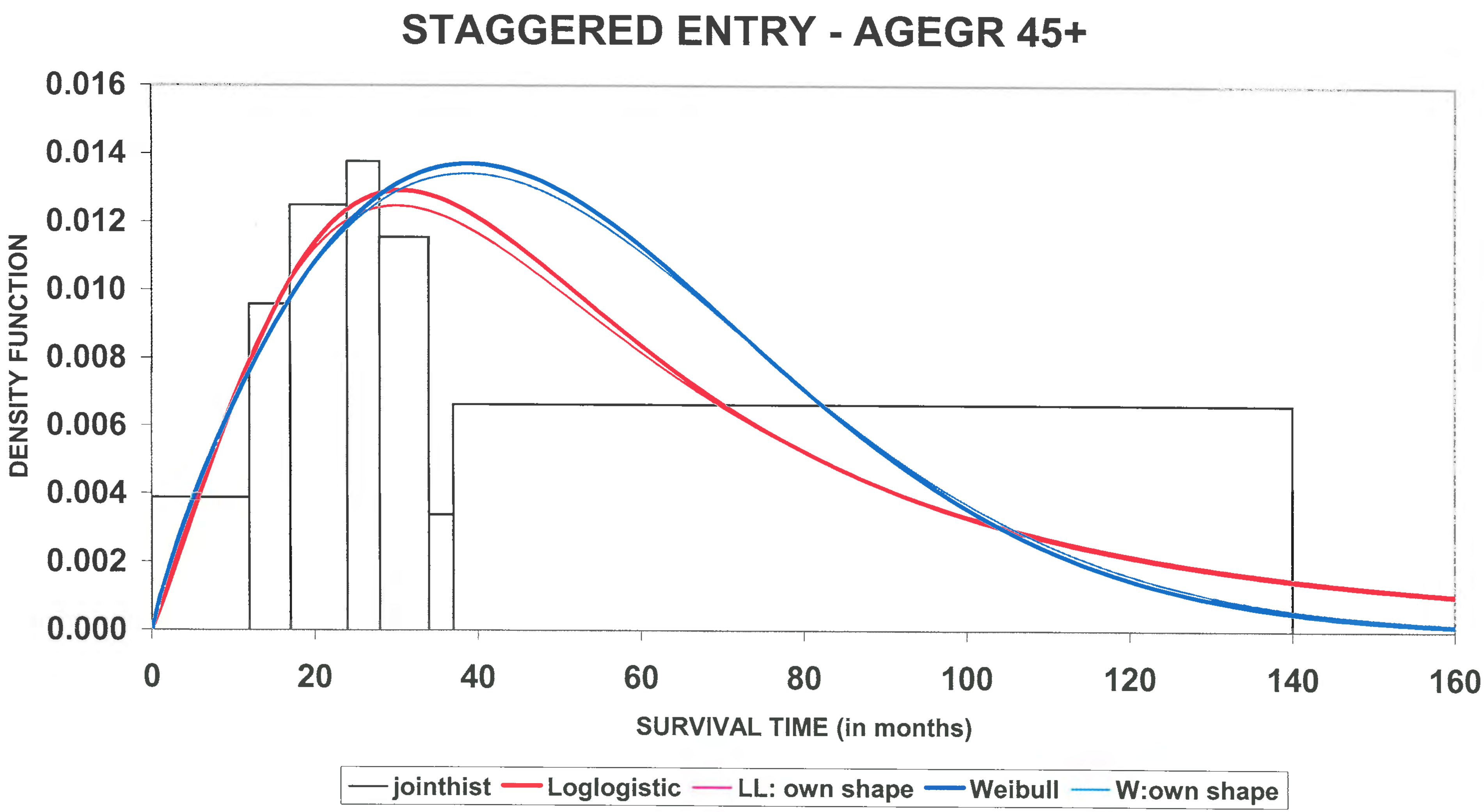

Figure 5.11: **Joint histogram and fitted survival distributions for age group [45+)**

5.3.3 Deriving of indices and risk scores from log-logistic regression model

Once the parameters of the log-logistic baseline distribution and log-logistic age group distributions have been estimated, estimated hazard and survivor functions, odds of a lapse, odds ratios and hazard ratios at time t can be calculated.

The odds ratio for age group [18;35) is the relative odds of a lapse at time t of a policy, with the age of the policyholder in [18;35), compared to a policy with the baseline characteristics. The odds ratios for the three age groups result in a set of indices, showing the effect of each age group on the baseline odds of a lapse at time t.

The hazard ratio for age group [18;35) is the relative hazard rate of a lapse at time t of a policy, with the age of the policyholder in [18;35), compared to a policy with the baseline characteristics. The hazard ratios for the three age groups result in a set of risk scores, showing the effect of each age group on the baseline hazard rate of a lapse at time t.

Percentiles of the four log-logistic survival distributions can also be estimated.

The calculations of estimated hazard and survivor functions, odds of a lapse, odds ratios and hazard ratios are illustrated on the following five pages.

The survival curves and the graphs of the hazard rates of the fitted log-logistic age group models are shown in Figure 5.12 and Figure 5.13 with measure of comparison the baseline curves. These survival curves can be described as graphs of the covariate-adjusted survivor functions, and the other graphs as graphs of the covariate-adjusted hazard rates.

The effects of the agegroups on the baseline distribution are clearly depicted in these figures. It is evident from these two figures that the policyholders in the age group 45+ have the lowest risk for their policies to lapse. Note the the survival curve of this age group lies above the baseline survival curve in Figure 5.12, while the curve of the hazard rates for this age group lies the furthest distance beneath the baseline curve of hazard rates. Similarly age group [18;35) has the highest risk for their policies to lapse.

Estimated Hazard Function at log-logistic regression model

Shape remains constant

$$\widehat{h}_0(t) = \frac{\widehat{\lambda}_0 \cdot \widehat{\alpha} \cdot t^{\widehat{\alpha}-1}}{(1+\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}})}$$

$$\widehat{h}_0(t) = \frac{e^{-7.98175} \cdot 2.0663843 \cdot t^{2.0663843-1}}{(1+e^{-7.98175} \cdot t^{2.0663843})}$$

$$\widehat{h}_0(12) = 0.009443$$
$$\widehat{h}_0(24) = 0.0168323$$

$$\widehat{h}_{A_1}(t) = \frac{\widehat{\lambda}_{A_1} \cdot \widehat{\alpha} \cdot t^{\widehat{\alpha}-1}}{(1+\widehat{\lambda}_{A_1} \cdot t^{\widehat{\alpha}})}$$

$$\widehat{h}_{A_1}(t) = \frac{e^{-7.800792} \cdot 2.0663843 \cdot t^{2.0663843-1}}{(1+e^{-7.800792} \cdot t^{2.0663843})}$$

$$\widehat{h}_{A_1}(12) = 0.0111944$$
$$\widehat{h}_{A_1}(24) = 0.0194182$$

$$\widehat{h}_{A_2}(t) = \frac{\widehat{\lambda}_{A_2} \cdot \widehat{\alpha} \cdot t^{\widehat{\alpha}-1}}{(1+\widehat{\lambda}_{A_2} \cdot t^{\widehat{\alpha}})}$$

$$\widehat{h}_{A_2}(t) = \frac{e^{-8.016725} \cdot 2.0663843 \cdot t^{2.0663843-1}}{(1+e^{-8.016725} \cdot t^{2.0663843})}$$

$$\widehat{h}_{A_2}(12) = 0.0091356$$
$$\widehat{h}_{A_2}(24) = 0.0163637$$

$$\widehat{h}_{A_3}(t) = \frac{\widehat{\lambda}_{A_3} \cdot \widehat{\alpha} \cdot t^{\widehat{\alpha}-1}}{(1+\widehat{\lambda}_{A_3} \cdot t^{\widehat{\alpha}})}$$

$$\widehat{h}_{A_3}(t) = \frac{e^{-8.127733} \cdot 2.0663843 \cdot t^{2.0663843-1}}{(1+e^{-8.127733} \cdot t^{2.0663843})}$$

$$\widehat{h}_{A_3}(12) = 0.0082216$$
$$\widehat{h}_{A_3}(24) = 0.0149428$$

Shape alters

$$\widehat{h}_0(t) = \frac{\widehat{\lambda}_0 \cdot \widehat{\alpha}_0 \cdot t^{\widehat{\alpha}_0-1}}{(1+\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}_0})}$$

$$\widehat{h}_0(t) = \frac{e^{-7.943357} \cdot 2.0597767 \cdot t^{2.0597767-1}}{(1+e^{-7.943357} \cdot t^{2.0597767})}$$

$$\widehat{h}_0(12) = 0.0096102$$
$$\widehat{h}_0(24) = 0.0170145$$

$$\widehat{h}_{A_1}(t) = \frac{\widehat{\lambda}_{A_1} \cdot \widehat{\alpha}_{A_1} \cdot t^{\widehat{\alpha}_{A_1}-1}}{(1+\widehat{\lambda}_{A_1} \cdot t^{\widehat{\alpha}_{A_1}})}$$

$$\widehat{h}_{A_1}(t) = \frac{e^{-8.139369} \cdot 2.1680640 \cdot t^{2.1680640-1}}{(1+e^{-8.139369} \cdot t^{2.1680640})}$$

$$\widehat{h}_{A_1}(12) = 0.0108363$$
$$\widehat{h}_{A_1}(24) = 0.0201312$$

$$\widehat{h}_{A_2}(t) = \frac{\widehat{\lambda}_{A_2} \cdot \widehat{\alpha}_{A_2} \cdot t^{\widehat{\alpha}_{A_2}-1}}{(1+\widehat{\lambda}_{A_2} \cdot t^{\widehat{\alpha}_{A_2}})}$$

$$\widehat{h}_{A_2}(t) = \frac{e^{-7.786381} \cdot 1.9974967 \cdot t^{1.9974967-1}}{(1+e^{-7.786381} \cdot t^{1.9974967})}$$

$$\widehat{h}_{A_2}(12) = 0.0093391$$
$$\widehat{h}_{A_2}(24) = 0.015965$$

$$\widehat{h}_{A_3}(t) = \frac{\widehat{\lambda}_{A_3} \cdot \widehat{\alpha}_{A_3} \cdot t^{\widehat{\alpha}_{A_3}-1}}{(1+\widehat{\lambda}_{A_3} \cdot t^{\widehat{\alpha}_{A_3}})}$$

$$\widehat{h}_{A_3}(t) = \frac{e^{-7.904321} \cdot 1.9995073 \cdot t^{1.9995073-1}}{(1+e^{-7.904321} \cdot t^{1.9995073})}$$

$$\widehat{h}_{A_3}(12) = 0.0084005$$
$$\widehat{h}_{A_3}(24) = 0.0145896$$

Estimated Survivor Function at log-logistic regression model

Shape remains constant

$$\widehat{S}_0(t) = \frac{1}{1+\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}}}$$

$$\widehat{S}_0(t) = \frac{1}{1+e^{-7.981750} \, . \, t^{2.066384}}$$

$$\begin{aligned} \widehat{S}_0(12) &= 0.9451622 \\ \widehat{S}_0(24) &= 0.8045013 \end{aligned}$$

$$\widehat{S}_{A_1}(t) = \frac{1}{1+\widehat{\lambda}_{A_1} \cdot t^{\widehat{\alpha}}}$$

$$\widehat{S}_{A_1}(t) = \frac{1}{1+e^{-7.800792} \, . \, t^{2.066384}}$$

$$\begin{aligned} \widehat{S}_{A_1}(12) &= 0.9349915 \\ \widehat{S}_{A_1}(24) &= 0.7744674 \end{aligned}$$

$$\widehat{S}_{A_2}(t) = \frac{1}{1+\widehat{\lambda}_{A_2} \cdot t^{\widehat{\alpha}}}$$

$$\widehat{S}_{A_2}(t) = \frac{1}{1+e^{-8.016725} \, . \, t^{2.066384}}$$

$$\begin{aligned} \widehat{S}_{A_2}(12) &= 0.946947 \\ \widehat{S}_{A_2}(24) &= 0.809944 \end{aligned}$$

$$\widehat{S}_{A_3}(t) = \frac{1}{1+\widehat{\lambda}_{A_3} \cdot t^{\widehat{\alpha}}}$$

$$\widehat{S}_{A_3}(t) = \frac{1}{1+e^{-8.127733} \, . \, t^{2.066384}}$$

$$\begin{aligned} \widehat{S}_{A_3}(12) &= 0.9522551 \\ \widehat{S}_{A_3}(24) &= 0.8264469 \end{aligned}$$

Shape alters

$$\widehat{S}_0(t) = \frac{1}{1+\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}_0}}$$

$$\widehat{S}_0(t) = \frac{1}{1+e^{-7.943357} \, . \, t^{2.0597767}}$$

$$\begin{aligned} \widehat{S}_0(12) &= 0.9440121 \\ \widehat{S}_0(24) &= 0.8017512 \end{aligned}$$

$$\widehat{S}_{A_1}(t) = \frac{1}{1+\widehat{\lambda}_{A_1} \cdot t^{\widehat{\alpha}_{A_1}}}$$

$$\widehat{S}_{A_1}(t) = \frac{1}{1+e^{-8.139369} \, . \, t^{2.168064}}$$

$$\begin{aligned} \widehat{S}_{A_1}(12) &= 0.9400224 \\ \widehat{S}_{A_1}(24) &= 0.7771517 \end{aligned}$$

$$\widehat{S}_{A_2}(t) = \frac{1}{1+\widehat{\lambda}_{A_2} \cdot t^{\widehat{\alpha}_{A_2}}}$$

$$\widehat{S}_{A_2}(t) = \frac{1}{1+e^{-7.786381} \, . \, t^{1.997497}}$$

$$\begin{aligned} \widehat{S}_{A_2}(12) &= 0.9438949 \\ \widehat{S}_{A_2}(24) &= 0.8081802 \end{aligned}$$

$$\widehat{S}_{A_3}(t) = \frac{1}{1+\widehat{\lambda}_{A_3} \cdot t^{\widehat{\alpha}_{A_3}}}$$

$$\widehat{S}_{A_3}(t) = \frac{1}{1+e^{-7.904321} \, . \, t^{1.999507}}$$

$$\begin{aligned} \widehat{S}_{A_3}(12) &= 0.9495848 \\ \widehat{S}_{A_3}(24) &= 0.8248819 \end{aligned}$$

Estimated Odds at log-logistic regression model

Shape remains constant

$$\widehat{odds}_0(t) = \frac{1-\widehat{S}_0(t)}{\widehat{S}_0(t)} = \widehat{\lambda}_0 \cdot t^{\widehat{\alpha}}$$

$$\widehat{odds}_0(t) = e^{-7.981750} \cdot t^{2.066384}$$

$$\widehat{odds}_0(12) = 0.0580194$$
$$\widehat{odds}_0(24) = 0.243006$$

$$\widehat{odds}_{A_1}(t) = \frac{1-\widehat{S}_{A_1}(t)}{\widehat{S}_{A_1}(t)} = \widehat{\lambda}_{A_1} \cdot t^{\widehat{\alpha}}$$

$$\widehat{odds}_{A_1}(t) = e^{-7.800792} \cdot t^{2.066384}$$

$$\widehat{odds}_{A_1}(12) = 0.0695284$$
$$\widehat{odds}_{A_1}(24) = 0.2912099$$

$$\widehat{odds}_{A_2}(t) = \frac{1-\widehat{S}_{A_2}(t)}{\widehat{S}_{A_2}(t)} = \widehat{\lambda}_{A_2} \cdot t^{\widehat{\alpha}}$$

$$\widehat{odds}_{A_2}(t) = e^{-8.016725} \cdot t^{2.066384}$$

$$\widehat{odds}_{A_2}(12) = 0.056025$$
$$\widehat{odds}_{A_2}(24) = 0.234654$$

$$\widehat{odds}_{A_3}(t) = \frac{1-\widehat{S}_{A_3}(t)}{\widehat{S}_{A_3}(t)} = \widehat{\lambda}_{A_3} \cdot t^{\widehat{\alpha}}$$

$$\widehat{odds}_{A_3}(t) = e^{-8.127733} \cdot t^{2.066384}$$

$$\widehat{odds}_{A_3}(12) = 0.0501388$$
$$\widehat{odds}_{A_3}(24) = 0.2099991$$

Shape alters

$$\widehat{odds}_0(t) = \frac{1-\widehat{S}_0(t)}{\widehat{S}_0(t)} = \widehat{\lambda}_0 \cdot t^{\widehat{\alpha}_0}$$

$$\widehat{odds}_0(t) = e^{-7.943357} \cdot t^{2.059777}$$

$$\widehat{odds}_0(12) = 0.0593084$$
$$\widehat{odds}_0(24) = 0.2472697$$

$$\widehat{odds}_{A_1}(t) = \frac{1-\widehat{S}_{A_1}(t)}{\widehat{S}_{A_1}(t)} = \widehat{\lambda}_{A_1} \cdot t^{\widehat{\alpha}_{A_1}}$$

$$\widehat{odds}_{A_1}(t) = e^{-8.139369} \cdot t^{2.168064}$$

$$\widehat{odds}_{A_1}(12) = 0.0638045$$
$$\widehat{odds}_{A_1}(24) = 0.2867500$$

$$\widehat{odds}_{A_2}(t) = \frac{1-\widehat{S}_{A_2}(t)}{\widehat{S}_{A_2}(t)} = \widehat{\lambda}_{A_2} \cdot t^{\widehat{\alpha}_{A_2}}$$

$$\widehat{odds}_{A_2}(t) = e^{-7.786381} \cdot t^{1.997497}$$

$$\widehat{odds}_{A_2}(12) = 0.05944$$
$$\widehat{odds}_{A_2}(24) = 0.237348$$

$$\widehat{odds}_{A_3}(t) = \frac{1-\widehat{S}_{A_3}(t)}{\widehat{S}_{A_3}(t)} = \widehat{\lambda}_{A_3} \cdot t^{\widehat{\alpha}_{A_3}}$$

$$\widehat{odds}_{A_3}(t) = e^{-7.904321} \cdot t^{1.999507}$$

$$\widehat{odds}_{A_3}(12) = 0.0530918$$
$$\widehat{odds}_{A_3}(24) = 0.2122948$$

Estimated Odds Ratio at log-logistic regression model

Shape remains constant

$$\widehat{oddsratio}_{A_1}(t) = \frac{\widehat{odds}_{A_1}(t)}{\widehat{odds}_0(t)} = \frac{\widehat{\lambda}_{A_1} \cdot t^{\widehat{\alpha}}}{\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}}}$$

$$\begin{aligned}\widehat{oddsratio}_{A_1}(12) &= \frac{0.0695284}{0.0580194} \\ &= 1.198365 \\ \widehat{oddsratio}_{A_1}(24) &= \frac{0.2912099}{0.243006} \\ &= 1.198365\end{aligned}$$

$$\widehat{oddsratio}_{A_2}(t) = \frac{\widehat{odds}_{A_2}(t)}{\widehat{odds}_0(t)} = \frac{\widehat{\lambda}_{A_2} \cdot t^{\widehat{\alpha}}}{\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}}}$$

$$\begin{aligned}\widehat{oddsratio}_{A_2}(12) &= \frac{0.0560253}{0.0580194} \\ &= 0.965629 \\ \widehat{oddsratio}_{A_2}(24) &= \frac{0.2346538}{0.243006} \\ &= 0.965629\end{aligned}$$

$$\widehat{oddsratio}_{A_3}(t) = \frac{\widehat{odds}_{A_3}(t)}{\widehat{odds}_0(t)} = \frac{\widehat{\lambda}_{A_3} \cdot t^{\widehat{\alpha}}}{\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}}}$$

$$\begin{aligned}\widehat{oddsratio}_{A_3}(12) &= \frac{0.0501388}{0.0580194} \\ &= 0.864172 \\ \widehat{oddsratio}_{A_3}(24) &= \frac{0.2099991}{0.243006} \\ &= 0.864172\end{aligned}$$

Odds ratio of an age group is constant over time

Odds ratio is called an index

One set of indices, irrespective of time

Shape alters

$$\widehat{oddsratio}_{A_1}(t) = \frac{\widehat{odds}_{A_1}(t)}{\widehat{odds}_0(t)} = \frac{\widehat{\lambda}_{A_1} \cdot t^{\widehat{\alpha}_{A_1}}}{\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}_0}}$$

$$\begin{aligned}\widehat{oddsratio}_{A_1}(12) &= \frac{0.0638045}{0.0593084} \\ &= 1.075808 \\ \widehat{oddsratio}_{A_1}(24) &= \frac{0.286750}{0.2472697} \\ &= 1.159665\end{aligned}$$

$$\widehat{oddsratio}_{A_2}(t) = \frac{\widehat{odds}_{A_2}(t)}{\widehat{odds}_0(t)} = \frac{\widehat{\lambda}_{A_2} \cdot t^{\widehat{\alpha}_{A_2}}}{\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}_0}}$$

$$\begin{aligned}\widehat{oddsratio}_{A_2}(12) &= \frac{0.05944}{0.0593084} \\ &= 1.002219 \\ \widehat{oddsratio}_{A_2}(24) &= \frac{0.237348}{0.2472697} \\ &= 0.959874\end{aligned}$$

$$\widehat{oddsratio}_{A_3}(t) = \frac{\widehat{odds}_{A_3}(t)}{\widehat{odds}_0(t)} = \frac{\widehat{\lambda}_{A_3} \cdot t^{\widehat{\alpha}_{A_3}}}{\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}_0}}$$

$$\begin{aligned}\widehat{oddsratio}_{A_3}(12) &= \frac{0.0530918}{0.0593084} \\ &= 0.895182 \\ \widehat{oddsratio}_{A_3}(24) &= \frac{0.2122948}{0.2472697} \\ &= 0.8585556\end{aligned}$$

Odds ratio of an age group depends on time

Odds ratio is called an index

Two sets of indices, one for t=12 and one for t=24

Estimated Hazard Ratio at log-logistic regression model

Shape remains constant

$$\widehat{hazardratio}_{A_1}(t) = \frac{\widehat{h}_{A_1}(t)}{\widehat{h}_0(t)} = \frac{\widehat{\lambda}_{A_1}}{\widehat{\lambda}_0} \cdot \frac{(1+\widehat{\lambda}_0 t^{\widehat{\alpha}})}{(1+\widehat{\lambda}_{A_1} t^{\widehat{\alpha}})}$$

$$\begin{aligned} \widehat{hazardratio}_{A_1}(12) &= \tfrac{0.0111944}{0.009443} \\ &= 1.1854696 \\ \widehat{hazardratio}_{A_1}(24) &= \tfrac{0.0194182}{0.0168323} \\ &= 1.1536273 \end{aligned}$$

$$\widehat{hazardratio}_{A_2}(t) = \frac{\widehat{h}_{A_2}(t)}{\widehat{h}_0(t)} = \frac{\widehat{\lambda}_{A_2}}{\widehat{\lambda}_0} \cdot \frac{(1+\widehat{\lambda}_0 t^{\widehat{\alpha}})}{(1+\widehat{\lambda}_{A_2} t^{\widehat{\alpha}})}$$

$$\begin{aligned} \widehat{hazardratio}_{A_2}(12) &= \tfrac{0.0087902}{0.009443} \\ &= 0.967453 \\ \widehat{hazardratio}_{A_2}(24) &= \tfrac{0.0157604}{0.0168323} \\ &= 0.9721618 \end{aligned}$$

$$\widehat{hazardratio}_{A_3}(t) = \frac{\widehat{h}_{A_3}(t)}{\widehat{h}_0(t)} = \frac{\widehat{\lambda}_{A_3}}{\widehat{\lambda}_0} \cdot \frac{(1+\widehat{\lambda}_0 t^{\widehat{\alpha}})}{(1+\widehat{\lambda}_{A_3} t^{\widehat{\alpha}})}$$

$$\begin{aligned} \widehat{hazardratio}_{A_3}(12) &= \tfrac{0.0080242}{0.009443} \\ &= 0.8706574 \\ \widehat{hazardratio}_{A_3}(24) &= \tfrac{0.014387}{0.0168323} \\ &= 0.8877456 \end{aligned}$$

Hazard ratio of an age group depends on time

Hazard ratio is called a risk score

Two sets of risk scores, one for t=12 and one for t=24

Shape alters

$$\widehat{hazardratio}_{A_1}(t) = \frac{\widehat{h}_{A_1}(t)}{\widehat{h}_0(t)} = \frac{\widehat{\lambda}_{A_1}}{\widehat{\lambda}_0} \cdot \frac{(1+\widehat{\lambda}_0 t^{\widehat{\alpha}_0})}{(1+\widehat{\lambda}_{A_1} t^{\widehat{\alpha}_{A_1}})}$$

$$\begin{aligned} \widehat{hazardratio}_{A_1}(12) &= \tfrac{0.0108363}{0.0096102} \\ &= 1.1275801 \\ \widehat{hazardratio}_{A_1}(24) &= \tfrac{0.0201312}{0.0170145} \\ &= 1.1831797 \end{aligned}$$

$$\widehat{hazardratio}_{A_2}(t) = \frac{\widehat{h}_{A_2}(t)}{\widehat{h}_0(t)} = \frac{\widehat{\lambda}_{A_2}}{\widehat{\lambda}_0} \cdot \frac{(1+\widehat{\lambda}_0 t^{\widehat{\alpha}_0})}{(1+\widehat{\lambda}_{A_2} t^{\widehat{\alpha}_{A_2}})}$$

$$\begin{aligned} \widehat{hazardratio}_{A_2}(12) &= \tfrac{0.0093391}{0.0096102} \\ &= 0.9717948 \\ \widehat{hazardratio}_{A_2}(24) &= \tfrac{0.015965}{0.0170145} \\ &= 0.9383156 \end{aligned}$$

$$\widehat{hazardratio}_{A_3}(t) = \frac{\widehat{h}_{A_3}(t)}{\widehat{h}_0(t)} = \frac{\widehat{\lambda}_{A_3}}{\widehat{\lambda}_0} \cdot \frac{(1+\widehat{\lambda}_0 t^{\widehat{\alpha}_0})}{(1+\widehat{\lambda}_{A_3} t^{\widehat{\alpha}_{A_3}})}$$

$$\begin{aligned} \widehat{hazardratio}_{A_3}(12) &= \tfrac{0.0084005}{0.0096102} \\ &= 0.8741186 \\ \widehat{hazardratio}_{A_3}(24) &= \tfrac{0.0145896}{0.0170145} \\ &= 0.8574789 \end{aligned}$$

Hazard ratio of an age group depends on time

Hazard ratio is called a risk score

Two sets of risk scores, one for t=12 and one for t=24

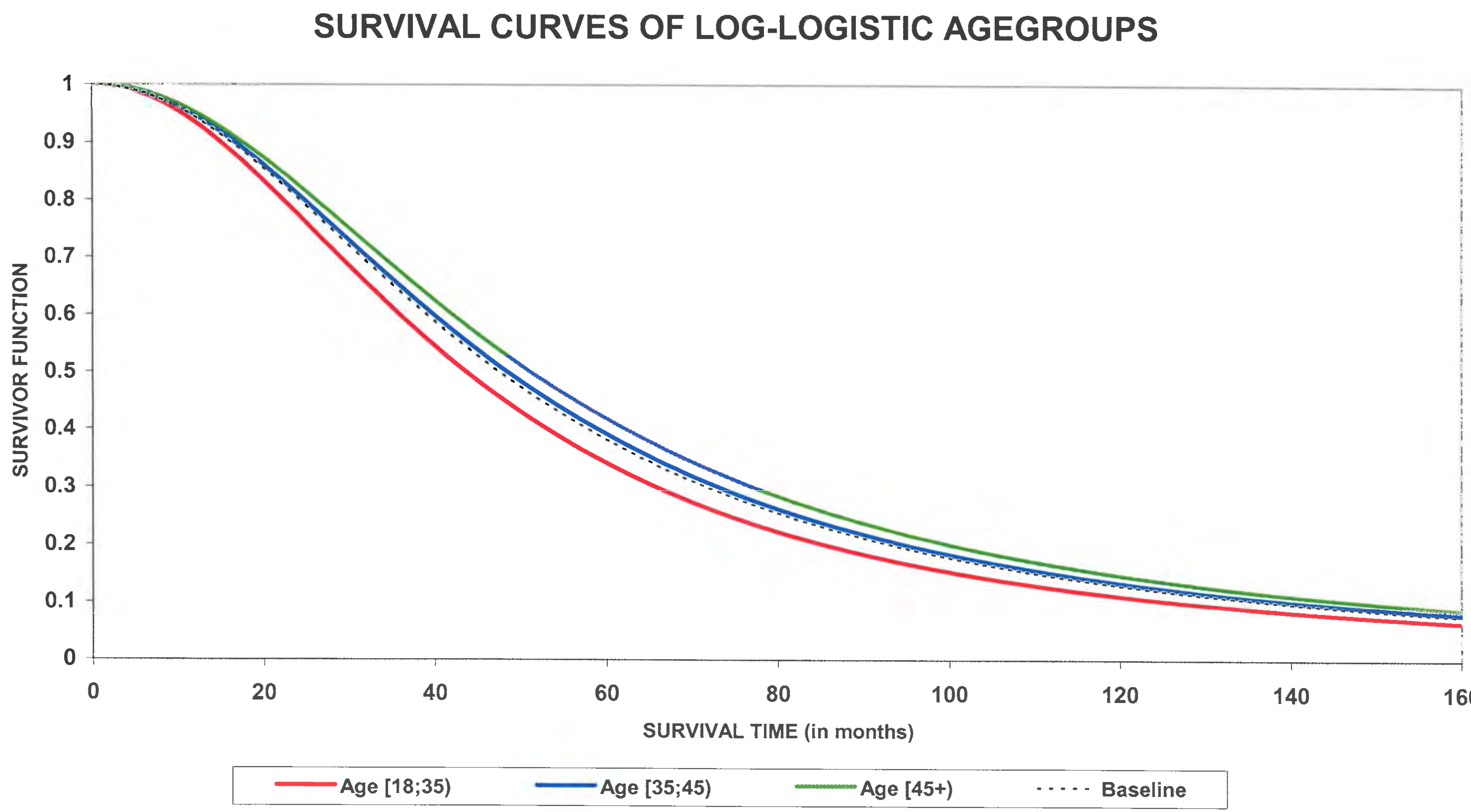

Figure 5.12: **Survival curves of fitted log-logistic age group models**

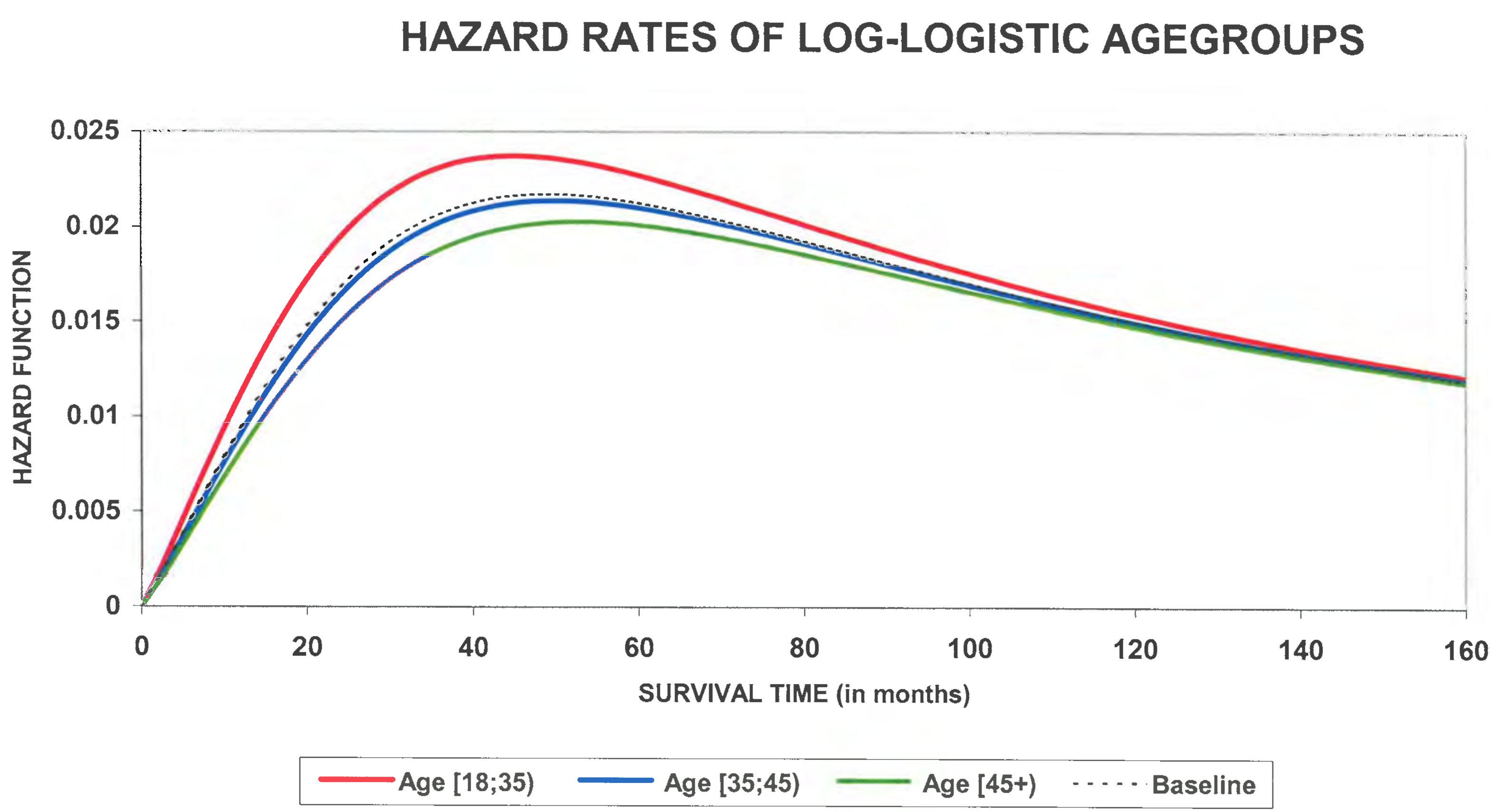

Figure 5.13: **Graphs of hazard rates of fitted log-logistic age group models**

For a **constant shape parameter** in the log-logistic distributions, the indices (estimated odds ratios) may be obtained also from the exponent of the $\widehat{\beta}$-values in the log-logistic regression model, for example

$$\begin{aligned} e^{\widehat{\beta}_{A_1}} &= e^{0.180958} &= 1.198365 \\ e^{\widehat{\beta}_{A_2}} &= e^{-0.034975} &= 0.965629 \\ e^{\widehat{\beta}_{A_3}} &= e^{-0.145983} &= 0.864172. \end{aligned}$$

The indices of the three age groups, estimated from the log-logistic regression model, are compared to the indices, obtained from the logit model, in Table 5.13.

Table 5.13: **Comparison of indices: log-logistic regression model and logit model**

		Log-logistic regression model				Logit model	
		Shape remains constant		Shape alters			
Effect	n	Index		Index		Index	
		t=12	t=24	t=12	t=24	t=12	t=24
Baseline odds	10077	0.058019	0.243006	0.059308	0.247270	0.0543	0.2570
Age [18;35)	3644	1.198365	1.198365	1.075808	1.159665	1.1321	1.1370
Age [35;45)	3425	0.965629	0.965629	1.002219	0.959874	0.9679	0.9817
Age [45+)	3008	0.864172	0.864172	0.898556	0.858556	0.9126	0.8960

The index of age group [18;35) of 1.198365 shows the effect of this age group on the baseline odds of a lapse. This effect is multiplicative on the baseline odds of a lapse. Thus the effect of age group [18;35) is to increase the baseline odds of a lapse by a factor 1.198365 .

From Table 5.13 follows that one log-logistic regression model provides indices for any time value, while a new logit model has to be built for a fixed time value, say t=12 months, conditional on a restricted experimental design where all the policies must have an exposure of at least one year when investigating the lapses of policies in the first year. There is no such restrictions in the more general experimental design for the log-logistic regression model where all the policies can be used in the analysis, even those policies with inception dates very close to the cut-off point.

Predicted indices from the log-logistic regression model, for varying time values, are shown in Table 5.14 (constant shape) and in Table 5.15 (shape alters).

Predicted risk scores from the log-logistic regression model, for varying time values, are shown in Table 5.16 (constant shape) and in Table 5.17 (shape alters).

Table 5.13: **Predicted indices from log-logistic regression model (constant shape)**

	Predicted indices from log-logistic regression model									
Effect	t=6	t=12	t=18	t=24	t=30	t=36	t=42	t=48	t=54	t=60
Baseline odds	0.013853	0.058019	0.134105	0.243006	0.385363	0.561680	0.772373	1.017796	1.298259	1.614039
Age [18;35)	1.198365	1.198365	1.198365	1.198365	1.198365	1.198365	1.198365	1.198365	1.198365	1.198365
Age [35;45)	0.965629	0.965629	0.965629	0.965629	0.965629	0.965629	0.965629	0.965629	0.965629	0.965629
Age [45+)	0.864172	0.864172	0.864172	0.864172	0.864172	0.864172	0.864172	0.864172	0.864172	0.864172

Table 5.14: **Predicted indices from log-logistic regression model (shape alters)**

	Predicted indices from log-logistic regression model									
Effect	t=6	t=12	t=18	t=24	t=30	t=36	t=42	t=48	t=54	t=60
Baseline odds	0.014225	0.059308	0.136718	0.247270	0.391547	0.570006	0.783034	1.030921	1.313978	1.632444
Age [18;35)	0.998015	1.075808	1.124096	1.159665	1.188028	1.211716	1.232113	1.250058	1.266104	1.280632
Age [35;45)	1.046431	1.002219	0.977227	0.959874	0.946627	0.935939	0.926996	0.919319	0.912600	0.906631
Age [45+)	0.933371	0.898556	0.873571	0.858556	0.847086	0.837829	0.830081	0.823428	0.817603	0.812428

Table 5.15: **Predicted risk scores from log-logistic regression model (constant shape)**

Effect	Predicted risk scores from log-logistic regression model									
	t=6	t=12	t=18	t=24	t=30	t=36	t=42	t=48	t=54	t=60
Baseline hazard rate	0.004706	0.009443	0.013575	0.016832	0.01916	0.020645	0.021440	0.021715	0.021616	0.021265
Age [18;35)	1.195126	1.185470	1.170900	1.153627	1.135699	1.118561	1.103015	1.089366	1.077614	1.067604
Age [35;45)	0.966083	0.967453	0.969570	0.972162	0.974951	0.977716	0.980313	0.982666	0.984749	0.986566
Age [45+)	0.865779	0.870657	0.878279	0.887746	0.898105	0.908557	0.918542	0.927734	0.935988	0.943282

Table 5.16: **Predicted risk scores from log-logistic regression model (shape alters)**

Effect	Predicted risk scores from log-logistic regression model									
	t=6	t=12	t=18	t=24	t=30	t=36	t=42	t=48	t=54	t=60
Baseline hazard rate	0.004815	0.009610	0.013763	0.017014	0.019319	0.020773	0.021537	0.021783	0.021660	0.021289
Age [18;35)	1.050512	1.127580	1.165792	1.183180	1.187651	1.184381	1.176920	1.167574	1.157727	1.148149
Age [35;45)	1.014131	0.971795	0.950282	0.938316	0.932001	0.929252	0.928743	0.929593	0.932223	0.933254
Age [45+)	0.906908	0.874119	0.861105	0.857479	0.859272	0.864196	0.870770	0.878032	0.885382	0.892466

For a **constant shape parameter** in the log-logistic distributions, the pth percentile of the baseline lifetime distribution can be calculated from

$$\left(\frac{1}{\widehat{\lambda}_0}\cdot\frac{p}{100-p}\right)^{\dfrac{1}{\widehat{\alpha}}}$$

and that of the age group distributions from

$$\left(\frac{1}{\widehat{\lambda}_{A_i}}\cdot\frac{p}{100-p}\right)^{\dfrac{1}{\widehat{\alpha}}} \qquad i=1,2,3.$$

Note that the latter is equal to the pth percentile of the baseline distribution multiplied by the specific index to the power $-\frac{1}{\widehat{\alpha}}$.

For a **shape parameter that alters**, the formulaes change to

$$\left(\frac{1}{\widehat{\lambda}_0}\cdot\frac{p}{100-p}\right)^{\dfrac{1}{\widehat{\alpha}_0}} \quad \text{and} \quad \left(\frac{1}{\widehat{\lambda}_{A_i}}\cdot\frac{p}{100-p}\right)^{\dfrac{1}{\widehat{\alpha}_{A_i}}} \qquad i=1,2,3.$$

The estimated percentiles of the baseline and the age group log-logistic distributions for a constant shape parameter as well as for different shape parameters are reported in Table 5.18.

Table 5.18: **Lifetime percentiles estimated from log-logistic regression model**

	Log-logistic regression model							
	Lifetime distribution: constant shape				Lifetime distribution: shape alters			
Lifetime percentile	Baseline	Age group [18;35)	Age group [35;45)	Age group [45+)	Baseline	Age group [18;35)	Age group [35;45)	Age group [45+)
P5	11.45	10.49	11.64	12.28	11.32	10.98	11.29	11.95
P10	16.43	15.05	16.71	17.64	16.27	15.50	16.41	17.36
P20	24.33	22.29	24.75	26.11	24.13	22.53	24.63	26.04
P25	27.97	25.62	28.44	30.01	27.74	25.73	28.45	30.07
P30	31.58	28.93	32.12	33.89	31.34	28.89	32.26	34.10
P40	39.11	35.83	39.78	41.97	38.84	35.42	40.25	42.54
P50	47.59	43.60	48.40	51.07	47.29	42.70	49.31	52.10
P60	57.91	53.05	58.90	62.15	57.58	51.48	60.40	63.81
P70	71.71	65.70	72.94	76.96	71.36	63.12	75.36	79.59
P75	80.99	74.20	82.37	86.92	80.62	70.88	85.46	90.25
P80	93.09	85.28	94.68	99.90	92.71	80.93	98.70	104.21
P90	137.82	126.27	140.18	147.91	137.43	117.64	148.12	156.34
P95	197.87	181.27	201.24	212.35	197.53	166.05	215.32	227.17

The median time to a lapse of a policy, over all three age groups, is 47.59 months. The baseline odds of a lapse at 47.59 months is 1, that means that $P(T >$47.59 months$)=P(T <$47.59 months).

5.3.4 Deriving of indices and risk scores from Weibull regression model

Once the parameters of the Weibull baseline distribution and Weibull age group distributions have been estimated, estimated hazard and survivor functions, odds of a lapse, odds ratios and hazard ratios at time t can be calculated.

The odds ratio for age group [18;35) is the relative odds of a lapse at time t of a policy, with the age of the policyholder in [18;35), compared to a policy with the baseline characteristics. The odds ratios for the three age groups result in a set of indices, showing the effect of each age group on the baseline odds of a lapse at time t.

The hazard ratio for age group [18;35) is the relative hazard rate of a lapse at time t of a policy, with the age of the policyholder in [18;35), compared to a policy with the baseline characteristics. The hazard ratios for the three age groups result in a set of risk scores, showing the effect of each age group on the baseline hazard rate of a lapse at time t.

Percentiles of the four Weibull survival distributions can also be estimated.

The calculations of estimated hazard and survivor functions, odds of a lapse, odds ratios and hazard ratios are illustrated on the following five pages.

The survival curves and the graphs of the hazard rates of the fitted Weibull age group models are shown in Figure 5.14 and Figure 5.15.

Estimated Hazard Function at Weibull regression model

Shape remains constant

$$\boxed{\widehat{h}_0(t) = \widehat{\lambda}_0 \cdot \widehat{\alpha} \cdot t^{\widehat{\alpha}-1}}$$

$$\widehat{h}_0(t) = e^{-7.404312} \cdot 1.842334 \cdot t^{1.842334-1}$$

$$\begin{aligned}\widehat{h}_0(12) &= 0.0090938\\ \widehat{h}_0(24) &= 0.0163048\end{aligned}$$

$$\boxed{\widehat{h}_{A_1}(t) = \widehat{\lambda}_{A_1} \cdot \widehat{\alpha} \cdot t^{\widehat{\alpha}-1}}$$

$$\widehat{h}_{A_1}(t) = e^{-7.245223} \cdot 1.842334 \cdot t^{1.842334-1}$$

$$\begin{aligned}\widehat{h}_{A_1}(12) &= 0.010662\\ \widehat{h}_{A_1}(24) &= 0.0191164\end{aligned}$$

$$\boxed{\widehat{h}_{A_2}(t) = \widehat{\lambda}_{A_2} \cdot \widehat{\alpha} \cdot t^{\widehat{\alpha}-1}}$$

$$\widehat{h}_{A_2}(t) = e^{-7.438269} \cdot 1.842334 \cdot t^{1.842334-1}$$

$$\begin{aligned}\widehat{h}_{A_2}(12) &= 0.0087902\\ \widehat{h}_{A_2}(24) &= 0.0157604\end{aligned}$$

$$\boxed{\widehat{h}_{A_3}(t) = \widehat{\lambda}_{A_3} \cdot \widehat{\alpha} \cdot t^{\widehat{\alpha}-1}}$$

$$\widehat{h}_{A_3}(t) = e^{-7.529445} \cdot 1.842334 \cdot t^{1.842334-1}$$

$$\begin{aligned}\widehat{h}_{A_3}(12) &= 0.0080242\\ \widehat{h}_{A_3}(24) &= 0.014387\end{aligned}$$

Shape alters

$$\boxed{\widehat{h}_0(t) = \widehat{\lambda}_0 \cdot \widehat{\alpha}_0 t^{\widehat{\alpha}_0-1}}$$

$$\widehat{h}_0(t) = e^{-7.381423} \cdot 1.838073 \cdot t^{1.838073-1}$$

$$\begin{aligned}\widehat{h}_0(12) &= 0.0091851\\ \widehat{h}_0(24) &= 0.0164199\end{aligned}$$

$$\boxed{\widehat{h}_{A_1}(t) = \widehat{\lambda}_{A_1} \cdot \widehat{\alpha}_{A_1} \cdot t^{\widehat{\alpha}_{A_1}-1}}$$

$$\widehat{h}_{A_1}(t) = e^{-7.456598} \cdot 1.904217 \cdot t^{1.904217-1}$$

$$\begin{aligned}\widehat{h}_{A_1}(12) &= 0.0104033\\ \widehat{h}_{A_1}(24) &= 0.0194701\end{aligned}$$

$$\boxed{\widehat{h}_{A_2}(t) = \widehat{\lambda}_{A_2} \cdot \widehat{\alpha}_{A_2} \cdot t^{\widehat{\alpha}_{A_2}-1}}$$

$$\widehat{h}_{A_2}(t) = e^{-7.438269} \cdot 1.842334 \cdot t^{1.842334-1}$$

$$\begin{aligned}\widehat{h}_{A_2}(12) &= 0.0089654\\ \widehat{h}_{A_2}(24) &= 0.0155084\end{aligned}$$

$$\boxed{\widehat{h}_{A_3}(t) = \widehat{\lambda}_{A_3} \cdot \widehat{\alpha}_{A_3} \cdot t^{\widehat{\alpha}_{A_3}-1}}$$

$$\widehat{h}_{A_3}(t) = e^{-7.426139} \cdot 1.811986 \cdot t^{1.811986-1}$$

$$\begin{aligned}\widehat{h}_{A_3}(12) &= 0.0081153\\ \widehat{h}_{A_3}(24) &= 0.0142474\end{aligned}$$

Estimated Survivor Function at Weibull regression model

Shape remains constant

$$\widehat{S}_0(t) = \exp(-\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}})$$

$$\widehat{S}_0(t) = \exp(-e^{-7.404312} \cdot t^{1.842334})$$

$$\widehat{S}_0(12) = 0.9424876$$
$$\widehat{S}_0(24) = 0.8086397$$

$$\widehat{S}_{A_1}(t) = \exp(-\widehat{\lambda}_{A_1} \cdot t^{\widehat{\alpha}})$$

$$\widehat{S}_{A_1}(t) = \exp(-e^{-7.245223} \cdot t^{1.842334})$$

$$\widehat{S}_{A_1}(12) = 0.9329098$$
$$\widehat{S}_{A_1}(24) = 0.7795573$$

$$\widehat{S}_{A_2}(t) = \exp(-\widehat{\lambda}_{A_2} \cdot t^{\widehat{\alpha}})$$

$$\widehat{S}_{A_2}(t) = \exp(-e^{-7.438269} \cdot t^{1.842334})$$

$$\widehat{S}_{A_2}(12) = 0.9443533$$
$$\widehat{S}_{A_2}(24) = 0.8143945$$

$$\widehat{S}_{A_3}(t) = \exp(-\widehat{\lambda}_{A_3} \cdot t^{\widehat{\alpha}})$$

$$\widehat{S}_{A_3}(t) = \exp(-e^{-7.529445} \cdot t^{1.842334})$$

$$\widehat{S}_{A_3}(12) = 0.9490768$$
$$\widehat{S}_{A_3}(24) = 0.8290963$$

Shape alters

$$\widehat{S}_0(t) = \exp(-\widehat{\lambda}_0 \cdot t^{\widehat{\alpha_0}})$$

$$\widehat{S}_0(t) = \exp(-e^{-7.381423} \cdot t^{1.8380729})$$

$$\widehat{S}_0(12) = 0.9417969$$
$$\widehat{S}_0(24) = 0.8070283$$

$$\widehat{S}_{A_1}(t) = \exp(-\widehat{\lambda}_{A_1} \cdot t^{\widehat{\alpha_{A_1}}})$$

$$\widehat{S}_{A_1}(t) = \exp(-e^{-7.456598} \cdot t^{1.904217})$$

$$\widehat{S}_{A_1}(12) = 0.9365433$$
$$\widehat{S}_{A_1}(24) = 0.7823969$$

$$\widehat{S}_{A_2}(t) = \exp(-\widehat{\lambda}_{A_2} \cdot t^{\widehat{\alpha_{A_2}}})$$

$$\widehat{S}_{A_2}(t) = \exp(-e^{-7.261531} \cdot t^{1.790610})$$

$$\widehat{S}_{A_2}(12) = 0.9416866$$
$$\widehat{S}_{A_2}(24) = 0.8123183$$

$$\widehat{S}_{A_3}(t) = \exp(-\widehat{\lambda}_{A_3} \cdot t^{\widehat{\alpha_{A_3}}})$$

$$\widehat{S}_{A_3}(t) = \exp(-e^{-7.426139} \cdot t^{1.811986})$$

$$\widehat{S}_{A_3}(12) = 0.9476747$$
$$\widehat{S}_{A_3}(24) = 0.8280275$$

Estimated Odds at Weibull regression model

Shape remains constant

$$\widehat{odds}_0(t) = \frac{1-\widehat{S}_0(t)}{\widehat{S}_0(t)} = \exp(\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}-1})$$

$$\widehat{odds}_0(t) = \exp(e^{-7.404312} \cdot t^{1.842334-1})$$

$$\widehat{odds}_0(12) = 0.0610219$$
$$\widehat{odds}_0(24) = 0.2366447$$

$$\widehat{odds}_{A_1}(t) = \frac{1-\widehat{S}_{A_1}(t)}{\widehat{S}_{A_1}(t)} = \exp(\widehat{\lambda}_{A_1} \cdot t^{\widehat{\alpha}-1})$$

$$\widehat{odds}_{A_1}(t) = \exp(e^{-7.245223} \cdot t^{1.842334-1})$$

$$\widehat{odds}_{A_1}(12) = 0.071915$$
$$\widehat{odds}_{A_1}(24) = 0.2827793$$

$$\widehat{odds}_{A_2}(t) = \frac{1-\widehat{S}_{A_2}(t)}{\widehat{S}_{A_2}(t)} = \exp(\widehat{\lambda}_{A_2} \cdot t^{\widehat{\alpha}-1})$$

$$\widehat{odds}_{A_2}(t) = \exp(e^{-7.438269} \cdot t^{1.842334-1})$$

$$\widehat{odds}_{A_2}(12) = 0.0589258$$
$$\widehat{odds}_{A_2}(24) = 0.2279062$$

$$\widehat{odds}_{A_3}(t) = \frac{1-\widehat{S}_{A_3}(t)}{\widehat{S}_{A_3}(t)} = \exp(\widehat{\lambda}_{A_3} \cdot t^{\widehat{\alpha}-1})$$

$$\widehat{odds}_{A_3}(t) = \exp(e^{-7.529445} \cdot t^{1.842334-1})$$

$$\widehat{odds}_{A_3}(12) = 0.0536555$$
$$\widehat{odds}_{A_3}(24) = 0.2061326$$

Shape alters

$$\widehat{odds}_0(t) = \frac{1-\widehat{S}_0(t)}{\widehat{S}_0(t)} = \exp(\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}_0-1})$$

$$\widehat{odds}_0(t) = \exp(e^{-7.381423} \cdot t^{1.8380729-1})$$

$$\widehat{odds}_0(12) = 0.0618001$$
$$\widehat{odds}_0(24) = 0.2391139$$

$$\widehat{odds}_{A_1}(t) = \frac{1-\widehat{S}_{A_1}(t)}{\widehat{S}_{A_1}(t)} = \exp(\widehat{\lambda}_{A_1} \cdot t^{\widehat{\alpha}_{A_1}-1})$$

$$\widehat{odds}_{A_1}(t) = \exp(e^{-7.456598} \cdot t^{1.904217-1})$$

$$\widehat{odds}_{A_1}(12) = 0.0677563$$
$$\widehat{odds}_{A_1}(24) = 0.2781237$$

$$\widehat{odds}_{A_2}(t) = \frac{1-\widehat{S}_{A_2}(t)}{\widehat{S}_{A_2}(t)} = \exp(\widehat{\lambda}_{A_2} \cdot t^{\widehat{\alpha}_{A_2}-1})$$

$$\widehat{odds}_{A_2}(t) = \exp(e^{-7.261531} \cdot t^{1.790610-1})$$

$$\widehat{odds}_{A_2}(12) = 0.0619244$$
$$\widehat{odds}_{A_2}(24) = 0.2310445$$

$$\widehat{odds}_{A_3}(t) = \frac{1-\widehat{S}_{A_3}(t)}{\widehat{S}_{A_3}(t)} = \exp(\widehat{\lambda}_{A_3} \cdot t^{\widehat{\alpha}_{A_3}-1})$$

$$\widehat{odds}_{A_3}(t) = \exp(e^{-7.426139} \cdot t^{1.811986-1})$$

$$\widehat{odds}_{A_3}(12) = 0.0552145$$
$$\widehat{odds}_{A_3}(24) = 0.2076894$$

Estimated Odds Ratio at Weibull regression model

Shape remains constant

$$\widehat{oddsratio}_{A_1}(t) = \frac{\widehat{odds}_{A_1}(t)}{\widehat{odds}_0(t)} = \frac{\exp(\widehat{\lambda}_{A_1} \cdot t^{\widehat{\alpha}-1})}{\exp(\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}-1})}$$

$$\begin{aligned} \widehat{oddsratio}_{A_1}(12) &= \tfrac{0.071915}{0.0610219} \\ &= 1.1785109 \\ \widehat{oddsratio}_{A_1}(24) &= \tfrac{0.2827793}{0.2366447} \\ &= 1.1949532 \end{aligned}$$

$$\widehat{oddsratio}_{A_2}(t) = \frac{\widehat{odds}_{A_2}(t)}{\widehat{odds}_0(t)} = \frac{\exp(\widehat{\lambda}_{A_2} \cdot t^{\widehat{\alpha}-1})}{\exp(\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}-1})}$$

$$\begin{aligned} \widehat{oddsratio}_{A_2}(12) &= \tfrac{0.0589258}{0.0610219} \\ &= 0.9656487 \\ \widehat{oddsratio}_{A_2}(24) &= \tfrac{0.2279062}{0.2366447} \\ &= 0.9630732 \end{aligned}$$

$$\widehat{oddsratio}_{A_3}(t) = \frac{\widehat{odds}_{A_3}(t)}{\widehat{odds}_0(t)} = \frac{\exp(\widehat{\lambda}_{A_3} \cdot t^{\widehat{\alpha}-1})}{\exp(\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}-1})}$$

$$\begin{aligned} \widehat{oddsratio}_{A_3}(12) &= \tfrac{0.0536555}{0.0610219} \\ &= 0.8792827 \\ \widehat{oddsratio}_{A_3}(24) &= \tfrac{0.2061326}{0.2366447} \\ &= 0.8710636 \end{aligned}$$

Odds ratio of an age group depends on time

Odds ratio is called an index

Two sets of indices, one for t=12 and one for t=24

Shape alters

$$\widehat{oddsratio}_{A_1}(t) = \frac{\widehat{odds}_{A_1}(t)}{\widehat{odds}_0(t)} = \frac{\exp(\widehat{\lambda}_{A_1} \cdot t^{\widehat{\alpha}_{A_1}-1})}{\exp(\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}_0-1})}$$

$$\begin{aligned} \widehat{oddsratio}_{A_1}(12) &= \tfrac{0.0677563}{0.0618001} \\ &= 1.0963792 \\ \widehat{oddsratio}_{A_1}(24) &= \tfrac{0.2781237}{0.2391139} \\ &= 1.1631435 \end{aligned}$$

$$\widehat{oddsratio}_{A_2}(t) = \frac{\widehat{odds}_{A_2}(t)}{\widehat{odds}_0(t)} = \frac{\exp(\widehat{\lambda}_{A_2} \cdot t^{\widehat{\alpha}_{A_2}-1})}{\exp(\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}_0-1})}$$

$$\begin{aligned} \widehat{oddsratio}_{A_2}(12) &= \tfrac{0.0619244}{0.0618001} \\ &= 1.0020119 \\ \widehat{oddsratio}_{A_2}(24) &= \tfrac{0.2310445}{0.2391139} \\ &= 0.9662529 \end{aligned}$$

$$\widehat{oddsratio}_{A_3}(t) = \frac{\widehat{odds}_{A_3}(t)}{\widehat{odds}_0(t)} = \frac{\exp(\widehat{\lambda}_{A_3} \cdot t^{\widehat{\alpha}_{A_3}-1})}{\exp(\widehat{\lambda}_0 \cdot t^{\widehat{\alpha}_0-1})}$$

$$\begin{aligned} \widehat{oddsratio}_{A_3}(12) &= \tfrac{0.0552145}{0.0618001} \\ &= 0.8934366 \\ \widehat{oddsratio}_{A_3}(24) &= \tfrac{0.2076894}{0.2391139} \\ &= 0.8685792 \end{aligned}$$

Odds ratio of an age group depends on time

Odds ratio is called an index

Two sets of indices, one for t=12 and one for t=24

Estimated Hazard Ratio at Weibull regression model

Shape remains constant

$$\widehat{hazardratio}_{A_1}(t) = \frac{\widehat{h}_{A_1}(t)}{\widehat{h}_0(t)} = \frac{\widehat{\lambda}_{A_1}\widehat{\alpha}\cdot t^{\widehat{\alpha}-1}}{\widehat{\lambda}_0\widehat{\alpha}\cdot t^{\widehat{\alpha}-1}}$$

$$\begin{aligned}\widehat{hazardratio}_{A_1}(12) &= \frac{0.010662}{0.0090938} \\ &= 1.1724432 \\ \widehat{hazardratio}_{A_1}(24) &= \frac{0.0191164}{0.0163048} \\ &= 1.1724432\end{aligned}$$

$$\widehat{hazardratio}_{A_2}(t) = \frac{\widehat{h}_{A_2}(t)}{\widehat{h}_0(t)} = \frac{\widehat{\lambda}_{A_2}\widehat{\alpha}\cdot t^{\widehat{\alpha}-1}}{\widehat{\lambda}_0\widehat{\alpha}\cdot t^{\widehat{\alpha}-1}}$$

$$\begin{aligned}\widehat{hazardratio}_{A_2}(12) &= \frac{0.0087902}{0.0090938} \\ &= 0.9666133 \\ \widehat{hazardratio}_{A_2}(24) &= \frac{0.0157604}{0.0163048} \\ &= 0.9666133\end{aligned}$$

$$\widehat{hazardratio}_{A_3}(t) = \frac{\widehat{h}_{A_3}(t)}{\widehat{h}_0(t)} = \frac{\widehat{\lambda}_{A_3}\widehat{\alpha}\cdot t^{\widehat{\alpha}-1}}{\widehat{\lambda}_0\widehat{\alpha}\cdot t^{\widehat{\alpha}-1}}$$

$$\begin{aligned}\widehat{hazardratio}_{A_3}(12) &= \frac{0.0080242}{0.0090938} \\ &= 0.8823795 \\ \widehat{hazardratio}_{A_3}(24) &= \frac{0.014387}{0.0163048} \\ &= 0.8823795\end{aligned}$$

Hazard ratio of an age group is constant over time

Hazard ratio is called a risk score

One set of risk scores, irrespective of time

Shape alters

$$\widehat{hazardratio}_{A_1}(t) = \frac{\widehat{h}_{A_1}(t)}{\widehat{h}_0(t)} = \frac{\widehat{\lambda}_{A_1}\widehat{\alpha}_{A_1}\cdot t^{\widehat{\alpha}_{A_1}-1}}{\widehat{\lambda}_0\widehat{\alpha}_0\cdot t^{\widehat{\alpha}_0-1}}$$

$$\begin{aligned}\widehat{hazardratio}_{A_1}(12) &= \frac{0.0104033}{0.0091851} \\ &= 1.1326275 \\ \widehat{hazardratio}_{A_1}(24) &= \frac{0.0194701}{0.0164199} \\ &= 1.1857647\end{aligned}$$

$$\widehat{hazardratio}_{A_2}(t) = \frac{\widehat{h}_{A_2}(t)}{\widehat{h}_0(t)} = \frac{\widehat{\lambda}_{A_2}\widehat{\alpha}_{A_2}\cdot t^{\widehat{\alpha}_{A_2}-1}}{\widehat{\lambda}_0\widehat{\alpha}_0\cdot t^{\widehat{\alpha}_0-1}}$$

$$\begin{aligned}\widehat{hazardratio}_{A_2}(12) &= \frac{0.0089654}{0.0091851} \\ &= 0.9760801 \\ \widehat{hazardratio}_{A_2}(24) &= \frac{0.0155084}{0.0164199} \\ &= 0.9444907\end{aligned}$$

$$\widehat{hazardratio}_{A_3}(t) = \frac{\widehat{h}_{A_3}(t)}{\widehat{h}_0(t)} = \frac{\widehat{\lambda}_{A_3}\widehat{\alpha}_{A_3}\cdot t^{\widehat{\alpha}_{A_3}-1}}{\widehat{\lambda}_0\widehat{\alpha}_0\cdot t^{\widehat{\alpha}_0-1}}$$

$$\begin{aligned}\widehat{hazardratio}_{A_3}(12) &= \frac{0.0081153}{0.0091851} \\ &= 0.8835268 \\ \widehat{hazardratio}_{A_3}(24) &= \frac{0.0142474}{0.0164199} \\ &= 0.8676945\end{aligned}$$

Hazard ratio of an age group depends on time

Hazard ratio is called a risk score

Two sets of risk scores, one for t=12 and one for t=24

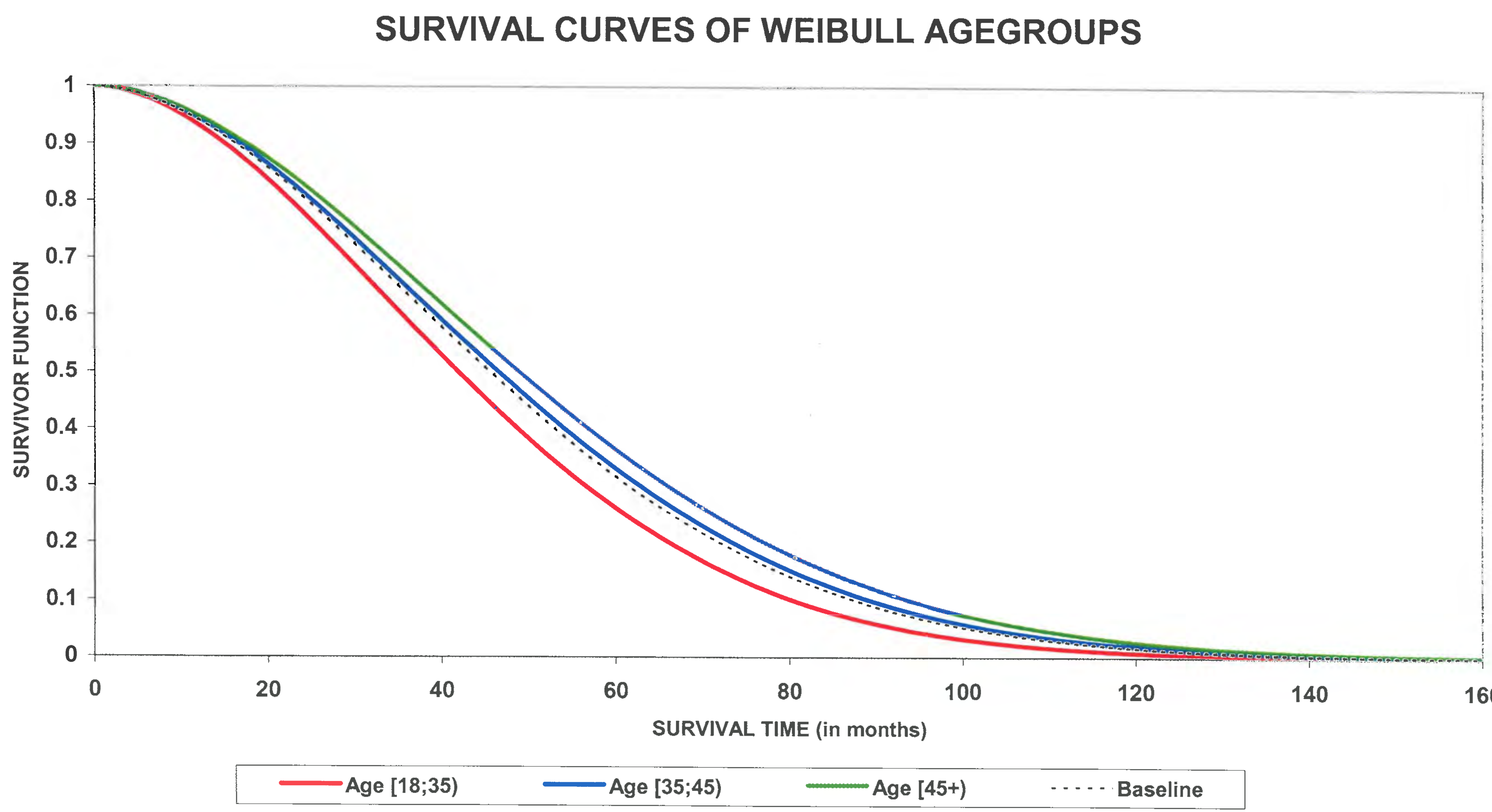

Figure 5.14: **Survival curves of fitted Weibull age group models**

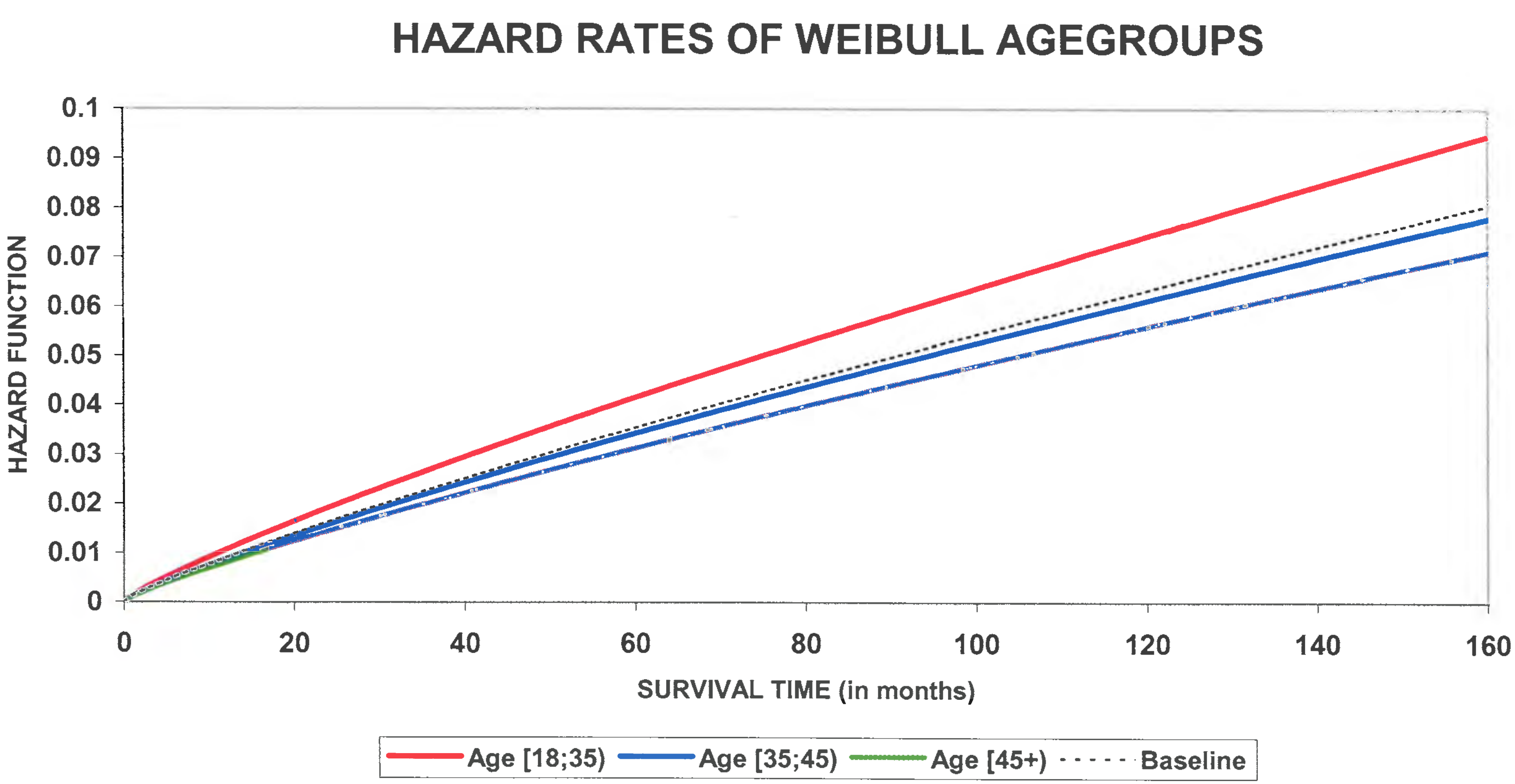

Figure 5.15: **Graphs of hazard rates of fitted Weibull age group models**

For a **constant shape parameter** in the Weibull distributions, the risk scores (estimated hazard ratios) may be obtained also from the exponent of the $\widehat{\beta}$-values in the Weibull regression model, for example

$$\begin{aligned} e^{\widehat{\beta}_{A_1}} &= e^{0.1590898} = 1.1724432 \\ e^{\widehat{\beta}_{A_2}} &= e^{-0.033957} = 0.9666133 \\ e^{\widehat{\beta}_{A_3}} &= e^{-0.125133} = 0.8823795. \end{aligned}$$

The indices (estimated odds ratios) of the three age groups, estimated from the Weibull regression model, are compared to the indices, obtained from the logit model, in Table 5.19. The index of age group [45+) of 0.879283 shows the effect of this age group on the baseline

Table 5.19: **Comparison of indices: Weibull regression model and logit model**

		Weibull regression model				Logit model	
		Shape remains constant		Shape alters			
Effect	n	Index		Index		Index	
		t=12	t=24	t=12	t=24	t=12	t=24
Baseline odds	10077	0.061022	0.236645	0.061800	0.239114	0.0543	0.2570
Age [18;35)	3644	1.178511	1.194953	1.096379	1.163143	1.1321	1.1370
Age [35;45)	3425	0.965649	0.963073	1.002012	0.966253	0.9679	0.9817
Age [45+)	3008	0.879283	0.871064	0.893437	0.868579	0.9126	0.8960

odds of a lapse. This effect is multiplicative on the baseline odds of a lapse. Thus the effect of age group [45+) is to decrease the baseline odds of a lapse by a factor 0.879283 .

From Table 5.19 follows that one Weibull regression model provides indices for any time value, while a new logitmodel has to be built for a fixed time value, say t=12 months, conditional on a restricted experimental design where all the policies must have an exposure of at least one year when investigating the lapses of policies in the first year. There is no such restrictions in the more general experimental design for the Weibull regression model where all the policies can be used in the analysis, even those policies with inception dates very close to the cut-off point.

Predicted indices from the Weibull regression model, for varying time values, are shown in Table 5.20 (constant shape) and in Table 5.21 (shape alters).

Predicted risk scores from the Weibull regression model, for varying time values, are shown in Table 5.22 (constant shape) and in Table 5.23 (shape alters).

Table 5.19: **Predicted indices from Weibull regression model (constant shape)**

Effect	Predicted indices from Weibull regression model									
	t=6	t=12	t=18	t=24	t=30	t=36	t=42	t=48	t=54	t=60
Baseline odds	0.016655	0.061022	0.133171	0.236645	0.377685	0.565662	0.814022	1.141810	1.575977	2.154842
Age [18;35)	1.174119	1.178511	1.185439	1.194953	1.207213	1.222464	1.241024	1.263282	1.289700	1.320802
Age [35;45)	0.966346	0.965649	0.964557	0.963073	0.961187	0.958881	0.956128	0.952901	0.94917	0.944903
Age [45+)	0.881521	0.879283	0.875789	0.871064	0.86509	0.857832	0.849241	0.839262	0.827842	0.814934

Table 5.20: **Predicted indices from Weibull regression model (shape alters)**

Effect	Predicted indices from Weibull regression model									
	t=6	t=12	t=18	t=24	t=30	t=36	t=42	t=48	t=54	t=60
Baseline odds	0.016913	0.061800	0.134678	0.239114	0.381412	0.571040	0.821586	1.152299	1.590420	2.174686
Age [18;35)	1.044681	1.096379	1.131960	1.163143	1.194129	1.227247	1.264219	1.306589	1.355927	1.413938
Age [35;45)	1.035773	1.002012	0.981769	0.966253	0.952697	0.939827	0.926922	0.913510	0.899255	0.883902
Age [45+)	0.911930	0.893437	0.880373	0.868579	0.856722	0.844171	0.830554	0.815621	0.799192	0.781140

Table 5.21: **Predicted risk scores from Weibull regression model (constant shape)**

Effect	Predicted risk scores from Weibull regression model									
	t=6	t=12	t=18	t=24	t=30	t=36	t=42	t=48	t=54	t=60
Baseline hazard rate	0.005072	0.009094	0.012796	0.016305	0.019676	0.022943	0.026124	0.029234	0.032283	0.035279
Age [18;35)	1.172443	1.172443	1.172443	1.172443	1.172443	1.172443	1.172443	1.172443	1.172443	1.172443
Age [35;45)	0.966613	0.966613	0.966613	0.966613	0.966613	0.966613	0.966613	0.966613	0.966613	0.966613
Age [45+)	0.882380	0.882380	0.882380	0.882380	0.882380	0.882380	0.882380	0.882380	0.882380	0.882380

Table 5.22: **Predicted risk scores from Weibull regression model (shape alters)**

Effect	Predicted risk scores from Weibull regression model									
	t=6	t=12	t=18	t=24	t=30	t=36	t=42	t=48	t=54	t=60
Baseline hazard rate	0.005138	0.009185	0.012902	0.016420	0.019796	0.023065	0.026245	0.029353	0.032398	0.035389
Age [18;35)	1.081872	1.132627	1.163415	1.185765	1.203396	1.217996	1.230479	1.241395	1.251104	1.259853
Age [35;45)	1.008726	0.976080	0.957475	0.944491	0.934540	0.926488	0.919734	0.913924	0.908829	0.904295
Age [45+)	0.899648	0.883527	0.874231	0.867694	0.862658	0.858565	0.855119	0.852146	0.849532	0.847200

For a **constant shape parameter** in the Weibull distributions, the pth percentile of the baseline lifetime distribution can be calculated from

$$\frac{1}{\widehat{\lambda}_0}\cdot\ln\frac{100}{100-p})^{\frac{1}{\widehat{\alpha}}}$$

and that of the age group distributions from

$$\frac{1}{\widehat{\lambda}_{A_i}}\cdot\ln\frac{100}{100-p})^{\frac{1}{\widehat{\alpha}}} \quad i=1,2,3.$$

Note that the latter is equal to the pth percentile of the baseline distribution multiplied by the specific index to the power $-\frac{1}{\widehat{\alpha}}$.

For a **shape parameter that alters**, the formulaes change to

$$\frac{1}{\widehat{\lambda}_0}\cdot\ln\frac{100}{100-p})^{\frac{1}{\widehat{\alpha}_0}} \quad \text{and} \quad \frac{1}{\widehat{\lambda}_{A_i}}\cdot\ln\frac{100}{100-p})^{\frac{1}{\widehat{\alpha}_{A_i}}} \quad i=1,2,3.$$

The estimated percentiles of the baseline and the age group Weibull distributions for a constant shape parameter as well as for different shape parameters are reported in Table 5.24.

Table 5.24: **Lifetime percentiles estimated from Weibull regression model**

	Weibull regression model							
	Lifetime distribution: constant shape				Lifetime distribution: shape alters			
Lifetime percentile	Baseline	Age group [18;35)	Age group [35;45)	Age group [45+)	Baseline	Age group [18;35)	Age group [35;45)	Age group [45+)
P5	11.10	10.18	11.30	11.88	11.02	10.55	10.98	11.69
P10	16.40	15.05	16.71	17.56	16.31	15.39	16.42	17.40
P20	24.65	22.61	25.11	26.38	24.53	22.83	24.97	26.32
P25	28.30	25.95	28.82	30.28	28.16	26.09	28.78	30.29
P30	31.80	29.17	32.39	34.03	31.66	29.21	32.45	34.10
P40	38.64	35.45	39.36	41.36	38.49	35.27	39.65	41.58
P50	45.61	41.83	46.45	48.81	45.44	41.40	47.02	49.21
P60	53.06	48.67	54.05	56.79	52.89	47.94	54.95	57.40
P70	61.54	56.45	62.69	65.87	61.36	55.33	64.01	66.74
P75	66.44	60.94	67.67	71.11	66.26	59.58	69.25	72.14
P80	72.04	66.08	73.38	77.11	71.86	64.44	75.27	78.33
P90	87.50	80.26	89.13	93.65	87.32	77.77	91.94	95.45
P95	100.94	92.59	102.82	108.03	100.76	89.30	106.49	110.37

The median time to a lapse of a policy, over all three age groups, is 45.61 months. The baseline odds of a lapse at 45.61 months is 1, that means that $P(T>$45.61 months)$=P(T<$45.61 months). The lowest estimated percentile lifetime values in the third column of Table 5.24 again confirm the highest risk of a policy to lapse if the policyholder is in the youngest agegroup.

5.3.5 The fitting of a regression model with a continuous predictor

Consider AGE as a continuous predictor that can be categorized into three age groups. The ordinal covariate Z takes on the values

z=1 for the age group [18;35)
z=2 for the age group [35;45)
z=3 for age group [45+).

The midpoints of the age group intervals can also be used as values of the continuous predictor AGE, that means

$z = \frac{18+34}{2} = 26$ for age group $[18;35)$
$z = \frac{35+44}{2} = 39.5$ for age group $[35;45)$
$z = \frac{45+59}{2} = 52$ for age group $[45+)$
if 60 months is assumed to be an upper limit for the open interval.

A log-logistic as well as a Weibull regression model are fitted to the grouped survival data with known z-values. From the estimated regression parameters, survival model parameters can be found for each age group as well as for the baseline distribution.

The estimated regression coefficients of these two regression models are reported in Table 5.25.

Table 5.25: **Fitting a regression model (constant shape) to grouped data with one continuous predictor**

		Regression model			
Effect	Maximum likelihood estimates	Log-logistic		Weibull	
		z-values 1 2 3	z-values 26 39.5 52	z-values 1 2 3	z-values 26 39.5 52
Baseline mean	$\widehat{\ln\lambda_0} = \ln\widehat{\lambda}_0$	-7.647250	-7.477800	-7.111259	-6.962854
Constant shape	$\widehat{\alpha}$	2.066059	2.066104	1.841998	1.842030
Age	$\widehat{\beta}$	-0.166957	-0.012856	-0.146264	-0.011261

The estimated lambda parameters of the three survival distributions for the three age groups then are

$$\widehat{\lambda}_{Age1} = \exp(\ln \widehat{\lambda}_0 + \widehat{\beta} * 1)$$
$$\widehat{\lambda}_{Age2} = \exp(\ln \widehat{\lambda}_0 + \widehat{\beta} * 2)$$
$$\widehat{\lambda}_{Age3} = \exp(\ln \widehat{\lambda}_0 + \widehat{\beta} * 3)$$

or

$$\widehat{\lambda}_{Age1} = \exp(\ln \widehat{\lambda}_0 + \widehat{\beta} * 26)$$
$$\widehat{\lambda}_{Age2} = \exp(\ln \widehat{\lambda}_0 + \widehat{\beta} * 39.5)$$
$$\widehat{\lambda}_{Age3} = \exp(\ln \widehat{\lambda}_0 + \widehat{\beta} * 52)$$

with the same estimated alpha parameter $\widehat{\alpha}$. These parameters are summarized for each age group in Table 5.26.

Table 5.26: **Parameters of a survival model (constant shape) for each age group**

AGE group	Maximum likelihood estimates	**Survival model** Log-logistic z-values 1 2 3	Log-logistic z-values 26 39.5 52	Weibull z-values 1 2 3	Weibull z-values 26 39.5 52
Age [18;35)	$\ln \widehat{\lambda}_z$	-7.814208	-7.490656	-7.257523	-7.800792
	$\widehat{\alpha}$	2.0660588	2.0661037	1.841998	1.8421299
Age [35;45)	$\ln \widehat{\lambda}_z$	-7.901165	-7.503511	-7.403787	-8.016725
	$\widehat{\alpha}$	2.0660588	2.0661037	1.841998	1.8421299
Age [45+)	$\ln \widehat{\lambda}_z$	-8.148122	-7.516367	-7.550051	-8.127733
	$\widehat{\alpha}$	2.0660588	2.0661037	1.841998	1.8421299
Baseline	$\ln \widehat{\lambda}_0$	-7.647250	-7.477800	-7.111259	-6.962854
	$\widehat{\alpha}$	2.0660588	2.0661037	1.841998	1.8421299

5.3.6 A survival model for each combination of levels of two risk factors

Consider two risk factors AGE and SCORE where AGE has three levels [18;35), [35;45) and [45+) years and SCORE has three levels 'Low', 'Medium' and 'High'.

A cross tabulation of AGE and SCORE for the 10077 observations are given in Table 5.27.

Table 5.27: **Cross table of Age and Score**

	Score			
	Low	Medium	High	Total
Age [18;35)	833	1758	1053	3644
Age [35;45)	769	1546	1110	3425
Age [45+)	813	1541	654	3008
Total	2415	4845	2817	10077

A regression model is fitted to the grouped survival data where each policy has information on the entry period, age level and score level. The grouped lifetimes of the policies with staggered entry as well as the concomitant information on AGE and SCORE are given in Table 5.28.

The combined frequency vector is

$$\boldsymbol{f}' = (\boldsymbol{f}'_{111}, \boldsymbol{f}'_{211}, \boldsymbol{f}'_{311}, \boldsymbol{f}'_{411}, \boldsymbol{f}'_{112}, \boldsymbol{f}'_{212}, \boldsymbol{f}'_{312}, \boldsymbol{f}'_{412}, \boldsymbol{f}'_{113}, \boldsymbol{f}'_{213}, \boldsymbol{f}'_{313}, \boldsymbol{f}'_{413}, \boldsymbol{f}'_{121}, \boldsymbol{f}'_{221}, \boldsymbol{f}'_{321}, \boldsymbol{f}'_{421}, \boldsymbol{f}'_{122}, \boldsymbol{f}'_{222}, \boldsymbol{f}'_{322}, \boldsymbol{f}'_{422}, \boldsymbol{f}'_{123}, \boldsymbol{f}'_{223}, \boldsymbol{f}'_{323}, \boldsymbol{f}'_{423}, \boldsymbol{f}'_{131}, \boldsymbol{f}'_{231}, \boldsymbol{f}'_{331}, \boldsymbol{f}'_{431}, \boldsymbol{f}'_{132}, \boldsymbol{f}'_{232}, \boldsymbol{f}'_{332}, \boldsymbol{f}'_{432}, \boldsymbol{f}'_{133}, \boldsymbol{f}'_{233}, \boldsymbol{f}'_{333}, \boldsymbol{f}'_{433}).$$

$\boldsymbol{f}_{ilm}$ is the frequency vector for the i^{th} entry group, the l^{th} AGE level and the m^{th} SCORE level

$i = 1, 2, 3, 4$ and $l = 1, 2, 3$ and $m = 1, 2, 3$.

Table 5.28: **Multi-dimensional frequency table of grouped data set with two risk factors**

Entry	Age	Score	Lifetime intervals						
March 98			[0, 12)	[12, 17)	[17, 24)	[24, 28)	[28, 34)	[34, 37)	[37, ∞)
	[18;35)	Low	12	34	51	39	57	11	59
		Medium	10	12	22	19	32	4	418
		High	7	13	22	15	19	0	165
	[35;45)	Low	13	14	45	27	33	4	66
		Medium	4	22	22	8	25	4	297
		High	4	14	24	10	17	5	190
	[45+)	Low	10	25	29	17	46	2	116
		Medium	6	13	28	16	16	5	273
		High	0	11	11	6	5	0	82
June 98			[0, 12)	[12, 17)	[17, 24)	[24, 28)	[28, 34)	[34, ∞)	
	[18;35)	Low	22	25	58	53	40	45	
		Medium	10	26	32	20	29	379	
		High	9	24	13	19	14	204	
	[35;45)	Low	24	24	28	30	25	106	
		Medium	12	20	14	17	16	409	
		High	13	18	19	19	13	238	
	[45+)	Low	13	15	32	19	17	107	
		Medium	11	13	22	17	12	319	
		High	4	1	11	6	6	117	
Nov 98			[0, 12)	[12, 17)	[17, 24)	[24, 28)	[28, ∞)		
	[18;35)	Low	34	16	50	23	54		
		Medium	19	2	32	24	317		
		High	15	16	17	10	199		
	[35;45)	Low	19	18	38	16	75		
		Medium	16	14	25	10	263		
		High	5	12	20	7	195		
	[45+)	Low	28	16	22	12	98		
		Medium	13	0	24	4	323		
		High	5	5	14	11	150		
March 99			[0, 12)	[12, 17)	[17, 24)	[24, ∞)			
	[18;35)	Low	40	30	30	50			
		Medium	9	14	27	301			
		High	22	16	12	222			
	[35;45)	Low	24	30	29	81			
		Medium	14	15	12	307			
		High	16	16	27	228			
	[45+)	Low	20	22	28	119			
		Medium	19	12	26	369			
		High	11	11	16	171			

$\boldsymbol{f}_{111} = (12, 34, 51, 39, 57, 11, 59)'$

$\boldsymbol{f}_{112} = (10, 12, 22, 19, 32, 4, 418)'$

$\boldsymbol{f}_{113} = (7, 13, 22, 15, 19, 0, 165)'$

$\boldsymbol{f}_{121} = (13, 14, 45, 27, 33, 4, 66)'$

$\boldsymbol{f}_{122} = (4, 22, 22, 8, 25, 4, 297)'$

$\boldsymbol{f}_{123} = (4, 14, 24, 10, 17, 5, 190)'$

$\boldsymbol{f}_{131} = (10, 25, 29, 17, 46, 2, 116)'$

$\boldsymbol{f}_{132} = (6, 13, 28, 16, 16, 5, 273)'$

$\boldsymbol{f}_{133} = (0, 11, 11, 6, 5, 0, 82)'$

$\boldsymbol{f}_{211} = (22, 25, 58, 53, 40, 45)'$

$\boldsymbol{f}_{212} = (10, 26, 32, 20, 29, 379)'$

$\boldsymbol{f}_{213} = (9, 24, 13, 19, 14, 204)'$

$\boldsymbol{f}_{221} = (24, 24, 28, 30, 25, 106)'$

$\boldsymbol{f}_{222} = (12, 20, 14, 17, 16, 409)'$

$\boldsymbol{f}_{223} = (13, 18, 19, 19, 13, 238)'$

$\boldsymbol{f}_{231} = (13, 15, 32, 19, 17, 107)'$

$\boldsymbol{f}_{232} = (11, 13, 22, 17, 12, 319)'$

$\boldsymbol{f}_{233} = (4, 1, 11, 6, 6, 117)'$

$\boldsymbol{f}_{311} = (34, 16, 50, 23, 54)'$

$\boldsymbol{f}_{312} = (19, 2, 32, 24, 317)'$

$\boldsymbol{f}_{313} = (15, 16, 17, 10, 199)'$

$\boldsymbol{f}_{321} = (19, 18, 38, 16, 75)'$

$\boldsymbol{f}_{322} = (16, 14, 25, 10, 263)'$

$\boldsymbol{f}_{323} = (5, 12, 20, 7, 195)'$

$\boldsymbol{f}_{331} = (28, 16, 22, 12, 98)'$

$\boldsymbol{f}_{332} = (13, 0, 24, 4, 323)'$

$\boldsymbol{f}_{333} = (5, 5, 14, 11, 150)'$

$\boldsymbol{f}_{411} = (40, 30, 30, 50)'$

$\boldsymbol{f}_{412} = (9, 14, 27, 301)'$

$\boldsymbol{f}_{413} = (22, 16, 12, 222)'$

$\boldsymbol{f}_{421} = (24, 30, 29, 81)'$

$\boldsymbol{f}_{422} = (14, 15, 12, 307)'$

$\boldsymbol{f}_{423} = (16, 16, 27, 228)'$
$\boldsymbol{f}_{431} = (20, 22, 28, 119)'$
$\boldsymbol{f}_{432} = (19, 12, 26, 369)'$
$\boldsymbol{f}_{433} = (11, 11, 16, 171)'$

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

$$\boldsymbol{x}_1 = \begin{pmatrix} 12 \\ 17 \\ 24 \\ 28 \\ 34 \\ 37 \end{pmatrix} \quad \boldsymbol{x}_2 = \begin{pmatrix} 12 \\ 17 \\ 24 \\ 28 \\ 34 \end{pmatrix} \quad \boldsymbol{x}_3 = \begin{pmatrix} 12 \\ 17 \\ 24 \\ 28 \end{pmatrix} \quad \text{and} \quad \boldsymbol{x}_4 = \begin{pmatrix} 12 \\ 17 \\ 24 \end{pmatrix}.$$

From the estimated regression parameters, survival model parameters can be found for each combination of the levels of AGE and SCORE as well as for the baseline distribution. The estimated regression coefficients of the regression model with two risk factors AGE and SCORE are reported in Table 5.29.

Table 5.29: **Fitting a regression model with two risk factors**

Effect	Maximum likelihood estimates	**Regression model** Log-logistic	Weibull
Baseline mean	$\widehat{\ln \lambda_0} = \ln \widehat{\lambda}_0$	-8.550810	-7.709833
Age [18;35)	$\widehat{\beta}_{A_1}$	0.205367	0.212709
Age [35;45)	$\widehat{\beta}_{A_2}$	-0.011853	-0.014725
Age [45+)	$\widehat{\beta}_{A_3}$	-0.193514	-0.197984
Score 'Low'	$\widehat{\beta}_{B_1}$	1.047861	0.897721
Score 'Medium'	$\widehat{\beta}_{B_2}$	-0.714941	-0.612472
Score 'High'	$\widehat{\beta}_{B_3}$	-0.332746	-0.285249
Constant shape	$\widehat{\alpha}$	2.249510	1.938292

The estimated lambda parameters of the nine survival distributions for the nine combinations of AGE and SCORE levels are

$$\begin{aligned}
\widehat{\lambda}_{A_1B_1} &= \exp(\ln \widehat{\lambda}_0 + \widehat{\beta}_{A_1} + \widehat{\beta}_{B_1}) \\
\widehat{\lambda}_{A_1B_2} &= \exp(\ln \widehat{\lambda}_0 + \widehat{\beta}_{A_1} + \widehat{\beta}_{B_2}) \\
\widehat{\lambda}_{A_1B_3} &= \exp(\ln \widehat{\lambda}_0 + \widehat{\beta}_{A_1} + \widehat{\beta}_{B_3}) \\
\widehat{\lambda}_{A_2B_1} &= \exp(\ln \widehat{\lambda}_0 + \widehat{\beta}_{A_2} + \widehat{\beta}_{B_1}) \\
\widehat{\lambda}_{A_2B_2} &= \exp(\ln \widehat{\lambda}_0 + \widehat{\beta}_{A_2} + \widehat{\beta}_{B_2}) \\
\widehat{\lambda}_{A_2B_3} &= \exp(\ln \widehat{\lambda}_0 + \widehat{\beta}_{A_2} + \widehat{\beta}_{B_3}) \\
\widehat{\lambda}_{A_3B_1} &= \exp(\ln \widehat{\lambda}_0 + \widehat{\beta}_{A_3} + \widehat{\beta}_{B_1}) \\
\widehat{\lambda}_{A_3B_2} &= \exp(\ln \widehat{\lambda}_0 + \widehat{\beta}_{A_3} + \widehat{\beta}_{B_2}) \\
\widehat{\lambda}_{A_3B_3} &= \exp(\ln \widehat{\lambda}_0 + \widehat{\beta}_{A_3} + \widehat{\beta}_{B_3})
\end{aligned}$$

with the same estimated alpha parameter $\widehat{\alpha}$. These parameters are summarized for each combination of AGE and SCORE levels in Table 5.30.

Table 5.30: **Parameters of a survival model for each combination of AGE and SCORE levels**

Combination of Age and Score	Maximum likelihood estimates	Survival model: Log-logistic	Survival model: Weibull
Age [18;35) and Low score	$\ln \widehat{\lambda}_{A_1B_1}$	-7.297757	-6.599403
	$\widehat{\alpha}$	2.249510	1.938292
Age [18;35) and Medium score	$\ln \widehat{\lambda}_{A_1B_2}$	-9.060383	-8.109596
	$\widehat{\alpha}$	2.249510	1.938292
Age [18;35) and High score	$\ln \widehat{\lambda}_{A_1B_3}$	-8.678188	-7.782373
	$\widehat{\alpha}$	2.249510	1.938292
Age [35;45) and Low score	$\ln \widehat{\lambda}_{A_2B_1}$	-7.514976	-6.826837
	$\widehat{\alpha}$	2.249510	1.938292
Age [35;45) and Medium score	$\ln \widehat{\lambda}_{A_2B_2}$	-9.277603	-8.337030
	$\widehat{\alpha}$	2.249510	1.938292
Age [35;45) and High score	$\ln \widehat{\lambda}_{A_2B_3}$	-8.895408	-8.009807
	$\widehat{\alpha}$	2.249510	1.938292
Age [45+) and Low score	$\ln \widehat{\lambda}_{A_3B_1}$	-7.696638	-7.010096
	$\widehat{\alpha}$	2.249510	1.938292
Age [45+) and Medium score	$\ln \widehat{\lambda}_{A_3B_2}$	-9.459265	-8.520288
	$\widehat{\alpha}$	2.249510	1.938292
Age [45+) and High score	$\ln \widehat{\lambda}_{A_3B_3}$	-9.077070	-8.193066
	$\widehat{\alpha}$	2.249510	1.938292
Baseline	$\ln \widehat{\lambda}_0$	-8.55081	-7.709833
	$\widehat{\alpha}$	2.249510	1.938292

A joint histogram to the data of each combination of AGE and SCORE levels **over the four entry groups** is needed to make a graphical representation of the fitted models for each combination of AGE and SCORE levels. Table 5.31 gives the nine sets of fitted joint frequencies for the nine combinations of AGE and SCORE levels. This fitting was done by maximum likelihood estimation subject to constraints imposed by the experimental design. The Wald test and discrepancy value measure the goodness-of-fit.

Table 5.31: **Fitted joint frequency distributions for the nine combinations of AGE and SCORE levels**

Interval number	Interval of survival times	Fitted Joint Frequencies Age [18;35) Low score	Age [35;45) Low score	Age [45+) Low score
first	$[0, 12)$	108	80	71
second	$[12, 17)$	105	86	78
third	$[17, 24)$	189	140	111
fourth	$[24, 28)$	130.43421	90.690721	61.44444
fifth	$[28, 34)$	137.52303	92.281787	107.52778
sixth	$[34, 37)$	25.621006	16.001571	6.508945
seventh	$[37, \infty)$	137.42176	264.02592	377.51883
Wald		87.06	38.99	35.20
Discrepancy		0.1045	0.0507	0.0432974

Interval number	Interval of survival times	Fitted Joint Frequencies Age [18;35) Medium score	Age [35;45) Medium score	Age [45+) Medium score
first	$[0, 12)$	48	46	49
second	$[12, 17)$	54	71	38.000364
third	$[17, 24)$	113	73	100
fourth	$[24, 28)$	66.789123	43.685567	67.905775
fifth	$[28, 34)$	104.46504	71.64433	57.617021
sixth	$[34, 37)$	13.00233	16.48731	22.094914
seventh	$[37, \infty)$	1358.7435	1224.1828	1206.3823
Wald		34.26	31.51	20.50
Discrepancy		0.0195	0.0204	0.0133

Interval number	Interval of survival times	Fitted Joint Frequencies Age [18;35) High score	Age [35;45) High score	Age [45+) High score
first	$[0, 12)$	53	38	20.000115
second	$[12, 17)$	69	60	28
third	$[17, 24)$	64	90	52
fourth	$[24, 28)$	67.610092	54.345528	29.945937
fifth	$[28, 34)$	65.62156	56.219512	27.450442
sixth	$[34, 37)$	0	20.806025	0.0006965
seventh	$[37, \infty)$	733.76835	790.62893	496.60304
Wald		20.07	14.81	22.08
Discrepancy		0.0191	0.0133428	0.0338

Figure 5.16 shows the fitted joint histogram and the fitted survival distributions for age group [18;35) and low score.

Figure 5.17 shows the fitted joint histogram and the fitted survival distributions for age group [18;35) and medium score.

Figure 5.18 shows the fitted joint histogram and the fitted survival distributions for age group [18;35) and high score.

Figure 5.19 shows the fitted joint histogram and the fitted survival distributions for age group [35;45) and low score.

Figure 5.20 shows the fitted joint histogram and the fitted survival distributions for age group [35;45) and medium score.

Figure 5.21 shows the fitted joint histogram and the fitted survival distributions for age group [35;45) and high score.

Figure 5.22 shows the fitted joint histogram and the fitted survival distributions for age group [45+) and low score.

Figure 5.23 shows the fitted joint histogram and the fitted survival distributions for age group [45+) and medium score.

Figure 5.24 shows the fitted joint histogram and the fitted survival distributions for age group [45+) and high score.

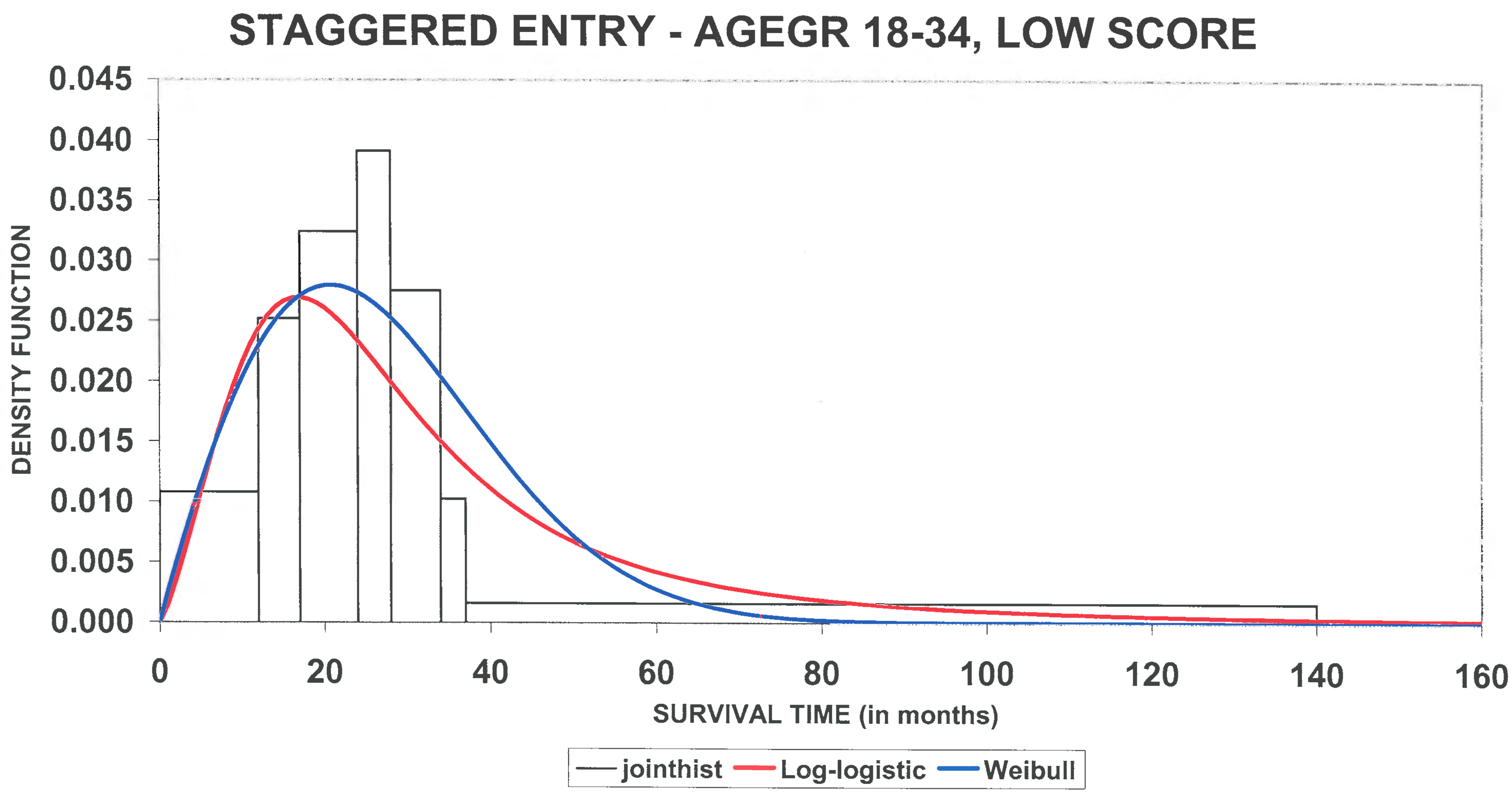

Figure 5.16: **Joint histogram and fitted survival distributions for age group [18;35) and low score**

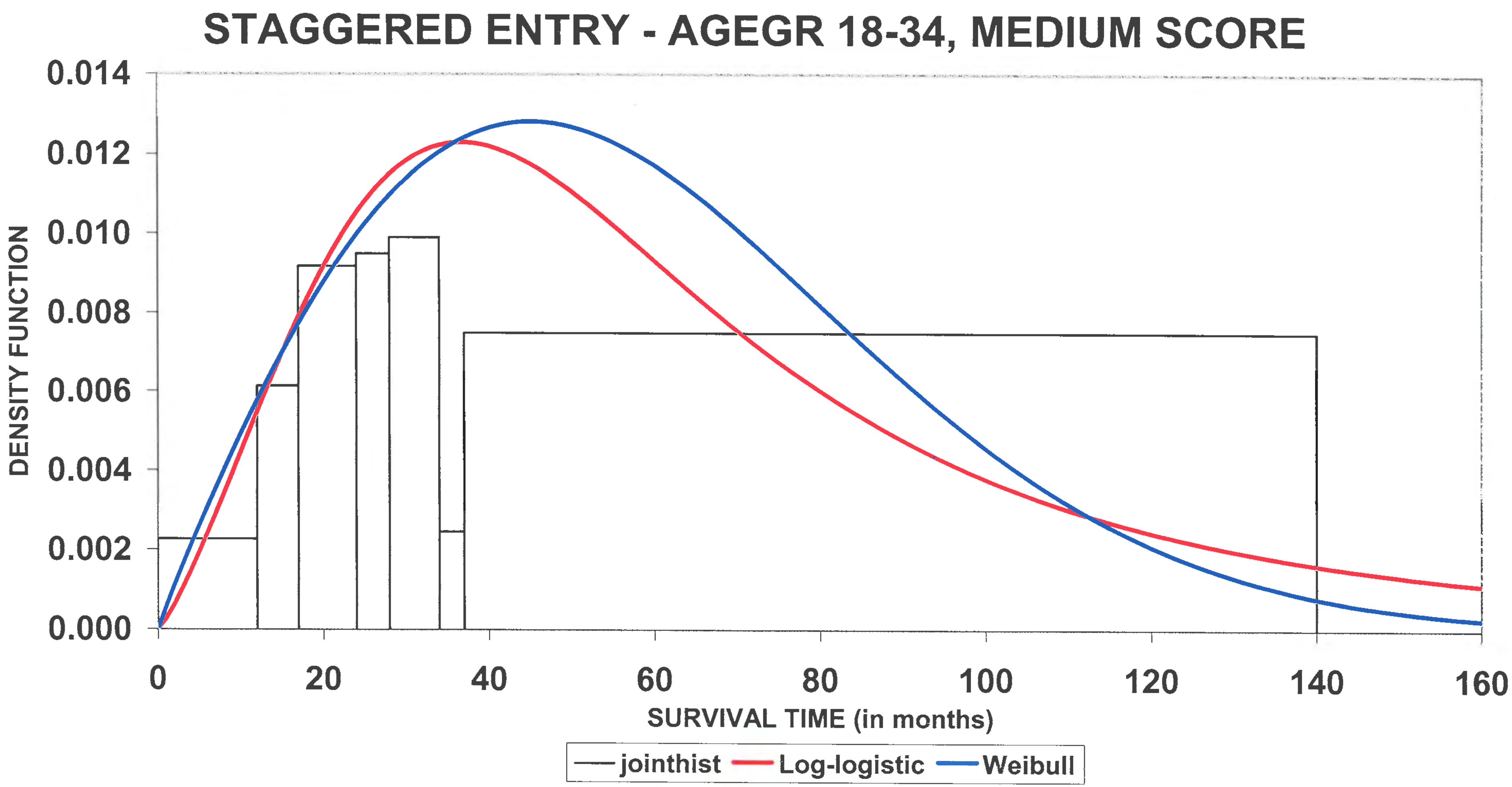

Figure 5.17: **Joint histogram and fitted survival distributions for age group [18;35) and medium score**

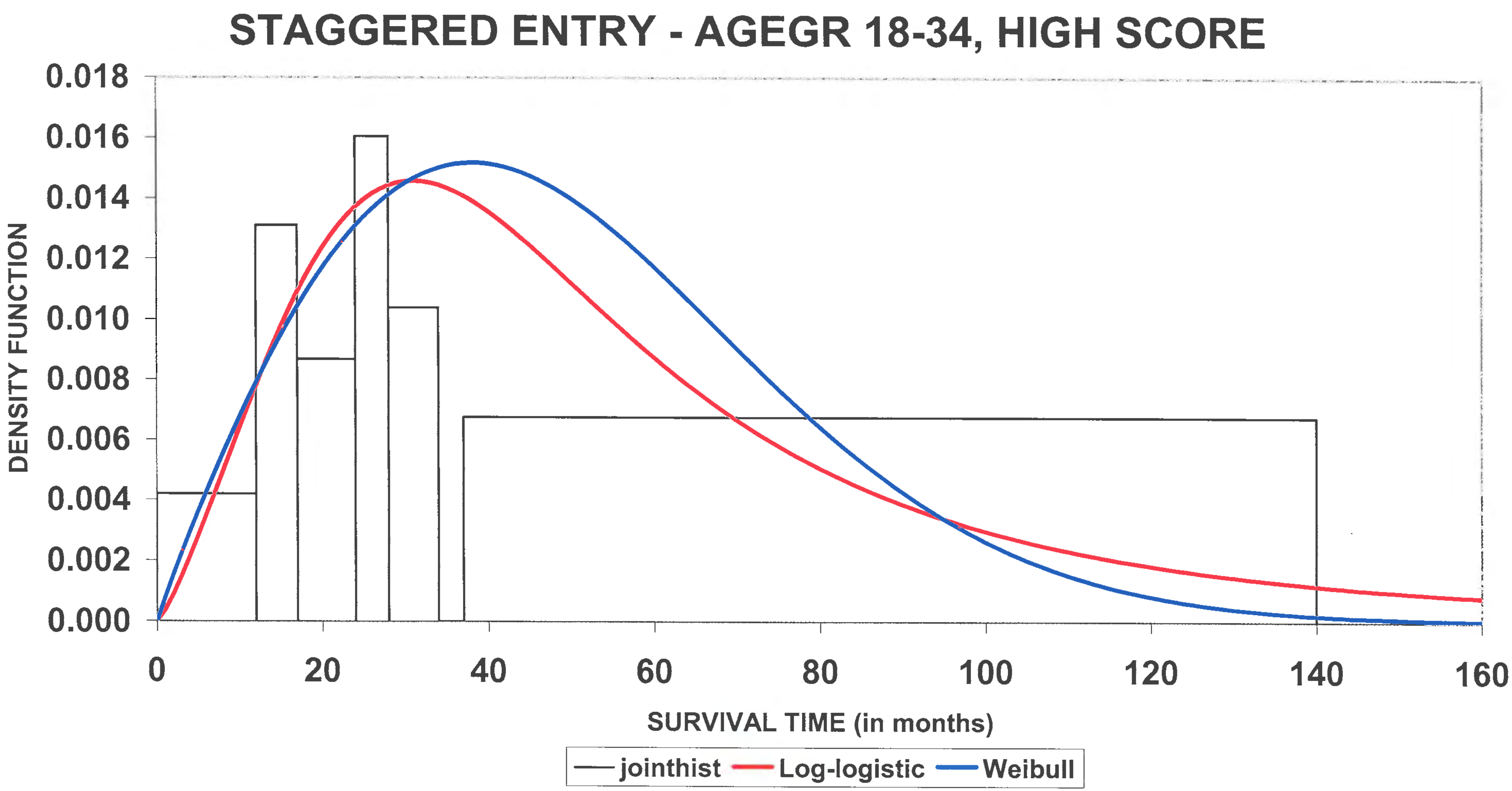

Figure 5.18: **Joint histogram and fitted survival distributions for age group [18;35) and high score**

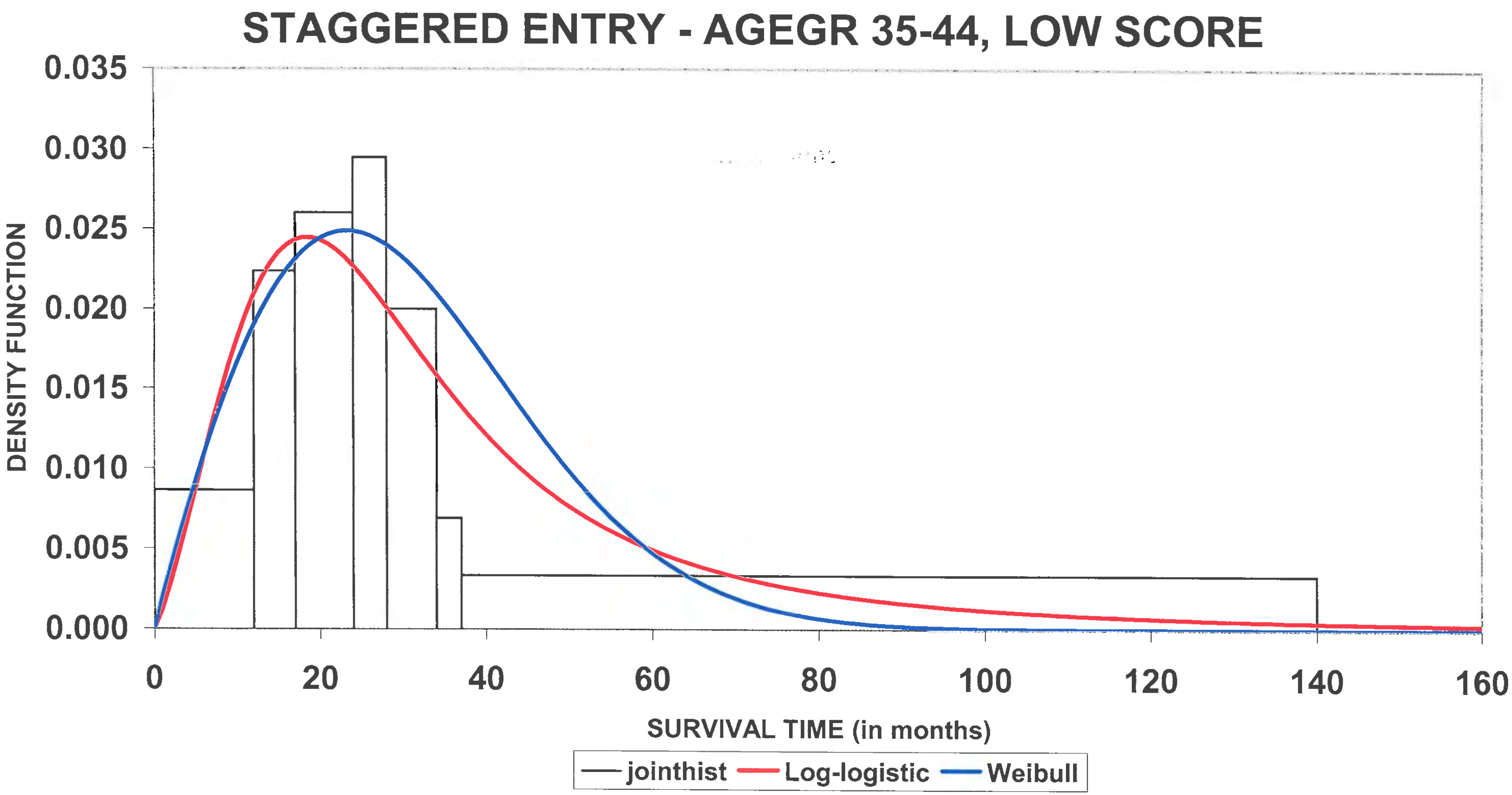

Figure 5.19: **Joint histogram and fitted survival distributions for age group [35;45) and low score**

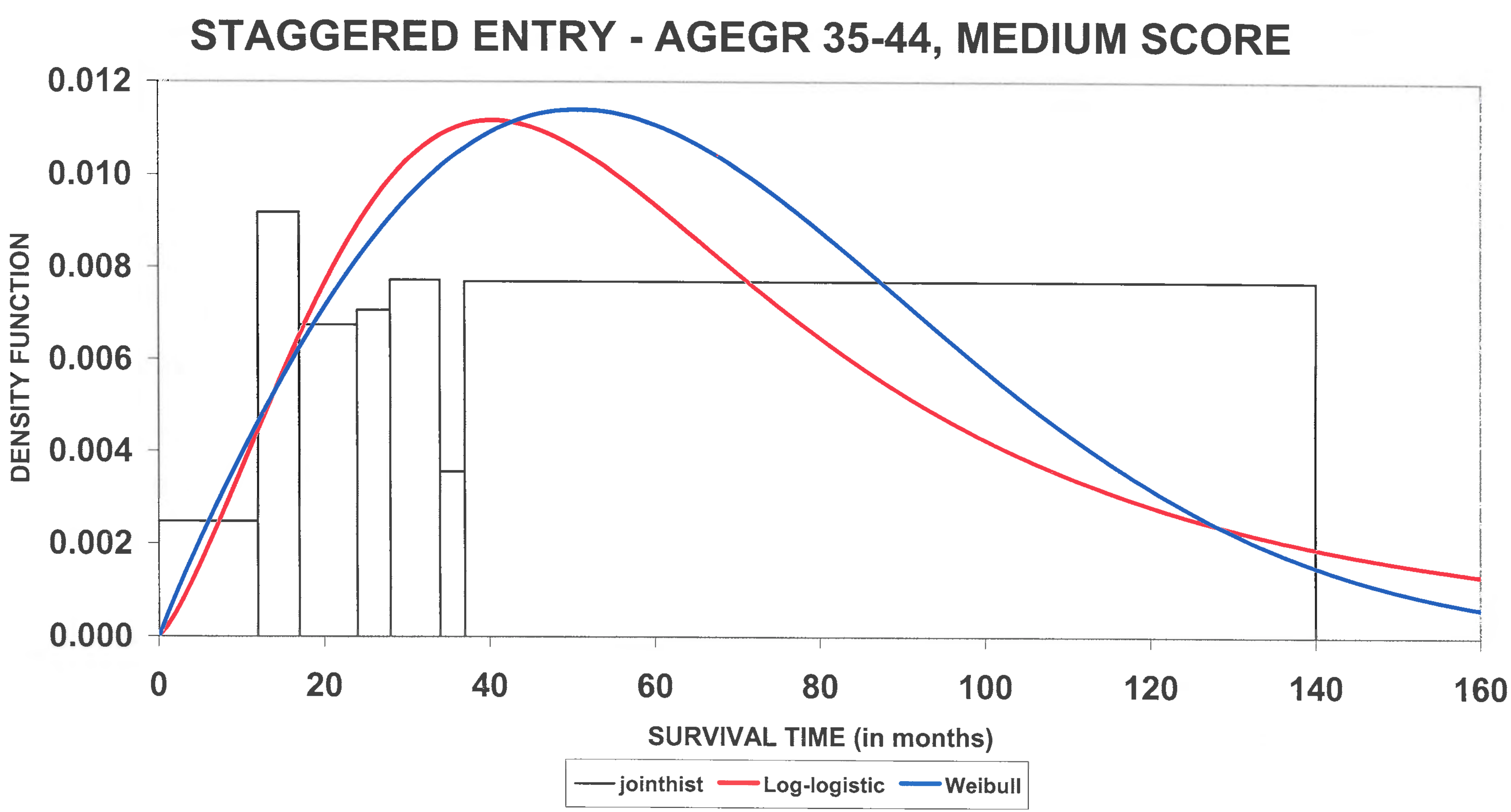

Figure 5.20: **Joint histogram and fitted survival distributions for age group [35;45) and medium score**

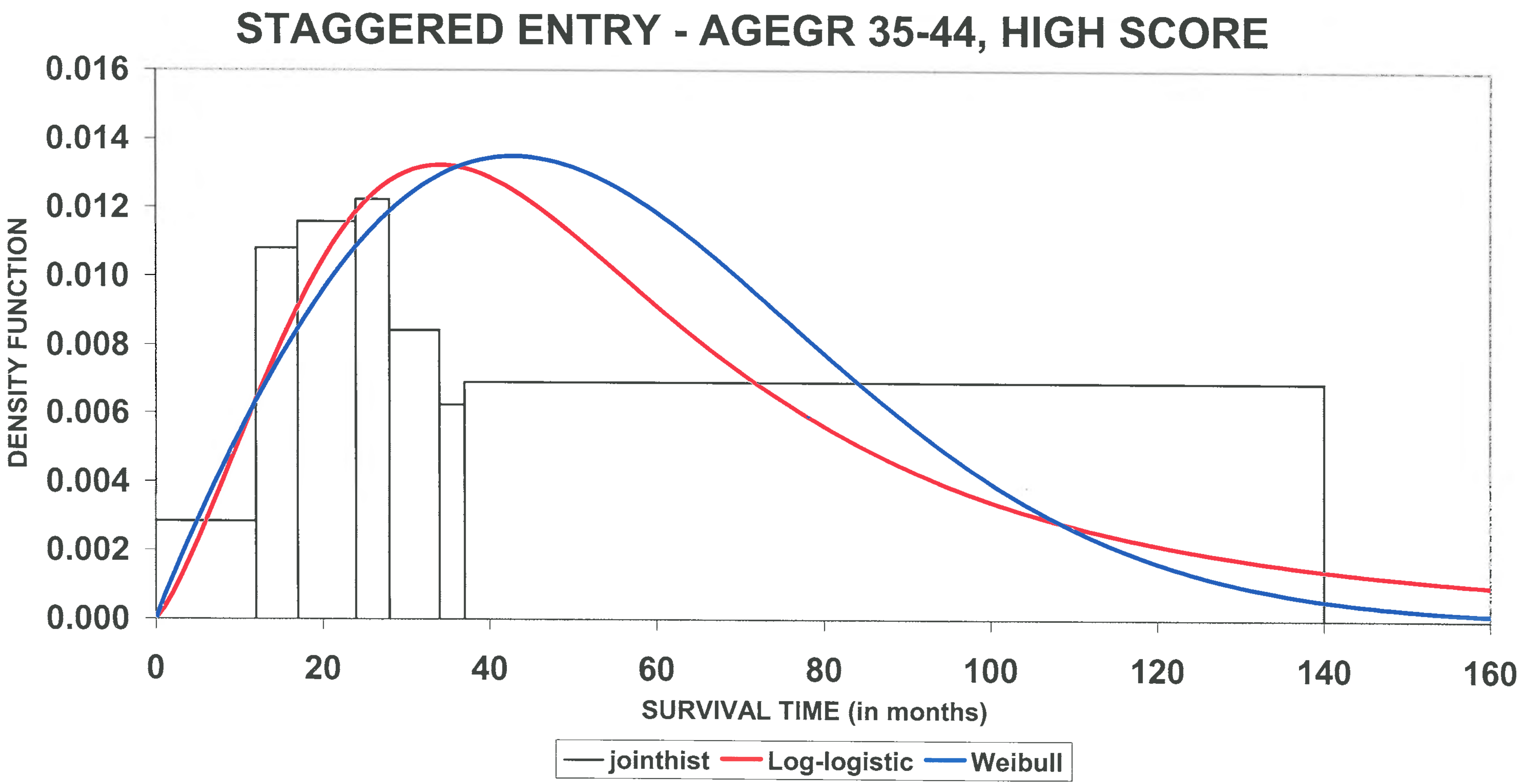

Figure 5.21: Joint histogram and fitted survival distributions for age group [35;45) and high score

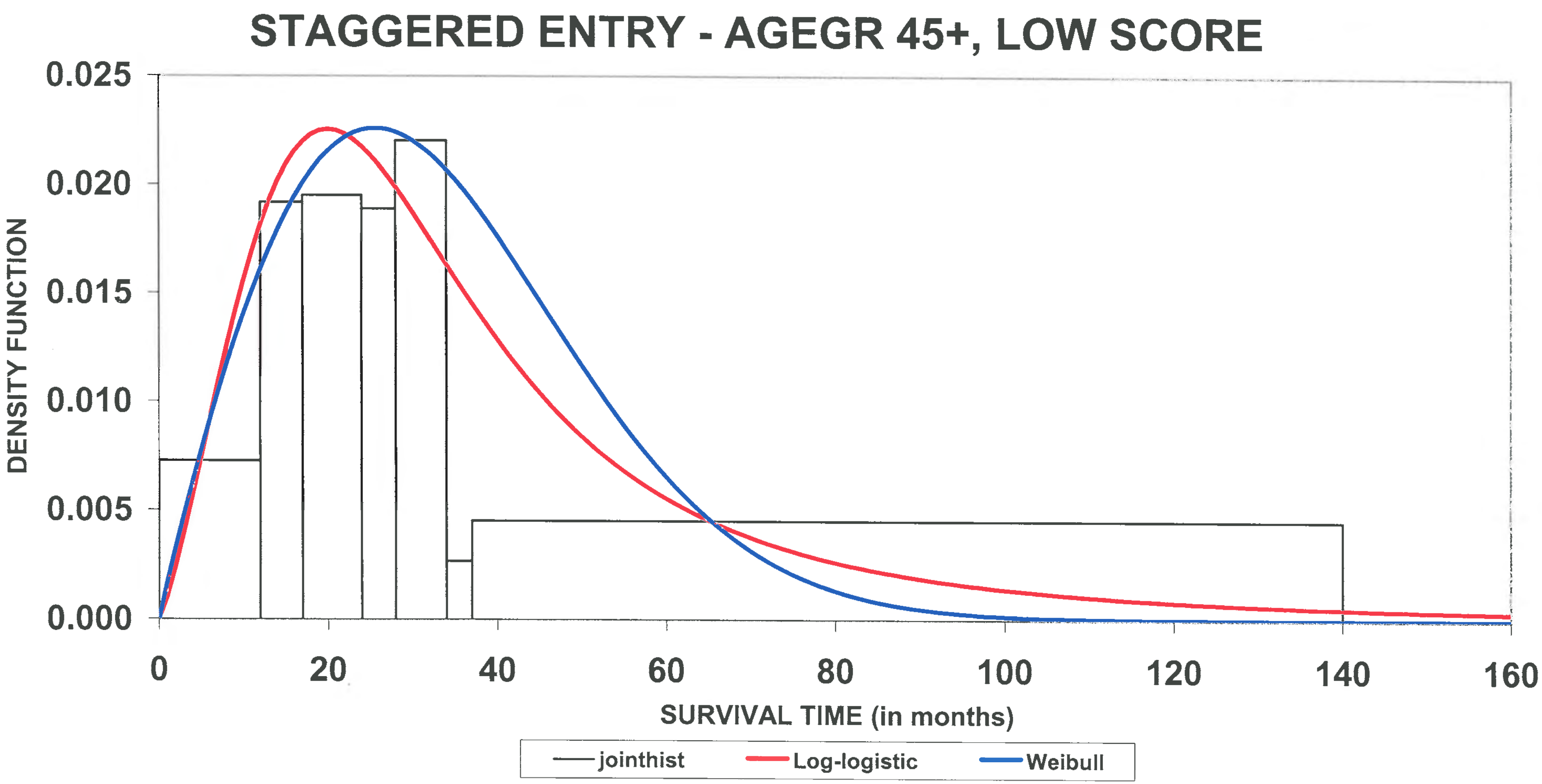

Figure 5.22: **Joint histogram and fitted survival distributions for age group [45+) and low score**

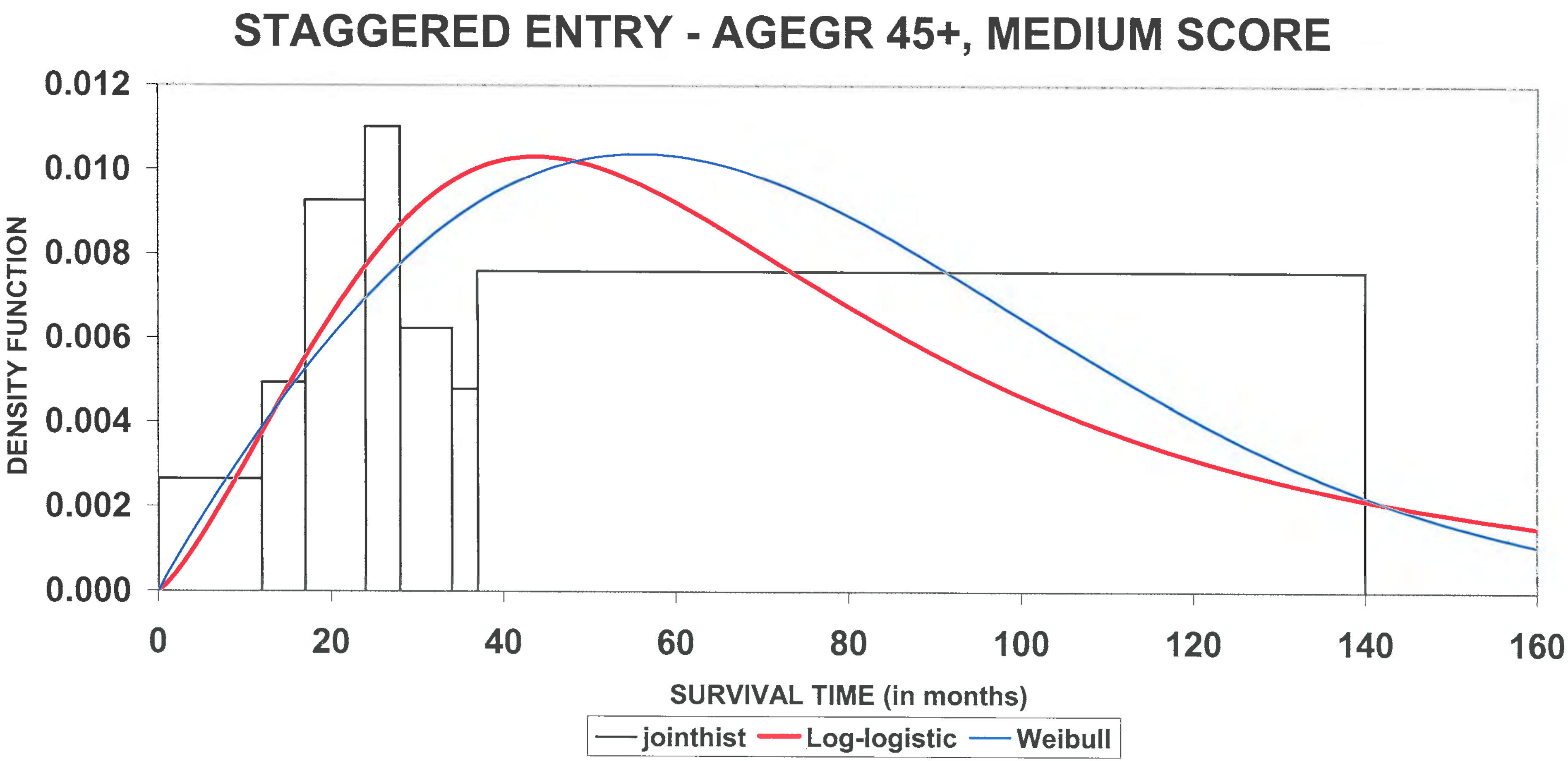

Figure 5.23: **Joint histogram and fitted survival distributions for age group [45+) and medium score**

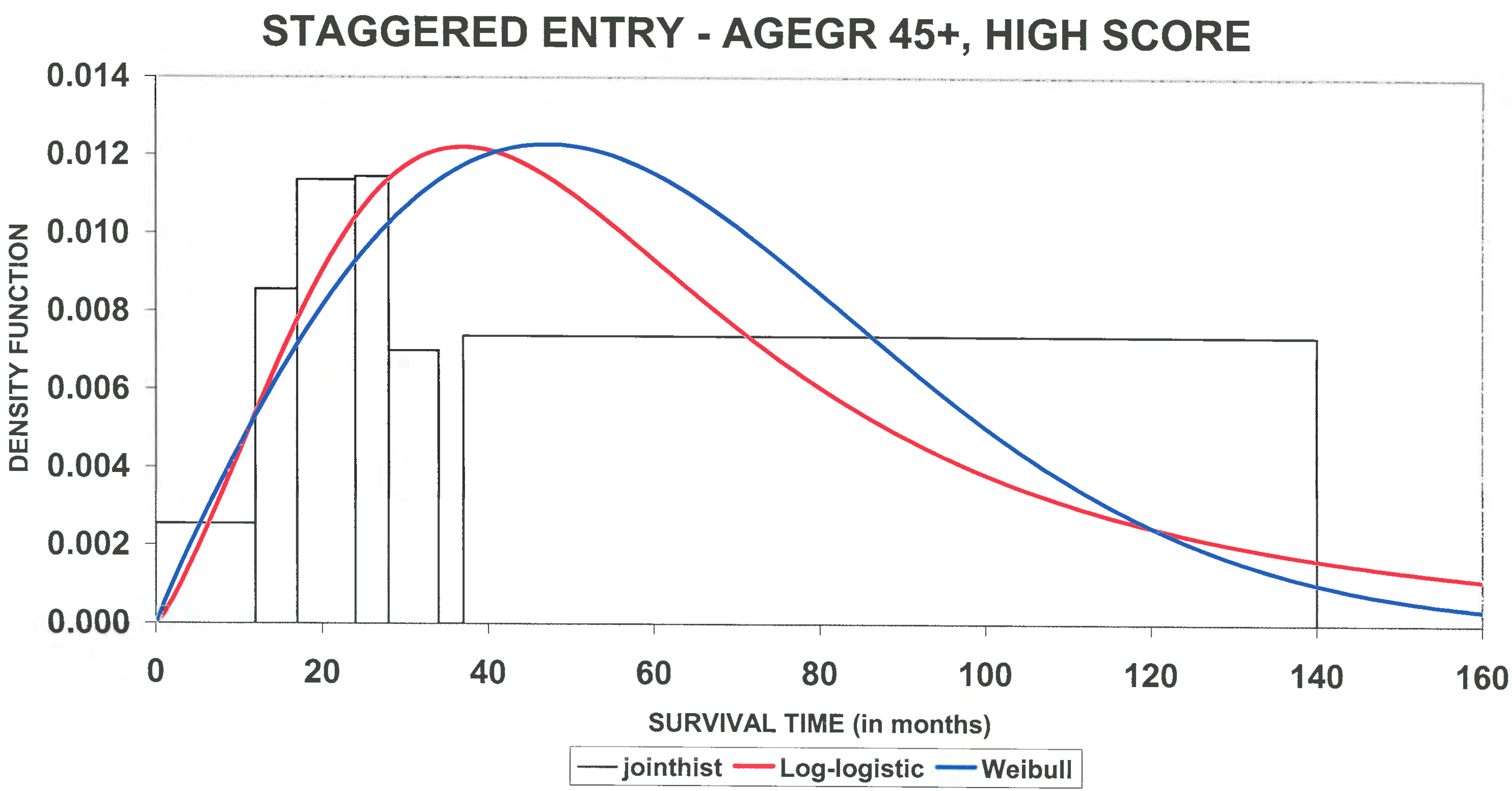

Figure 5.24: **Joint histogram and fitted survival distributions for age group [45+) and high score**

5.3.7 Relationship between the indices of the regression and logit model

Once the parameters of the baseline survival distribution and the nine age-score distributions have been estimated, estimated hazard and survivor functions, odds of a lapse, odds ratios and hazard ratios at time t can be calculated in a similar way as at the regression model with one risk factor (constant shape).

The odds ratio for age group [35;45) and a medium score is the relative odds of a lapse at time t of a policy, where the age of the policyholder is in [35;45) years and the policyholder has a medium score , compared to a policy with the baseline characteristics.

As an example , the odds ratio for a lapse of a policy at time t is calculated if the age of the policyholder is in the age group [18;35) years and the policyholder has a low score.

$$\boxed{\widehat{oddsratio}_{A_1B_1}(t) = \frac{\widehat{odds}_{A_1B_1}(t)}{\widehat{odds}_0(t)}}$$

where

$$\boxed{\widehat{odds}_{A_1B_1}(t) = \frac{1-\widehat{S}_{A_1B_1}(t)}{\widehat{S}_{A_1B_1}(t)} = \widehat{\lambda}_{A_1B_1} \cdot t^{\widehat{\alpha}}}$$

$$\Rightarrow \widehat{odds}_{A_1B_1}(12) = e^{-7.297757} \cdot 12^{2.249510} = 0.181240$$

$$\begin{aligned}\Rightarrow \widehat{oddsratio}_{A_1B_1}(12) &= \frac{0.181240}{0.051768} \\ &= 3.501004\end{aligned}$$

This odds ratio of 3.5 is called an index and shows the effect of age group [18;35) and a low score on the baseline odds of a lapse at time t. This effect is multiplicative on the baseline odds of a lapse. Thus the effect of the combination of this age group and this score group is to increase the baseline odds of a lapse by a factor 3.5 .

The other eight indices for the log-logistic regression model can be calculated in a similar way.

The relationship between the indices of the nine age-score combinations, obtained from the log-logistic model, must be compared to the six 'indices', obtained from the **loglinear logit model** for the three age levels and the three score levels.

Recall that the loglinear logit model models

$$\ln(\text{odds of a lapse}) = \mu + \lambda_i^{AGE} + \lambda_j^{SCORE}$$

where

$$\begin{aligned} \mu &= \text{the overall mean effect, over all AGE levels and SCORE levels} \\ \lambda_i^{AGE} &= \text{effect of the } i^{th} \text{ level of AGE} \\ \lambda_j^{SCORE} &= \text{effect of the } j^{th} \text{ level of SCORE.} \end{aligned}$$

The odds of a lapse then can be modelled as

$$\begin{aligned} \text{odds of lapse} &= e^{\mu+\lambda_i^{AGE}+\lambda_j^{SCORE}} \\ &= e^{\mu} \cdot e^{\lambda_i^{AGE}} \cdot e^{\lambda_j^{SCORE}} \\ &= \text{geometric mean odds} \cdot \text{index}_{AGE_i} \cdot \text{index}_{SCORE_j} \end{aligned}$$

i=1,2,3 and j=1,2,3.

The six 'indices' obtained from the logit model for each age level and for each score level are given in Table 5.32.

Table 5.32: **Logit model indices for three age levels and three score levels obtained from the logit model**

Effect	n	Logit model Index t=12	Logit model Index t=24
Baseline odds	10077	0.0537	0.2694
Age [18;35)	3644	1.1558	1.1745
Age [35;45)	3425	0.9844	0.9981
Age [45+),	3008	0.8790	0.8530
Low score	2415	2.2622	2.5756
Medium score	4845	0.5757	0.5234
High score	2817	0.7678	0.7418

The odds of a lapse of a policy in the first year , with the policyholder in the age group [18;35) and a low score, is calculated from the logit model as the product of the baseline odds and the index of age group [18;35) of 1.1558 and the index of score group 'Low' of 2.2622 for the first year (t=12), this means

$$odds_{A_1B_1}(12) = 0.0537 \times 1.1558 \times 2.2622 = 0.1404$$

$$\Rightarrow P(\text{lapse of this policy}) = \frac{odds}{1+odds} = \frac{0.1404}{1.1404} = 0.1231$$

Thus the odds ratio (relative odds of a lapse of this policy) is calculated by

$$oddsratio_{A_1B_1}(12) = \frac{odds_{A_1B_1}(12)}{baselineodds} = \frac{0.1404}{0.0537} = 2.614525$$

It is clear that this odds ratio can easily be found by multiplication of the two indices from the logit model, that is

$$oddsratio_{A_1B_1}(12) \doteq index_{A1}(12) \times index_{B1}(12) = 1.1558 \times 2.2622 = 2.614651$$

The odds ratio shows the effect of the combination of this age group and this score group on the baseline odds of a lapse. This effect is multiplicative on the baseline odds of a lapse. Thus the effect of the combination of this age group and this score group is to increase the baseline odds of a lapse by a factor 2.614 .

In the context of survival analysis, this odds ratio can be called an **index** for age group [18;35) and a low score. The odds ratios for the nine age-score combinations result in a set of indices, showing the effect of each combination of age group and score on the baseline odds of a lapse at time t.

The odds ratios (indices) of the nine age-score groups, estimated from the log-logistic or Weibull regression model, are compared to the odds ratios, obtained from the logit model, in Table 5.33.

Table 5.31: **Comparison of odds ratios (indices): log-logistic and Weibull regression models and logit model**

Effect	n	Log-logistic regression model					Weibull regression model					Logit model	
		Odds ratio					Odds ratio					Odds ratio	
		t=6	t=12	t=24	t=36	t=60	t=6	t=12	t=24	t=36	t=60	t=12	t=24
Baseline odds	10077	0.01	0.05	0.25	0.62	1.93	0.01	0.06	0.24	0.59	2.5	0.05	0.27
Age [18;35), Low score	833	3.5	3.5	3.5	3.5	3.5	3.08	3.22	3.83	5.25	17.6	2.6	3.0
Age [35;45), Low score	769	2.8	2.8	2.8	2.8	2.8	2.44	2.52	2.84	3.51	7.88	2.2	2.6
Age [45+), Low score	813	2.3	2.3	2.3	2.3	2.3	2.03	2.07	2.25	2.62	4.59	2.0	2.2
Age [18;35), Medium score	1758	0.6	0.6	0.6	0.6	0.6	0.67	0.66	0.65	0.62	0.53	0.7	0.6
Age [35;45), Medium score	1546	0.5	0.5	0.5	0.5	0.5	0.53	0.53	0.51	0.48	0.38	0.6	0.6
Age [45+), Medium score	1541	0.4	0.4	0.4	0.4	0.4	0.44	0.44	0.42	0.39	0.30	0.5	0.5
Age [18;35), High score	1053	0.9	0.9	0.9	0.9	0.9	0.93	0.93	0.92	0.91	0.88	0.9	0.9
Age [35;45), High score	1110	0.7	0.7	0.7	0.7	0.7	0.74	0.74	0.72	0.69	0.61	0.8	0.7
Age [45+), High score	654	0.6	0.6	0.6	0.6	0.6	0.62	0.61	0.59	0.56	0.47	0.7	0.6

It is clear from Table 5.33 that the odds ratios are constant over time at the log-logistic regression model, but the odds ratios do not remain constant over time at the Weibull regression model.

From Table 5.33 follows that one log-logistic regression model provides odds ratios (indices) for any time value, while a new logitmodel has to be built for a fixed time value, say t=12 months, conditional on a restricted experimental design where all the policies must have an exposure of at least one year when investigating the lapses of policies in the first year. There is no such restrictions in the more general experimental design for the log-logistic regression model where all the policies can be used in the analysis, even those policies with inception dates very close to the cut-off point.

The same argument holds for the Weibull regression model, except that the odds ratios do not remain constant over time.

5.3.8 Median lifetimes of the nine survival distributions

The median lifetimes (in months) of the nine survival distributions can also be estimated and compared with the baseline median. The medians are reported in Table 5.34. It is

Table 5.34: **Median lifetimes of the nine survival distributions**

	Regression model	
	Log-logistic	Weibull
Effect	median lifetime	median lifetime
Baseline	44.75	44.19
Age [18;35), Low score	25.64	24.92
Age [35;45), Low score	28.24	28.02
Age [45+), Low score	30.61	30.80
Age [18;35), Medium score	56.13	54.31
Age [35;45), Medium score	61.82	61.08
Age [45+), Medium score	67.02	67.13
Age [18;35), High score	47.36	45.88
Age [35;45), High score	52.16	51.59
Age [45+), High score	56.55	56.70

evident from Table 5.34 that the log-logistic and Weibull models deliver the same results. The estimated median values of the nine combinations of age and score levels suggest that the policy of a policyholder with a low score, coming from any age group, has a high risk to lapse. The policy of a policyholder in agegroup 45+ with a medium score has the lowest risk to lapse, lower than the combination 45+ and a high score.

Chapter 6

RESUMÉ

This thesis focuses on the analysis of insurance policy lifetimes in the form of **grouped data**, where the lifetime of a policy is measured from the inception date (entry month) up to the lapsing date (month in which policy lapsed) or a pre-determined cut-off date. Data from the insurance industry are extensive data sets with very large sample sizes. The focus in this thesis is on the estimation of lifetime distributions, based on a large sample of discrete lifetimes of policies that are grouped into intervals of lifetimes. The aim of the research is the statistical modelling of **parametric** survival distributions of **grouped survival data** of long- and shortterm policies in the insurance industry, by means of a method of maximum likelihood estimation **subject to constraints**. This scenario has become extremely important, not only for application in the actuarial context, but also in other fields.

In this thesis, the analysis takes account of the actual lifetime (duration) of the policy rather than just recording the fact that the policy lapsed or was still alive after (say) twelve months. In other words, the response variable, lifetime, is a continuous one, and the whole distribution of lifetimes can be used. A **general experimental design** admits that all the policies can be used in the analysis, even those policies with inception dates very close to the cut-off point.

Special attention has been given to **staggered entry** of policies, where policies written in different months or time-periods have different entry times.

The methodology of maximum likelihood estimation subject to constraints, used in this thesis, leads to **explicit expressions** for the estimates of the parameters, as well as for

approximated variances and covariances of the estimates, which gives **exact** maximum likelihood estimates of the parameters. This makes direct extension to more complex designs feasible.

Once the parameters of the survival distributions have been estimated, estimated hazard and survivor functions, odds of a lapse, **odds ratios** and **hazard ratios** at time t can be directly calculated, as well as estimated percentiles for the fitted survival distributions. These estimates form the statistical foundation for scientific decisionmaking with respect to actuarial design, maintenance and marketing of insurance policies.

Parametric regression models are fitted and important indicators of the effect of the covariates are defined such as risk scores (hazard ratios) and indices (odds ratios). This is in contrast to the famous semiparametric Cox's proportional hazards model. David Oakes states in the chapter on Survival Analysis in [39] that "following the incorporation of software for fitting proportional hazards models into packages such as BMDP and SAS, this model of Cox, for better or worse, became standard for the analysis of survival data. But the assumption of proportional hazards has no compelling mathematical justification and is often found to be false in applications."

It is generally assumed in the actuarial industry that incorrect assumptions regarding lifetime distributions have severe implications with respect to lapse probabilities and estimated income. For this reason non-parametric and semi-parametric models are standard practice. The implications of a proper and sound parametric model are far-reaching for the characterization of lapse probabilities and income, which are the corner stones of the insurance industry. It directly determines lapse indices in terms of risk factors, including the period of lapsing under consideration. It also determines the behaviour of such indices over time. It can also be assumed that the use of grouped data for determining lifetime distributions is more robust with respect to wrong assumptions than continuous data.

In this way a contribution is made to the global handling of lapse indices and risk scores. For example, if a log-logistic distribution holds, the indices are constant, independent of the lapsing period. If a Weibull distribution holds, the risk scores are constant, independent of the lapsing period. In any event, if the lifetime distribution is known, the lapse indices for any time period are known.

A complete exposition of these structural relationships and practical implications of it must be investigated further.

Although the methodology in this thesis is developed specifically for the insurance industry, it may be applied in the normal context of research and scientific decisionmaking, that includes for example survival distributions for the **medical, biological, engineering, econometric and sociological sciences**.

The potential for extending the methodology to other realistic practical application is unlimited. This can be to the advance of the insurance industry in general. The models should reflect an interactive adaptability for direct application in practice by salesforce (the marketing people on ground level), as well as for actuarial planning.

Appendix A

COMPUTER PROGRAMS

The SAS/IML programs appear under the appropriate chapter heading.

A.1 Chapter 3: Maximum Likelihood Estimation

A.1.1 A Fixed Censoring Time Standard Program using PROC LIFEREG

Program for fitting a single survival distribution to grouped survival data

```
options nodate pagesize=500 pageno=1;
libname hsbc 'c:\hsbc1\sd2';

title1 'Fitting of a single survival model: the standard SAS method';
title2 'Fixed censoring time';

data fin;
input lower upper freq;
cards;
 .   12    66
12   17   158
17   24   254
24   28   157
28   34   250
34   37    35
37    .  1666
;

proc lifereg data=fin covout outest=ops;
model (lower, upper) =  / dist=weibull;       *default is Weibull;
model (lower, upper) =  / dist=llogistic;
model (lower, upper) =  / dist=lnormal;
weight freq;
output out=weib cdf=cdf predicted=months
quantiles = 0.02 to 0.98 by .02 control=c;
title 'Fit single Weibull curve (SAS method)';
title2 'Fixed censoring time';

data par;
set ops;
if _N_=1;
lambdaSAS=exp(-intercept/_scale_);
```

```
alphaSAS=1/_scale_;
oualphaSAS=-intercept/_scale_;

proc print data=par;
var lambdaSAS oualphaSAS alphaSAS;
run;
```

A.1.2 Staggered Entry of Policies Standard Program using PROC LIFEREG

Program for fitting one Weibull/log-logistic/lognormal distribution to the four histograms of the entry groups

```
options nodate pagesize=500 pageno=1;
libname hsbc 'c:\hsbc1\sd2';

title1 'STAGGERED ENTRY OF POLICIES AT FOUR ENTRY TIMES';
title2 'Fit one survival distribution to the four histograms of the entry groups';
title3 'Standard SAS method';

data fin;
input lower upper freq;
cards;
 .   12     66
12   17    158
17   24    254
24   28    157
28   34    250
34   37     35
37    .   1666
 .   12    118
12   17    166
17   24    229
24   28    200
28   34    172
34    .   1924
 .   12    154
12   17     99
17   24    242
24   28    117
28    .   1674
 .   12    175
12   17    166
17   24    207
24    .   1848
;

proc lifereg data=fin covout outest=ops;
model (lower, upper) =  / dist=weibull;       *default is Weibull;
model (lower, upper) =  / dist=llogistic;
model (lower, upper) =  / dist=lnormal;
weight freq;
output out=weib cdf=cdf predicted=months
quantiles = 0.02 to 0.98 by .02 control=c;
title 'Fit single Weibull curve (SAS method)';
title2 'Four entry dates';

data par;
set ops;
if _N_=1;
lambdaSAS=exp(-intercept/_scale_);
alphaSAS=1/_scale_;
oualphaSAS=-intercept/_scale_;

proc print data=par;
var lambdaSAS oualphaSAS alphaSAS;
run;
```

A.2 Chapter 3: M L Estimation subject to Constraints

A.2.1 A Fixed Censoring Time - IML Programs

1. **Program for fitting a Weibull distribution to grouped survival data**

```
proc iml worksize=60;
reset nolog;

**********Frequency vector;
f={66,158,254,157,250,35,1666};

**********Vector of upper boundaries;
x={12,17,24,28,34,37};

**********Relative frequency vector;
n=f[+];
k=nrow(f);
d=k-1;
p=f/n;

**********Design matrix and matrix orthogonal to design matrix;

S1=J(d,1,1)@cusum(J(1,k,1));
S2=J(1,k,1)@cusum(J(d,1,1));
S=S1<=S2;  print S;

X1=J(d,1,1)||log(x);        print X1;
C=I(d)-X1*inv(X1`*X1)*X1`;  print C;
projmatrix=X1*inv(X1`*X1)*X1`;  print projmatrix;

**********ITERATIVE PROCEDURE (double iterations over m and p);
*****starting value for m;
m=p;  print p;
ms=S*m; ps=ms;
p0=p;

*****iteration over m;
itr=0;
verskil1=1;
i=0;
do while (verskil1>1e-6);
i=i+1;
p=p0;
Gm=-C*diag(1/(log(1-ms)))*diag(1/(1-ms))*S;              *Weibull;

**********iteration over p;
     verskil=1;
     j=0;
     do while (verskil>1e-6);
     j=j+1;
     p1=p;
     ps=S*p;
     g=C*log(-log(1-ps));                                *Weibull;
     Gp=-C*diag(1/(log(1-ps)))*diag(1/(1-ps))*S;         *Weibull;
*********covariance matrix;
sig=(1/n)*(diag(m)-m*m`);
V=sig;
***************************;
     p=p-(Gm*V)`*ginv(Gp*V*Gm`)*g;                       *Weibull;
     verskil=sqrt((p-p1)`*(p-p1));
     print i j p m;
     end;
verskil1=sqrt((p-m)`*(p-m));
m=p;ms=S*m;     print m;
end;

**********Parameter vector for linear model;
par=inv(x1`*x1)*x1`*log(-log(1-ms));  print par;         *Weibull;

**********Parameters for Weibull model;
lambda=exp(par[1]);
alpha=par[2];

print 'Weibull parameters: MLE subject to constraints';
print 'lambda=' lambda  'alpha=' alpha;

**********Compute Wald statistic;
p=p0;
ps=S*p;
Gp=-C*diag(1/(log(1-ps)))*diag(1/(1-ps))*S;              *Weibull;
g=C*log(-log(1-ps));                                     *Weibull;
******covariance matrix;
sig=(1/n)*(diag(p)-p*p`);
```

```
V=sig;
************************;
wald=g'*ginv(Gp*V*Gp')*g; nu=eigval(C); nu=nu[+];
discr=wald/n;
prob=1-probchi(wald,nu) ;
alpha=par[2];
Gini=1-0.5##(1/alpha);
print 'Measure of fit';
print 'Wald='wald 'Discrepancy=' discr;
print 'prob=' prob 'degrees of freedom=' nu 'Gini=' Gini;
```

2. **Program for fitting a log-logistic distribution to grouped survival data**

```
proc iml worksize=60;
reset nolog;

**********Frequency vector;
f={66,158,254,157,250,35,1666};

**********Vector of upper boundaries;
x={12,17,24,28,34,37};

**********Relative frequency vector;
n=f[+];
k=nrow(f);
d=k-1;
p=f/n;

**********Design matrix and matrix orthogonal to design matrix;

S1=J(d,1,1)@cusum(J(1,k,1));
S2=J(1,k,1)@cusum(J(d,1,1));
S=S1<=S2;  print S;

X1=J(d,1,1)||log(x);         print X1;
C=I(d)-X1*inv(X1'*X1)*X1';   print C;
projmatrix=X1*inv(X1'*X1)*X1';  print projmatrix;

**********ITERATIVE PROCEDURE (double iterations over m and p);
*****starting value for m;
m=p;  print p;
ms=S*m; ps=ms;
p0=p;

*****iteration over m;
itr=0;
verskil1=1;
i=0;
do while (verskil1>0.00000001);
i=i+1;
p=p0;
Gm=C*(diag(1/ms)+diag(1/(1-ms)))*S;                        *loglogistic;

**********iteration over p;
     verskil=1;
     j=0;
     do while (verskil>0.00000001);
     j=j+1;
     p1=p;
     ps=S*p;
     g=C*(log(ps)-log(1-ps));                              *loglogistic;
     Gp=C*(diag(1/ps)+diag(1/(1-ps)))*S;                   *loglogistic;
*********covariance matrix;
sig=(1/n)*(diag(m)-m*m');
V=sig;
**************************;
     p=p-(Gm*V)'*ginv(Gp*V*Gm')*g;                         *loglogistic;
     verskil=sqrt((p-p1)'*(p-p1));
     print i j p m;
     end;
verskil1=sqrt((p-m)'*(p-m));
m=p;ms=S*m;     print m;
end;

**********Parameter vector for linear model;
par=inv(x1'*x1)*x1'*(log(ms)-log(1-ms));  print par;       *loglogistic;

**********Parameters for loglogistic model;
lambda=exp(par[1]);
alpha=par[2];

print 'Loglogistic parameters: MLE subject to constraints';
print 'lambda=' lambda  'alpha=' alpha;

**********Compute Wald statistic;
p=p0;
ps=S*p;
Gp=C*(diag(1/ps)+diag(1/(1-ps)))*S;                        *loglogistic;
g=C*(log(ps)-log(1-ps));                                   *loglogistic;
******covariance matrix;
```

```
sig=(1/n)*(diag(p)-p*p');
V=sig;
*************************;
wald=g'*ginv(Gp*V*Gp')*g; nu=eigval(C); nu=nu[+];
discr=wald/n;
prob=1-probchi(wald,nu) ;
alpha=par[2];
Gini=1-0.5##(1/alpha);
print 'Measure of fit';
print 'Wald='wald 'Discrepancy=' discr;
print 'prob=' prob 'degrees of freedom=' nu 'Gini=' Gini;
```

3. Program for fitting a lognormal distribution to grouped survival data

```
proc iml worksize=60;
reset nolog;

**********Frequency vector;
f={66,158,254,157,250,35,1666};

**********Vector of upper boundaries;
x={12,17,24,28,34,37};
x=log(x);

**********Relative frequency vector;
n=f[+];
k=nrow(f);
d=k-1;
p=f/n;

pi=(gamma(0.5))##2;

**********Design matrix and matrix orthogonal to design matrix;

S1=J(d,1,1)@cusum(J(1,k,1));
S2=J(1,k,1)@cusum(J(d,1,1));
S=S1<=S2;  print S;

X1=J(d,1,1)||log(x);         print X1;
C=I(d)-X1*inv(X1'*X1)*X1';  print C;
projmatrix=X1*inv(X1'*X1)*X1';  print projmatrix;

**********ITERATIVE PROCEDURE (double iterations over m and p);
*****starting value for m;
m=p;  print p;
ms=S*m; ps=ms;
p0=p;

*****iteration over m;
itr=0;
verskil1=1;
i=0;
do while (verskil1>1e-6);
i=i+1;
p=p0;
arg1=2;
arg2=3;
par=inv( J(2,1,1)||probit(ms[arg1]//ms[arg2]) )*(x[arg1]//x[arg2]);
mu=par[1];sigma=par[2];
Gm=C*diag(sqrt(2#pi)/(exp(-(x-mu)#(x-mu)/2/sigma/sigma)))*S; *lognormal;

**********iteration over p;
     verskil=1;
     j=0;
     do while (verskil>1e-6);
     j=j+1;
     p1=p;
     ps=S*p;
     g=C*probit(ps);                                        *lognormal;
     parp=inv( J(2,1,1)||probit(ps[arg1]//ps[arg2]) )*(x[arg1]//x[arg2]);
     mup=parp[1];sigmap=parp[2];
     Gp=C*diag(sqrt(2#pi)/(exp(-(x-mup)#(x-mup)/2/sigmap/sigmap)))*S; *lognormal;
*****covariance matrix;
       sig=(1/n)*(diag(m)-m*m');
       V=sig;
**********************;
     p=p-(Gm*V)'*ginv(Gp*V*Gm')*g;                          *lognormal;
     verskil=sqrt((p-p1)'*(p-p1));
     end;
verskil1=sqrt((p-m)'*(p-m));
m=p;ms=S*m;      print m;
end;

**********Parameter vector for linear model;
par=inv( J(2,1,1)||probit(ms[arg1]//ms[arg2]) )*(x[arg1]//x[arg2]);

**********Parameters for Lognormal model;
mu=par[1]; sigma=par[2];

print 'Lognormal parameters: MLE subject to constraints';
```

```
print 'mu=' mu  'sigma=' sigma;

**********Compute Wald statistic;
p=p0;
ps=S*p;
  Gp=C*diag(sqrt(2#pi)/(exp(-(x-mup)#(x-mup)/2/sigmap/sigmap)))*S;
  g=C*probit(ps);
*****covariance matrix;
sig=(1/n)*(diag(p)-p*p`);
V=sig;
**********************;
wald=(g)`*ginv(Gp*V*Gp`)*g; nu=eigval(C); nu=nu[+];
discr=wald/n;
prob=1-probchi(wald,nu) ;
print 'Measure of fit';
print 'Wald='wald 'Discrepancy=' discr;
print 'prob=' prob 'degrees of freedom=' nu;
```

A.2.2 Staggered Entry of Policies - IML Programs

1. **Program for fitting one survival distribution to the four histograms**

Program for fitting one Weibull/log-logistic distribution to the four histograms of the entry groups

```
title1 'STAGGERED ENTRY OF POLICIES AT FOUR ENTRY TIMES';
title2 'Fit one survival distribution to the four histograms of the entry groups';
title3 'Constraints: specified model';

proc iml worksize= 60;
reset nolog;
options pagesize=500;

**********Frequency vector;
f1={66,158,254,157,250,35,1666};

**********Vector of upper boundaries;
x1={12,17,24,28,34,37};

**********Frequency vector;
f2={118,166,229,200,172,1924};

**********Vector of upper boundaries;
x2={12,17,24,28,34};

**********Frequency vector;
f3={154,99,242,117,1674};

**********Vector of upper boundaries;
x3={12,17,24,28};

**********Frequency vector;
f4={175,166,207,1848};

**********Vector of upper boundaries;
x4={12,17,24};

**********Relative frequency vectors;
n1=f1[+]; n2=f2[+]; n3=f3[+]; n4=f4[+]; n=n1+n2+n3+n4;
k1=nrow(f1); d1=k1-1;
k2=nrow(f2); d2=k2-1;
k3=nrow(f3); d3=k3-1;
k4=nrow(f4); d4=k4-1;
k=k1+k2+k3+k4;
d=d1+d2+d3+d4;
p1=f1/n1; p2=f2/n2; p3=f3/n3; p4=f4/n4;
p=p1//p2//p3//p4;

**********Design matrix and matrix orthogonal to design matrix;

S1=J(d1,1,1)@cusum(J(1,k1,1));
S2=J(1,k1,1)@cusum(J(d1,1,1));
S1=S1<=S2;
S2=S1[1:d2,1:d1];
S3=S1[1:d3,1:d2];
S4=S1[1:d4,1:d3];
S=block(S1,S2,S3,S4);

lx1=J(d1,1,1)||log(x1);
lx2=J(d2,1,1)||log(x2);
lx3=J(d3,1,1)||log(x3);
lx4=J(d4,1,1)||log(x4);

xc=lx1//lx2//lx3//lx4;
C=I(d)-xc*inv(xc`*xc)*xc`;

**********ITERATIVE PROCEDURE (double iterations over m and p);
*****starting value for m;
m=p;  print p;
ms=S*m; ps=ms;
p0=p;

*****iteration over m;
itr=0;
verskil1=1;
i=0;
do while (verskil1>1e-6);
i=i+1;
p=p0;
Gm=-C*diag(1/(log(1-ms)))*diag(1/(1-ms))*S;            *Weibull;
*Gm=C*(diag(1/ms)+diag(1/(1-ms)))*S;                    *Loglogistic;

**********iteration over p;
     verskil=1;
```

```
      j=0;
      do while (verskil>1e-6);
      j=j+1;
      p1=p;
      ps=S*p;
      g=C*log(-log(1-ps));                                    *Weibull;
     *g=C*(log(ps)-log(1-ps));                                *Loglogistic;
      Gp=-C*diag(1/(log(1-ps)))*diag(1/(1-ps))*S;             *Weibull;
     *Gp=C*(diag(1/ps)+diag(1/(1-ps)))*S;                     *loglogistic;
**********covariance matrix;
m1=m[1:k1];
m2=m[k1+1:k1+k2];
m3=m[k1+k2+1:k1+k2+k3];
m4=m[k1+k2+k3+1:k1+k2+k3+k4];

sig1=(1/n1)*(diag(m1)-m1*m1`);
sig2=(1/n2)*(diag(m2)-m2*m2`);
sig3=(1/n3)*(diag(m3)-m3*m3`);
sig4=(1/n4)*(diag(m4)-m4*m4`);
sig=block(sig1,sig2,sig3,sig4);
V=sig;
****************************;
      p=p-(Gm*V)`*ginv(Gp*V*Gm`)*g;                           *Weibull;
     *p=p-(Gm*V)`*ginv(Gp*V*Gm`)*g;                           *loglogistic;
      verskil=sqrt((p-p1)`*(p-p1));
      end;
verskil1=sqrt((p-m)`*(p-m));
m=p;ms=S*m;
end;
print m; print i j;

**********Parameter vector for linear model;
par=inv(xc`*xc)*xc`*log(-log(1-ms));                           *Weibull;
*par=inv(xc`*xc)*xc`*(log(ms)-log(1-ms));                      *Loglogisties;
print par;

**********Parameters for Weibull(*Loglogistic) model;
oualpha=par[1];
lambda=exp(par[1]);
alpha=par[2];

print 'Weibull(*Loglogistic) parameters: MLE subject to constraints';
print 'lambda=' lambda oualpha 'alpha=' alpha;

**********Hazard and Survival function, Odds***********************************;
whaz12=(lambda*alpha*12**(alpha-1))/(1+lambda*12**alpha);
whaz24=(lambda*alpha*24**(alpha-1))/(1+lambda*24**alpha);
wsurv12=(1+lambda*12**alpha)**(-1);
wsurv24=(1+lambda*24**alpha)**(-1);
wodds12=(1-wsurv12)/wsurv12;
wodds24=(1-wsurv24)/wsurv24;

           /*
llhaz12=(lambda*alpha*12**(alpha-1))/(1+lambda*12**alpha);
llhaz24=(lambda*alpha*24**(alpha-1))/(1+lambda*24**alpha);
llsurv12=(1+lambda*12**alpha)**(-1);
llsurv24=(1+lambda*24**alpha)**(-1);
llodds12=(1-llsurv12)/llsurv12;
llodds24=(1-llsurv24)/llsurv24;
           */

print whaz12 whaz24 wsurv12 wsurv24 wodds12 wodds24;
*print llhaz12 llhaz24 llsurv12 llsurv24 llodds12 llodds24;

**********Percentiles**********************************************************;
wmedian=((1/lambda)#log(2))##(1/alpha);

 wperc5=((1/lambda)#log(100/(100- 5)))##(1/alpha);
wperc10=((1/lambda)#log(100/(100-10)))##(1/alpha);
wperc20=((1/lambda)#log(100/(100-20)))##(1/alpha);
wperc25=((1/lambda)#log(100/(100-25)))##(1/alpha);
wperc30=((1/lambda)#log(100/(100-30)))##(1/alpha);
wperc40=((1/lambda)#log(100/(100-40)))##(1/alpha);
wperc50=((1/lambda)#log(100/(100-50)))##(1/alpha);
wperc60=((1/lambda)#log(100/(100-60)))##(1/alpha);
wperc70=((1/lambda)#log(100/(100-70)))##(1/alpha);
wperc75=((1/lambda)#log(100/(100-75)))##(1/alpha);
wperc80=((1/lambda)#log(100/(100-80)))##(1/alpha);
wperc90=((1/lambda)#log(100/(100-90)))##(1/alpha);
wperc95=((1/lambda)#log(100/(100-95)))##(1/alpha);

                /*
lmedian=(1/lambda)##(1/alpha);

 llperc5=((1/lambda)#( 5/(100- 5)))##(1/alpha);
llperc10=((1/lambda)#(10/(100-10)))##(1/alpha);
llperc20=((1/lambda)#(20/(100-20)))##(1/alpha);
llperc25=((1/lambda)#(25/(100-25)))##(1/alpha);
llperc30=((1/lambda)#(30/(100-30)))##(1/alpha);
```

```
llperc40=((1/lambda)#(40/(100-40)))##(1/alpha);
llperc50=((1/lambda)#(50/(100-50)))##(1/alpha);
llperc60=((1/lambda)#(60/(100-60)))##(1/alpha);
llperc70=((1/lambda)#(70/(100-70)))##(1/alpha);
llperc75=((1/lambda)#(75/(100-75)))##(1/alpha);
llperc80=((1/lambda)#(80/(100-80)))##(1/alpha);
llperc90=((1/lambda)#(90/(100-90)))##(1/alpha);
llperc95=((1/lambda)#(95/(100-95)))##(1/alpha);

                 */

print wmedian;
print wperc5  wperc10 wperc20 wperc25 wperc30;
print wperc40 wperc50 wperc60 wperc70;
print wperc75 wperc80 wperc90 wperc95;
*print lmedian;
*print llperc5  llperc10 llperc20 llperc25 llperc30;
*print llperc40 llperc50 llperc60 llperc70;
*print llperc75 llperc80 llperc90 llperc95;

**********Compute Wald statistic*********************************************;
p=p0;
ps=S*p;

Gp=-C*diag(1/(log(1-ps)))*diag(1/(1-ps))*S;               *Weibull;
*Gp=C*diag(1/ps+1/(1-ps))*S;                              *Loglogistic;
g=C*log(-log(1-ps));                                      *Weibull;
*g=C*(log(ps)-log(1-ps));                                 *Loglogistic;

******covariance matrix;
p1=p[1:k1];
p2=p[k1+1:k1+k2];
p3=p[k1+k2+1:k1+k2+k3];
p4=p[k1+k2+k3+1:k1+k2+k3+k4];

sig1=(1/n1)*(diag(p1)-p1*p1`);
sig2=(1/n2)*(diag(p2)-p2*p2`);
sig3=(1/n3)*(diag(p3)-p3*p3`);
sig4=(1/n4)*(diag(p4)-p4*p4`);
sig=block(sig1,sig2,sig3,sig4);
V=sig;
************************;
wald=g`*ginv(Gp*V*Gp`)*g; nu=eigval(C); nu=nu[+];
discr=wald/n;
prob=1-probchi(wald,nu) ;
alpha=par[2];
Gini=1-0.5##(1/alpha);
print 'Measure of fit';
print 'Wald='wald 'Discrepancy=' discr;
print 'prob=' prob 'degrees of freedom=' nu 'Gini=' Gini;
```

Program for fitting one lognormal distribution to the four histograms of the entry groups

```
title1 'STAGGERED ENTRY OF POLICIES AT FOUR ENTRY TIMES';
title2 'Fit one lognormal distribution to the four entry groups';
title3 'Constraints: specified model';

proc iml worksize= 60;
reset nolog;
options pagesize=500;

**********Frequency vector;
f1={66,158,254,157,250,35,1666};

**********Vector of upper boundaries;
x1={12,17,24,28,34,37};

**********Frequency vector;
f2={118,166,229,200,172,1924};

**********Vector of upper boundaries;
x2={12,17,24,28,34};

**********Frequency vector;
f3={154,99,242,117,1674};

**********Vector of upper boundaries;
x3={12,17,24,28};

**********Frequency vector;
f4={175,166,207,1848};

**********Vector of upper boundaries;
x4={12,17,24};

**********Relative frequency vectors;
n1=f1[+]; n2=f2[+]; n3=f3[+]; n4=f4[+]; n=n1+n2+n3+n4;
k1=nrow(f1); d1=k1-1;
```

```
k2=nrow(f2); d2=k2-1;
k3=nrow(f3); d3=k3-1;
k4=nrow(f4); d4=k4-1;
k=k1+k2+k3+k4;
d=d1+d2+d3+d4;
p1=f1/n1; p2=f2/n2; p3=f3/n3; p4=f4/n4;
p=p1//p2//p3//p4;

**********Design matrix and matrix orthogonal to design matrix;

S1=J(d1,1,1)@cusum(J(1,k1,1));
S2=J(1,k1,1)@cusum(J(d1,1,1));
S1=S1<=S2;
S2=S1[1:d2,1:d1];
S3=S1[1:d3,1:d2];
S4=S1[1:d4,1:d3];
S=block(S1,S2,S3,S4);    print S;

lx1=J(d1,1,1)||log(x1);
lx2=J(d2,1,1)||log(x2);
lx3=J(d3,1,1)||log(x3);
lx4=J(d4,1,1)||log(x4);

xc=lx1//lx2//lx3//lx4;      print xc;
C=I(d)-xc*inv(xc`*xc)*xc`;

**********ITERATIVE PROCEDURE (double iterations over m and p);
*****starting value for m;
m=p;  print p;
ms=S*m; ps=ms;
p0=p;

*****iteration over m;
itr=0;
verskil1=1;
i=0;
do while (verskil1>1e-6);
i=i+1;
p=p0;
arg1=2;
arg2=3;
par=inv( J(2,1,1)||probit(ms[arg1]//ms[arg2]) )*(x[arg1]//x[arg2]);
mu=par[1];sigma=par[2];
Gm=C*diag(sqrt(2#pi)/(exp(-(x-mu)#(x-mu)/2/sigma/sigma)))*S;          *lognormal;

**********iteration over p;
     verskil=1;
     j=0;
     do while (verskil>1e-6);
     j=j+1;
     p1=p;
     ps=S*p;
     g=C*probit(ps);                                             *lognormal;
     parp=inv( J(2,1,1)||probit(ps[arg1]//ps[arg2]) )*(x[arg1]//x[arg2]);
     mup=parp[1];sigmap=parp[2];
     Gp=C*diag(sqrt(2#pi)/(exp(-(x-mup)#(x-mup)/2/sigmap/sigmap)))*S; *lognormal;
**********covariance matrix;
m1=m[1:k1];
m2=m[k1+1:k1+k2];
m3=m[k1+k2+1:k1+k2+k3];
m4=m[k1+k2+k3+1:k1+k2+k3+k4];

sig1=(1/n1)*(diag(m1)-m1*m1`);
sig2=(1/n2)*(diag(m2)-m2*m2`);
sig3=(1/n3)*(diag(m3)-m3*m3`);
sig4=(1/n4)*(diag(m4)-m4*m4`);
sig=block(sig1,sig2,sig3,sig4);
V=sig;
****************************;
     p=p-(Gm*V)`*ginv(Gp*V*Gm`)*g;                               *lognormal;
     verskil=sqrt((p-p1)`*(p-p1));
     end;
verskil1=sqrt((p-m)`*(p-m));
m=p;ms=S*m;
end;
print m; print i j;

**********Parameter vector for linear model;
par=inv( J(2,1,1)||probit(ms[arg1]//ms[arg2]) )*(x[arg1]//x[arg2]);  *lognormal;

**********Parameters for Lognormal model;
mu=par[1]; sigma=par[2];

print 'Lognormal parameters: MLE subject to constraints';
print 'mu=' mu  'sigma=' sigma;

**********Compute Wald statistic;
p=p0;
ps=S*p;
```

```
Gp=C*diag(sqrt(2#pi)/(exp(-(x-mup)#(x-mup)/2/sigmap/sigmap)))*S;    *lognormal;
g=C*probit(ps);                                                     *lognormal;
******covariance matrix;
p1=p[1:k1];
p2=p[k1+1:k1+k2];
p3=p[k1+k2+1:k1+k2+k3];
p4=p[k1+k2+k3+1:k1+k2+k3+k4];

sig1=(1/n1)*(diag(p1)-p1*p1`);
sig2=(1/n2)*(diag(p2)-p2*p2`);
sig3=(1/n3)*(diag(p3)-p3*p3`);
sig4=(1/n4)*(diag(p4)-p4*p4`);
sig=block(sig1,sig2,sig3,sig4);
V=sig;
************************;
wald=g`*ginv(Gp*V*Gp`)*g; nu=eigval(C); nu=nu[+];
discr=wald/n;
prob=1-probchi(wald,nu) ;
alpha=par[2];
Gini=1-0.5##(1/alpha);
print 'Measure of fit';
print 'Wald='wald 'Discrepancy=' discr;
print 'prob=' prob 'degrees of freedom=' nu 'Gini=' Gini;
```

2. **Programs for fitting four survival distributions to the four histograms and then set the four sets of parameters equal**

Program for fitting four Weibull/log-logistic distributions to the four histograms and then set the lambda's equal and the alpha's equal

```
title1 'STAGGERED ENTRY OF POLICIES AT FOUR ENTRY TIMES';
title2 'Fit four survival distributions to the four entry groups';
title3 'Restrictions: specified model AND set lambda's equal and alpha's equal';

proc iml worksize= 60;
reset nolog;
options pagesize=500;

**********Frequency vector;
f1={66,158,254,157,250,35,1666};

**********Vector of upper boundaries;
x1={12,17,24,28,34,37};

**********Frequency vector;
f2={118,166,229,200,172,1924};

**********Vector of upper boundaries;
x2={12,17,24,28,34};

**********Frequency vector;
f3={154,99,242,117,1674};

**********Vector of upper boundaries;
x3={12,17,24,28};

**********Frequency vector;
f4={175,166,207,1848};

**********Vector of upper boundaries;
x4={12,17,24};

**********Relative frequency vectors;
n1=f1[+]; n2=f2[+]; n3=f3[+]; n4=f4[+]; n=n1+n2+n3+n4;
k1=nrow(f1); d1=k1-1;
k2=nrow(f2); d2=k2-1;
k3=nrow(f3); d3=k3-1;
k4=nrow(f4); d4=k4-1;
k=k1+k2+k3+k4;
d=d1+d2+d3+d4;
p1=f1/n1; p2=f2/n2; p3=f3/n3; p4=f4/n4;
p=p1//p2//p3//p4;

**********Design matrix and matrix orthogonal to design matrix;

S1=J(d1,1,1)@cusum(J(1,k1,1));
S2=J(1,k1,1)@cusum(J(d1,1,1));
S1=S1<=S2;
S2=S1[1:d2,1:d1];
S3=S1[1:d3,1:d2];
S4=S1[1:d4,1:d3];
S=block(S1,S2,S3,S4);   print S;

lx1=J(d1,1,1)||log(x1);
lx2=J(d2,1,1)||log(x2);
lx3=J(d3,1,1)||log(x3);
```

```
lx4=J(d4,1,1)||log(x4);

xc=block(lx1,lx2,lx3,lx4);
C=I(d)-xc*inv(xc'*xc)*xc';
CP=C;
C=C//({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} *inv(xc'*xc)*xc');
print xc C;

**********ITERATIVE PROCEDURE (double iterations over m and p);
*****starting value for m;
m=p;  print p;
ms=S*m; ps=ms;
p0=p;

*****iteration over m;
itr=0;
verskil1=1;
i=0;
do while (verskil1>1e-6);
i=i+1;
p=p0;
Gm=-C*diag(1/(log(1-ms)))*diag(1/(1-ms))*S;                    *Weibull;
*Gm=C*(diag(1/ms)+diag(1/(1-ms)))*S;                           *Loglogistic;

**********iteration over p;
     verskil=1;
     j=0;
     do while (verskil>1e-6);
     j=j+1;
     p1=p;
     ps=S*p;
     g=C*log(-log(1-ps));                                       *Weibull;
    *g=C*(log(ps)-log(1-ps));                                   *Loglogistic;
     Gp=-C*diag(1/(log(1-ps)))*diag(1/(1-ps))*S;                *Weibull;
    *Gp=C*(diag(1/ps)+diag(1/(1-ps)))*S;                        *Loglogistic;
**********covariance matrix;
m1=m[1:k1];
m2=m[k1+1:k1+k2];
m3=m[k1+k2+1:k1+k2+k3];
m4=m[k1+k2+k3+1:k1+k2+k3+k4];

sig1=(1/n1)*(diag(m1)-m1*m1');
sig2=(1/n2)*(diag(m2)-m2*m2');
sig3=(1/n3)*(diag(m3)-m3*m3');
sig4=(1/n4)*(diag(m4)-m4*m4');
sig=block(sig1,sig2,sig3,sig4);
V=sig;
****************************;
     p=p-(Gm*V)'*ginv(Gp*V*Gm')*g;                              *Weibull;
    *p=p-(Gm*V)'*ginv(Gp*V*Gm')*g;                              *Loglogistic;
     verskil=sqrt((p-p1)'*(p-p1));
     end;
verskil1=sqrt((p-m)'*(p-m));
m=p;ms=S*m;
end;
print m; print i j;

**********Parameter vector for linear model;
par=inv(xc'*xc)*xc'*log(-log(1-ms));  print par;               *Weibull;
*par=inv(xc'*xc)*xc'*(log(ms)-log(1-ms));                       *Loglogistic;

**********Parameters for Weibull(*Loglogistic) model;
oualpha=par[1];
lambda=exp(par[1]);
alpha=par[2];

print 'Weibull(*Loglogistic) parameters: MLE subject to constraints';
print 'lambda=' lambda oualpha 'alpha=' alpha;

**********Compute Wald statistic;
p=p0;
ps=S*p;

Gp=-C*diag(1/(log(1-ps)))*diag(1/(1-ps))*S;                    *Weibull;
*Gp=C*diag(1/ps+1/(1-ps))*S;                                    *Loglogistic;
g=C*log(-log(1-ps));                                            *Weibull;
*g=C*(log(ps)-log(1-ps));                                       *Loglogistic;

******covariance matrix;
p1=p[1:k1];
p2=p[k1+1:k1+k2];
p3=p[k1+k2+1:k1+k2+k3];
p4=p[k1+k2+k3+1:k1+k2+k3+k4];

sig1=(1/n1)*(diag(p1)-p1*p1');
sig2=(1/n2)*(diag(p2)-p2*p2');
sig3=(1/n3)*(diag(p3)-p3*p3');
sig4=(1/n4)*(diag(p4)-p4*p4');
```

```
sig=block(sig1,sig2,sig3,sig4);
V=sig;
*************************;
wald=g`*ginv(Gp*V*Gp`)*g; nu=eigval(CP); nu=nu[+]+4;
discr=wald/n;
prob=1-probchi(wald,nu) ;
alpha=par[2];
Gini=1-0.5##(1/alpha);
print 'Measure of fit';
print 'Wald='wald 'Discrepancy=' discr;
print 'prob=' prob 'degrees of freedom=' nu 'Gini=' Gini;
```

Program for fitting four lognormal distributions to the four histograms and then set the mu's equal and the sigma's equal

```
title1 'STAGGERED ENTRY OF POLICIES AT FOUR ENTRY TIMES';
title2 'Fit four lognormal distributions to the four entry groups';
title3 'Restrictions: specified model AND set mu's equal and sigma's equal';

proc iml worksize= 60;
reset nolog;
options pagesize=500;

**********Frequency vector;
f1={66,158,254,157,250,35,1666};

**********Vector of upper boundaries;
x1={12,17,24,28,34,37};

**********Frequency vector;
f2={118,166,229,200,172,1924};

**********Vector of upper boundaries;
x2={12,17,24,28,34};

**********Frequency vector;
f3={154,99,242,117,1674};

**********Vector of upper boundaries;
x3={12,17,24,28};

**********Frequency vector;
f4={175,166,207,1848};

**********Vector of upper boundaries;
x4={12,17,24};

**********Relative frequency vectors;
n1=f1[+]; n2=f2[+]; n3=f3[+]; n4=f4[+]; n=n1+n2+n3+n4;
k1=nrow(f1); d1=k1-1;
k2=nrow(f2); d2=k2-1;
k3=nrow(f3); d3=k3-1;
k4=nrow(f4); d4=k4-1;
k=k1+k2+k3+k4;
d=d1+d2+d3+d4;
p1=f1/n1; p2=f2/n2; p3=f3/n3; p4=f4/n4;
p=p1//p2//p3//p4;

**********Design matrix and matrix orthogonal to design matrix;

S1=J(d1,1,1)@cusum(J(1,k1,1));
S2=J(1,k1,1)@cusum(J(d1,1,1));
S1=S1<=S2;
S2=S1[1:d2,1:d1];
S3=S1[1:d3,1:d2];
S4=S1[1:d4,1:d3];
S=block(S1,S2,S3,S4);    print S;

lx1=J(d1,1,1)||log(x1);
lx2=J(d2,1,1)||log(x2);
lx3=J(d3,1,1)||log(x3);
lx4=J(d4,1,1)||log(x4);

xc=block(lx1,lx2,lx3,lx4);
C=I(d)-xc*inv(xc`*xc)*xc`;
CP=C;
C=C//({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} *inv(xc`*xc)*xc`);
print xc C;

**********ITERATIVE PROCEDURE (double iterations over m and p);
*****starting value for m;
m=p;  print p;
ms=S*m; ps=ms;
p0=p;

*****iteration over m;
itr=0;
verskil1=1;
```

```
i=0;
do while (verskil1>1e-6);
i=i+1;
p=p0;
arg1=2;
arg2=3;
par=inv( J(2,1,1)||probit(ms[arg1]//ms[arg2]) )*(x[arg1]//x[arg2]);
mu=par[1];sigma=par[2];
Gm=C*diag(sqrt(2#pi)/(exp(-(x-mu)#(x-mu)/2/sigma/sigma)))*S;         *lognormal;

**********iteration over p;
     verskil=1;
     j=0;
     do while (verskil>1e-6);
     j=j+1;
     p1=p;
     ps=S*p;
     g=C*probit(ps);                                             *lognormal;
     parp=inv( J(2,1,1)||probit(ps[arg1]//ps[arg2]) )*(x[arg1]//x[arg2]);
     mup=parp[1];sigmap=parp[2];
     Gp=C*diag(sqrt(2#pi)/(exp(-(x-mup)#(x-mup)/2/sigmap/sigmap)))*S; *lognormal;
**********covariance matrix;
m1=m[1:k1];
m2=m[k1+1:k1+k2];
m3=m[k1+k2+1:k1+k2+k3];
m4=m[k1+k2+k3+1:k1+k2+k3+k4];

sig1=(1/n1)*(diag(m1)-m1*m1`);
sig2=(1/n2)*(diag(m2)-m2*m2`);
sig3=(1/n3)*(diag(m3)-m3*m3`);
sig4=(1/n4)*(diag(m4)-m4*m4`);
sig=block(sig1,sig2,sig3,sig4);
V=sig;
****************************;
     p=p-(Gm*V)`*ginv(Gp*V*Gm`)*g;                                *lognormal;
     verskil=sqrt((p-p1)`*(p-p1));
     end;
verskil1=sqrt((p-m)`*(p-m));
m=p;ms=S*m;
end;
print m; print i j;

**********Parameter vector for linear model;
par=inv( J(2,1,1)||probit(ms[arg1]//ms[arg2]) )*(x[arg1]//x[arg2]);  *lognormal;

**********Parameters for Lognormal model;
mu=par[1]; sigma=par[2];

print 'Lognormal parameters: MLE subject to constraints';
print 'mu=' mu  'sigma=' sigma;

**********Compute Wald statistic;
p=p0;
ps=S*p;

Gp=C*diag(sqrt(2#pi)/(exp(-(x-mup)#(x-mup)/2/sigmap/sigmap)))*S;   *lognormal;
g=C*probit(ps);                                                    *lognormal;
******covariance matrix;
p1=p[1:k1];
p2=p[k1+1:k1+k2];
p3=p[k1+k2+1:k1+k2+k3];
p4=p[k1+k2+k3+1:k1+k2+k3+k4];
sig1=(1/n1)*(diag(p1)-p1*p1`);
sig2=(1/n2)*(diag(p2)-p2*p2`);
sig3=(1/n3)*(diag(p3)-p3*p3`);
sig4=(1/n4)*(diag(p4)-p4*p4`);
sig=block(sig1,sig2,sig3,sig4);
V=sig;
*************************;
wald=g`*ginv(Gp*V*Gp`)*g; nu=eigval(CP); nu=nu[+]+4;
discr=wald/n;
prob=1-probchi(wald,nu) ;
alpha=par[2];
Gini=1-0.5##(1/alpha);
print 'Measure of fit';
print 'Wald='wald 'Discrepancy=' discr;
print 'prob=' prob 'degrees of freedom=' nu 'Gini=' Gini;
```

3. **Program for fitting a joint histogram to the four histograms of the entry groups**

```
title1 'STAGGERED ENTRY OF POLICIES AT FOUR ENTRY TIMES';
title2 'Fit a joint histogram to the four histograms of the entry groups';

proc iml worksize= 60;
reset nolog;
options pagesize=500;

**********Frequency vector;
f1={66,158,254,157,250,35,1666};
```

```
**********Vector of upper boundaries;
x1={12,17,24,28,34,37};

**********Frequency vector;
f2={118,166,229,200,172,1924};

**********Vector of upper boundaries;
x2={12,17,24,28,34};

**********Frequency vector;
f3={154,99,242,117,1674};

**********Vector of upper boundaries;
x3={12,17,24,28};

**********Frequency vector;
f4={175,166,207,1848};

**********Vector of upper boundaries;
x4={12,17,24};

**********Relative frequency vectors;
n1=f1[+]; n2=f2[+]; n3=f3[+]; n4=f4[+]; n=n1+n2+n3+n4;
k1=nrow(f1); d1=k1-1;
k2=nrow(f2); d2=k2-1;
k3=nrow(f3); d3=k3-1;
k4=nrow(f4); d4=k4-1;
k=k1+k2+k3+k4;
d=d1+d2+d3+d4;
p1=f1/n1; p2=f2/n2; p3=f3/n3; p4=f4/n4;
p=p1//p2//p3//p4;
**********Constraints imposed by the experimental design;
Gm=
(I(d4)     ||J(d4,1,0)||J(d4,1,0)||J(d4,1,0)||J(d4,1,0)||-I(d4)     ||J(d4,1,0)||
(I(d4)     ||J(d4,1,0)||J(d4,1,0)||J(d4,1,0)||J(d4,1,0)||J(d4,d4,0)||J(d4,1,0)||
(I(d4)     ||J(d4,1,0)||J(d4,1,0)||J(d4,1,0)||J(d4,1,0)||J(d4,d4,0)||J(d4,1,0)||
(J(1,d4,0)||         1||         1||         1||         1||J(1,d4,0) ||        -1||
(J(1,d4,0)||         1||         1||         1||         1||J(1,d4,0) ||         0||
(J(1,d4,0)||         1||         1||         1||         1||J(1,d4,0) ||         0||
(J(1,d4,0)||         0||         1||         0||         0||J(1,d4,0) ||         0||
(J(1,d4,0)||         1||         0||         0||         0||J(1,d4,0) ||        -1||

J(d4,1,0)||J(d4,1,0)||J(d4,d4,0)||J(d4,1,0)||J(d4,1,0)||J(d4,d4,0)||J(d4,1,0))//
J(d4,1,0)||J(d4,1,0)||-I(d4)     ||J(d4,1,0)||J(d4,1,0)||J(d4,d4,0)||J(d4,1,0))//
J(d4,1,0)||J(d4,1,0)||J(d4,d4,0)||J(d4,1,0)||J(d4,1,0)||-I(d4)     ||J(d4,1,0))//
       -1||       -1||J(1,d4,0) ||        0||        0||J(1,d4,0) ||        0)//
        0||        0||J(1,d4,0) ||       -1||       -1||J(1,d4,0) ||        0)//
        0||        0||J(1,d4,0) ||        0||        0||J(1,d4,0) ||       -1)//
       -1||        0||J(1,d4,0) ||        0||        0||J(1,d4,0) ||        0)//
        0||        0||J(1,d4,0) ||        0||        0||J(1,d4,0) ||        0);

*print Gm;

**********starting value for m;
m=p; Gp=Gm;
p0=p;
**********iteration over m;
verskil=1;
i=0;
do while (verskil>1e-6);
i=i+1;
p=p0;
g=Gm*p;
**********covariance matrix;
m1=m[1:k1];
m2=m[k1+1:k1+k2];
m3=m[k1+k2+1:k1+k2+k3];
m4=m[k1+k2+k3+1:k1+k2+k3+k4];

sig1=(1/n1)*(diag(m1)-m1*m1`);
sig2=(1/n2)*(diag(m2)-m2*m2`);
sig3=(1/n3)*(diag(m3)-m3*m3`);
sig4=(1/n4)*(diag(m4)-m4*m4`);
sig=block(sig1,sig2,sig3,sig4);
V=sig;
****************************;
p=p-(Gm*V)`*ginv(Gp*V*Gm`)*g;

*****Define joint frequencies;
p1=p[1:k1];
p2=p[k1+1:k1+k2];
p3=p[k1+k2+1:k1+k2+k3];
p4=p[k1+k2+k3+1:k1+k2+k3+k4];
p=p1//p2//p3//p4;

np1=n1#p1; np2=n2#p2; np3=n3#p3; np4=n4#p4; np=n#p;
****************************;

*print i p m np1 np2 np3 np4 np;
```

```
verskil=sqrt((p-m)`*(p-m));
m=p; m1=p1; m2=p2; m3=p3; m4=p4;
end;

*****Print frequencies of joint histogram;
print 'Frequencies of joint histogram=' np;
print 'Relative Frequencies of joint histogram=' p;

**********Compute Wald statistic;
******covariance matrix;
p1=p[1:k1];
p2=p[k1+1:k1+k2];
p3=p[k1+k2+1:k1+k2+k3];
p4=p[k1+k2+k3+1:k1+k2+k3+k4];

sig1=(1/n1)*(diag(p1)-p1*p1`);
sig2=(1/n2)*(diag(p2)-p2*p2`);
sig3=(1/n3)*(diag(p3)-p3*p3`);
sig4=(1/n4)*(diag(p4)-p4*p4`);
sig=block(sig1,sig2,sig3,sig4);
V=sig;
************************;
Gp=Gm;
wald=g`*ginv(Gp*V*Gp`)*g;
discr=wald/n;

print 'Wald=' wald 'Discrepancy=' discr;
```

A.3 Chapter 3: Simulation Studies

A.3.1 Program to simulate continuous right-censored lifetime data

1. **Program to generate continuous right-censored lifetime data from the Weibull(Loglogistic) distribution and to run simulations with the technique of MLE under constraints as well as the standard technique of MLE (1000 samples of size 100 from Weib(Logl)(0.15;0.5) - censored at 50)**

```
proc iml worksize=6000 symsize=2000;
reset noname nocenter;

**********Contents of module*************************************************;

start mod_est(x,f) global(lambda,alpha) ;

*****Relative frequencies;
n=100;
k=nrow(f);
d=k-1;
p=f/n;

*****Design matrix and matrix orthogonal to design matrix;
S1=J(d,1,1)@cusum(J(1,k,1));
S2=J(1,k,1)@cusum(J(d,1,1));
S=S1<=S2;

x1=J(d,1,1)||log(x);
C=I(d)-x1*inv(x1`*x1)*x1`;

**********ITERATIVE PROCEDURE (double iterations over m and p);
*****starting value for m;
  m=p;
  ms=S*m; ps=ms;
  p0=p;

*****iteration over m;
itr=0;
verskil1=1;
i=0;
do while (verskil1>1e-6);
i=i+1;
p=p0;
Gm=-C*diag(1/(log(1-ms)))*diag(1/(1-ms))*S;                *Weibull;
*Gm=C*(diag(1/ms)+diag(1/(1-ms)))*S;                        *Loglogistic;

*****iteration over p;
   verskil=1;
   j=0;
   do while (verskil>1e-6);
   j=j+1;
   p1=p;
   ps=S*p;
   g=C*log(-log(1-ps));                                     *Weibull;
  *g=C*(log(ps)-log(1-ps));                                 *Loglogistic;
   Gp=-C*diag(1/(log(1-ps)))*diag(1/(1-ps))*S;              *Weibull;
  *Gp=C*(diag(1/ps)+diag(1/(1-ps)))*S;                      *Loglogistic;
*****covariance matrix;
sig=(1/n)*(diag(m)-m*m`);
V=sig;
***********************;
   p=p-(Gm*V)`*ginv(Gp*V*Gm`)*g;                            *Weibull;
  *p=p-(Gm*V)`*ginv(Gp*V*Gm`)*g;                            *Loglogistic;
   verskil=sqrt((p-p1)`*(p-p1));
   end;
verskil1=sqrt((p-m)`*(p-m));
m=p;ms=S*m;
end;

******Parameter vector for linear model;
par=inv(x1`*x1)*x1`*log(-log(1-ms));                        *Weibull;
*par=inv(xc`*xc)*xc`*(log(ms)-log(1-ms));                   *Loglogistic;

**********Parameters for Weibull(*Loglogistic) model;
oualpha=par[1];
lambda=exp(par[1]);
alpha=par[2];

finish mod_est;

*****************************************************************;
**********Simulate 1000 samples of size 100 from Weib(Logl)(0.15;0.5)
```

```
                         - censored at 50;

lambda=0.15;
alpha=0.5;
n=100;
z=1000;
DT=((1/lambda)#(-log(1-ranuni(J(n,z,0)))))##(1/alpha);     *Weibull;
*FT=ranuni(J(n,z,0));
*DT=(FT/(lambda#(1-FT)))##(1/alpha);                        *Loglogistic;
DT=DT><J(n,z,50);                                           *censored at 50;

**********Define class boundaries and a frequency vector for each sample
          of continuous values, then run the module and store the 1000 estimates
          of lambda and alpha in a file PARMS;

filename parms 'c:\sim\sd2\wiml100a.sd2';                   *Weibull;
*filename parms 'c:\sim\sd2\lliml100a.sd2';                 *Loglogistic;
file parms;

do w=1 to z;
T=DT[,w];
B=T;
T[rank(T),]=B;
spnr=J(n,1,1)#w;

Y=(T=J(n,1,50));
nc=Y[+];
if nc=0 then nc=1e-4;
nl=n-nc;
ox=T[1:nl];
nc=1e-4<>nc;
f=J(nl,1,1)//nc;
perccens=(nc/n)#100;
             nr=nrow(ox); nr1=nr-1;
             x1=ox[1:nr1];x2=ox[2:nr];
             x=(x1+x2)/2;
             x=x//50;

run mod_est(x,f);
put lambda +3 alpha +3 perccens;
end;

closefile parms;

**********Put simulated data in a format that can be inputted in SAS
                   as a thousand continuous data sets;
%macro subgr(a);
%do i=1 %to &a;
name={"spnr" "time" "cens"};
di=DT[,&i];
cens=(di<J(n,1,50));
spnr=J(n,1,&i);
di=spnr||di||cens;
create d&i from di [colname=name];
append from di;
%end;
%mend subgr;
%subgr(1000);

**********Repeat the LIFEREG procedure of SAS one thousand times to get
          estimates for the intercept and scale parameters;

%macro mac(stel);
%do i=1 %to &stel;

proc lifereg data=d&i noprint outest=out&i  (keep=intercept _scale_);
model time*cens(0)= ;                            *Weibull;
*model time*cens(0)= / d=llogistic;              *Loglogistic;
run;

%end;
%mend mac;
%mac(1000);

**********Append the thousand estimates;

%macro ind(stel);
 %do i=2 %to &stel;

proc append base=out1 data=out&i;
run;

%end;
%mend ind;
%ind(1000);

**************************************************************;
**********Sampling distribution of parameter estimates*********;
```

```
*****IML estimates;
data spv_iml;
infile 'c:\sim\sd2\wiml100a.sd2';                          *Weibull;
*infile 'c:\sim\sd2\liml100a.sd2';                         *Loglogistic;
input lambda alpha;
title1 'Sampling distribution: 1000 samples of size 100 from
        Weib(Logl)(0.15;0.5) - censored at 50';
title2 'IML method';

proc univariate data=spv_iml normal plot;
var lambda alpha;
run;

*****SAS estimates;
data sim.wsas100a (keep=lambda alpha);                     *Weibull;
*data sim.lsas100a (keep=lambda alpha);                    *Loglogistic;
set out1;
lambda=exp(-intercept/_scale_);
alpha=1/_scale_;

title1 'Sampling distribution: 1000 samples of size 100 from
        Weib(Logl)(0.15;0.5) - censored at 50';
title2 'SAS method';

proc univariate data=sim.wsas100a normal plot;             *Weibull;
*proc univariate data=sim.lsas100a normal plot;            *Loglogistic;
var lambda alpha;
run;
```

2. **Program to generate continuous right-censored lifetime data from the log-normal distribution and to run simulations with the technique of MLE under constraints as well as the standard technique of MLE (1000 samples of size 200 from Lognormal -normal(2;0.5) - censored at 8)**

```
proc iml worksize=60 symsize=2000;
reset noname nocenter;

**********Contents of module**************************************************;

start mod_est(x,f,nl) global(mu,sigma) ;

*****Relative frequencies;
x=log(x);
n=200;
k=nrow(f);
d=k-1;
pi=(gamma(0.5))**2;
p=f/n;

*****Design matrix and matrix orthogonal to design matrix;
S1=J(d,1,1)@cusum(J(1,k,1));
S2=J(1,k,1)@cusum(J(d,1,1));
S=S1<=S2;

x1=J(d,1,1)||x;
C=I(d)-x1*inv(x1`*x1)*x1`;

**********ITERATIVE PROCEDURE (double iterations over m and p);
*****starting value for m;
  m=p;
  ms=S*m; ps=ms;
  p0=p;

*****iteration over m;
itr=0;
verskil1=1;
i=0;
do while (verskil1>1e-6);
i=i+1;
p=p0;
arg1=15;
arg2=nl-15;
par=inv( J(2,1,1)||probit(ms[arg1]//ms[arg2]) )*(x[arg1]//x[arg2]);
mu=par[1];sigma=par[2];
Gm=C*diag(sqrt(2#pi)/(exp(-(x-mu)#(x-mu)/2/sigma/sigma)))*S;

*****iteration over p;
   verskil=1;
   j=0;
   do while (verskil>1e-6);
   j=j+1;
   p1=p;
   ps=S*p;
   g=C*probit(ps);
   parp=inv( J(2,1,1)||probit(ps[arg1]//ps[arg2]) )*(x[arg1]//x[arg2]);
   mup=parp[1];sigmap=parp[2];
   Gp=C*diag(sqrt(2#pi)/(exp(-(x-mup)#(x-mup)/2/sigmap/sigmap)))*S;
*****covariance matrix;
```

```
sig=(1/n)*(diag(m)-m*m');
V=sig;
**********************;
   p=p-(Gm*V)'*ginv(Gp*V*Gm')*g;
   verskil=sqrt((p-p1)'*(p-p1));
   end;
verskil1=sqrt((p-m)'*(p-m));
m=p;ms=S*m;
end;

******Parameter vector for linear model;
par=inv( J(2,1,1)||probit(ms[arg1]//ms[arg2]) )*(x[arg1]//x[arg2]);

**********Parameters for Lognormal model;
mu=par[1];
sigma=par[2];

finish mod_est;

***************************************************************;
**********Simulate 1000 samples of size 200 from Lognormal(2;(0.5)^2)
                         - censored at 8;

n=200;
z=1000;
mean=2;
stddev=0.5;
AA=mean+stddev#normal(J(n,z,0));
DT=exp(AA);
DT=DT><J(n,z,8);      *8=censor point;

**********Define class boundaries and a frequency vector for each sample
          of continuous values, then run the module and store the
          1000 estimates of lambda and alpha in a file PARMS;

filename parms 'c:\sim\sd2\lnik200a.sd2';
file parms;

do w=1 to z;
T=DT[,w];
B=T;
T[rank(T),]=B;
spnr=J(n,1,1)#w;

Y=(T=J(n,1,8));          *8=censor point;
nc=Y[+];
if nc=0 then nc=1e-4;
nl=n-nc;
ox=T[1:nl];
nc=1e-4<>nc;
f=J(nl,1,1)//nc;
perccens=(nc/n)#100;
             nr=nrow(ox); nr1=nr-1;
             x1=ox[1:nr1];x2=ox[2:nr];
             x=(x1+x2)/2;
             x=x//8;    *8=censor point;
*print f x nl;
run mod_est(x,f,nl);
put (mu) +3 (sigma) +3  (perccens) +3 (nl);
end;

closefile parms;

**********Put simulated data in a format that can be inputted in SAS
                   as a thousand continuous data sets;
%macro subgr(a);
%do i=1 %to &a;
name={"spnr" "time" "cens"};
di=DT[,&i];
cens=(di<J(n,1,8));      *8=censor point;
spnr=J(n,1,&i);
di=spnr||di||cens;
create d&i from di [colname=name];
append from di;
%end;
%mend subgr;
%subgr(1000);

**********Repeat the LIFEREG procedure of SAS one thousand times to get
          estimates for the intercept and scale parameters;

%macro mac(stel);
%do i=1 %to &stel;

proc lifereg data=d&i noprint outest=out&i  (keep=intercept _scale_);
model time*cens(0)= / d=lnormal;
run;

%end;
```

```
%mend mac;
%mac(1000);

**********Append the thousand estimates;

%macro ind(stel);
 %do i=2 %to &stel;

proc append base=out1 data=out&i;
run;

%end;
%mend ind;
%ind(1000);

**************************************************************;
**********Sampling distribution of parameter estimates*********;

*****IML estimates;
data spv_iml;
infile 'c:\sim\sd2\lnik200a.sd2';
input mu sigma;
*ods html body='c:\sim\lnik200a.htm';
title1 'Sampling distribution: 1000 samples of size 200 from lognormal
        - normal(2;0.5) - censored at 8';
title2 'IML method';

proc univariate data=spv_iml normal plot;
var mu sigma;
run;

*****SAS estimates;

data sim.lnsk200a (keep=mu sigma);
set out1;
mu=intercept;
sigma=_scale_;
title1 'Sampling distribution: 1000 samples of size 200 from lognormal
        - normal(2;0.5) - censored at 8';
title2 'SAS method';

proc univariate data=sim.lnsk200a normal plot;
var mu sigma;
run;
```

A.3.2 Program to simulate grouped right-censored lifetime data

1. **Program to generate right-censored grouped lifetime data from the Weibull (loglogistic) distribution and to run simulations with the technique of MLE under constraints as well as the standard technique of MLE (1000 samples of size 2000 from Weib(30;1.8) - censored at 0.15 (grouped into 5 classes))**

```
proc iml worksize=6000 symsize=2000;
reset noname nocenter;

**********Contents of module**************************************************;

start mod_est(x,f) global(lambda,alpha) ;

*****Relative frequencies;
n=2000;
k=nrow(f);
d=k-1;
p=f/n;

*****Design matrix and matrix orthogonal to design matrix;
S1=J(d,1,1)@cusum(J(1,k,1));
S2=J(1,k,1)@cusum(J(d,1,1));
S=S1<=S2;

x1=J(d,1,1)||log(x);
C=I(d)-x1*inv(x1`*x1)*x1`;

**********ITERATIVE PROCEDURE (double iterations over m and p);
*****starting value for m;
  m=p;
  ms=S*m; ps=ms;
  p0=p;

*****iteration over m;
itr=0;
verskil1=1;
i=0;
```

```
do while (verskil1>1e-6);
i=i+1;
p=p0;
Gm=-C*diag(1/(log(1-ms)))*diag(1/(1-ms))*S;                     *Weibull;
*Gm=C*(diag(1/ms)+diag(1/(1-ms)))*S;                            *Loglogistic;

*****iteration over p;
   verskil=1;
   j=0;
   do while (verskil>1e-6);
   j=j+1;
   p1=p;
   ps=S*p;
   g=C*log(-log(1-ps));                                         *Weibull;
  *g=C*(log(ps)-log(1-ps));                                     *Loglogistic;
   Gp=-C*diag(1/(log(1-ps)))*diag(1/(1-ps))*S;                  *Weibull;
  *Gp=C*(diag(1/ps)+diag(1/(1-ps)))*S;                          *Loglogistic;
*****covariance matrix;
sig=(1/n)*(diag(m)-m*m`);
V=sig;
**********************;
   p=p-(Gm*V)`*ginv(Gp*V*Gm`)*g;                                *Weibull;
  *p=p-(Gm*V)`*ginv(Gp*V*Gm`)*g;                                *Loglogistic;
   verskil=sqrt((p-p1)`*(p-p1));
   end;
verskil1=sqrt((p-m)`*(p-m));
m=p;ms=S*m;
end;

******Parameter vector for linear model;
par=inv(x1`*x1)*x1`*log(-log(1-ms));                            *Weibull;
*par=inv(xc`*xc)*xc`*(log(ms)-log(1-ms));                       *Loglogistic;

**********Parameters for Weibull(*Loglogistic) model;
oualpha=par[1];
lambda=exp(par[1]);
alpha=par[2];

finish mod_est;

*****************************************************************;
**********Simulate 1000 samples of size 2000 from Weib(Logl)(30;1.8) -
          censored at 0.15 and then define lower and upper class boundaries
          for 5 class intervals;

lambda=30;
alpha=1.8;
n=2000;
z=1000;
TT=((1/lambda)#(-log(1-ranuni(J(n,z,0)))))##(1/alpha);         *Weibull;
*FT=ranuni(J(n,z,0));
*TT=(FT/(lambda#(1-FT)))##(1/alpha);                            *Loglogistic;
x1={0,0.08,0.10,0.12,0.15};    *lower boundaries;              *censored at 0.15;
x2={0.08,0.10,0.12,0.15,100}; *upper boundaries;
k=5;
*****************************************************************;
**********calculate the frequency vector for each sample, then run the
module and store the 1000 estimates of lambda and alpha in a file PARMS;

filename parms 'c:\sim\sd2\wigtwdc1.sd2';                       *Weibull;
*filename parms 'c:\sim\sd2\ligtwdc1.sd2';                      *Loglogistic;
file parms;

do w=1 to z;
T=TT[w,];
A=(J(k,1,1)*T)<=(x2*J(1,n,1));
B=(J(k,1,1)*T)>=(x1*J(1,n,1));
E=A=B;
f=E*J(n,1,1);
perccens=(f[5,1]/n)#100;
x=x2[1:4,];

run mod_est(x,f);
put lambda +3 alpha +3 perccens;
end;

closefile parms;

**********Put simulated data in a format that can be inputted in SAS
                   as a thousand grouped data sets;
%macro subgr(a);
%do i=1 %to &a;
name={"lower" "upper" "frek"};
lower=.//x1[2:5,];
upper=x2[1:4,]//.;

T=TT[&i,];
A=(J(k,1,1)*T)<=(x2*J(1,n,1));
B=(J(k,1,1)*T)>=(x1*J(1,n,1));
```

```
E=A=B;
frek=E*J(n,1,1);
di=lower||upper||frek;
create d&i from di [colname=name];
append from di;
%end;
%mend subgr;
%subgr(1000);

**********Repeat the LIFEREG procedure of SAS for grouped data one thousand
          times to get estimates for the intercept and scale parameters;

%macro mac(stel);
%do i=1 %to &stel;

proc lifereg data=d&i noprint outest=out&i  (keep=intercept _scale_);
model (lower,upper)= ;                       *Weibull;
*model (lower,upper)= / d=llogistic;         *Loglogistic;
weight frek;
run;

%end;
%mend mac;
%mac(1000);

**********Append the thousand estimates;

%macro ind(stel);
 %do i=2 %to &stel;

proc append base=out1 data=out&i;
run;

%end;
%mend ind;
%ind(1000);

*****************************************************************;
**********Sampling distribution of parameter estimates*********;

*****IML estimates;

data spv_iml;
infile 'c:\sim\sd2\wigtwdc1.sd2';                          *Weibull;
*infile 'c:\sim\sd2\ligtwdc1.sd2';                         *Loglogistic;
input lambda alpha;
title1 'Sampling distribution: 1000 samples of size 2000 from
        Weib(Logl)(30;1.8) - censored at 0.15 (5 classes)';
title2 'IML method';

proc univariate data=spv_iml normal plot;
var lambda alpha;
run;

*****SAS estimates;
data sim.wsgtwdc1 (keep=lambda alpha);                     *Weibull;
*data sim.lsgtwdc1 (keep=lambda alpha);                    *Loglogistic;
set out1;
lambda=exp(-intercept/_scale_);
alpha=1/_scale_;

title1 'Sampling distribution: 1000 samples of size 2000 from Weib(30;1.8)
        - censored at 0.15 (5 classes)';
title2 'SAS method';

proc univariate data=sim.wsgtwdc1 normal plot;             *Weibull;
*proc univariate data=sim.lsgtwdc1 normal plot;            *Loglogistic;
var lambda alpha;
run;
```

2. **Program to generate right-censored grouped lifetime data from the lognormal distribution and to run simulations with the technique of MLE under constraints as well as the standard technique of MLE (1000 samples of size 2000 from Lognormal - normal(2;0.5) - censored at 8 (grouped into 5 classes))**

```
proc iml worksize=60 symsize=2000;
reset noname nocenter;

**********Contents of module*************************************************;

start mod_est(x,f) global(mu,sigma) ;

*****Relative frequencies;
x=log(x);
n=2000;
k=nrow(f);
d=k-1;
pi=(gamma(0.5))**2;
```

```
p=f/n; ;

*****Design matrix and matrix orthogonal to design matrix;
S1=J(d,1,1)@cusum(J(1,k,1));
S2=J(1,k,1)@cusum(J(d,1,1));
S=S1<=S2;

x1=J(d,1,1)||x;
C=I(d)-x1*inv(x1`*x1)*x1`;

**********ITERATIVE PROCEDURE (double iterations over m and p);
*****starting value for m;
  m=p;
  ms=S*m; ps=ms;
  p0=p;

*****iteration over m;
itr=0;
verskil1=1;
i=0;
do while (verskil1>1e-6);
i=i+1;
p=p0;
arg1=2;
arg2=3;
par=inv( J(2,1,1)||probit(ms[arg1]//ms[arg2]) )*(x[arg1]//x[arg2]);
mu=par[1];sigma=par[2];
Gm=C*diag(sqrt(2#pi)/(exp(-(x-mu)#(x-mu)/2/sigma/sigma)))*S;

*****iteration over p;
   verskil=1;
   j=0;
   do while (verskil>1e-6);
   j=j+1;
   p1=p;
   ps=S*p;
   g=C*probit(ps);
   parp=inv( J(2,1,1)||probit(ps[arg1]//ps[arg2]) )*(x[arg1]//x[arg2]);
   mup=parp[1];sigmap=parp[2];
   Gp=C*diag(sqrt(2#pi)/(exp(-(x-mup)#(x-mup)/2/sigmap/sigmap)))*S;
*****covariance matrix;
sig=(1/n)*(diag(m)-m*m`);
V=sig;
***********************;
   p=p-(Gm*V)`*ginv(Gp*V*Gm`)*g;
   verskil=sqrt((p-p1)`*(p-p1));
   end;
verskil1=sqrt((p-m)`*(p-m));
m=p;ms=S*m;
end;

******Parameter vector for linear model;
par=inv( J(2,1,1)||probit(ms[arg1]//ms[arg2]) )*(x[arg1]//x[arg2]);

**********Parameters for Lognormal model;
mu=par[1];
sigma=par[2];

finish mod_est;

*************************************************************************;
**********Simulate 1000 samples of size 2000 from Lognormal(2;(0.5)^2) -
censored at 8 and then define lower and upper boundaries for 5 class intervals;

n=2000;
z=1000;
mean=2;
stddev=0.5;
AA=mean+stddev#normal(J(z,n,0));
DT=exp(AA);
DT=DT><J(z,n,8);       *8=censor point;
x1={0,4,6,7,8};         *lower boundaries;  *censored at 8;
x2={4,6,7,8,10000};     *upper boundaries;
k=5;

*****************************************************************;
**********calculate the frequency vector for each sample, then run
the module and store the 1000 estimates of lambda and alpha in a file PARMS;

filename parms 'c:\sim\sd2\lnigtwa1.sd2';
file parms;

do w=1 to z;
T=DT[w,];
A=(J(k,1,1)*T)< (x2*J(1,n,1));
B=(J(k,1,1)*T)>=(x1*J(1,n,1));
E=A=B;
f=E*J(n,1,1);
perccens=(f[5,1]/n)#100;
```

```
x=x2[1:4,];

run mod_est(x,f);
put mu +3 sigma +3  perccens;
end;

closefile parms;

**********Put simulated data in a format that can be inputted in SAS
                   as a thousand grouped data sets;
%macro subgr(a);
%do i=1 %to &a;
name={"lower" "upper" "frek"};
lower=.//x1[2:5,];
upper=x2[1:4,]//.;

T=DT[&i,];
A=(J(k,1,1)*T)< (x2*J(1,n,1));
B=(J(k,1,1)*T)>=(x1*J(1,n,1));
E=A=B;
frek=E*J(n,1,1);
di=lower||upper||frek;
create d&i from di [colname=name];
append from di;
%end;
%mend subgr;
%subgr(1000);

**********Repeat the LIFEREG procedure of SAS for grouped data one thousand
          times to get estimates for the intercept and scale parameters;

%macro mac(stel);
%do i=1 %to &stel;

proc lifereg data=d&i noprint outest=out&i  (keep=intercept _scale_);
model (lower,upper)= / d=lnormal;
weight frek;
run;

%end;
%mend mac;
%mac(1000);

**********Append the thousand estimates;

%macro ind(stel);
 %do i=2 %to &stel;

proc append base=out1 data=out&i;
run;

%end;
%mend ind;
%ind(1000);

****************************************************************;
**********Sampling distribution of parameter estimates*********;

*****IML estimates;

data spv_iml;
infile 'c:\sim\sd2\lnigtwa1.sd2';
input mu sigma;
title1 'Sampling distribution: 1000 samples of size 2000 from lognormal
        - normal(2;0.5) - censored at 8 (5 classes)';
title2 'IML method';

proc univariate data=spv_iml normal plot;
var mu sigma;
run;

*****SAS estimates;

data sim.lnsgtwa1 (keep=mu sigma);
set out1;
mu=intercept;
sigma=_scale_;
title1 'Sampling distribution: 1000 samples of size 2000 from lognormal
        - normal(2;0.5) - censored at 8 (5 classes)';
title2 'SAS method';

proc univariate data=sim.lnsgtwa1 normal plot;
var mu sigma;
run;
```

A.4 Chapter 4: Maximum Likelihood Estimation

A.4.1 Staggered Entry of Policies Standard Programs using PROC LIFEREG

1. **Program for fitting a log-logistic/Weibull regression model with one predictor to grouped survival data**

```
options nodate pagesize=500 pageno=1;
libname hsbc 'c:\hsbc1\sd2';

title 'AGE:three levels (define two dummies)';

data agegr;
input lower upper  agegr1 agegr2  freq;
*input lower upper  agegr $ freq;
*at class statement: use one column A B C A B C etc.;
cards;
 .   12  1  0   29
 .   12  0  1   21
 .   12 -1 -1   16
12   17  1  0   59
12   17  0  1   50
12   17 -1 -1   49
17   24  1  0   95
17   24  0  1   91
17   24 -1 -1   68
24   28  1  0   73
24   28  0  1   45
24   28 -1 -1   39
28   34  1  0  108
28   34  0  1   75
28   34 -1 -1   67
34   37  1  0   15
34   37  0  1   13
34   37 -1 -1    7
37    .  1  0  642
37    .  0  1  553
37    . -1 -1  471
 .   12  1  0   41
 .   12  0  1   49
 .   12 -1 -1   28
12   17  1  0   75
12   17  0  1   62
12   17 -1 -1   29
17   24  1  0  103
17   24  0  1   61
17   24 -1 -1   65
24   28  1  0   92
24   28  0  1   66
24   28 -1 -1   42
28   34  1  0   83
28   34  0  1   54
28   34 -1 -1   35
34    .  1  0  628
34    .  0  1  753
34    . -1 -1  543
 .   12  1  0   68
 .   12  0  1   40
 .   12 -1 -1   46
12   17  1  0   34
12   17  0  1   44
12   17 -1 -1   21
17   24  1  0   99
17   24  0  1   83
17   24 -1 -1   60
24   28  1  0   57
24   28  0  1   33
24   28 -1 -1   27
28    .  1  0  570
28    .  0  1  533
28    . -1 -1  571
 .   12  1  0   71
 .   12  0  1   54
 .   12 -1 -1   50
12   17  1  0   60
12   17  0  1   61
12   17 -1 -1   45
17   24  1  0   69
17   24  0  1   68
17   24 -1 -1   70
24    .  1  0  573
```

```
24    .  0  1  616
24    . -1 -1  659
;

data add;
c=1;
run;

data fin;
set agegr add;

proc lifereg data=fin covout;
*class agegr;
*model (lower, upper) = agegr / dist=llogistic;
model (lower, upper) = agegr1 agegr2 / dist=llogistic;
weight freq;
output out=llog cdf=cdf predicted=months
quantiles = 0.02 to 0.98 by .02
control=c;
title 'Fit Loglogistic curve (SAS method) to HSBC: MPC policies - predictor AGE';
title2 'Four entry dates';
run;

data par;
intercept=3.86266;
_scale_=0.48394;
theta1=-0.08757;
theta2= 0.01693;
theta3= -(theta1+theta2);
if _N_=1;

oualphaSASage1=-(intercept+theta1*1+theta2*0)/_scale_;
lambdaSASage1=exp(-(intercept+theta1*1+theta2*0)/_scale_);
oualphaSASage2=-(intercept+theta1*0+theta2*1)/_scale_;
lambdaSASage2=exp(-(intercept+theta1*0+theta2*1)/_scale_);
oualphaSASage3=-(intercept+theta1*(-1)+theta2*(-1))/_scale_;
lambdaSASage3=exp(-(intercept+theta1*(-1)+theta2*(-1))/_scale_);
alphaSAS=1/_scale_;
oubetaSAS=1/_scale_;
oualphaBASELINESAS=-(intercept/_scale_);
lambdaBASELINESAS=exp(-(intercept/_scale_));

proc print data=par;
var oualphaSASage1 oualphaSASage2 oualphaSASage3
    lambdaSASage1 lambdaSASage2 lambdaSASage3
    alphaSAS oubetaSAS
    oualphaBASELINESAS lambdaBASELINESAS;
run;
```

2. Program for fitting a log-logistic/Weibull regression model with two predictors to grouped survival data

```
options nodate pagesize=500 pageno=1;
libname hsbc 'c:\hsbc1\sd2';

title1 'AGE:three levels (define two dummies)';
title2 'SCORE:three levels (define two dummies)';

data agegr;
input lower upper  agegr1 agegr2  score1 score2 freq;
*input lower upper  agegr $ freq;
*at class statement: use one column of A B C A B C etc.;
cards;
 .   12  1  0  1  0 12
12   17  1  0  1  0 34
17   24  1  0  1  0 51
24   28  1  0  1  0 39
28   34  1  0  1  0 57
34   37  1  0  1  0 11
37    .  1  0  1  0 59

 .   12  1  0  0  1 10
12   17  1  0  0  1 12
17   24  1  0  0  1 22
24   28  1  0  0  1 19
28   34  1  0  0  1 32
34   37  1  0  0  1  4
37    .  1  0  0  1 418

 .   12  1  0 -1 -1  7
12   17  1  0 -1 -1 13
17   24  1  0 -1 -1 22
24   28  1  0 -1 -1 15
28   34  1  0 -1 -1 19
34   37  1  0 -1 -1  0
37    .  1  0 -1 -1 165

 .   12  0  1  1  0 13
12   17  0  1  1  0 14
```

```
17  24  0  1  1  0 45
24  28  0  1  1  0 27
28  34  0  1  1  0 33
34  37  0  1  1  0  4
37   .  0  1  1  0  66

 .  12  0  1  0  1  4
12  17  0  1  0  1 22
17  24  0  1  0  1 22
24  28  0  1  0  1  8
28  34  0  1  0  1 25
34  37  0  1  0  1  4
37   .  0  1  0  1 297

 .  12  0  1 -1 -1  4
12  17  0  1 -1 -1 14
17  24  0  1 -1 -1 24
24  28  0  1 -1 -1 10
28  34  0  1 -1 -1 17
34  37  0  1 -1 -1  5
37   .  0  1 -1 -1 190

 .  12 -1 -1  1  0 10
12  17 -1 -1  1  0 25
17  24 -1 -1  1  0 29
24  28 -1 -1  1  0 17
28  34 -1 -1  1  0 46
34  37 -1 -1  1  0  2
37   . -1 -1  1  0 116

 .  12 -1 -1  0  1  6
12  17 -1 -1  0  1 13
17  24 -1 -1  0  1 28
24  28 -1 -1  0  1 16
28  34 -1 -1  0  1 16
34  37 -1 -1  0  1  5
37   . -1 -1  0  1 273

 .  12 -1 -1 -1 -1  0
12  17 -1 -1 -1 -1 11
17  24 -1 -1 -1 -1 11
24  28 -1 -1 -1 -1  6
28  34 -1 -1 -1 -1  5
34  37 -1 -1 -1 -1  0
37   . -1 -1 -1 -1  82

 .  12  1  0  1  0 22
12  17  1  0  1  0 25
17  24  1  0  1  0 58
24  28  1  0  1  0 53
28  34  1  0  1  0 40
34   .  1  0  1  0  45

 .  12  1  0  0  1 10
12  17  1  0  0  1 26
17  24  1  0  0  1 32
24  28  1  0  0  1 20
28  34  1  0  0  1 29
34   .  1  0  0  1 379

 .  12  1  0 -1 -1  9
12  17  1  0 -1 -1 24
17  24  1  0 -1 -1 13
24  28  1  0 -1 -1 19
28  34  1  0 -1 -1 14
34   .  1  0 -1 -1 204

 .  12  0  1  1  0 24
12  17  0  1  1  0 24
17  24  0  1  1  0 28
24  28  0  1  1  0 30
28  34  0  1  1  0 25
34   .  0  1  1  0 106

 .  12  0  1  0  1 12
12  17  0  1  0  1 20
17  24  0  1  0  1 14
24  28  0  1  0  1 17
28  34  0  1  0  1 16
34   .  0  1  0  1 409

 .  12  0  1 -1 -1 13
12  17  0  1 -1 -1 18
17  24  0  1 -1 -1 19
24  28  0  1 -1 -1 19
28  34  0  1 -1 -1 13
34   .  0  1 -1 -1 238

 .  12 -1 -1  1  0 13
```

```
12   17 -1 -1  1  0 15
17   24 -1 -1  1  0 32
24   28 -1 -1  1  0 19
28   34 -1 -1  1  0 17
34    . -1 -1  1  0 107

 .   12 -1 -1  0  1 11
12   17 -1 -1  0  1 13
17   24 -1 -1  0  1 22
24   28 -1 -1  0  1 17
28   34 -1 -1  0  1 12
34    . -1 -1  0  1 319

 .   12 -1 -1 -1 -1  4
12   17 -1 -1 -1 -1  1
17   24 -1 -1 -1 -1 11
24   28 -1 -1 -1 -1  6
28   34 -1 -1 -1 -1  6
34    . -1 -1 -1 -1 117

 .   12  1  0  1  0 34
12   17  1  0  1  0 16
17   24  1  0  1  0 50
24   28  1  0  1  0 23
28    .  1  0  1  0  54

 .   12  1  0  0  1 19
12   17  1  0  0  1  2
17   24  1  0  0  1 32
24   28  1  0  0  1 24
28    .  1  0  0  1 317

 .   12  1  0 -1 -1 15
12   17  1  0 -1 -1 16
17   24  1  0 -1 -1 17
24   28  1  0 -1 -1 10
28    .  1  0 -1 -1 199

 .   12  0  1  1  0 19
12   17  0  1  1  0 18
17   24  0  1  1  0 38
24   28  0  1  1  0 16
28    .  0  1  1  0  75

 .   12  0  1  0  1 16
12   17  0  1  0  1 14
17   24  0  1  0  1 25
24   28  0  1  0  1 10
28    .  0  1  0  1 263

 .   12  0  1 -1 -1  5
12   17  0  1 -1 -1 12
17   24  0  1 -1 -1 20
24   28  0  1 -1 -1  7
28    .  0  1 -1 -1 195

 .   12 -1 -1  1  0 28
12   17 -1 -1  1  0 16
17   24 -1 -1  1  0 22
24   28 -1 -1  1  0 12
28    . -1 -1  1  0  98

 .   12 -1 -1  0  1 13
12   17 -1 -1  0  1  0
17   24 -1 -1  0  1 24
24   28 -1 -1  0  1  4
28    . -1 -1  0  1 323

 .   12 -1 -1 -1 -1  5
12   17 -1 -1 -1 -1  5
17   24 -1 -1 -1 -1 14
24   28 -1 -1 -1 -1 11
28    . -1 -1 -1 -1 150

 .   12  1  0  1  0 40
12   17  1  0  1  0 30
17   24  1  0  1  0 30
24    .  1  0  1  0  50

 .   12  1  0  0  1  9
12   17  1  0  0  1 14
17   24  1  0  0  1 27
24    .  1  0  0  1 301

 .   12  1  0 -1 -1 22
12   17  1  0 -1 -1 16
17   24  1  0 -1 -1 12
24    .  1  0 -1 -1 222
```

```
 .   12  0  1  1  0 24
12   17  0  1  1  0 30
17   24  0  1  1  0 29
24    .  0  1  1  0  81

 .   12  0  1  0  1 14
12   17  0  1  0  1 15
17   24  0  1  0  1 12
24    .  0  1  0  1 307

 .   12  0  1 -1 -1 16
12   17  0  1 -1 -1 16
17   24  0  1 -1 -1 27
24    .  0  1 -1 -1 228

 .   12 -1 -1  1  0 20
12   17 -1 -1  1  0 22
17   24 -1 -1  1  0 28
24    . -1 -1  1  0 119

 .   12 -1 -1  0  1 19
12   17 -1 -1  0  1 12
17   24 -1 -1  0  1 26
24    . -1 -1  0  1 369

 .   12 -1 -1 -1 -1 11
12   17 -1 -1 -1 -1 11
17   24 -1 -1 -1 -1 16
24    . -1 -1 -1 -1 171
;

data add;
c=1;
run;

data fin;
set agegr add;

proc lifereg data=fin covout;
*class agegr;
*model (lower, upper) = agegr / dist=llogistic;
model (lower, upper) = agegr1 agegr2 score1 score2 / dist=llogistic;
weight freq;
output out=llog cdf=cdf predicted=months
quantiles = 0.02 to 0.98 by .02
control=c;
title 'Fit Loglogistic curve (SAS method) to HSBC: MPC policies - predictor AGE';
title2 'Four entry dates';
run;

data par;
intercept=3.80119;
_scale_=0.44454;
thetaA1=-0.09129;
thetaA2= 0.0052689;
thetaA3= -(thetaA1+thetaA2);
thetaS1=-0.46574;
thetaS2= 0.31782;
thetaS3= -(thetaS1+thetaS2);
if _N_=1;

oualphaSASa1s1=-(intercept+thetaA1*1+thetaA2*0+thetaS1*1+thetaS2*0)/_scale_;
lambdaSASa1s1=exp(-(intercept+thetaA1*1+thetaA2*0+thetaS1*1+thetaS2*0)/_scale_);
oualphaSASa1s2=-(intercept+thetaA1*1+thetaA2*0+thetaS1*0+thetaS2*1)/_scale_;
lambdaSASa1s2=exp(-(intercept+thetaA1*1+thetaA2*0+thetaS1*0+thetaS2*1)/_scale_);
oualphaSASa1s3=-(intercept+thetaA1*1+thetaA2*0+thetaS1*(-1)+thetaS2*(-1))/_scale_;
lambdaSASa1s3=exp(-(intercept+thetaA1*1+thetaA2*0+thetaS1*(-1)+thetaS2*(-1))/_scale_);
oualphaSASa2s1=-(intercept+thetaA1*0+thetaA2*1+thetaS1*1+thetaS2*0)/_scale_;
lambdaSASa2s1=exp(-(intercept+thetaA1*0+thetaA2*1+thetaS1*1+thetaS2*0)/_scale_);
oualphaSASa2s2=-(intercept+thetaA1*0+thetaA2*1+thetaS1*0+thetaS2*1)/_scale_;
lambdaSASa2s2=exp(-(intercept+thetaA1*0+thetaA2*1+thetaS1*0+thetaS2*1)/_scale_);
oualphaSASa2s3=-(intercept+thetaA1*0+thetaA2*1+thetaS1*(-1)+thetaS2*(-1))/_scale_;
lambdaSASa2s3=exp(-(intercept+thetaA1*0+thetaA2*1+thetaS1*(-1)+thetaS2*(-1))/_scale_);
oualphaSASa3s1=-(intercept+thetaA1*(-1)+thetaA2*(-1)+thetaS1*1+thetaS2*0)/_scale_;
lambdaSASa3s1=exp(-(intercept+thetaA1*(-1)+thetaA2*(-1)+thetaS1*1+thetaS2*0)/_scale_);
oualphaSASa3s2=-(intercept+thetaA1*(-1)+thetaA2*(-1)+thetaS1*0+thetaS2*1)/_scale_;
lambdaSASa3s2=exp(-(intercept+thetaA1*(-1)+thetaA2*(-1)+thetaS1*0+thetaS2*1)/_scale_);
oualphaSASa3s3=-(intercept+thetaA1*(-1)+thetaA2*(-1)+thetaS1*(-1)+thetaS2*(-1))/_scale_;
lambdaSASa3s3=exp(-(intercept+thetaA1*(-1)+thetaA2*(-1)+thetaS1*(-1)+thetaS2*(-1))/_scale_);
alphaSAS=1/_scale_;
oubetaSAS=1/_scale_;
oualphaBASELINESAS=-(intercept/_scale_);
lambdaBASELINESAS=exp(-(intercept/_scale_));

proc print data=par;
var oualphaSASa1s1 oualphaSASa1s2 oualphaSASa1s3
    lambdaSASa1s1 lambdaSASa1s2 lambdaSASa1s3;
proc print data=par;
var oualphaSASa2s1 oualphaSASa2s2 oualphaSASa2s3
```

```
    lambdaSASa2s1 lambdaSASa2s2 lambdaSASa2s3;
proc print data=par;
var oualphaSASa3s1 oualphaSASa3s2 oualphaSASa3s3
    lambdaSASa3s1 lambdaSASa3s2 lambdaSASa3s3;
proc print data=par;
var alphaSAS oubetaSAS oualphaBASELINESAS lambdaBASELINESAS;
run;
```

A.5 Chapter 4: M L Estimation subject to Constraints

A.5.1 Staggered Entry of Policies - IML Programs

1. **Program for fitting a log-logistic regression model (constant shape, one covariate)**

```
title1 'Fitting of regression model with one covariate';
title2 'Staggered entry: constant shape parameter';

proc iml worksize= 60;
reset nolog;
options pagesize=500;

**********Frequency vector (first entry,3 agegroups);
                f11={29,59,95,73,108,15,642};
                f12={21,50,91,45,75,13,553};
                f13={16,49,68,39, 67, 7,471};

**********Vector of upper boundaries;
                x1={12,17,24,28,34,37};

**********Frequency vector (second entry,3 agegroups);
                f21={41,75,103,92,83,628};
                f22={49,62,61,66,54,753};
                f23={28,29,65,42,35,543};

**********Vector of upper boundaries;
                x2={12,17,24,28,34};

**********Frequency vector (third entry,3 agegroups);
                f31={68,34,99,57,570};
                f32={40,44,83,33,533};
                f33={46,21,60,27,571};

**********Vector of upper boundaries;
                x3={12,17,24,28};

**********Frequency vector (fourth entry,3 agegroups);
                f41={71,60,69,573};
                f42={54,61,68,616};
                f43={50,45,70,659};

**********Vector of upper boundaries;
                x4={12,17,24};

**********Relative frequency vectors;
n11=f11[+];    n21=f21[+];    n31=f31[+];    n41=f41[+];    n1=n11+n21+n31+n41;
n12=f12[+];    n22=f22[+];    n32=f32[+];    n42=f42[+];    n2=n12+n22+n32+n42;
n13=f13[+];    n23=f23[+];    n33=f33[+];    n43=f43[+];    n3=n13+n23+n33+n43;
                                                              n=n1+n2+n3;
k1=nrow(f11); d1=k1-1;
k2=nrow(f21); d2=k2-1;
k3=nrow(f31); d3=k3-1;
k4=nrow(f41); d4=k4-1;
k=k1+k2+k3+k4;
d=d1+d2+d3+d4;

p11=f11/n11;    p21=f21/n21;    p31=f31/n31;    p41=f41/n41;    p1=p11//p21//p31//p41;
p12=f12/n12;    p22=f22/n22;    p32=f32/n32;    p42=f42/n42;    p2=p12//p22//p32//p42;
p13=f13/n13;    p23=f23/n23;    p33=f33/n33;    p43=f43/n43;    p3=p13//p23//p33//p43;
                                                                p=p1//p2//p3;

**********Design matrix and matrix orthogonal to design matrix;

S1=J(d1,1,1)@cusum(J(1,k1,1));
S2=J(1,k1,1)@cusum(J(d1,1,1));
S1=S1<=S2;
S2=S1[1:d2,1:d1];
S3=S1[1:d3,1:d2];
S4=S1[1:d4,1:d3];
```

```
S=block(S1,S2,S3,S4);
S=I(3)@S;

AA=J(54,1,1);       *54=3 times((7-1)+(6-1)+(5-1)+(4-1))=3 times 18=54;
BB=J(3,1,1);
CC=(I(2)//J(1,2,-1));
DD=J(18,1,1);            *d=18;
EE=J(9,1,1);

lx=BB@(log(x1)//log(x2)//log(x3)//log(x4));

xc=AA||(CC@DD)||lx;  print xc;

C=I(3#d)-xc*inv(xc`*xc)*xc`;

**********ITERATIVE PROCEDURE (double iterations over m and p);
*****starting value for m;
m=p;
ms=S*m; ps=ms;
p0=p;

*****iteration over m;
itr=0;
verskil1=1;
i=0;
do while (verskil1>0.00000001);
i=i+1;
p=p0;
*Gm=-C*diag(1/(log(1-ms)))*diag(1/(1-ms))*S;              *Weibull;
Gm=C*diag(1/ms+1/(1-ms))*S;                                *log-logistic;

**********iteration over p;
     verskil=1;
     j=0;
     do while (verskil>0.00000001);
     j=j+1;
     p1=p;
     ps=S*p;
    *g=C*log(-log(1-ps));                                  *Weibull;
    g=C*(log(ps)-log(1-ps));                               *log-logistic;
    *Gp=-C*diag(1/(log(1-ps)))*diag(1/(1-ps))*S;           *Weibull;
    Gp=C*(diag(1/ps)+diag(1/(1-ps)))*S;                    *log-logistic;
*****************covariance matrix*********************************;
m11=m[1:k1]; m21=m[k1+1:k1+k2];
m31=m[k1+k2+1:k1+k2+k3]; m41=m[k1+k2+k3+1:k1+k2+k3+k4];

m12=m[k+1:k+k1]; m22=m[k+k1+1:k+k1+k2];
m32=m[k+k1+k2+1:k+k1+k2+k3]; m42=m[k+k1+k2+k3+1:k+k1+k2+k3+k4];

m13=m[2#k+1:2#k+k1]; m23=m[2#k+k1+1:2#k+k1+k2];
m33=m[2#k+k1+k2+1:2#k+k1+k2+k3]; m43=m[2#k+k1+k2+k3+1:2#k+k1+k2+k3+k4];

sig11=(1/n11)*(diag(m11)-m11*m11`);
sig21=(1/n21)*(diag(m21)-m21*m21`);
sig31=(1/n31)*(diag(m31)-m31*m31`);
sig41=(1/n41)*(diag(m41)-m41*m41`);
sig1=block(sig11,sig21,sig31,sig41);

sig12=(1/n12)*(diag(m12)-m12*m12`);
sig22=(1/n22)*(diag(m22)-m22*m22`);
sig32=(1/n32)*(diag(m32)-m32*m32`);
sig42=(1/n42)*(diag(m42)-m42*m42`);
sig2=block(sig12,sig22,sig32,sig42);

sig13=(1/n13)*(diag(m13)-m13*m13`);
sig23=(1/n23)*(diag(m23)-m23*m23`);
sig33=(1/n33)*(diag(m33)-m33*m33`);
sig43=(1/n43)*(diag(m43)-m43*m43`);
sig3=block(sig13,sig23,sig33,sig43);

sig=block(sig1,sig2,sig3);
V=sig;
*********************************************************************;
    *p=p-(Gm*V)`*ginv(Gp*V*Gm`)*g;                        *Weibull;
    p=p-(Gm*V)`*ginv(Gp*V*Gm`)*g;                         *log-logistic;
     verskil=sqrt((p-p1)`*(p-p1));
     end;
verskil1=sqrt((p-m)`*(p-m));
m=p;ms=S*m;
end;
print m; print i j;

**********Parameter vector for linear model;
*par=inv(xc`*xc)*xc`*log(-log(1-ms));                      *Weibull;
par=inv(xc`*xc)*xc`*(log(ms)-log(1-ms));                   *log-logistic;

print par[format=E20.];

**********Regression coefficients;
```

```
oualphaBASELINE=par[1];
betaA1=par[2];
betaA2=par[3];
betaA3=-(par[2]+par[3]);
lambdaBASELINE=exp(par[1]);
lambdaA1=exp(par[1]+par[2]);
lambdaA2=exp(par[1]+par[3]);
lambdaA3=exp(par[1]+betaA3);
alpha=par[4];

**********Indices, lambda's and constant alpha for age levels******************;
indexA1=exp(betaA1);        *constant shape parameter;
indexA2=exp(betaA2);
indexA3=exp(betaA3);

lambdaA1=lambdaBASELINE#indexA1;   *same as lambdaA1=exp(par[1]+par[2]);
lambdaA2=lambdaBASELINE#indexA2;   *same as lambdaA2=exp(par[1]+par[3]);
lambdaA3=lambdaBASELINE#indexA3;   *same as lambdaA3=exp(par[1]+betaA3);
oualphaA1=log(lambdaA1);           *same as oualphaA1=oualphaBASELINE+betaA1;
oualphaA2=log(lambdaA2);
oualphaA3=log(lambdaA3);

print ' Loglogistic parameters, beta effects and indices for age levels A1 A2 A3';

print 'lambdaBASELINE=' lambdaBASELINE[format=E20.]
'oualphaBASELINE=' oualphaBASELINE alpha;
print 'lambda(age=A1)=' lambdaA1[format=E20.]
'lambda(age=A2)=' lambdaA2[format=E20.]
'lambda(age=A3)=' lambdaA3[format=E20.];
print 'alpha=' alpha[format=E20.] oualphaA1 oualphaA2 oualphaA3;
print 'beta(age=A1)=' betaA1 'beta(age=A2)=' betaA2 'beta(age=A3)=' betaA3
'index(age=A1)=' indexA1[format=E20.] 'index(age=A2)=' indexA2[format=E20.]
'index(age=A3)=' indexA3[format=E20.];

**********Hazard ratio and Odds ratio*******************************************;
hazBASELINE12=(lambdaBASELINE*alpha*12**(alpha-1))/(1+lambdaBASELINE*12**alpha);
hazBASELINE24=(lambdaBASELINE*alpha*24**(alpha-1))/(1+lambdaBASELINE*24**alpha);
survBASELINE12=(1+lambdaBASELINE*12**alpha)**(-1);
survBASELINE24=(1+lambdaBASELINE*24**alpha)**(-1);
oddsBASELINE12=(1-survBASELINE12)/survBASELINE12;
oddsBASELINE24=(1-survBASELINE24)/survBASELINE24;

hazAGE_A1_12=(lambdaA1*alpha*12**(alpha-1))/(1+lambdaA1*12**alpha);
hazAGE_A1_24=(lambdaA1*alpha*24**(alpha-1))/(1+lambdaA1*24**alpha);
survAGE_A1_12=(1+lambdaA1*12**alpha)**(-1);
survAGE_A1_24=(1+lambdaA1*24**alpha)**(-1);
oddsAGE_A1_12=(1-survAGE_A1_12)/survAGE_A1_12;
oddsAGE_A1_24=(1-survAGE_A1_24)/survAGE_A1_24;

hazratioAGE_A1_12=hazAGE_A1_12/hazBASELINE12;
hazratioAGE_A1_24=hazAGE_A1_24/hazBASELINE24;
oddsratioAGE_A1_12=oddsAGE_A1_12/oddsBASELINE12;
oddsratioAGE_A1_24=oddsAGE_A1_24/oddsBASELINE24;

print hazBASELINE12 hazBASELINE24 survBASELINE12 survBASELINE24
      oddsBASELINE12 oddsBASELINE24;

print hazAGE_A1_12 hazAGE_A1_24;
print survAGE_A1_12 survAGE_A1_24;
print oddsAGE_A1_12 oddsAGE_A1_24;
print hazratioAGE_A1_12 hazratioAGE_A1_24 oddsratioAGE_A1_12 oddsratioAGE_A1_24;
********************************************************************************;
hazAGE_A2_12=(lambdaA2*alpha*12**(alpha-1))/(1+lambdaA2*12**alpha);
hazAGE_A2_24=(lambdaA2*alpha*24**(alpha-1))/(1+lambdaA2*24**alpha);
survAGE_A2_12=(1+lambdaA2*12**alpha)**(-1);
survAGE_A2_24=(1+lambdaA2*24**alpha)**(-1);
oddsAGE_A2_12=(1-survAGE_A2_12)/survAGE_A2_12;
oddsAGE_A2_24=(1-survAGE_A2_24)/survAGE_A2_24;

hazratioAGE_A2_12=hazAGE_A2_12/hazBASELINE12;
hazratioAGE_A2_24=hazAGE_A2_24/hazBASELINE24;
oddsratioAGE_A2_12=oddsAGE_A2_12/oddsBASELINE12;
oddsratioAGE_A2_24=oddsAGE_A2_24/oddsBASELINE24;

print hazAGE_A2_12 hazAGE_A2_24;
print survAGE_A2_12 survAGE_A2_24;
print oddsAGE_A2_12 oddsAGE_A2_24;
print hazratioAGE_A2_12 hazratioAGE_A2_24 oddsratioAGE_A2_12 oddsratioAGE_A2_24;
********************************************************************************;
hazAGE_A3_12=(lambdaA3*alpha*12**(alpha-1))/(1+lambdaA3*12**alpha);
hazAGE_A3_24=(lambdaA3*alpha*24**(alpha-1))/(1+lambdaA3*24**alpha);
survAGE_A3_12=(1+lambdaA3*12**alpha)**(-1);
survAGE_A3_24=(1+lambdaA3*24**alpha)**(-1);
oddsAGE_A3_12=(1-survAGE_A3_12)/survAGE_A3_12;
oddsAGE_A3_24=(1-survAGE_A3_24)/survAGE_A3_24;

hazratioAGE_A3_12=hazAGE_A3_12/hazBASELINE12;
hazratioAGE_A3_24=hazAGE_A3_24/hazBASELINE24;
oddsratioAGE_A3_12=oddsAGE_A3_12/oddsBASELINE12;
```

```
oddsratioAGE_A3_24=oddsAGE_A3_24/oddsBASELINE24;

print hazAGE_A3_12 hazAGE_A3_24;
print survAGE_A3_12 survAGE_A3_24;
print oddsAGE_A3_12 oddsAGE_A3_24;
print hazratioAGE_A3_12 hazratioAGE_A3_24 oddsratioAGE_A3_12 oddsratioAGE_A3_24;

**********Median Lifetime******************************************************;
medianBASELINE=(1/lambdaBASELINE)##(1/alpha);
medianA1=(1/lambdaA1)##(1/alpha);
medianA2=(1/lambdaA2)##(1/alpha);
medianA3=(1/lambdaA3)##(1/alpha);

 perc5BASELINE=((1/lambdaBASELINE)#( 5/(100- 5)))##(1/alpha);
perc10BASELINE=((1/lambdaBASELINE)#(10/(100-10)))##(1/alpha);
perc20BASELINE=((1/lambdaBASELINE)#(20/(100-20)))##(1/alpha);
perc25BASELINE=((1/lambdaBASELINE)#(25/(100-25)))##(1/alpha);
perc30BASELINE=((1/lambdaBASELINE)#(30/(100-30)))##(1/alpha);
perc40BASELINE=((1/lambdaBASELINE)#(40/(100-40)))##(1/alpha);
perc50BASELINE=((1/lambdaBASELINE)#(50/(100-50)))##(1/alpha);
perc60BASELINE=((1/lambdaBASELINE)#(60/(100-60)))##(1/alpha);
perc70BASELINE=((1/lambdaBASELINE)#(70/(100-70)))##(1/alpha);
perc75BASELINE=((1/lambdaBASELINE)#(75/(100-75)))##(1/alpha);
perc80BASELINE=((1/lambdaBASELINE)#(80/(100-80)))##(1/alpha);
perc90BASELINE=((1/lambdaBASELINE)#(90/(100-90)))##(1/alpha);
perc95BASELINE=((1/lambdaBASELINE)#(95/(100-95)))##(1/alpha);

 perc5A1=((1/lambdaA1)#( 5/(100- 5)))##(1/alpha);
perc10A1=((1/lambdaA1)#(10/(100-10)))##(1/alpha);
perc20A1=((1/lambdaA1)#(20/(100-20)))##(1/alpha);
perc25A1=((1/lambdaA1)#(25/(100-25)))##(1/alpha);
perc30A1=((1/lambdaA1)#(30/(100-30)))##(1/alpha);
perc40A1=((1/lambdaA1)#(40/(100-40)))##(1/alpha);
perc50A1=((1/lambdaA1)#(50/(100-50)))##(1/alpha);
perc60A1=((1/lambdaA1)#(60/(100-60)))##(1/alpha);
perc70A1=((1/lambdaA1)#(70/(100-70)))##(1/alpha);
perc75A1=((1/lambdaA1)#(75/(100-75)))##(1/alpha);
perc80A1=((1/lambdaA1)#(80/(100-80)))##(1/alpha);
perc90A1=((1/lambdaA1)#(90/(100-90)))##(1/alpha);
perc95A1=((1/lambdaA1)#(95/(100-95)))##(1/alpha);

 perc5A2=((1/lambdaA2)#( 5/(100- 5)))##(1/alpha);
perc10A2=((1/lambdaA2)#(10/(100-10)))##(1/alpha);
perc20A2=((1/lambdaA2)#(20/(100-20)))##(1/alpha);
perc25A2=((1/lambdaA2)#(25/(100-25)))##(1/alpha);
perc30A2=((1/lambdaA2)#(30/(100-30)))##(1/alpha);
perc40A2=((1/lambdaA2)#(40/(100-40)))##(1/alpha);
perc50A2=((1/lambdaA2)#(50/(100-50)))##(1/alpha);
perc60A2=((1/lambdaA2)#(60/(100-60)))##(1/alpha);
perc70A2=((1/lambdaA2)#(70/(100-70)))##(1/alpha);
perc75A2=((1/lambdaA2)#(75/(100-75)))##(1/alpha);
perc80A2=((1/lambdaA2)#(80/(100-80)))##(1/alpha);
perc90A2=((1/lambdaA2)#(90/(100-90)))##(1/alpha);
perc95A2=((1/lambdaA2)#(95/(100-95)))##(1/alpha);

 perc5A3=((1/lambdaA3)#( 5/(100- 5)))##(1/alpha);
perc10A3=((1/lambdaA3)#(10/(100-10)))##(1/alpha);
perc20A3=((1/lambdaA3)#(20/(100-20)))##(1/alpha);
perc25A3=((1/lambdaA3)#(25/(100-25)))##(1/alpha);
perc30A3=((1/lambdaA3)#(30/(100-30)))##(1/alpha);
perc40A3=((1/lambdaA3)#(40/(100-40)))##(1/alpha);
perc50A3=((1/lambdaA3)#(50/(100-50)))##(1/alpha);
perc60A3=((1/lambdaA3)#(60/(100-60)))##(1/alpha);
perc70A3=((1/lambdaA3)#(70/(100-70)))##(1/alpha);
perc75A3=((1/lambdaA3)#(75/(100-75)))##(1/alpha);
perc80A3=((1/lambdaA3)#(80/(100-80)))##(1/alpha);
perc90A3=((1/lambdaA3)#(90/(100-90)))##(1/alpha);
perc95A3=((1/lambdaA3)#(95/(100-95)))##(1/alpha);

print perc5BASELINE perc10BASELINE perc20BASELINE perc25BASELINE perc30BASELINE;
print perc40BASELINE perc50BASELINE perc60BASELINE perc70BASELINE;
print perc75BASELINE perc80BASELINE perc90BASELINE perc95BASELINE;

print perc5A1 perc10A1 perc20A1 perc25A1 perc30A1;
print perc40A1 perc50A1 perc60A1 perc70A1;
print perc75A1 perc80A1 perc90A1 perc95A1;

print perc5A2 perc10A2 perc20A2 perc25A2 perc30A2;
print perc40A2 perc50A2 perc60A2 perc70A2;
print perc75A2 perc80A2 perc90A2 perc95A2;

print perc5A3 perc10A3 perc20A3 perc25A3 perc30A3;
print perc40A3 perc50A3 perc60A3 perc70A3;
print perc75A3 perc80A3 perc90A3 perc95A3;

print 'medianlifetimeBASELINE=' medianBASELINE;
print 'medianlifetime(age=A1)=' medianA1
      'medianlifetime(age=A2)=' medianA2
      'medianlifetime(age=A3)=' medianA3;
```

2. Program for fitting a log-logistic regression model (shape alters, one covariate)

```
title1 'Fitting of regression model with one covariate';
title2 'Staggered entry: shape parameter alters';

proc iml worksize= 60;
reset nolog;
options pagesize=500;

**********Frequency vector (first entry,3 agegroups);
                f11={29,59,95,73,108,15,642};
                f12={21,50,91,45,75,13,553};
                f13={16,49,68,39, 67, 7,471};

**********Vector of upper boundaries;
                x1={12,17,24,28,34,37};

**********Frequency vector (second entry,3 agegroups);
                f21={41,75,103,92,83,628};
                f22={49,62,61,66,54,753};
                f23={28,29,65,42,35,543};

**********Vector of upper boundaries;
                x2={12,17,24,28,34};

**********Frequency vector (third entry,3 agegroups);
                f31={68,34,99,57,570};
                f32={40,44,83,33,533};
                f33={46,21,60,27,571};

**********Vector of upper boundaries;
                x3={12,17,24,28};

**********Frequency vector (fourth entry,3 agegroups);
                f41={71,60,69,573};
                f42={54,61,68,616};
                f43={50,45,70,659};

**********Vector of upper boundaries;
                x4={12,17,24};

**********Relative frequency vectors;
n11=f11[+];    n21=f21[+];    n31=f31[+];    n41=f41[+];    n1=n11+n21+n31+n41;
n12=f12[+];    n22=f22[+];    n32=f32[+];    n42=f42[+];    n2=n12+n22+n32+n42;
n13=f13[+];    n23=f23[+];    n33=f33[+];    n43=f43[+];    n3=n13+n23+n33+n43;
                                                            n=n1+n2+n3;
k1=nrow(f11); d1=k1-1;
k2=nrow(f21); d2=k2-1;
k3=nrow(f31); d3=k3-1;
k4=nrow(f41); d4=k4-1;
k=k1+k2+k3+k4;
d=d1+d2+d3+d4;

p11=f11/n11;    p21=f21/n21;    p31=f31/n31;    p41=f41/n41;    p1=p11//p21//p31//p41;
p12=f12/n12;    p22=f22/n22;    p32=f32/n32;    p42=f42/n42;    p2=p12//p22//p32//p42;
p13=f13/n13;    p23=f23/n23;    p33=f33/n33;    p43=f43/n43;    p3=p13//p23//p33//p43;
                                                                p=p1//p2//p3;

**********Design matrix and matrix orthogonal to design matrix;

S1=J(d1,1,1)@cusum(J(1,k1,1));
S2=J(1,k1,1)@cusum(J(d1,1,1));
S1=S1<=S2;
S2=S1[1:d2,1:d1];
S3=S1[1:d3,1:d2];
S4=S1[1:d4,1:d3];
S=block(S1,S2,S3,S4);
S=I(3)@S;

AA=J(54,1,1);       *54=3 times((7-1)+(6-1)+(5-1)+(4-1))=3 times 18=54;
BB=J(3,1,1);
CC=(I(2)//J(1,2,-1));
DD=J(18,1,1);            *d=18;
EE=J(9,1,1);

lx=log(x1)//log(x2)//log(x3)//log(x4);

xc=AA||(CC@DD)||I(3)@lx;  print xc;

C=I(3#d)-xc*inv(xc`*xc)*xc`;

**********ITERATIVE PROCEDURE (double iterations over m and p);
*****starting value for m;
m=p;
ms=S*m; ps=ms;
p0=p;

*****iteration over m;
itr=0;
```

```
verskil1=1;
i=0;
do while (verskil1>0.00000001);
i=i+1;
p=p0;
*Gm=-C*diag(1/(log(1-ms)))*diag(1/(1-ms))*S;            *Weibull;
Gm=C*diag(1/ms+1/(1-ms))*S;                             *log-logistic;

**********iteration over p;
     verskil=1;
     j=0;
     do while (verskil>0.00000001);
     j=j+1;
     p1=p;
     ps=S*p;
    *g=C*log(-log(1-ps));                                *Weibull;
    g=C*(log(ps)-log(1-ps));                             *log-logistic;
    *Gp=-C*diag(1/(log(1-ps)))*diag(1/(1-ps))*S;         *Weibull;
    Gp=C*(diag(1/ps)+diag(1/(1-ps)))*S;                  *log-logistic;
*****************covariance matrix*********************************;
m11=m[1:k1]; m21=m[k1+1:k1+k2];
m31=m[k1+k2+1:k1+k2+k3]; m41=m[k1+k2+k3+1:k1+k2+k3+k4];

m12=m[k+1:k+k1]; m22=m[k+k1+1:k+k1+k2];
m32=m[k+k1+k2+1:k+k1+k2+k3]; m42=m[k+k1+k2+k3+1:k+k1+k2+k3+k4];

m13=m[2#k+1:2#k+k1]; m23=m[2#k+k1+1:2#k+k1+k2];
m33=m[2#k+k1+k2+1:2#k+k1+k2+k3]; m43=m[2#k+k1+k2+k3+1:2#k+k1+k2+k3+k4];

sig11=(1/n11)*(diag(m11)-m11*m11`);
sig21=(1/n21)*(diag(m21)-m21*m21`);
sig31=(1/n31)*(diag(m31)-m31*m31`);
sig41=(1/n41)*(diag(m41)-m41*m41`);
sig1=block(sig11,sig21,sig31,sig41);

sig12=(1/n12)*(diag(m12)-m12*m12`);
sig22=(1/n22)*(diag(m22)-m22*m22`);
sig32=(1/n32)*(diag(m32)-m32*m32`);
sig42=(1/n42)*(diag(m42)-m42*m42`);
sig2=block(sig12,sig22,sig32,sig42);

sig13=(1/n13)*(diag(m13)-m13*m13`);
sig23=(1/n23)*(diag(m23)-m23*m23`);
sig33=(1/n33)*(diag(m33)-m33*m33`);
sig43=(1/n43)*(diag(m43)-m43*m43`);
sig3=block(sig13,sig23,sig33,sig43);

sig=block(sig1,sig2,sig3);
V=sig;
*****************************************************************;
    *p=p-(Gm*V)`*ginv(Gp*V*Gm`)*g;                      *Weibull;
    p=p-(Gm*V)`*ginv(Gp*V*Gm`)*g;                       *log-logistic;
     verskil=sqrt((p-p1)`*(p-p1));
     end;
verskil1=sqrt((p-m)`*(p-m));
m=p;ms=S*m;
end;
print m; print i j;

**********Parameter vector for linear model;
*par=inv(xc`*xc)*xc`*log(-log(1-ms));                     *Weibull;
par=inv(xc`*xc)*xc`*(log(ms)-log(1-ms)); print par;       *log-logistic;

print par[format=E20.];

**********Regression coefficients;
oualphaBASELINE=par[1];
lambdaBASELINE=exp(par[1]);
betaA1=par[2];
betaA2=par[3];
betaA3=-(par[2]+par[3]);
lambdaA1=exp(par[1]+par[2]);
lambdaA2=exp(par[1]+par[3]);
lambdaA3=exp(par[1]+betaA3);
alphaA1=par[4];
alphaA2=par[5];
alphaA3=par[6];
alphaBASELINE=(n1#alphaA1+n2#alphaA2+n3#alphaA3)/n;

**********Lambda's and alpha's for age levels*****************;
lambdaA1=exp(par[1]+par[2]);
lambdaA2=exp(par[1]+par[3]);
lambdaA3=exp(par[1]+betaA3);
oualphaA1=log(lambdaA1);
oualphaA2=log(lambdaA2);
oualphaA3=log(lambdaA3);

print ' Loglogistic parameters and beta effects for age levels A1 A2 A3';
print 'lambdaBASELINE=' lambdaBASELINE[format=E20.]
```

```
      'oualphaBASELINE=' oualphaBASELINE alphaBASELINE;
print 'lambda(age=A1)=' lambdaA1[format=E20.]
      'lambda(age=A2)=' lambdaA2[format=E20.]
      'lambda(age=A3)=' lambdaA3[format=E20.];
print oualphaA1 oualphaA2 oualphaA3;
print 'alpha(age=A1)=' alphaA1[format=E20.]
      'alpha(age=A2)=' alphaA2[format=E20.]
      'alpha(age=A3)=' alphaA3[format=E20.];

print 'beta(age=A1)=' betaA1 'beta(age=A2)=' betaA2 'beta(age=A3)=' betaA3;

**********Hazard ratio and Odds ratio******************************************;
hazBASELINE12=(lambdaBASELINE*alphaBASELINE*12**(alphaBASELINE-1))/(1+lambdaBASELINE*12**alphaBASELINE);
hazBASELINE24=(lambdaBASELINE*alphaBASELINE*24**(alphaBASELINE-1))/(1+lambdaBASELINE*24**alphaBASELINE);
survBASELINE12=(1+lambdaBASELINE*12**alphaBASELINE)**(-1);
survBASELINE24=(1+lambdaBASELINE*24**alphaBASELINE)**(-1);
oddsBASELINE12=(1-survBASELINE12)/survBASELINE12;
oddsBASELINE24=(1-survBASELINE24)/survBASELINE24;

hazAGE_A1_12=(lambdaA1*alphaA1*12**(alphaA1-1))/(1+lambdaA1*12**alphaA1);
hazAGE_A1_24=(lambdaA1*alphaA1*24**(alphaA1-1))/(1+lambdaA1*24**alphaA1);
survAGE_A1_12=(1+lambdaA1*12**alphaA1)**(-1);
survAGE_A1_24=(1+lambdaA1*24**alphaA1)**(-1);
oddsAGE_A1_12=(1-survAGE_A1_12)/survAGE_A1_12;
oddsAGE_A1_24=(1-survAGE_A1_24)/survAGE_A1_24;

hazratioAGE_A1_12=hazAGE_A1_12/hazBASELINE12;
hazratioAGE_A1_24=hazAGE_A1_24/hazBASELINE24;
oddsratioAGE_A1_12=oddsAGE_A1_12/oddsBASELINE12;
oddsratioAGE_A1_24=oddsAGE_A1_24/oddsBASELINE24;

print hazBASELINE12 hazBASELINE24 survBASELINE12 survBASELINE24
      oddsBASELINE12 oddsBASELINE24;

print hazAGE_A1_12 hazAGE_A1_24;
print survAGE_A1_12 survAGE_A1_24;
print oddsAGE_A1_12 oddsAGE_A1_24;
print hazratioAGE_A1_12 hazratioAGE_A1_24 oddsratioAGE_A1_12 oddsratioAGE_A1_24;
******************************************************************************;
hazAGE_A2_12=(lambdaA2*alphaA2*12**(alphaA2-1))/(1+lambdaA2*12**alphaA2);
hazAGE_A2_24=(lambdaA2*alphaA2*24**(alphaA2-1))/(1+lambdaA2*24**alphaA2);
survAGE_A2_12=(1+lambdaA2*12**alphaA2)**(-1);
survAGE_A2_24=(1+lambdaA2*24**alphaA2)**(-1);
oddsAGE_A2_12=(1-survAGE_A2_12)/survAGE_A2_12;
oddsAGE_A2_24=(1-survAGE_A2_24)/survAGE_A2_24;

hazratioAGE_A2_12=hazAGE_A2_12/hazBASELINE12;
hazratioAGE_A2_24=hazAGE_A2_24/hazBASELINE24;
oddsratioAGE_A2_12=oddsAGE_A2_12/oddsBASELINE12;
oddsratioAGE_A2_24=oddsAGE_A2_24/oddsBASELINE24;

print hazAGE_A2_12 hazAGE_A2_24;
print survAGE_A2_12 survAGE_A2_24;
print oddsAGE_A2_12 oddsAGE_A2_24;
print hazratioAGE_A2_12 hazratioAGE_A2_24 oddsratioAGE_A2_12 oddsratioAGE_A2_24;
******************************************************************************;
hazAGE_A3_12=(lambdaA3*alphaA3*12**(alphaA3-1))/(1+lambdaA3*12**alphaA3);
hazAGE_A3_24=(lambdaA3*alphaA3*24**(alphaA3-1))/(1+lambdaA3*24**alphaA3);
survAGE_A3_12=(1+lambdaA3*12**alphaA3)**(-1);
survAGE_A3_24=(1+lambdaA3*24**alphaA3)**(-1);
oddsAGE_A3_12=(1-survAGE_A3_12)/survAGE_A3_12;
oddsAGE_A3_24=(1-survAGE_A3_24)/survAGE_A3_24;

hazratioAGE_A3_12=hazAGE_A3_12/hazBASELINE12;
hazratioAGE_A3_24=hazAGE_A3_24/hazBASELINE24;
oddsratioAGE_A3_12=oddsAGE_A3_12/oddsBASELINE12;
oddsratioAGE_A3_24=oddsAGE_A3_24/oddsBASELINE24;

print hazAGE_A3_12 hazAGE_A3_24;
print survAGE_A3_12 survAGE_A3_24;
print oddsAGE_A3_12 oddsAGE_A3_24;
print hazratioAGE_A3_12 hazratioAGE_A3_24 oddsratioAGE_A3_12 oddsratioAGE_A3_24;

**********Median Lifetime*****************************************************;
medianBASELINE=(1/lambdaBASELINE)##(1/alphaBASELINE);
medianA1=(1/lambdaA1)##(1/alphaA1);
medianA2=(1/lambdaA2)##(1/alphaA2);
medianA3=(1/lambdaA3)##(1/alphaA3);

 perc5BASELINE=((1/lambdaBASELINE)#( 5/(100- 5))##(1/alphaBASELINE);
perc10BASELINE=((1/lambdaBASELINE)#(10/(100-10))##(1/alphaBASELINE);
perc20BASELINE=((1/lambdaBASELINE)#(20/(100-20))##(1/alphaBASELINE);
perc25BASELINE=((1/lambdaBASELINE)#(25/(100-25))##(1/alphaBASELINE);
perc30BASELINE=((1/lambdaBASELINE)#(30/(100-30))##(1/alphaBASELINE);
perc40BASELINE=((1/lambdaBASELINE)#(40/(100-40))##(1/alphaBASELINE);
perc50BASELINE=((1/lambdaBASELINE)#(50/(100-50))##(1/alphaBASELINE);
perc60BASELINE=((1/lambdaBASELINE)#(60/(100-60))##(1/alphaBASELINE);
perc70BASELINE=((1/lambdaBASELINE)#(70/(100-70))##(1/alphaBASELINE);
perc75BASELINE=((1/lambdaBASELINE)#(75/(100-75))##(1/alphaBASELINE);
```

```
perc80BASELINE=((1/lambdaBASELINE)#(80/(100-80)))##(1/alphaBASELINE);
perc90BASELINE=((1/lambdaBASELINE)#(90/(100-90)))##(1/alphaBASELINE);
perc95BASELINE=((1/lambdaBASELINE)#(95/(100-95)))##(1/alphaBASELINE);

 perc5A1=((1/lambdaA1)#( 5/(100- 5)))##(1/alphaA1);
perc10A1=((1/lambdaA1)#(10/(100-10)))##(1/alphaA1);
perc20A1=((1/lambdaA1)#(20/(100-20)))##(1/alphaA1);
perc25A1=((1/lambdaA1)#(25/(100-25)))##(1/alphaA1);
perc30A1=((1/lambdaA1)#(30/(100-30)))##(1/alphaA1);
perc40A1=((1/lambdaA1)#(40/(100-40)))##(1/alphaA1);
perc50A1=((1/lambdaA1)#(50/(100-50)))##(1/alphaA1);
perc60A1=((1/lambdaA1)#(60/(100-60)))##(1/alphaA1);
perc70A1=((1/lambdaA1)#(70/(100-70)))##(1/alphaA1);
perc75A1=((1/lambdaA1)#(75/(100-75)))##(1/alphaA1);
perc80A1=((1/lambdaA1)#(80/(100-80)))##(1/alphaA1);
perc90A1=((1/lambdaA1)#(90/(100-90)))##(1/alphaA1);
perc95A1=((1/lambdaA1)#(95/(100-95)))##(1/alphaA1);

 perc5A2=((1/lambdaA2)#( 5/(100- 5)))##(1/alphaA2);
perc10A2=((1/lambdaA2)#(10/(100-10)))##(1/alphaA2);
perc20A2=((1/lambdaA2)#(20/(100-20)))##(1/alphaA2);
perc25A2=((1/lambdaA2)#(25/(100-25)))##(1/alphaA2);
perc30A2=((1/lambdaA2)#(30/(100-30)))##(1/alphaA2);
perc40A2=((1/lambdaA2)#(40/(100-40)))##(1/alphaA2);
perc50A2=((1/lambdaA2)#(50/(100-50)))##(1/alphaA2);
perc60A2=((1/lambdaA2)#(60/(100-60)))##(1/alphaA2);
perc70A2=((1/lambdaA2)#(70/(100-70)))##(1/alphaA2);
perc75A2=((1/lambdaA2)#(75/(100-75)))##(1/alphaA2);
perc80A2=((1/lambdaA2)#(80/(100-80)))##(1/alphaA2);
perc90A2=((1/lambdaA2)#(90/(100-90)))##(1/alphaA2);
perc95A2=((1/lambdaA2)#(95/(100-95)))##(1/alphaA2);

 perc5A3=((1/lambdaA3)#( 5/(100- 5)))##(1/alphaA3);
perc10A3=((1/lambdaA3)#(10/(100-10)))##(1/alphaA3);
perc20A3=((1/lambdaA3)#(20/(100-20)))##(1/alphaA3);
perc25A3=((1/lambdaA3)#(25/(100-25)))##(1/alphaA3);
perc30A3=((1/lambdaA3)#(30/(100-30)))##(1/alphaA3);
perc40A3=((1/lambdaA3)#(40/(100-40)))##(1/alphaA3);
perc50A3=((1/lambdaA3)#(50/(100-50)))##(1/alphaA3);
perc60A3=((1/lambdaA3)#(60/(100-60)))##(1/alphaA3);
perc70A3=((1/lambdaA3)#(70/(100-70)))##(1/alphaA3);
perc75A3=((1/lambdaA3)#(75/(100-75)))##(1/alphaA3);
perc80A3=((1/lambdaA3)#(80/(100-80)))##(1/alphaA3);
perc90A3=((1/lambdaA3)#(90/(100-90)))##(1/alphaA3);
perc95A3=((1/lambdaA3)#(95/(100-95)))##(1/alphaA3);

print perc5BASELINE perc10BASELINE perc20BASELINE perc25BASELINE perc30BASELINE;
print perc40BASELINE perc50BASELINE perc60BASELINE perc70BASELINE;
print perc75BASELINE perc80BASELINE perc90BASELINE perc95BASELINE;

print perc5A1 perc10A1 perc20A1 perc25A1 perc30A1;
print perc40A1 perc50A1 perc60A1 perc70A1;
print perc75A1 perc80A1 perc90A1 perc95A1;

print perc5A2 perc10A2 perc20A2 perc25A2 perc30A2;
print perc40A2 perc50A2 perc60A2 perc70A2;
print perc75A2 perc80A2 perc90A2 perc95A2;

print perc5A3 perc10A3 perc20A3 perc25A3 perc30A3;
print perc40A3 perc50A3 perc60A3 perc70A3;
print perc75A3 perc80A3 perc90A3 perc95A3;

print 'medianlifetimeBASELINE=' medianBASELINE;
print 'medianlifetime(age=A1)=' medianA1
      'medianlifetime(age=A2)=' medianA2
      'medianlifetime(age=A3)=' medianA3;
```

3. Program for fitting a Weibull regression model (constant shape, one covariate)

```
title1 'Fitting of regression model with one covariate';
title2 'Staggered entry: constant shape parameter';

proc iml worksize= 60;
reset nolog;
options pagesize=500;

**********Frequency vector (first entry,3 agegroups);
                f11={29,59,95,73,108,15,642};
                f12={21,50,91,45,75,13,553};
                f13={16,49,68,39, 67, 7,471};

**********Vector of upper boundaries;
                x1={12,17,24,28,34,37};

**********Frequency vector (second entry,3 agegroups);
                f21={41,75,103,92,83,628};
                f22={49,62,61,66,54,753};
                f23={28,29,65,42,35,543};
```

```
**********Vector of upper boundaries;
                x2={12,17,24,28,34};

**********Frequency vector (third entry,3 agegroups);
                f31={68,34,99,57,570};
                f32={40,44,83,33,533};
                f33={46,21,60,27,571};

**********Vector of upper boundaries;
                x3={12,17,24,28};

**********Frequency vector (fourth entry,3 agegroups);
                f41={71,60,69,573};
                f42={54,61,68,616};
                f43={50,45,70,659};

**********Vector of upper boundaries;
                x4={12,17,24};

**********Relative frequency vectors;
n11=f11[+];    n21=f21[+];    n31=f31[+];    n41=f41[+];    n1=n11+n21+n31+n41;
n12=f12[+];    n22=f22[+];    n32=f32[+];    n42=f42[+];    n2=n12+n22+n32+n42;
n13=f13[+];    n23=f23[+];    n33=f33[+];    n43=f43[+];    n3=n13+n23+n33+n43;
                                                              n=n1+n2+n3;
k1=nrow(f11); d1=k1-1;
k2=nrow(f21); d2=k2-1;
k3=nrow(f31); d3=k3-1;
k4=nrow(f41); d4=k4-1;
k=k1+k2+k3+k4;
d=d1+d2+d3+d4;

p11=f11/n11;    p21=f21/n21;    p31=f31/n31;    p41=f41/n41;    p1=p11//p21//p31//p41;
p12=f12/n12;    p22=f22/n22;    p32=f32/n32;    p42=f42/n42;    p2=p12//p22//p32//p42;
p13=f13/n13;    p23=f23/n23;    p33=f33/n33;    p43=f43/n43;    p3=p13//p23//p33//p43;
                                                                p=p1//p2//p3;

**********Design matrix and matrix orthogonal to design matrix;

S1=J(d1,1,1)@cusum(J(1,k1,1));
S2=J(1,k1,1)@cusum(J(d1,1,1));
S1=S1<=S2;
S2=S1[1:d2,1:d1];
S3=S1[1:d3,1:d2];
S4=S1[1:d4,1:d3];
S=block(S1,S2,S3,S4);
S=I(3)@S;

AA=J(54,1,1);        *54=3 times((7-1)+(6-1)+(5-1)+(4-1))=3 times 18=54;
BB=J(3,1,1);
CC=(I(2)//J(1,2,-1));
DD=J(18,1,1);             *d=18;
EE=J(9,1,1);

lx=BB@(log(x1)//log(x2)//log(x3)//log(x4));

xc=AA||(CC@DD)||lx;  print xc;

C=I(3#d)-xc*inv(xc`*xc)*xc`;

**********ITERATIVE PROCEDURE (double iterations over m and p);
*****starting value for m;
m=p;
ms=S*m; ps=ms;
p0=p;

*****iteration over m;
itr=0;
verskil1=1;
i=0;
do while (verskil1>0.00000001);
i=i+1;
p=p0;
Gm=-C*diag(1/(log(1-ms)))*diag(1/(1-ms))*S;              *Weibull;
*Gm=C*diag(1/ms+1/(1-ms))*S;                             *log-logistic;

**********iteration over p;
     verskil=1;
     j=0;
     do while (verskil>0.00000001);
     j=j+1;
     p1=p;
     ps=S*p;
    g=C*log(-log(1-ps));                                  *Weibull;
    *g=C*(log(ps)-log(1-ps));                             *Loglogistic;
    Gp=-C*diag(1/(log(1-ps)))*diag(1/(1-ps))*S;           *Weibull;
    *Gp=C*(diag(1/ps)+diag(1/(1-ps)))*S;                  *loglogistic;
******************covariance matrix**********************************;
m11=m[1:k1]; m21=m[k1+1:k1+k2];
m31=m[k1+k2+1:k1+k2+k3]; m41=m[k1+k2+k3+1:k1+k2+k3+k4];
```

```
m12=m[k+1:k+k1]; m22=m[k+k1+1:k+k1+k2];
m32=m[k+k1+k2+1:k+k1+k2+k3]; m42=m[k+k1+k2+k3+1:k+k1+k2+k3+k4];

m13=m[2#k+1:2#k+k1]; m23=m[2#k+k1+1:2#k+k1+k2];
m33=m[2#k+k1+k2+1:2#k+k1+k2+k3]; m43=m[2#k+k1+k2+k3+1:2#k+k1+k2+k3+k4];

sig11=(1/n11)*(diag(m11)-m11*m11`);
sig21=(1/n21)*(diag(m21)-m21*m21`);
sig31=(1/n31)*(diag(m31)-m31*m31`);
sig41=(1/n41)*(diag(m41)-m41*m41`);
sig1=block(sig11,sig21,sig31,sig41);

sig12=(1/n12)*(diag(m12)-m12*m12`);
sig22=(1/n22)*(diag(m22)-m22*m22`);
sig32=(1/n32)*(diag(m32)-m32*m32`);
sig42=(1/n42)*(diag(m42)-m42*m42`);
sig2=block(sig12,sig22,sig32,sig42);

sig13=(1/n13)*(diag(m13)-m13*m13`);
sig23=(1/n23)*(diag(m23)-m23*m23`);
sig33=(1/n33)*(diag(m33)-m33*m33`);
sig43=(1/n43)*(diag(m43)-m43*m43`);
sig3=block(sig13,sig23,sig33,sig43);

sig=block(sig1,sig2,sig3);
V=sig;
**********************************************************************;
    p=p-(Gm*V)`*ginv(Gp*V*Gm`)*g;                        *Weibull;
    *p=p-(Gm*V)`*ginv(Gp*V*Gm`)*g;                       *loglogistic;
     verskil=sqrt((p-p1)`*(p-p1));
     end;
verskil1=sqrt((p-m)`*(p-m));
m=p;ms=S*m;
end;
print m; print i j;

**********Parameter vector for linear model;
par=inv(xc`*xc)*xc`*log(-log(1-ms));                       *Weibull;
*par=inv(xc`*xc)*xc`*(log(ms)-log(1-ms)); print par;       *log-logistic;

print par[format=E20.];

**********Regression coefficients;
oualphaBASELINE=par[1];
betaA1=par[2];
betaA2=par[3];
betaA3=-(par[2]+par[3]);
lambdaBASELINE=exp(par[1]);
lambdaA1=exp(par[1]+par[2]);
lambdaA2=exp(par[1]+par[3]);
lambdaA3=exp(par[1]+betaA3);
alpha=par[4];

**********Risk scores, lambda's and constant alpha for age levels*****************;
riskscoreA1=exp(betaA1);      *constant shape parameter;
riskscoreA2=exp(betaA2);
riskscoreA3=exp(betaA3);

lambdaA1=lambdaBASELINE#riskscoreA1;
lambdaA2=lambdaBASELINE#riskscoreA2;
lambdaA3=lambdaBASELINE#riskscoreA3;
oualphaA1=log(lambdaA1);
oualphaA2=log(lambdaA2);
oualphaA3=log(lambdaA3);

print ' Weibull parameters, beta effects and risk scores for age levels A1 A2 A3';
print 'lambdaBASELINE=' lambdaBASELINE[format=E20.]
'oualphaBASELINE=' oualphaBASELINE
'lambda(age=A1)=' lambdaA1[format=E20.]
'lambda(age=A2)=' lambdaA2[format=E20.]
'lambda(age=A3)=' lambdaA3[format=E20.]
'alpha=' alpha[format=E20.] oualphaA1 oualphaA2 oualphaA3
'beta(age=A1)=' betaA1 'beta(age=A2)=' betaA2 'beta(age=A3)=' betaA3
'riskscore(age=A1)=' riskscoreA1[format=E20.] 'riskscore(age=A2)=' riskscoreA2[format=E20.]
'riskscore(age=A3)=' riskscoreA3[format=E20.];

**********Hazard ratio and Odds ratio******************************************;
hazBASELINE12=lambdaBASELINE*alpha*12**(alpha-1);
hazBASELINE24=lambdaBASELINE*alpha*24**(alpha-1);
survBASELINE12=exp(-lambdaBASELINE*12**alpha);
survBASELINE24=exp(-lambdaBASELINE*24**alpha);
oddsBASELINE12=(1-survBASELINE12)/survBASELINE12;
oddsBASELINE24=(1-survBASELINE24)/survBASELINE24;

hazAGE_A1_12=lambdaA1*alpha*12**(alpha-1);
hazAGE_A1_24=lambdaA1*alpha*24**(alpha-1);
survAGE_A1_12=exp(-lambdaA1*12**alpha);
survAGE_A1_24=exp(-lambdaA1*24**alpha);
```

```
oddsAGE_A1_12=(1-survAGE_A1_12)/survAGE_A1_12;
oddsAGE_A1_24=(1-survAGE_A1_24)/survAGE_A1_24;

hazratioAGE_A1_12=hazAGE_A1_12/hazBASELINE12;
hazratioAGE_A1_24=hazAGE_A1_24/hazBASELINE24;
oddsratioAGE_A1_12=oddsAGE_A1_12/oddsBASELINE12;
oddsratioAGE_A1_24=oddsAGE_A1_24/oddsBASELINE24;

print hazBASELINE12 hazBASELINE24 survBASELINE12 survBASELINE24
      oddsBASELINE12 oddsBASELINE24;

print hazAGE_A1_12 hazAGE_A1_24;
print survAGE_A1_12 survAGE_A1_24;
print oddsAGE_A1_12 oddsAGE_A1_24;
print hazratioAGE_A1_12 hazratioAGE_A1_24 oddsratioAGE_A1_12 oddsratioAGE_A1_24;
********************************************************************************;
hazAGE_A2_12=lambdaA2*alpha*12**(alpha-1);
hazAGE_A2_24=lambdaA2*alpha*24**(alpha-1);
survAGE_A2_12=exp(-lambdaA2*12**alpha);
survAGE_A2_24=exp(-lambdaA2*24**alpha);
oddsAGE_A2_12=(1-survAGE_A2_12)/survAGE_A2_12;
oddsAGE_A2_24=(1-survAGE_A2_24)/survAGE_A2_24;

hazratioAGE_A2_12=hazAGE_A2_12/hazBASELINE12;
hazratioAGE_A2_24=hazAGE_A2_24/hazBASELINE24;
oddsratioAGE_A2_12=oddsAGE_A2_12/oddsBASELINE12;
oddsratioAGE_A2_24=oddsAGE_A2_24/oddsBASELINE24;

print hazAGE_A2_12 hazAGE_A2_24;
print survAGE_A2_12 survAGE_A2_24;
print oddsAGE_A2_12 oddsAGE_A2_24;
print hazratioAGE_A2_12 hazratioAGE_A2_24 oddsratioAGE_A2_12 oddsratioAGE_A2_24;
********************************************************************************;
hazAGE_A3_12=lambdaA3*alpha*12**(alpha-1);
hazAGE_A3_24=lambdaA3*alpha*24**(alpha-1);
survAGE_A3_12=exp(-lambdaA3*12**alpha);
survAGE_A3_24=exp(-lambdaA3*24**alpha);
oddsAGE_A3_12=(1-survAGE_A3_12)/survAGE_A3_12;
oddsAGE_A3_24=(1-survAGE_A3_24)/survAGE_A3_24;

hazratioAGE_A3_12=hazAGE_A3_12/hazBASELINE12;
hazratioAGE_A3_24=hazAGE_A3_24/hazBASELINE24;
oddsratioAGE_A3_12=oddsAGE_A3_12/oddsBASELINE12;
oddsratioAGE_A3_24=oddsAGE_A3_24/oddsBASELINE24;

print hazAGE_A3_12 hazAGE_A3_24;
print survAGE_A3_12 survAGE_A3_24;
print oddsAGE_A3_12 oddsAGE_A3_24;
print hazratioAGE_A3_12 hazratioAGE_A3_24 oddsratioAGE_A3_12 oddsratioAGE_A3_24;

**********Median Lifetime*******************************************************;
medianBASELINE=((1/lambdaBASELINE)#log(2))##(1/alpha);
medianA1=((1/lambdaA1)#log(2))##(1/alpha);
medianA2=((1/lambdaA2)#log(2))##(1/alpha);
medianA3=((1/lambdaA3)#log(2))##(1/alpha);

 perc5BASELINE=((1/lambdaBASELINE)#log(100/(100- 5))##(1/alpha);
perc10BASELINE=((1/lambdaBASELINE)#log(100/(100-10)))##(1/alpha);
perc20BASELINE=((1/lambdaBASELINE)#log(100/(100-20)))##(1/alpha);
perc25BASELINE=((1/lambdaBASELINE)#log(100/(100-25)))##(1/alpha);
perc30BASELINE=((1/lambdaBASELINE)#log(100/(100-30)))##(1/alpha);
perc40BASELINE=((1/lambdaBASELINE)#log(100/(100-40)))##(1/alpha);
perc50BASELINE=((1/lambdaBASELINE)#log(100/(100-50)))##(1/alpha);
perc60BASELINE=((1/lambdaBASELINE)#log(100/(100-60)))##(1/alpha);
perc70BASELINE=((1/lambdaBASELINE)#log(100/(100-70)))##(1/alpha);
perc75BASELINE=((1/lambdaBASELINE)#log(100/(100-75)))##(1/alpha);
perc80BASELINE=((1/lambdaBASELINE)#log(100/(100-80)))##(1/alpha);
perc90BASELINE=((1/lambdaBASELINE)#log(100/(100-90)))##(1/alpha);
perc95BASELINE=((1/lambdaBASELINE)#log(100/(100-95)))##(1/alpha);

 perc5A1=((1/lambdaA1)#log(100/(100- 5)))##(1/alpha);
perc10A1=((1/lambdaA1)#log(100/(100-10)))##(1/alpha);
perc20A1=((1/lambdaA1)#log(100/(100-20)))##(1/alpha);
perc25A1=((1/lambdaA1)#log(100/(100-25)))##(1/alpha);
perc30A1=((1/lambdaA1)#log(100/(100-30)))##(1/alpha);
perc40A1=((1/lambdaA1)#log(100/(100-40)))##(1/alpha);
perc50A1=((1/lambdaA1)#log(100/(100-50)))##(1/alpha);
perc60A1=((1/lambdaA1)#log(100/(100-60)))##(1/alpha);
perc70A1=((1/lambdaA1)#log(100/(100-70)))##(1/alpha);
perc75A1=((1/lambdaA1)#log(100/(100-75)))##(1/alpha);
perc80A1=((1/lambdaA1)#log(100/(100-80)))##(1/alpha);
perc90A1=((1/lambdaA1)#log(100/(100-90)))##(1/alpha);
perc95A1=((1/lambdaA1)#log(100/(100-95)))##(1/alpha);

 perc5A2=((1/lambdaA2)#log(100/(100- 5)))##(1/alpha);
perc10A2=((1/lambdaA2)#log(100/(100-10)))##(1/alpha);
perc20A2=((1/lambdaA2)#log(100/(100-20)))##(1/alpha);
perc25A2=((1/lambdaA2)#log(100/(100-25)))##(1/alpha);
perc30A2=((1/lambdaA2)#log(100/(100-30)))##(1/alpha);
```

```
perc40A2=((1/lambdaA2)#log(100/(100-40)))##(1/alpha);
perc50A2=((1/lambdaA2)#log(100/(100-50)))##(1/alpha);
perc60A2=((1/lambdaA2)#log(100/(100-60)))##(1/alpha);
perc70A2=((1/lambdaA2)#log(100/(100-70)))##(1/alpha);
perc75A2=((1/lambdaA2)#log(100/(100-75)))##(1/alpha);
perc80A2=((1/lambdaA2)#log(100/(100-80)))##(1/alpha);
perc90A2=((1/lambdaA2)#log(100/(100-90)))##(1/alpha);
perc95A2=((1/lambdaA2)#log(100/(100-95)))##(1/alpha);

 perc5A3=((1/lambdaA3)#log(100/(100- 5)))##(1/alpha);
perc10A3=((1/lambdaA3)#log(100/(100-10)))##(1/alpha);
perc20A3=((1/lambdaA3)#log(100/(100-20)))##(1/alpha);
perc25A3=((1/lambdaA3)#log(100/(100-25)))##(1/alpha);
perc30A3=((1/lambdaA3)#log(100/(100-30)))##(1/alpha);
perc40A3=((1/lambdaA3)#log(100/(100-40)))##(1/alpha);
perc50A3=((1/lambdaA3)#log(100/(100-50)))##(1/alpha);
perc60A3=((1/lambdaA3)#log(100/(100-60)))##(1/alpha);
perc70A3=((1/lambdaA3)#log(100/(100-70)))##(1/alpha);
perc75A3=((1/lambdaA3)#log(100/(100-75)))##(1/alpha);
perc80A3=((1/lambdaA3)#log(100/(100-80)))##(1/alpha);
perc90A3=((1/lambdaA3)#log(100/(100-90)))##(1/alpha);
perc95A3=((1/lambdaA3)#log(100/(100-95)))##(1/alpha);

print perc5BASELINE perc10BASELINE perc20BASELINE perc25BASELINE perc30BASELINE;
print perc40BASELINE perc50BASELINE perc60BASELINE perc70BASELINE;
print perc75BASELINE perc80BASELINE perc90BASELINE perc95BASELINE;

print perc5A1 perc10A1 perc20A1 perc25A1 perc30A1;
print perc40A1 perc50A1 perc60A1 perc70A1;
print perc75A1 perc80A1 perc90A1 perc95A1;

print perc5A2 perc10A2 perc20A2 perc25A2 perc30A2;
print perc40A2 perc50A2 perc60A2 perc70A2;
print perc75A2 perc80A2 perc90A2 perc95A2;

print perc5A3 perc10A3 perc20A3 perc25A3 perc30A3;
print perc40A3 perc50A3 perc60A3 perc70A3;
print perc75A3 perc80A3 perc90A3 perc95A3;

print 'medianlifetimeBASELINE=' medianBASELINE;
print 'medianlifetime(age=A1)=' medianA1
      'medianlifetime(age=A2)=' medianA2
      'medianlifetime(age=A3)=' medianA3;
```

4. Program for fitting a Weibull regression model (shape alters, one covariate)

```
title1 'Fitting of regression model with one covariate';
title2 'Staggered entry: shape parameter alters';

proc iml worksize= 60;
reset nolog;
options pagesize=500;

**********Frequency vector (first entry,3 agegroups);
                f11={29,59,95,73,108,15,642};
                f12={21,50,91,45,75,13,553};
                f13={16,49,68,39, 67, 7,471};

**********Vector of upper boundaries;
                x1={12,17,24,28,34,37};

**********Frequency vector (second entry,3 agegroups);
                f21={41,75,103,92,83,628};
                f22={49,62,61,66,54,753};
                f23={28,29,65,42,35,543};

**********Vector of upper boundaries;
                x2={12,17,24,28,34};

**********Frequency vector (third entry,3 agegroups);
                f31={68,34,99,57,570};
                f32={40,44,83,33,533};
                f33={46,21,60,27,571};

**********Vector of upper boundaries;
                x3={12,17,24,28};

**********Frequency vector (fourth entry,3 agegroups);
                f41={71,60,69,573};
                f42={54,61,68,616};
                f43={50,45,70,659};

**********Vector of upper boundaries;
                x4={12,17,24};

**********Relative frequency vectors;
n11=f11[+];    n21=f21[+];    n31=f31[+];    n41=f41[+];    n1=n11+n21+n31+n41;
n12=f12[+];    n22=f22[+];    n32=f32[+];    n42=f42[+];    n2=n12+n22+n32+n42;
n13=f13[+];    n23=f23[+];    n33=f33[+];    n43=f43[+];    n3=n13+n23+n33+n43;
```

```
                                                            n=n1+n2+n3;
k1=nrow(f11); d1=k1-1;
k2=nrow(f21); d2=k2-1;
k3=nrow(f31); d3=k3-1;
k4=nrow(f41); d4=k4-1;
k=k1+k2+k3+k4;
d=d1+d2+d3+d4;

p11=f11/n11;   p21=f21/n21;   p31=f31/n31;   p41=f41/n41;   p1=p11//p21//p31//p41;
p12=f12/n12;   p22=f22/n22;   p32=f32/n32;   p42=f42/n42;   p2=p12//p22//p32//p42;
p13=f13/n13;   p23=f23/n23;   p33=f33/n33;   p43=f43/n43;   p3=p13//p23//p33//p43;
                                                            p=p1//p2//p3;

**********Design matrix and matrix orthogonal to design matrix;

S1=J(d1,1,1)@cusum(J(1,k1,1));
S2=J(1,k1,1)@cusum(J(d1,1,1));
S1=S1<=S2;
S2=S1[1:d2,1:d1];
S3=S1[1:d3,1:d2];
S4=S1[1:d4,1:d3];
S=block(S1,S2,S3,S4);
S=I(3)@S;

AA=J(54,1,1);       *54=3 times((7-1)+(6-1)+(5-1)+(4-1))=3 times 18=54;
BB=J(3,1,1);
CC=(I(2)//J(1,2,-1));
DD=J(18,1,1);           *d=18;
EE=J(9,1,1);

lx=log(x1)//log(x2)//log(x3)//log(x4);

xc=AA||(CC@DD)||I(3)@lx;  print xc;

C=I(3#d)-xc*inv(xc`*xc)*xc`;

**********ITERATIVE PROCEDURE (double iterations over m and p);
*****starting value for m;
m=p;
ms=S*m; ps=ms;
p0=p;

*****iteration over m;
itr=0;
verskil1=1;
i=0;
do while (verskil1>0.00000001);
i=i+1;
p=p0;
Gm=-C*diag(1/(log(1-ms)))*diag(1/(1-ms))*S;              *Weibull;
*Gm=C*diag(1/ms+1/(1-ms))*S;                               *log-logistic;

**********iteration over p;
     verskil=1;
     j=0;
     do while (verskil>0.00000001);
     j=j+1;
     p1=p;
     ps=S*p;
    g=C*log(-log(1-ps));                                  *Weibull;
    *g=C*(log(ps)-log(1-ps));                              *Loglogistic;
    Gp=-C*diag(1/(log(1-ps)))*diag(1/(1-ps))*S;           *Weibull;
    *Gp=C*(diag(1/ps)+diag(1/(1-ps)))*S;                   *loglogistic;
******************covariance matrix**********************************;
m11=m[1:k1]; m21=m[k1+1:k1+k2];
m31=m[k1+k2+1:k1+k2+k3]; m41=m[k1+k2+k3+1:k1+k2+k3+k4];

m12=m[k+1:k+k1]; m22=m[k+k1+1:k+k1+k2];
m32=m[k+k1+k2+1:k+k1+k2+k3]; m42=m[k+k1+k2+k3+1:k+k1+k2+k3+k4];

m13=m[2#k+1:2#k+k1]; m23=m[2#k+k1+1:2#k+k1+k2];
m33=m[2#k+k1+k2+1:2#k+k1+k2+k3]; m43=m[2#k+k1+k2+k3+1:2#k+k1+k2+k3+k4];

sig11=(1/n11)*(diag(m11)-m11*m11`);
sig21=(1/n21)*(diag(m21)-m21*m21`);
sig31=(1/n31)*(diag(m31)-m31*m31`);
sig41=(1/n41)*(diag(m41)-m41*m41`);
sig1=block(sig11,sig21,sig31,sig41);

sig12=(1/n12)*(diag(m12)-m12*m12`);
sig22=(1/n22)*(diag(m22)-m22*m22`);
sig32=(1/n32)*(diag(m32)-m32*m32`);
sig42=(1/n42)*(diag(m42)-m42*m42`);
sig2=block(sig12,sig22,sig32,sig42);

sig13=(1/n13)*(diag(m13)-m13*m13`);
sig23=(1/n23)*(diag(m23)-m23*m23`);
sig33=(1/n33)*(diag(m33)-m33*m33`);
sig43=(1/n43)*(diag(m43)-m43*m43`);
```

```
sig3=block(sig13,sig23,sig33,sig43);

sig=block(sig1,sig2,sig3);
V=sig;
*********************************************************************;
   p=p-(Gm*V)`*ginv(Gp*V*Gm`)*g;                           *Weibull;
   * p=p-(Gm*V)`*ginv(Gp*V*Gm`)*g;                           *loglogistic;
     verskil=sqrt((p-p1)`*(p-p1));
     end;
verskil1=sqrt((p-m)`*(p-m));
m=p;ms=S*m;
end;
print m; print i j;

**********Parameter vector for linear model;
par=inv(xc`*xc)*xc`*log(-log(1-ms));                         *Weibull;
*par=inv(xc`*xc)*xc`*(log(ms)-log(1-ms)); print par;          *log-logistic;

print par[format=E20.];

**********Regression coefficients;
oualphaBASELINE=par[1];
lambdaBASELINE=exp(par[1]);
betaA1=par[2];
betaA2=par[3];
betaA3=-(par[2]+par[3]);
lambdaA1=exp(par[1]+par[2]);
lambdaA2=exp(par[1]+par[3]);
lambdaA3=exp(par[1]+betaA3);
alphaA1=par[4];
alphaA2=par[5];
alphaA3=par[6];
alphaBASELINE=(n1#alphaA1+n2#alphaA2+n3#alphaA3)/n;

**********lambda's and alpha's for age levels*****************;
lambdaA1=lambdaBASELINE#exp(par[2]);
lambdaA2=lambdaBASELINE#exp(par[3]);
lambdaA3=lambdaBASELINE#exp(betaA3);
oualphaA1=log(lambdaA1);
oualphaA2=log(lambdaA2);
oualphaA3=log(lambdaA3);

print ' Weibull parameters, beta effects for age levels A1 A2 A3';
print 'lambdaBASELINE=' lambdaBASELINE[format=E20.]
      'oualphaBASELINE=' oualphaBASELINE alphaBASELINE;
print 'lambda(age=A1)=' lambdaA1[format=E20.]
      'lambda(age=A2)=' lambdaA2[format=E20.]
      'lambda(age=A3)=' lambdaA3[format=E20.];
print oualphaA1 oualphaA2 oualphaA3;
print 'alpha(age=A1)=' alphaA1[format=E20.]
      'alpha(age=A2)=' alphaA2[format=E20.]
      'alpha(age=A3)=' alphaA3[format=E20.];

print 'beta(age=A1)=' betaA1 'beta(age=A2)=' betaA2 'beta(age=A3)=' betaA3;

**********Hazard ratio and Odds ratio*****************************************;
hazBASELINE12=lambdaBASELINE*alphaBASELINE*12**(alphaBASELINE-1);
hazBASELINE24=lambdaBASELINE*alphaBASELINE*24**(alphaBASELINE-1);
survBASELINE12=exp(-lambdaBASELINE*12**alphaBASELINE);
survBASELINE24=exp(-lambdaBASELINE*24**alphaBASELINE);
oddsBASELINE12=(1-survBASELINE12)/survBASELINE12;
oddsBASELINE24=(1-survBASELINE24)/survBASELINE24;

hazAGE_A1_12=lambdaA1*alphaA1*12**(alphaA1-1);
hazAGE_A1_24=lambdaA1*alphaA1*24**(alphaA1-1);
survAGE_A1_12=exp(-lambdaA1*12**alphaA1);
survAGE_A1_24=exp(-lambdaA1*24**alphaA1);
oddsAGE_A1_12=(1-survAGE_A1_12)/survAGE_A1_12;
oddsAGE_A1_24=(1-survAGE_A1_24)/survAGE_A1_24;

hazratioAGE_A1_12=hazAGE_A1_12/hazBASELINE12;
hazratioAGE_A1_24=hazAGE_A1_24/hazBASELINE24;
oddsratioAGE_A1_12=oddsAGE_A1_12/oddsBASELINE12;
oddsratioAGE_A1_24=oddsAGE_A1_24/oddsBASELINE24;

print hazBASELINE12 hazBASELINE24 survBASELINE12 survBASELINE24
      oddsBASELINE12 oddsBASELINE24;

print hazAGE_A1_12 hazAGE_A1_24;
print survAGE_A1_12 survAGE_A1_24;
print oddsAGE_A1_12 oddsAGE_A1_24;
print hazratioAGE_A1_24 hazratioAGE_A1_24 oddsratioAGE_A1_12 oddsratioAGE_A1_24;
*******************************************************************************;
hazAGE_A2_12=lambdaA2*alphaA2*12**(alphaA2-1);
hazAGE_A2_24=lambdaA2*alphaA2*24**(alphaA2-1);
survAGE_A2_12=exp(-lambdaA2*12**alphaA2);
survAGE_A2_24=exp(-lambdaA2*24**alphaA2);
oddsAGE_A2_12=(1-survAGE_A2_12)/survAGE_A2_12;
oddsAGE_A2_24=(1-survAGE_A2_24)/survAGE_A2_24;
```

```
hazratioAGE_A2_12=hazAGE_A2_12/hazBASELINE12;
hazratioAGE_A2_24=hazAGE_A2_24/hazBASELINE24;
oddsratioAGE_A2_12=oddsAGE_A2_12/oddsBASELINE12;
oddsratioAGE_A2_24=oddsAGE_A2_24/oddsBASELINE24;

print hazAGE_A2_12 hazAGE_A2_24;
print survAGE_A2_12 survAGE_A2_24;
print oddsAGE_A2_12 oddsAGE_A2_24;
print hazratioAGE_A2_12 hazratioAGE_A2_24 oddsratioAGE_A2_12 oddsratioAGE_A2_24;
********************************************************************************;
hazAGE_A3_12=lambdaA3*alphaA3*12**(alphaA3-1);
hazAGE_A3_24=lambdaA3*alphaA3*24**(alphaA3-1);
survAGE_A3_12=exp(-lambdaA3*12**alphaA3);
survAGE_A3_24=exp(-lambdaA3*24**alphaA3);
oddsAGE_A3_12=(1-survAGE_A3_12)/survAGE_A3_12;
oddsAGE_A3_24=(1-survAGE_A3_24)/survAGE_A3_24;

hazratioAGE_A3_12=hazAGE_A3_12/hazBASELINE12;
hazratioAGE_A3_24=hazAGE_A3_24/hazBASELINE24;
oddsratioAGE_A3_12=oddsAGE_A3_12/oddsBASELINE12;
oddsratioAGE_A3_24=oddsAGE_A3_24/oddsBASELINE24;

print hazAGE_A3_12 hazAGE_A3_24;
print survAGE_A3_12 survAGE_A3_24;
print oddsAGE_A3_12 oddsAGE_A3_24;
print hazratioAGE_A3_12 hazratioAGE_A3_24 oddsratioAGE_A3_12 oddsratioAGE_A3_24;

**********Median Lifetime*******************************************************;
medianBASELINE=((1/lambdaBASELINE)#log(2))##(1/alphaBASELINE);
medianA1=((1/lambdaA1)#log(2))##(1/alphaA1);
medianA2=((1/lambdaA2)#log(2))##(1/alphaA2);
medianA3=((1/lambdaA3)#log(2))##(1/alphaA3);

 perc5BASELINE=((1/lambdaBASELINE)#log(100/(100- 5)))##(1/alphaBASELINE);
perc10BASELINE=((1/lambdaBASELINE)#log(100/(100-10)))##(1/alphaBASELINE);
perc20BASELINE=((1/lambdaBASELINE)#log(100/(100-20)))##(1/alphaBASELINE);
perc25BASELINE=((1/lambdaBASELINE)#log(100/(100-25)))##(1/alphaBASELINE);
perc30BASELINE=((1/lambdaBASELINE)#log(100/(100-30)))##(1/alphaBASELINE);
perc40BASELINE=((1/lambdaBASELINE)#log(100/(100-40)))##(1/alphaBASELINE);
perc50BASELINE=((1/lambdaBASELINE)#log(100/(100-50)))##(1/alphaBASELINE);
perc60BASELINE=((1/lambdaBASELINE)#log(100/(100-60)))##(1/alphaBASELINE);
perc70BASELINE=((1/lambdaBASELINE)#log(100/(100-70)))##(1/alphaBASELINE);
perc75BASELINE=((1/lambdaBASELINE)#log(100/(100-75)))##(1/alphaBASELINE);
perc80BASELINE=((1/lambdaBASELINE)#log(100/(100-80)))##(1/alphaBASELINE);
perc90BASELINE=((1/lambdaBASELINE)#log(100/(100-90)))##(1/alphaBASELINE);
perc95BASELINE=((1/lambdaBASELINE)#log(100/(100-95)))##(1/alphaBASELINE);

 perc5A1=((1/lambdaA1)#log(100/(100- 5)))##(1/alphaA1);
perc10A1=((1/lambdaA1)#log(100/(100-10)))##(1/alphaA1);
perc20A1=((1/lambdaA1)#log(100/(100-20)))##(1/alphaA1);
perc25A1=((1/lambdaA1)#log(100/(100-25)))##(1/alphaA1);
perc30A1=((1/lambdaA1)#log(100/(100-30)))##(1/alphaA1);
perc40A1=((1/lambdaA1)#log(100/(100-40)))##(1/alphaA1);
perc50A1=((1/lambdaA1)#log(100/(100-50)))##(1/alphaA1);
perc60A1=((1/lambdaA1)#log(100/(100-60)))##(1/alphaA1);
perc70A1=((1/lambdaA1)#log(100/(100-70)))##(1/alphaA1);
perc75A1=((1/lambdaA1)#log(100/(100-75)))##(1/alphaA1);
perc80A1=((1/lambdaA1)#log(100/(100-80)))##(1/alphaA1);
perc90A1=((1/lambdaA1)#log(100/(100-90)))##(1/alphaA1);
perc95A1=((1/lambdaA1)#log(100/(100-95)))##(1/alphaA1);

 perc5A2=((1/lambdaA2)#log(100/(100- 5)))##(1/alphaA2);
perc10A2=((1/lambdaA2)#log(100/(100-10)))##(1/alphaA2);
perc20A2=((1/lambdaA2)#log(100/(100-20)))##(1/alphaA2);
perc25A2=((1/lambdaA2)#log(100/(100-25)))##(1/alphaA2);
perc30A2=((1/lambdaA2)#log(100/(100-30)))##(1/alphaA2);
perc40A2=((1/lambdaA2)#log(100/(100-40)))##(1/alphaA2);
perc50A2=((1/lambdaA2)#log(100/(100-50)))##(1/alphaA2);
perc60A2=((1/lambdaA2)#log(100/(100-60)))##(1/alphaA2);
perc70A2=((1/lambdaA2)#log(100/(100-70)))##(1/alphaA2);
perc75A2=((1/lambdaA2)#log(100/(100-75)))##(1/alphaA2);
perc80A2=((1/lambdaA2)#log(100/(100-80)))##(1/alphaA2);
perc90A2=((1/lambdaA2)#log(100/(100-90)))##(1/alphaA2);
perc95A2=((1/lambdaA2)#log(100/(100-95)))##(1/alphaA2);

 perc5A3=((1/lambdaA3)#log(100/(100- 5)))##(1/alphaA3);
perc10A3=((1/lambdaA3)#log(100/(100-10)))##(1/alphaA3);
perc20A3=((1/lambdaA3)#log(100/(100-20)))##(1/alphaA3);
perc25A3=((1/lambdaA3)#log(100/(100-25)))##(1/alphaA3);
perc30A3=((1/lambdaA3)#log(100/(100-30)))##(1/alphaA3);
perc40A3=((1/lambdaA3)#log(100/(100-40)))##(1/alphaA3);
perc50A3=((1/lambdaA3)#log(100/(100-50)))##(1/alphaA3);
perc60A3=((1/lambdaA3)#log(100/(100-60)))##(1/alphaA3);
perc70A3=((1/lambdaA3)#log(100/(100-70)))##(1/alphaA3);
perc75A3=((1/lambdaA3)#log(100/(100-75)))##(1/alphaA3);
perc80A3=((1/lambdaA3)#log(100/(100-80)))##(1/alphaA3);
perc90A3=((1/lambdaA3)#log(100/(100-90)))##(1/alphaA3);
perc95A3=((1/lambdaA3)#log(100/(100-95)))##(1/alphaA3);
```

```
print perc5BASELINE perc10BASELINE perc20BASELINE perc25BASELINE perc30BASELINE;
print perc40BASELINE perc50BASELINE perc60BASELINE perc70BASELINE;
print perc75BASELINE perc80BASELINE perc90BASELINE perc95BASELINE;

print perc5A1 perc10A1 perc20A1 perc25A1 perc30A1;
print perc40A1 perc50A1 perc60A1 perc70A1;
print perc75A1 perc80A1 perc90A1 perc95A1;

print perc5A2 perc10A2 perc20A2 perc25A2 perc30A2;
print perc40A2 perc50A2 perc60A2 perc70A2;
print perc75A2 perc80A2 perc90A2 perc95A2;

print perc5A3 perc10A3 perc20A3 perc25A3 perc30A3;
print perc40A3 perc50A3 perc60A3 perc70A3;
print perc75A3 perc80A3 perc90A3 perc95A3;

print 'medianlifetimeBASELINE=' medianBASELINE;
print 'medianlifetime(age=A1)=' medianA1
      'medianlifetime(age=A2)=' medianA2
      'medianlifetime(age=A3)=' medianA3;
```

5. Program for fitting a log-logistic/Weibull regression model with a continuous predictor

```
title1 'Fitting of regression model with one ordinal predictor';
title2 'Staggered entry: constant shape parameter';

proc iml worksize= 60;
reset nolog;
options pagesize=500;

**********Frequency vector (first entry,3 agegroups);
                f11={29,59,95,73,108,15,642};
                f12={21,50,91,45,75,13,553};
                f13={16,49,68,39, 67, 7,471};

**********Vector of upper boundaries;
                x1={12,17,24,28,34,37};

**********Frequency vector (second entry,3 agegroups);
                f21={41,75,103,92,83,628};
                f22={49,62,61,66,54,753};
                f23={28,29,65,42,35,543};

**********Vector of upper boundaries;
                x2={12,17,24,28,34};

**********Frequency vector (third entry,3 agegroups);
                f31={68,34,99,57,570};
                f32={40,44,83,33,533};
                f33={46,21,60,27,571};

**********Vector of upper boundaries;
                x3={12,17,24,28};

**********Frequency vector (fourth entry,3 agegroups);
                f41={71,60,69,573};
                f42={54,61,68,616};
                f43={50,45,70,659};

**********Vector of upper boundaries;
                x4={12,17,24};

**********Relative frequency vectors;
n11=f11[+];    n21=f21[+];    n31=f31[+];    n41=f41[+];    n1=n11+n21+n31+n41;
n12=f12[+];    n22=f22[+];    n32=f32[+];    n42=f42[+];    n2=n12+n22+n32+n42;
n13=f13[+];    n23=f23[+];    n33=f33[+];    n43=f43[+];    n3=n13+n23+n33+n43;
                                                            n=n1+n2+n3;
k1=nrow(f11); d1=k1-1;
k2=nrow(f21); d2=k2-1;
k3=nrow(f31); d3=k3-1;
k4=nrow(f41); d4=k4-1;
k=k1+k2+k3+k4;
d=d1+d2+d3+d4;

p11=f11/n11;    p21=f21/n21;    p31=f31/n31;    p41=f41/n41;    p1=p11//p21//p31//p41;
p12=f12/n12;    p22=f22/n22;    p32=f32/n32;    p42=f42/n42;    p2=p12//p22//p32//p42;
p13=f13/n13;    p23=f23/n23;    p33=f33/n33;    p43=f43/n43;    p3=p13//p23//p33//p43;
                                                                p=p1//p2//p3;

**********Design matrix and matrix orthogonal to design matrix;

S1=J(d1,1,1)@cusum(J(1,k1,1));
S2=J(1,k1,1)@cusum(J(d1,1,1));
S1=S1<=S2;
S2=S1[1:d2,1:d1];
S3=S1[1:d3,1:d2];
S4=S1[1:d4,1:d3];
```

```
S=block(S1,S2,S3,S4);
S=I(3)@S;

AA=J(54,1,1);       *54=3 times((7-1)+(6-1)+(5-1)+(4-1))=3 times 18=54;
BB=J(3,1,1);
DD=J(18,1,1);       *d=18;

lx=log(x1)//log(x2)//log(x3)//log(x4);

z={1,2,3};
*z={26,39.5,52};    *midpoints of age intervals;

xc=(AA||BB@lx||(z@DD));  print xc;

C=I(3#d)-xc*inv(xc`*xc)*xc`;

**********ITERATIVE PROCEDURE (double iterations over m and p);
*****starting value for m;
m=p;
ms=S*m; ps=ms;
p0=p;

*****iteration over m;
itr=0;
verskil1=1;
i=0;
do while (verskil1>0.00000001);
i=i+1;
p=p0;
*Gm=-C*diag(1/(log(1-ms)))*diag(1/(1-ms))*S;               *Weibull;
Gm=C*diag(1/ms+1/(1-ms))*S;                                *log-logistic;

**********iteration over p;
     verskil=1;
     j=0;
     do while (verskil>0.00000001);
     j=j+1;
     p1=p;
     ps=S*p;
    *g=C*log(-log(1-ps));                                  *Weibull;
    g=C*(log(ps)-log(1-ps));                               *log-logistic;
    *Gp=-C*diag(1/(log(1-ps)))*diag(1/(1-ps))*S;           *Weibull;
    Gp=C*(diag(1/ps)+diag(1/(1-ps)))*S;                    *log-logistic;
******************covariance matrix***********************************;
m11=m[1:k1]; m21=m[k1+1:k1+k2];
m31=m[k1+k2+1:k1+k2+k3]; m41=m[k1+k2+k3+1:k1+k2+k3+k4];

m12=m[k+1:k+k1]; m22=m[k+k1+1:k+k1+k2];
m32=m[k+k1+k2+1:k+k1+k2+k3]; m42=m[k+k1+k2+k3+1:k+k1+k2+k3+k4];

m13=m[2#k+1:2#k+k1]; m23=m[2#k+k1+1:2#k+k1+k2];
m33=m[2#k+k1+k2+1:2#k+k1+k2+k3]; m43=m[2#k+k1+k2+k3+1:2#k+k1+k2+k3+k4];

sig11=(1/n11)*(diag(m11)-m11*m11`);
sig21=(1/n21)*(diag(m21)-m21*m21`);
sig31=(1/n31)*(diag(m31)-m31*m31`);
sig41=(1/n41)*(diag(m41)-m41*m41`);
sig1=block(sig11,sig21,sig31,sig41);

sig12=(1/n12)*(diag(m12)-m12*m12`);
sig22=(1/n22)*(diag(m22)-m22*m22`);
sig32=(1/n32)*(diag(m32)-m32*m32`);
sig42=(1/n42)*(diag(m42)-m42*m42`);
sig2=block(sig12,sig22,sig32,sig42);

sig13=(1/n13)*(diag(m13)-m13*m13`);
sig23=(1/n23)*(diag(m23)-m23*m23`);
sig33=(1/n33)*(diag(m33)-m33*m33`);
sig43=(1/n43)*(diag(m43)-m43*m43`);
sig3=block(sig13,sig23,sig33,sig43);

sig=block(sig1,sig2,sig3);
V=sig;
**********************************************************************;
    *p=p-(Gm*V)`*ginv(Gp*V*Gm`)*g;                         *Weibull;
    p=p-(Gm*V)`*ginv(Gp*V*Gm`)*g;                          *log-logistic;
     verskil=sqrt((p-p1)`*(p-p1));
     end;
verskil1=sqrt((p-m)`*(p-m));
m=p;ms=S*m;
end;
print m; print i j;

**********Parameter vector for linear model;
*par=inv(xc`*xc)*xc`*log(-log(1-ms));                      *Weibull;
par=inv(xc`*xc)*xc`*(log(ms)-log(1-ms)); print par;        *log-logistic;

print par[format=E20.];
```

```
**********Regression coefficients;
oualphaBASELINE=par[1];
lambdaBASELINE=exp(par[1]);

alpha=par[2];

beta=par[3]; *AGE ordinal;

print ' Loglogistic parameters and beta effect for age ';
*print ' Weibull parameters and beta effect for age ';

print 'lambdaBASELINE=' lambdaBASELINE[format=E20.]
'oualphaBASELINE=' oualphaBASELINE alpha;
print 'alpha=' alpha[format=E20.];
print 'beta=' beta[format=E20.];

**********lambda for each agegroup (z=1,2,3)*********************;
lambdaAGE1=exp(par[1]+(par[3]*1));
lambdaAGE2=exp(par[1]+(par[3]*2));
lambdaAGE3=exp(par[1]+(par[3]*3));

oualphaAGE1=log(lambdaAGE1);
oualphaAGE2=log(lambdaAGE2);
oualphaAGE3=log(lambdaAGE3);

print 'lambda(age=1)=' lambdaAGE1[format=E20.]
'lambda(age=2)=' lambdaAGE2[format=E20.]
'lambda(age=3)=' lambdaAGE3[format=E20.];
print oualphaAGE1 oualphaAGE2 oualphaAGE3;
```

6. Program for fitting a log-logistic regression model with two predictors

```
title1 'Fitting of regression model with two covariates';
title2 'Staggered entry: constant shape parameter';

proc iml worksize= 60;
reset nolog;
options pagesize=500;

**********Frequency vector (first entry,3 agegroups,3 scoregroups);
                f111={12,34,51,39,57,11, 59};   f111=f111<>1e-04;
                f112={10,12,22,19,32, 4,418};   f112=f112<>1e-04;
                f113={ 7,13,22,15,19, 0,165};   f113=f113<>1e-04;
                f121={13,14,45,27,33, 4, 66};   f121=f121<>1e-04;
                f122={ 4,22,22, 8,25, 4,297};   f122=f122<>1e-04;
                f123={ 4,14,24,10,17, 5,190};   f123=f123<>1e-04;
                f131={10,25,29,17,46, 2,116};   f131=f131<>1e-04;
                f132={ 6,13,28,16,16, 5,273};   f132=f132<>1e-04;
                f133={ 0,11,11, 6, 5, 0, 82};   f133=f133<>1e-04;

**********Vector of upper boundaries;
                x1={12,17,24,28,34,37};

**********Frequency vector (second entry,3 agegroups,3 scoregroups);
                f211={22,25,58,53,40, 45};      f211=f211<>1e-04;
                f212={10,26,32,20,29,379};      f212=f212<>1e-04;
                f213={ 9,24,13,19,14,204};      f213=f213<>1e-04;
                f221={24,24,28,30,25,106};      f221=f221<>1e-04;
                f222={12,20,14,17,16,409};      f222=f222<>1e-04;
                f223={13,18,19,19,13,238};      f223=f223<>1e-04;
                f231={13,15,32,19,17,107};      f231=f231<>1e-04;
                f232={11,13,22,17,12,319};      f232=f232<>1e-04;
                f233={ 4, 1,11, 6, 6,117};      f233=f233<>1e-04;

**********Vector of upper boundaries;
                x2={12,17,24,28,34};

**********Frequency vector (third entry,3 agegroups,3 scoregroups);
                f311={34,16,50,23, 54};         f311=f311<>1e-04;
                f312={19, 2,32,24,317};         f312=f312<>1e-04;
                f313={15,16,17,10,199};         f313=f313<>1e-04;
                f321={19,18,38,16, 75};         f321=f321<>1e-04;
                f322={16,14,25,10,263};         f322=f322<>1e-04;
                f323={ 5,12,20, 7,195};         f323=f323<>1e-04;
                f331={28,16,22,12, 98};         f331=f331<>1e-04;
                f332={13, 0,24, 4,323};         f332=f332<>1e-04;
                f333={ 5, 5,14,11,150};         f333=f333<>1e-04;

**********Vector of upper boundaries;
                x3={12,17,24,28};

**********Frequency vector (fourth entry,3 agegroups,3 scoregroups);
                f411={40,30,30, 50};            f411=f411<>1e-04;
                f412={ 9,14,27,301};            f412=f412<>1e-04;
                f413={22,16,12,222};            f413=f413<>1e-04;
                f421={24,30,29, 81};            f421=f421<>1e-04;
                f422={14,15,12,307};            f422=f422<>1e-04;
                f423={16,16,27,228};            f423=f423<>1e-04;
                f431={20,22,28,119};            f431=f431<>1e-04;
                f432={19,12,26,369};            f432=f432<>1e-04;
```

```
                f433={11,11,16,171};            f433=f433<>1e-04;

**********Vector of upper boundaries;
                x4={12,17,24};

**********Relative frequency vectors;
n111=f111[+]; n211=f211[+]; n311=f311[+]; n411=f411[+]; n11=n111+n211+n311+n411;
n112=f112[+]; n212=f212[+]; n312=f312[+]; n412=f412[+]; n12=n112+n212+n312+n412;
n113=f113[+]; n213=f213[+]; n313=f313[+]; n413=f413[+]; n13=n113+n213+n313+n413;
n121=f121[+]; n221=f221[+]; n321=f321[+]; n421=f421[+]; n21=n121+n221+n321+n421;
n122=f122[+]; n222=f222[+]; n322=f322[+]; n422=f422[+]; n22=n122+n222+n322+n422;
n123=f123[+]; n223=f223[+]; n323=f323[+]; n423=f423[+]; n23=n123+n223+n323+n423;
n131=f131[+]; n231=f231[+]; n331=f331[+]; n431=f431[+]; n31=n131+n231+n331+n431;
n132=f132[+]; n232=f232[+]; n332=f332[+]; n432=f432[+]; n32=n132+n232+n332+n432;
n133=f133[+]; n233=f233[+]; n333=f333[+]; n433=f433[+]; n33=n133+n233+n333+n433;

                                                        n1=n11+n12+n13;
                                                        n2=n21+n22+n23;
                                                        n3=n31+n32+n33;
                                                        n=n1+n2+n3;

k1=nrow(f111); d1=k1-1;
k2=nrow(f211); d2=k2-1;
k3=nrow(f311); d3=k3-1;
k4=nrow(f411); d4=k4-1;
k=k1+k2+k3+k4;
d=d1+d2+d3+d4;

p111=f111/n111; p211=f211/n211; p311=f311/n311; p411=f411/n411;
p112=f112/n112; p212=f212/n212; p312=f312/n312; p412=f412/n412;
p113=f113/n113; p213=f213/n213; p313=f313/n313; p413=f413/n413;
p121=f121/n121; p221=f221/n221; p321=f321/n321; p421=f421/n421;
p122=f122/n122; p222=f222/n222; p322=f322/n322; p422=f422/n422;
p123=f123/n123; p223=f223/n223; p323=f323/n323; p423=f423/n423;
p131=f131/n131; p231=f231/n231; p331=f331/n331; p431=f431/n431;
p132=f132/n132; p232=f232/n232; p332=f332/n332; p432=f432/n432;
p133=f133/n133; p233=f233/n233; p333=f333/n333; p433=f433/n433;

p11=p111//p211//p311//p411;
p12=p112//p212//p312//p412;
p13=p113//p213//p313//p413;
p1=p11//p12//p13;

p21=p121//p221//p321//p421;
p22=p122//p222//p322//p422;
p23=p123//p223//p323//p423;
p2=p21//p22//p23;

p31=p131//p231//p331//p431;
p32=p132//p232//p332//p432;
p33=p133//p233//p333//p433;
p3=p31//p32//p33;

p=p1//p2//p3;

**********Design matrix and matrix orthogonal to design matrix;

S1=J(d1,1,1)@cusum(J(1,k1,1));
S2=J(1,k1,1)@cusum(J(d1,1,1));
S1=S1<=S2;
S2=S1[1:d2,1:d1];
S3=S1[1:d3,1:d2];
S4=S1[1:d4,1:d3];
S=block(S1,S2,S3,S4);
S=I(3)@I(3)@S;

AA=J(162,1,1);          *162=9 times((7-1)+(6-1)+(5-1)+(4-1))=9 times 18=162;
BB=J(3,1,1);
CC=(I(2)//J(1,2,-1));
DD=J(18,1,1);           *3 times d1=3(6)=18;
EE=J(15,1,1);           *3 times d2=3(5)=15;
FF=J(12,1,1);           *3 times d3=3(4)=12;
GG=J(9,1,1);            *3 times d4=3(3)=9;

KK=J(9,1,1);            *9=3 times 3;
LL=J(54,1,1);           *54=18 times 3;

lx=BB@BB@(log(x1)//log(x2)//log(x3)//log(x4));

HH=BB@(CC@DD);

xc=AA||CC@LL||HH||lx;  print xc;

C=I(9#d)-xc*inv(xc`*xc)*xc`;

**********ITERATIVE PROCEDURE (double iterations over m and p);
*****starting value for m;
m=p;
ms=S*m; ps=ms;
```

```
p0=p;

*****iteration over m;
itr=0;
verskil1=1;
i=0;
do while (verskil1>0.00000001);
i=i+1;
p=p0;
*Gm=-C*diag(1/(log(1-ms)))*diag(1/(1-ms))*S;               *Weibull;
Gm=C*diag(1/ms+1/(1-ms))*S;                                 *log-logistic;

**********iteration over p;
     verskil=1;
     j=0;
     do while (verskil>0.00000001);
     j=j+1;
     p1=p;
     ps=S*p;
    *g=C*log(-log(1-ps));                                   *Weibull;
    g=C*(log(ps)-log(1-ps));                                *log-logistic;
    *Gp=-C*diag(1/(log(1-ps)))*diag(1/(1-ps))*S;            *Weibull;
    Gp=C*(diag(1/ps)+diag(1/(1-ps)))*S;                     *log-logistic;

******************covariance matrix**********************************;
m111=m[1:k1];
m211=m[k1+1:k1+k2];
m311=m[k1+k2+1:k1+k2+k3];
m411=m[k1+k2+k3+1:k1+k2+k3+k4];

m112=m[k+1:k+k1];
m212=m[k+k1+1:k+k1+k2];
m312=m[k+k1+k2+1:k+k1+k2+k3];
m412=m[k+k1+k2+k3+1:k+k1+k2+k3+k4];

m113=m[2#k+1:2#k+k1];
m213=m[2#k+k1+1:2#k+k1+k2];
m313=m[2#k+k1+k2+1:2#k+k1+k2+k3];
m413=m[2#k+k1+k2+k3+1:2#k+k1+k2+k3+k4];

m121=m[3#k+1:3#k+k1];
m221=m[3#k+k1+1:3#k+k1+k2];
m321=m[3#k+k1+k2+1:3#k+k1+k2+k3];
m421=m[3#k+k1+k2+k3+1:3#k+k1+k2+k3+k4];

m122=m[4#k+1:4#k+k1];
m222=m[4#k+k1+1:4#k+k1+k2];
m322=m[4#k+k1+k2+1:4#k+k1+k2+k3];
m422=m[4#k+k1+k2+k3+1:4#k+k1+k2+k3+k4];

m123=m[5#k+1:5#k+k1];
m223=m[5#k+k1+1:5#k+k1+k2];
m323=m[5#k+k1+k2+1:5#k+k1+k2+k3];
m423=m[5#k+k1+k2+k3+1:5#k+k1+k2+k3+k4];

m131=m[6#k+1:6#k+k1];
m231=m[6#k+k1+1:6#k+k1+k2];
m331=m[6#k+k1+k2+1:6#k+k1+k2+k3];
m431=m[6#k+k1+k2+k3+1:6#k+k1+k2+k3+k4];

m132=m[7#k+1:7#k+k1];
m232=m[7#k+k1+1:7#k+k1+k2];
m332=m[7#k+k1+k2+1:7#k+k1+k2+k3];
m432=m[7#k+k1+k2+k3+1:7#k+k1+k2+k3+k4];

m133=m[8#k+1:8#k+k1];
m233=m[8#k+k1+1:8#k+k1+k2];
m333=m[8#k+k1+k2+1:8#k+k1+k2+k3];
m433=m[8#k+k1+k2+k3+1:9#k];

sig111=(1/n111)*(diag(m111)-m111*m111`);
sig112=(1/n112)*(diag(m112)-m112*m112`);
sig113=(1/n113)*(diag(m113)-m113*m113`);
sig211=(1/n211)*(diag(m211)-m211*m211`);
sig212=(1/n212)*(diag(m212)-m212*m212`);
sig213=(1/n213)*(diag(m213)-m213*m213`);
sig311=(1/n311)*(diag(m311)-m311*m311`);
sig312=(1/n312)*(diag(m312)-m312*m312`);
sig313=(1/n313)*(diag(m313)-m313*m313`);
sig411=(1/n411)*(diag(m411)-m411*m411`);
sig412=(1/n412)*(diag(m412)-m412*m412`);
sig413=(1/n413)*(diag(m413)-m413*m413`);

sig11=block(sig111,sig211,sig311,sig411);
sig12=block(sig112,sig212,sig312,sig412);
sig13=block(sig113,sig213,sig313,sig413);


sig121=(1/n121)*(diag(m121)-m121*m121`);
```

```
sig122=(1/n122)*(diag(m122)-m122*m122`);
sig123=(1/n123)*(diag(m123)-m123*m123`);
sig221=(1/n221)*(diag(m221)-m221*m221`);
sig222=(1/n222)*(diag(m222)-m222*m222`);
sig223=(1/n223)*(diag(m223)-m223*m223`);
sig321=(1/n321)*(diag(m321)-m321*m321`);
sig322=(1/n322)*(diag(m322)-m322*m322`);
sig323=(1/n323)*(diag(m323)-m323*m323`);
sig421=(1/n421)*(diag(m421)-m421*m421`);
sig422=(1/n422)*(diag(m422)-m422*m422`);
sig423=(1/n423)*(diag(m423)-m423*m423`);

sig21=block(sig121,sig221,sig321,sig421);
sig22=block(sig122,sig222,sig322,sig422);
sig23=block(sig123,sig223,sig323,sig423);


sig131=(1/n131)*(diag(m131)-m131*m131`);
sig132=(1/n132)*(diag(m132)-m132*m132`);
sig133=(1/n133)*(diag(m133)-m133*m133`);
sig231=(1/n231)*(diag(m231)-m231*m231`);
sig232=(1/n232)*(diag(m232)-m232*m232`);
sig233=(1/n233)*(diag(m233)-m233*m233`);
sig331=(1/n331)*(diag(m331)-m331*m331`);
sig332=(1/n332)*(diag(m332)-m332*m332`);
sig333=(1/n333)*(diag(m333)-m333*m333`);
sig431=(1/n431)*(diag(m431)-m431*m431`);
sig432=(1/n432)*(diag(m432)-m432*m432`);
sig433=(1/n433)*(diag(m433)-m433*m433`);

sig31=block(sig131,sig231,sig331,sig431);
sig32=block(sig132,sig232,sig332,sig432);
sig33=block(sig133,sig233,sig333,sig433);

sig=block(sig11,sig12,sig13,sig21,sig22,sig23,sig31,sig32,sig33);
V=sig;
*********************************************************************;
    *p=p-(Gm*V)`*ginv(Gp*V*Gm`)*g;                        *Weibull;
    p=p-(Gm*V)`*ginv(Gp*V*Gm`)*g;                         *log-logistic;
     verskil=sqrt((p-p1)`*(p-p1));
     end;
verskil1=sqrt((p-m)`*(p-m));
m=p;ms=S*m;
end;
print m; print i j;

**********Parameter vector for linear model;
*par=inv(xc`*xc)*xc`*log(-log(1-ms));                      *Weibull;
par=inv(xc`*xc)*xc`*(log(ms)-log(1-ms));                   *log-logistic;

print par[format=E20.];

**********Regression coefficients;
oualphaBASELINE=par[1];
betaA1=par[2];
betaA2=par[3];
betaA3=-(par[2]+par[3]);
betaS1=par[4];
betaS2=par[5];
betaS3=-(par[4]+par[5]);
lambdaBASELINE=exp(par[1]);
alpha=par[6];

**********Indices, lambda's and constant alpha for age levels******************;
indexA1=exp(betaA1);          *constant shape parameter;
indexA2=exp(betaA2);
indexA3=exp(betaA3);
indexS1=exp(betaS1);
indexS2=exp(betaS2);
indexS3=exp(betaS3);

lambdaA1S1=lambdaBASELINE#indexA1#indexS1;
lambdaA1S2=lambdaBASELINE#indexA1#indexS2;
lambdaA1S3=lambdaBASELINE#indexA1#indexS3;
lambdaA2S1=lambdaBASELINE#indexA2#indexS1;
lambdaA2S2=lambdaBASELINE#indexA2#indexS2;
lambdaA2S3=lambdaBASELINE#indexA2#indexS3;
lambdaA3S1=lambdaBASELINE#indexA3#indexS1;
lambdaA3S2=lambdaBASELINE#indexA3#indexS2;
lambdaA3S3=lambdaBASELINE#indexA3#indexS3;

oualphaA1S1=oualphaBASELINE+betaA1+betaS1; *same as oualphaA1S1=log(lambdaA1S1);
oualphaA1S2=oualphaBASELINE+betaA1+betaS2;
oualphaA1S3=oualphaBASELINE+betaA1+betaS3;
oualphaA2S1=oualphaBASELINE+betaA2+betaS1;
oualphaA2S2=oualphaBASELINE+betaA2+betaS2;
oualphaA2S3=oualphaBASELINE+betaA2+betaS3;
oualphaA3S1=oualphaBASELINE+betaA3+betaS1;
oualphaA3S2=oualphaBASELINE+betaA3+betaS2;
```

```
oualphaA3S3=oualphaBASELINE+betaA3+betaS3;

print ' Loglogistic parameters, beta effects and indices: MLE subject to constraints';

print 'lambdaBASELINE=' lambdaBASELINE[format=E20.]  'oualphaBASELINE=' oualphaBASELINE;
print 'alpha=' alpha[format=E20.];

print 'lambda(age=A1,score=S1)=' lambdaA1S1[format=E20.]
      'lambda(age=A1,score=S2)=' lambdaA1S2[format=E20.]
      'lambda(age=A1,score=S3)=' lambdaA1S3[format=E20.];
print oualphaA1S1 oualphaA1S2 oualphaA1S3;

print 'lambda(age=A2,score=S1)=' lambdaA2S1[format=E20.]
      'lambda(age=A2,score=S2)=' lambdaA2S2[format=E20.]
      'lambda(age=A2,score=S3)=' lambdaA2S3[format=E20.];
print oualphaA2S1 oualphaA2S2 oualphaA2S3;

print 'lambda(age=A3,score=S1)=' lambdaA3S1[format=E20.]
      'lambda(age=A3,score=S2)=' lambdaA3S2[format=E20.]
      'lambda(age=A3,score=S3)=' lambdaA3S3[format=E20.];
print oualphaA3S1 oualphaA3S2 oualphaA3S3;

print 'beta(age=A1)=' betaA1 'beta(age=A2)=' betaA2 'beta(age=A3)=' betaA3;
print 'beta(score=S1)=' betaS1 'beta(score=S2)=' betaS2 'beta(score=S3)=' betaS3;
print 'index(age=A1)=' indexA1 'index(age=A2)=' indexA2 'index(age=A3)=' indexA3;
print 'index(score=S1)=' indexS1 'index(score=S2)=' indexS2 'index(score=S3)=' indexS3;

**********Hazard ratio and Odds ratio*******************************************;
hazBASELINE12=(lambdaBASELINE*alpha*12**(alpha-1))/(1+lambdaBASELINE*12**alpha);
hazBASELINE24=(lambdaBASELINE*alpha*24**(alpha-1))/(1+lambdaBASELINE*24**alpha);
survBASELINE12=(1+lambdaBASELINE*12**alpha)**(-1);
survBASELINE24=(1+lambdaBASELINE*24**alpha)**(-1);
oddsBASELINE12=(1-survBASELINE12)/survBASELINE12;
oddsBASELINE24=(1-survBASELINE24)/survBASELINE24;

hazA1S1_12=(lambdaA1S1*alpha*12**(alpha-1))/(1+lambdaA1S1*12**alpha);
hazA1S1_24=(lambdaA1S1*alpha*24**(alpha-1))/(1+lambdaA1S1*24**alpha);
hazA1S2_12=(lambdaA1S2*alpha*12**(alpha-1))/(1+lambdaA1S2*12**alpha);
hazA1S2_24=(lambdaA1S2*alpha*24**(alpha-1))/(1+lambdaA1S2*24**alpha);
hazA1S3_12=(lambdaA1S3*alpha*12**(alpha-1))/(1+lambdaA1S3*12**alpha);
hazA1S3_24=(lambdaA1S3*alpha*24**(alpha-1))/(1+lambdaA1S3*24**alpha);
hazA2S1_12=(lambdaA2S1*alpha*12**(alpha-1))/(1+lambdaA2S1*12**alpha);
hazA2S1_24=(lambdaA2S1*alpha*24**(alpha-1))/(1+lambdaA2S1*24**alpha);
hazA2S2_12=(lambdaA2S2*alpha*12**(alpha-1))/(1+lambdaA2S2*12**alpha);
hazA2S2_24=(lambdaA2S2*alpha*24**(alpha-1))/(1+lambdaA2S2*24**alpha);
hazA2S3_12=(lambdaA2S3*alpha*12**(alpha-1))/(1+lambdaA2S3*12**alpha);
hazA2S3_24=(lambdaA2S3*alpha*24**(alpha-1))/(1+lambdaA2S3*24**alpha);
hazA3S1_12=(lambdaA3S1*alpha*12**(alpha-1))/(1+lambdaA3S1*12**alpha);
hazA3S1_24=(lambdaA3S1*alpha*24**(alpha-1))/(1+lambdaA3S1*24**alpha);
hazA3S2_12=(lambdaA3S2*alpha*12**(alpha-1))/(1+lambdaA3S2*12**alpha);
hazA3S2_24=(lambdaA3S2*alpha*24**(alpha-1))/(1+lambdaA3S2*24**alpha);
hazA3S3_12=(lambdaA3S3*alpha*12**(alpha-1))/(1+lambdaA3S3*12**alpha);
hazA3S3_24=(lambdaA3S3*alpha*24**(alpha-1))/(1+lambdaA3S3*24**alpha);

survA1S1_12=(1+lambdaA1S1*12**alpha)**(-1);
survA1S1_24=(1+lambdaA1S1*24**alpha)**(-1);
survA1S2_12=(1+lambdaA1S2*12**alpha)**(-1);
survA1S2_24=(1+lambdaA1S2*24**alpha)**(-1);
survA1S3_12=(1+lambdaA1S3*12**alpha)**(-1);
survA1S3_24=(1+lambdaA1S3*24**alpha)**(-1);
survA2S1_12=(1+lambdaA2S1*12**alpha)**(-1);
survA2S1_24=(1+lambdaA2S1*24**alpha)**(-1);
survA2S2_12=(1+lambdaA2S2*12**alpha)**(-1);
survA2S2_24=(1+lambdaA2S2*24**alpha)**(-1);
survA2S3_12=(1+lambdaA2S3*12**alpha)**(-1);
survA2S3_24=(1+lambdaA2S3*24**alpha)**(-1);
survA3S1_12=(1+lambdaA3S1*12**alpha)**(-1);
survA3S1_24=(1+lambdaA3S1*24**alpha)**(-1);
survA3S2_12=(1+lambdaA3S2*12**alpha)**(-1);
survA3S2_24=(1+lambdaA3S2*24**alpha)**(-1);
survA3S3_12=(1+lambdaA3S3*12**alpha)**(-1);
survA3S3_24=(1+lambdaA3S3*24**alpha)**(-1);

oddsA1S1_12=(1-survA1S1_12)/survA1S1_12;
oddsA1S1_24=(1-survA1S1_24)/survA1S1_24;
oddsA1S2_12=(1-survA1S2_12)/survA1S2_12;
oddsA1S2_24=(1-survA1S2_24)/survA1S2_24;
oddsA1S3_12=(1-survA1S3_12)/survA1S3_12;
oddsA1S3_24=(1-survA1S3_24)/survA1S3_24;
oddsA2S1_12=(1-survA2S1_12)/survA2S1_12;
oddsA2S1_24=(1-survA2S1_24)/survA2S1_24;
oddsA2S2_12=(1-survA2S2_12)/survA2S2_12;
oddsA2S2_24=(1-survA2S2_24)/survA2S2_24;
oddsA2S3_12=(1-survA2S3_12)/survA2S3_12;
oddsA2S3_24=(1-survA2S3_24)/survA2S3_24;
oddsA3S1_12=(1-survA3S1_12)/survA3S1_12;
oddsA3S1_24=(1-survA3S1_24)/survA3S1_24;
oddsA3S2_12=(1-survA3S2_12)/survA3S2_12;
oddsA3S2_24=(1-survA3S2_24)/survA3S2_24;
```

```
oddsA3S3_12=(1-survA3S3_12)/survA3S3_12;
oddsA3S3_24=(1-survA3S3_24)/survA3S3_24;

hazratioA1S1_12=hazA1S1_12/hazBASELINE12;
hazratioA1S1_24=hazA1S1_24/hazBASELINE24;
hazratioA1S2_12=hazA1S2_12/hazBASELINE12;
hazratioA1S2_24=hazA1S2_24/hazBASELINE24;
hazratioA1S3_12=hazA1S3_12/hazBASELINE12;
hazratioA1S3_24=hazA1S3_24/hazBASELINE24;
hazratioA2S1_12=hazA2S1_12/hazBASELINE12;
hazratioA2S1_24=hazA2S1_24/hazBASELINE24;
hazratioA2S2_12=hazA2S2_12/hazBASELINE12;
hazratioA2S2_24=hazA2S2_24/hazBASELINE24;
hazratioA2S3_12=hazA2S3_12/hazBASELINE12;
hazratioA2S3_24=hazA2S3_24/hazBASELINE24;
hazratioA3S1_12=hazA3S1_12/hazBASELINE12;
hazratioA3S1_24=hazA3S1_24/hazBASELINE24;
hazratioA3S2_12=hazA3S2_12/hazBASELINE12;
hazratioA3S2_24=hazA3S2_24/hazBASELINE24;
hazratioA3S3_12=hazA3S3_12/hazBASELINE12;
hazratioA3S3_24=hazA3S3_24/hazBASELINE24;

oddsratioA1S1_12=oddsA1S1_12/oddsBASELINE12;
oddsratioA1S1_24=oddsA1S1_24/oddsBASELINE24;
oddsratioA1S2_12=oddsA1S2_12/oddsBASELINE12;
oddsratioA1S2_24=oddsA1S2_24/oddsBASELINE24;
oddsratioA1S3_12=oddsA1S3_12/oddsBASELINE12;
oddsratioA1S3_24=oddsA1S3_24/oddsBASELINE24;
oddsratioA2S1_12=oddsA2S1_12/oddsBASELINE12;
oddsratioA2S1_24=oddsA2S1_24/oddsBASELINE24;
oddsratioA2S2_12=oddsA2S2_12/oddsBASELINE12;
oddsratioA2S2_24=oddsA2S2_24/oddsBASELINE24;
oddsratioA2S3_12=oddsA2S3_12/oddsBASELINE12;
oddsratioA2S3_24=oddsA2S3_24/oddsBASELINE24;
oddsratioA3S1_12=oddsA3S1_12/oddsBASELINE12;
oddsratioA3S1_24=oddsA3S1_24/oddsBASELINE24;
oddsratioA3S2_12=oddsA3S2_12/oddsBASELINE12;
oddsratioA3S2_24=oddsA3S2_24/oddsBASELINE24;
oddsratioA3S3_12=oddsA3S3_12/oddsBASELINE12;
oddsratioA3S3_24=oddsA3S3_24/oddsBASELINE24;

print hazBASELINE12 hazBASELINE24 survBASELINE12 survBASELINE24 oddsBASELINE12 oddsBASELINE24;

print hazA1S1_12 hazA1S1_24 hazA1S2_12 hazA1S2_24 hazA1S3_12 hazA1S3_24;
print hazA2S1_12 hazA2S1_24 hazA2S2_12 hazA2S2_24 hazA2S3_12 hazA2S3_24;
print hazA3S1_12 hazA3S1_24 hazA3S2_12 hazA3S2_24 hazA3S3_12 hazA3S3_24;

print survA1S1_12 survA1S1_24 survA1S2_12 survA1S2_24 survA1S3_12 survA1S3_24;
print survA2S1_12 survA2S1_24 survA2S2_12 survA2S2_24 survA2S3_12 survA2S3_24;
print survA3S1_12 survA3S1_24 survA3S2_12 survA3S2_24 survA3S3_12 survA3S3_24;

print oddsA1S1_12 oddsA1S1_24 oddsA1S2_12 oddsA1S2_24 oddsA1S3_12 oddsA1S3_24;
print oddsA2S1_12 oddsA2S1_24 oddsA2S2_12 oddsA2S2_24 oddsA2S3_12 oddsA2S3_24;
print oddsA3S1_12 oddsA3S1_24 oddsA3S2_12 oddsA3S2_24 oddsA3S3_12 oddsA3S3_24;

print hazratioA1S1_12 hazratioA1S1_24 hazratioA1S2_12 hazratioA1S2_24 hazratioA1S3_12 hazratioA1S3_24;
print hazratioA2S1_12 hazratioA2S1_24 hazratioA2S2_12 hazratioA2S2_24 hazratioA2S3_12 hazratioA2S3_24;
print hazratioA3S1_12 hazratioA3S1_24 hazratioA3S2_12 hazratioA3S2_24 hazratioA3S3_12 hazratioA3S3_24;

print oddsratioA1S1_12 oddsratioA1S1_24 oddsratioA1S2_12 oddsratioA1S2_24 oddsratioA1S3_12 oddsratioA1S3_24;
print oddsratioA2S1_12 oddsratioA2S1_24 oddsratioA2S2_12 oddsratioA2S2_24 oddsratioA2S3_12 oddsratioA2S3_24;
print oddsratioA3S1_12 oddsratioA3S1_24 oddsratioA3S2_12 oddsratioA3S2_24 oddsratioA3S3_12 oddsratioA3S3_24;

**********Median Lifetime*****************************************************;
medianlifetime=(1/lambdaBASELINE)##(1/alpha);
medianA1S1=(1/lambdaA1S1)##(1/alpha);
medianA1S2=(1/lambdaA1S2)##(1/alpha);
medianA1S3=(1/lambdaA1S3)##(1/alpha);
medianA2S1=(1/lambdaA2S1)##(1/alpha);
medianA2S2=(1/lambdaA2S2)##(1/alpha);
medianA2S3=(1/lambdaA2S3)##(1/alpha);
medianA3S1=(1/lambdaA3S1)##(1/alpha);
medianA3S2=(1/lambdaA3S2)##(1/alpha);
medianA3S3=(1/lambdaA3S3)##(1/alpha);

print 'medianlifetime=' medianlifetime;
print 'medianlifetime(age=A1,score=S1)=' medianA1S1
      'medianlifetime(age=A1,score=S2)=' medianA1S2
      'medianlifetime(age=A1,score=S3)=' medianA1S3;
print 'medianlifetime(age=A2,score=S1)=' medianA2S1
      'medianlifetime(age=A2,score=S2)=' medianA2S2
      'medianlifetime(age=A2,score=S3)=' medianA2S3;
print 'medianlifetime(age=A3,score=S1)=' medianA3S1
      'medianlifetime(age=A3,score=S2)=' medianA3S2
      'medianlifetime(age=A3,score=S3)=' medianA3S3;
```

7. Program for fitting a Weibull regression model with two predictors

```
title1 'Fitting of regression model with two covariates';
title2 'Staggered entry: constant shape parameter';
```

```
proc iml worksize= 60;
reset nolog;
options pagesize=500;

**********Frequency vector (first entry,3 agegroups,3 scoregroups);
                f111={12,34,51,39,57,11, 59};    f111=f111<>1e-04;
                f112={10,12,22,19,32, 4,418};    f112=f112<>1e-04;
                f113={ 7,13,22,15,19, 0,165};    f113=f113<>1e-04;
                f121={13,14,45,27,33, 4, 66};    f121=f121<>1e-04;
                f122={ 4,22,22, 8,25, 4,297};    f122=f122<>1e-04;
                f123={ 4,14,24,10,17, 5,190};    f123=f123<>1e-04;
                f131={10,25,29,17,46, 2,116};    f131=f131<>1e-04;
                f132={ 6,13,28,16,16, 5,273};    f132=f132<>1e-04;
                f133={ 0,11,11, 6, 5, 0, 82};    f133=f133<>1e-04;

**********Vector of upper boundaries;
                x1={12,17,24,28,34,37};

**********Frequency vector (second entry,3 agegroups,3 scoregroups);
                f211={22,25,58,53,40, 45};       f211=f211<>1e-04;
                f212={10,26,32,20,29,379};       f212=f212<>1e-04;
                f213={ 9,24,13,19,14,204};       f213=f213<>1e-04;
                f221={24,24,28,30,25,106};       f221=f221<>1e-04;
                f222={12,20,14,17,16,409};       f222=f222<>1e-04;
                f223={13,18,19,19,13,238};       f223=f223<>1e-04;
                f231={13,15,32,19,17,107};       f231=f231<>1e-04;
                f232={11,13,22,17,12,319};       f232=f232<>1e-04;
                f233={ 4, 1,11, 6, 6,117};       f233=f233<>1e-04;

**********Vector of upper boundaries;
                x2={12,17,24,28,34};

**********Frequency vector (third entry,3 agegroups,3 scoregroups);
                f311={34,16,50,23, 54};          f311=f311<>1e-04;
                f312={19, 2,32,24,317};          f312=f312<>1e-04;
                f313={15,16,17,10,199};          f313=f313<>1e-04;
                f321={19,18,38,16, 75};          f321=f321<>1e-04;
                f322={16,14,25,10,263};          f322=f322<>1e-04;
                f323={ 5,12,20, 7,195};          f323=f323<>1e-04;
                f331={28,16,22,12, 98};          f331=f331<>1e-04;
                f332={13, 0,24, 4,323};          f332=f332<>1e-04;
                f333={ 5, 5,14,11,150};          f333=f333<>1e-04;

**********Vector of upper boundaries;
                x3={12,17,24,28};

**********Frequency vector (fourth entry,3 agegroups,3 scoregroups);
                f411={40,30,30, 50};             f411=f411<>1e-04;
                f412={ 9,14,27,301};             f412=f412<>1e-04;
                f413={22,16,12,222};             f413=f413<>1e-04;
                f421={24,30,29, 81};             f421=f421<>1e-04;
                f422={14,15,12,307};             f422=f422<>1e-04;
                f423={16,16,27,228};             f423=f423<>1e-04;
                f431={20,22,28,119};             f431=f431<>1e-04;
                f432={19,12,26,369};             f432=f432<>1e-04;
                f433={11,11,16,171};             f433=f433<>1e-04;

**********Vector of upper boundaries;
                x4={12,17,24};

**********Relative frequency vectors;
n111=f111[+]; n211=f211[+]; n311=f311[+]; n411=f411[+]; n11=n111+n211+n311+n411;
n112=f112[+]; n212=f212[+]; n312=f312[+]; n412=f412[+]; n12=n112+n212+n312+n412;
n113=f113[+]; n213=f213[+]; n313=f313[+]; n413=f413[+]; n13=n113+n213+n313+n413;
n121=f121[+]; n221=f221[+]; n321=f321[+]; n421=f421[+]; n21=n121+n221+n321+n421;
n122=f122[+]; n222=f222[+]; n322=f322[+]; n422=f422[+]; n22=n122+n222+n322+n422;
n123=f123[+]; n223=f223[+]; n323=f323[+]; n423=f423[+]; n23=n123+n223+n323+n423;
n131=f131[+]; n231=f231[+]; n331=f331[+]; n431=f431[+]; n31=n131+n231+n331+n431;
n132=f132[+]; n232=f232[+]; n332=f332[+]; n432=f432[+]; n32=n132+n232+n332+n432;
n133=f133[+]; n233=f233[+]; n333=f333[+]; n433=f433[+]; n33=n133+n233+n333+n433;

                                                        n1=n11+n12+n13;
                                                        n2=n21+n22+n23;
                                                        n3=n31+n32+n33;
                                                        n=n1+n2+n3;

k1=nrow(f111); d1=k1-1;
k2=nrow(f211); d2=k2-1;
k3=nrow(f311); d3=k3-1;
k4=nrow(f411); d4=k4-1;
k=k1+k2+k3+k4;
d=d1+d2+d3+d4;

p111=f111/n111; p211=f211/n211; p311=f311/n311; p411=f411/n411;
p112=f112/n112; p212=f212/n212; p312=f312/n312; p412=f412/n412;
p113=f113/n113; p213=f213/n213; p313=f313/n313; p413=f413/n413;
p121=f121/n121; p221=f221/n221; p321=f321/n321; p421=f421/n421;
p122=f122/n122; p222=f222/n222; p322=f322/n322; p422=f422/n422;
```

```
p123=f123/n123; p223=f223/n223; p323=f323/n323; p423=f423/n423;
p131=f131/n131; p231=f231/n231; p331=f331/n331; p431=f431/n431;
p132=f132/n132; p232=f232/n232; p332=f332/n332; p432=f432/n432;
p133=f133/n133; p233=f233/n233; p333=f333/n333; p433=f433/n433;

p11=p111//p211//p311//p411;
p12=p112//p212//p312//p412;
p13=p113//p213//p313//p413;
p1=p11//p12//p13;

p21=p121//p221//p321//p421;
p22=p122//p222//p322//p422;
p23=p123//p223//p323//p423;
p2=p21//p22//p23;

p31=p131//p231//p331//p431;
p32=p132//p232//p332//p432;
p33=p133//p233//p333//p433;
p3=p31//p32//p33;

p=p1//p2//p3;

**********Design matrix and matrix orthogonal to design matrix;

S1=J(d1,1,1)@cusum(J(1,k1,1));
S2=J(1,k1,1)@cusum(J(d1,1,1));
S1=S1<=S2;
S2=S1[1:d2,1:d1];
S3=S1[1:d3,1:d2];
S4=S1[1:d4,1:d3];
S=block(S1,S2,S3,S4);
S=I(3)@I(3)@S;

AA=J(162,1,1);           *162=9 times((7-1)+(6-1)+(5-1)+(4-1))=9 times 18=162;
BB=J(3,1,1);
CC=(I(2)//J(1,2,-1));
DD=J(18,1,1);            *3 times d1=3(6)=18;
EE=J(15,1,1);            *3 times d2=3(5)=15;
FF=J(12,1,1);            *3 times d3=3(4)=12;
GG=J(9,1,1);             *3 times d4=3(3)=9;

KK=J(9,1,1);             *9=3 times 3;
LL=J(54,1,1);            *54=18 times 3;

lx=BB@BB@(log(x1)//log(x2)//log(x3)//log(x4));

HH=BB@(CC@DD);

xc=AA||CC@LL||HH||lx;  print xc;

C=I(9#d)-xc*inv(xc`*xc)*xc`;

**********ITERATIVE PROCEDURE (double iterations over m and p);
*****starting value for m;
m=p;
ms=S*m; ps=ms;
p0=p;

*****iteration over m;
itr=0;
verskil1=1;
i=0;
do while (verskil1>0.00000001);
i=i+1;
p=p0;
Gm=-C*diag(1/(log(1-ms)))*diag(1/(1-ms))*S;                  *Weibull;
*Gm=C*diag(1/ms+1/(1-ms))*S;                                 *log-logistic;

**********iteration over p;
     verskil=1;
     j=0;
     do while (verskil>0.00000001);
     j=j+1;
     p1=p;
     ps=S*p;
    g=C*log(-log(1-ps));                                     *Weibull;
    *g=C*(log(ps)-log(1-ps));                                *log-logistic;
    Gp=-C*diag(1/(log(1-ps)))*diag(1/(1-ps))*S;              *Weibull;
    *Gp=C*(diag(1/ps)+diag(1/(1-ps)))*S;                     *log-logistic;

******************covariance matrix**********************************;
m111=m[1:k1];
m211=m[k1+1:k1+k2];
m311=m[k1+k2+1:k1+k2+k3];
m411=m[k1+k2+k3+1:k1+k2+k3+k4];

m112=m[k+1:k+k1];
m212=m[k+k1+1:k+k1+k2];
m312=m[k+k1+k2+1:k+k1+k2+k3];
```

```
m412=m[k+k1+k2+k3+1:k+k1+k2+k3+k4];

m113=m[2#k+1:2#k+k1];
m213=m[2#k+k1+1:2#k+k1+k2];
m313=m[2#k+k1+k2+1:2#k+k1+k2+k3];
m413=m[2#k+k1+k2+k3+1:2#k+k1+k2+k3+k4];

m121=m[3#k+1:3#k+k1];
m221=m[3#k+k1+1:3#k+k1+k2];
m321=m[3#k+k1+k2+1:3#k+k1+k2+k3];
m421=m[3#k+k1+k2+k3+1:3#k+k1+k2+k3+k4];

m122=m[4#k+1:4#k+k1];
m222=m[4#k+k1+1:4#k+k1+k2];
m322=m[4#k+k1+k2+1:4#k+k1+k2+k3];
m422=m[4#k+k1+k2+k3+1:4#k+k1+k2+k3+k4];

m123=m[5#k+1:5#k+k1];
m223=m[5#k+k1+1:5#k+k1+k2];
m323=m[5#k+k1+k2+1:5#k+k1+k2+k3];
m423=m[5#k+k1+k2+k3+1:5#k+k1+k2+k3+k4];

m131=m[6#k+1:6#k+k1];
m231=m[6#k+k1+1:6#k+k1+k2];
m331=m[6#k+k1+k2+1:6#k+k1+k2+k3];
m431=m[6#k+k1+k2+k3+1:6#k+k1+k2+k3+k4];

m132=m[7#k+1:7#k+k1];
m232=m[7#k+k1+1:7#k+k1+k2];
m332=m[7#k+k1+k2+1:7#k+k1+k2+k3];
m432=m[7#k+k1+k2+k3+1:7#k+k1+k2+k3+k4];

m133=m[8#k+1:8#k+k1];
m233=m[8#k+k1+1:8#k+k1+k2];
m333=m[8#k+k1+k2+1:8#k+k1+k2+k3];
m433=m[8#k+k1+k2+k3+1:9#k];

sig111=(1/n111)*(diag(m111)-m111*m111`);
sig112=(1/n112)*(diag(m112)-m112*m112`);
sig113=(1/n113)*(diag(m113)-m113*m113`);
sig211=(1/n211)*(diag(m211)-m211*m211`);
sig212=(1/n212)*(diag(m212)-m212*m212`);
sig213=(1/n213)*(diag(m213)-m213*m213`);
sig311=(1/n311)*(diag(m311)-m311*m311`);
sig312=(1/n312)*(diag(m312)-m312*m312`);
sig313=(1/n313)*(diag(m313)-m313*m313`);
sig411=(1/n411)*(diag(m411)-m411*m411`);
sig412=(1/n412)*(diag(m412)-m412*m412`);
sig413=(1/n413)*(diag(m413)-m413*m413`);

sig11=block(sig111,sig211,sig311,sig411);
sig12=block(sig112,sig212,sig312,sig412);
sig13=block(sig113,sig213,sig313,sig413);


sig121=(1/n121)*(diag(m121)-m121*m121`);
sig122=(1/n122)*(diag(m122)-m122*m122`);
sig123=(1/n123)*(diag(m123)-m123*m123`);
sig221=(1/n221)*(diag(m221)-m221*m221`);
sig222=(1/n222)*(diag(m222)-m222*m222`);
sig223=(1/n223)*(diag(m223)-m223*m223`);
sig321=(1/n321)*(diag(m321)-m321*m321`);
sig322=(1/n322)*(diag(m322)-m322*m322`);
sig323=(1/n323)*(diag(m323)-m323*m323`);
sig421=(1/n421)*(diag(m421)-m421*m421`);
sig422=(1/n422)*(diag(m422)-m422*m422`);
sig423=(1/n423)*(diag(m423)-m423*m423`);

sig21=block(sig121,sig221,sig321,sig421);
sig22=block(sig122,sig222,sig322,sig422);
sig23=block(sig123,sig223,sig323,sig423);


sig131=(1/n131)*(diag(m131)-m131*m131`);
sig132=(1/n132)*(diag(m132)-m132*m132`);
sig133=(1/n133)*(diag(m133)-m133*m133`);
sig231=(1/n231)*(diag(m231)-m231*m231`);
sig232=(1/n232)*(diag(m232)-m232*m232`);
sig233=(1/n233)*(diag(m233)-m233*m233`);
sig331=(1/n331)*(diag(m331)-m331*m331`);
sig332=(1/n332)*(diag(m332)-m332*m332`);
sig333=(1/n333)*(diag(m333)-m333*m333`);
sig431=(1/n431)*(diag(m431)-m431*m431`);
sig432=(1/n432)*(diag(m432)-m432*m432`);
sig433=(1/n433)*(diag(m433)-m433*m433`);

sig31=block(sig131,sig231,sig331,sig431);
sig32=block(sig132,sig232,sig332,sig432);
sig33=block(sig133,sig233,sig333,sig433);
```

```
sig=block(sig11,sig12,sig13,sig21,sig22,sig23,sig31,sig32,sig33);
V=sig;
**********************************************************************;
    p=p-(Gm*V)'*ginv(Gp*V*Gm')*g;                      *Weibull;
    *p=p-(Gm*V)'*ginv(Gp*V*Gm')*g;                     *log-logistic;
     verskil=sqrt((p-p1)'*(p-p1));
     end;
verskil1=sqrt((p-m)'*(p-m));
m=p;ms=S*m;
end;
print m; print i j;

**********Parameter vector for linear model;
par=inv(xc'*xc)*xc'*log(-log(1-ms));                    *Weibull;
*par=inv(xc'*xc)*xc'*(log(ms)-log(1-ms));               *log-logistic;

print par[format=E20.];

**********Regression coefficients;
oualphaBASELINE=par[1];
betaA1=par[2];
betaA2=par[3];
betaA3=-(par[2]+par[3]);
betaS1=par[4];
betaS2=par[5];
betaS3=-(par[4]+par[5]);
lambdaBASELINE=exp(par[1]);
alpha=par[6];

**********Indices, lambda's and constant alpha for age levels******************;
indexA1=exp(betaA1);          *constant shape parameter;
indexA2=exp(betaA2);
indexA3=exp(betaA3);
indexS1=exp(betaS1);
indexS2=exp(betaS2);
indexS3=exp(betaS3);

lambdaA1S1=lambdaBASELINE#indexA1#indexS1;
lambdaA1S2=lambdaBASELINE#indexA1#indexS2;
lambdaA1S3=lambdaBASELINE#indexA1#indexS3;
lambdaA2S1=lambdaBASELINE#indexA2#indexS1;
lambdaA2S2=lambdaBASELINE#indexA2#indexS2;
lambdaA2S3=lambdaBASELINE#indexA2#indexS3;
lambdaA3S1=lambdaBASELINE#indexA3#indexS1;
lambdaA3S2=lambdaBASELINE#indexA3#indexS2;
lambdaA3S3=lambdaBASELINE#indexA3#indexS3;

oualphaA1S1=oualphaBASELINE+betaA1+betaS1; *same as oualphaA1S1=log(lambdaA1S1);
oualphaA1S2=oualphaBASELINE+betaA1+betaS2;
oualphaA1S3=oualphaBASELINE+betaA1+betaS3;
oualphaA2S1=oualphaBASELINE+betaA2+betaS1;
oualphaA2S2=oualphaBASELINE+betaA2+betaS2;
oualphaA2S3=oualphaBASELINE+betaA2+betaS3;
oualphaA3S1=oualphaBASELINE+betaA3+betaS1;
oualphaA3S2=oualphaBASELINE+betaA3+betaS2;
oualphaA3S3=oualphaBASELINE+betaA3+betaS3;

print ' Weibull parameters, beta effects and indices: MLE subject to constraints';

print 'lambdaBASELINE=' lambdaBASELINE[format=E20.]  'oualphaBASELINE=' oualphaBASELINE;
print 'alpha=' alpha[format=E20.];

print 'lambda(age=A1,score=S1)=' lambdaA1S1[format=E20.]
      'lambda(age=A1,score=S2)=' lambdaA1S2[format=E20.]
      'lambda(age=A1,score=S3)=' lambdaA1S3[format=E20.];
print oualphaA1S1 oualphaA1S2 oualphaA1S3;

print 'lambda(age=A2,score=S1)=' lambdaA2S1[format=E20.]
      'lambda(age=A2,score=S2)=' lambdaA2S2[format=E20.]
      'lambda(age=A2,score=S3)=' lambdaA2S3[format=E20.];
print oualphaA2S1 oualphaA2S2 oualphaA2S3;

print 'lambda(age=A3,score=S1)=' lambdaA3S1[format=E20.]
      'lambda(age=A3,score=S2)=' lambdaA3S2[format=E20.]
      'lambda(age=A3,score=S3)=' lambdaA3S3[format=E20.];
print oualphaA3S1 oualphaA3S2 oualphaA3S3;

print 'beta(age=A1)=' betaA1 'beta(age=A2)=' betaA2 'beta(age=A3)=' betaA3;
print 'beta(score=S1)=' betaS1 'beta(score=S2)=' betaS2 'beta(score=S3)=' betaS3;
print 'index(age=A1)=' indexA1 'index(age=A2)=' indexA2 'index(age=A3)=' indexA3;
print 'index(score=S1)=' indexS1 'index(score=S2)=' indexS2 'index(score=S3)=' indexS3;

**********Hazard ratio and Odds ratio*******************************************;
hazBASELINE12=lambdaBASELINE*alpha*12**(alpha-1);
hazBASELINE24=lambdaBASELINE*alpha*24**(alpha-1);
survBASELINE12=exp(-lambdaBASELINE*12**alpha);
survBASELINE24=exp(-lambdaBASELINE*24**alpha);
oddsBASELINE12=(1-survBASELINE12)/survBASELINE12;
```

```
oddsBASELINE24=(1-survBASELINE24)/survBASELINE24;

hazA1S1_12=lambdaA1S1*alpha*12**(alpha-1);
hazA1S1_24=lambdaA1S1*alpha*24**(alpha-1);
hazA1S2_12=lambdaA1S2*alpha*12**(alpha-1);
hazA1S2_24=lambdaA1S2*alpha*24**(alpha-1);
hazA1S3_12=lambdaA1S3*alpha*12**(alpha-1);
hazA1S3_24=lambdaA1S3*alpha*24**(alpha-1);
hazA2S1_12=lambdaA2S1*alpha*12**(alpha-1);
hazA2S1_24=lambdaA2S1*alpha*24**(alpha-1);
hazA2S2_12=lambdaA2S2*alpha*12**(alpha-1);
hazA2S2_24=lambdaA2S2*alpha*24**(alpha-1);
hazA2S3_12=lambdaA2S3*alpha*12**(alpha-1);
hazA2S3_24=lambdaA2S3*alpha*24**(alpha-1);
hazA3S1_12=lambdaA3S1*alpha*12**(alpha-1);
hazA3S1_24=lambdaA3S1*alpha*24**(alpha-1);
hazA3S2_12=lambdaA3S2*alpha*12**(alpha-1);
hazA3S2_24=lambdaA3S2*alpha*24**(alpha-1);
hazA3S3_12=lambdaA3S3*alpha*12**(alpha-1);
hazA3S3_24=lambdaA3S3*alpha*24**(alpha-1);

survA1S1_12=exp(-lambdaA1S1*12**alpha);
survA1S1_24=exp(-lambdaA1S1*24**alpha);
survA1S2_12=exp(-lambdaA1S2*12**alpha);
survA1S2_24=exp(-lambdaA1S2*24**alpha);
survA1S3_12=exp(-lambdaA1S3*12**alpha);
survA1S3_24=exp(-lambdaA1S3*24**alpha);
survA2S1_12=exp(-lambdaA2S1*12**alpha);
survA2S1_24=exp(-lambdaA2S1*24**alpha);
survA2S2_12=exp(-lambdaA2S2*12**alpha);
survA2S2_24=exp(-lambdaA2S2*24**alpha);
survA2S3_12=exp(-lambdaA2S3*12**alpha);
survA2S3_24=exp(-lambdaA2S3*24**alpha);
survA3S1_12=exp(-lambdaA3S1*12**alpha);
survA3S1_24=exp(-lambdaA3S1*24**alpha);
survA3S2_12=exp(-lambdaA3S2*12**alpha);
survA3S2_24=exp(-lambdaA3S2*24**alpha);
survA3S3_12=exp(-lambdaA3S3*12**alpha);
survA3S3_24=exp(-lambdaA3S3*24**alpha);

oddsA1S1_12=(1-survA1S1_12)/survA1S1_12;
oddsA1S1_24=(1-survA1S1_24)/survA1S1_24;
oddsA1S2_12=(1-survA1S2_12)/survA1S2_12;
oddsA1S2_24=(1-survA1S2_24)/survA1S2_24;
oddsA1S3_12=(1-survA1S3_12)/survA1S3_12;
oddsA1S3_24=(1-survA1S3_24)/survA1S3_24;
oddsA2S1_12=(1-survA2S1_12)/survA2S1_12;
oddsA2S1_24=(1-survA2S1_24)/survA2S1_24;
oddsA2S2_12=(1-survA2S2_12)/survA2S2_12;
oddsA2S2_24=(1-survA2S2_24)/survA2S2_24;
oddsA2S3_12=(1-survA2S3_12)/survA2S3_12;
oddsA2S3_24=(1-survA2S3_24)/survA2S3_24;
oddsA3S1_12=(1-survA3S1_12)/survA3S1_12;
oddsA3S1_24=(1-survA3S1_24)/survA3S1_24;
oddsA3S2_12=(1-survA3S2_12)/survA3S2_12;
oddsA3S2_24=(1-survA3S2_24)/survA3S2_24;
oddsA3S3_12=(1-survA3S3_12)/survA3S3_12;
oddsA3S3_24=(1-survA3S3_24)/survA3S3_24;

hazratioA1S1_12=hazA1S1_12/hazBASELINE12;
hazratioA1S1_24=hazA1S1_24/hazBASELINE24;
hazratioA1S2_12=hazA1S2_12/hazBASELINE12;
hazratioA1S2_24=hazA1S2_24/hazBASELINE24;
hazratioA1S3_12=hazA1S3_12/hazBASELINE12;
hazratioA1S3_24=hazA1S3_24/hazBASELINE24;
hazratioA2S1_12=hazA2S1_12/hazBASELINE12;
hazratioA2S1_24=hazA2S1_24/hazBASELINE24;
hazratioA2S2_12=hazA2S2_12/hazBASELINE12;
hazratioA2S2_24=hazA2S2_24/hazBASELINE24;
hazratioA2S3_12=hazA2S3_12/hazBASELINE12;
hazratioA2S3_24=hazA2S3_24/hazBASELINE24;
hazratioA3S1_12=hazA3S1_12/hazBASELINE12;
hazratioA3S1_24=hazA3S1_24/hazBASELINE24;
hazratioA3S2_12=hazA3S2_12/hazBASELINE12;
hazratioA3S2_24=hazA3S2_24/hazBASELINE24;
hazratioA3S3_12=hazA3S3_12/hazBASELINE12;
hazratioA3S3_24=hazA3S3_24/hazBASELINE24;

oddsratioA1S1_12=oddsA1S1_12/oddsBASELINE12;
oddsratioA1S1_24=oddsA1S1_24/oddsBASELINE24;
oddsratioA1S2_12=oddsA1S2_12/oddsBASELINE12;
oddsratioA1S2_24=oddsA1S2_24/oddsBASELINE24;
oddsratioA1S3_12=oddsA1S3_12/oddsBASELINE12;
oddsratioA1S3_24=oddsA1S3_24/oddsBASELINE24;
oddsratioA2S1_12=oddsA2S1_12/oddsBASELINE12;
oddsratioA2S1_24=oddsA2S1_24/oddsBASELINE24;
oddsratioA2S2_12=oddsA2S2_12/oddsBASELINE12;
oddsratioA2S2_24=oddsA2S2_24/oddsBASELINE24;
oddsratioA2S3_12=oddsA2S3_12/oddsBASELINE12;
```

```
oddsratioA2S3_24=oddsA2S3_24/oddsBASELINE24;
oddsratioA3S1_12=oddsA3S1_12/oddsBASELINE12;
oddsratioA3S1_24=oddsA3S1_24/oddsBASELINE24;
oddsratioA3S2_12=oddsA3S2_12/oddsBASELINE12;
oddsratioA3S2_24=oddsA3S2_24/oddsBASELINE24;
oddsratioA3S3_12=oddsA3S3_12/oddsBASELINE12;
oddsratioA3S3_24=oddsA3S3_24/oddsBASELINE24;

print hazBASELINE12 hazBASELINE24 survBASELINE12 survBASELINE24 oddsBASELINE12 oddsBASELINE24;

print hazA1S1_12 hazA1S1_24 hazA1S2_12 hazA1S2_24 hazA1S3_12 hazA1S3_24;
print hazA2S1_12 hazA2S1_24 hazA2S2_12 hazA2S2_24 hazA2S3_12 hazA2S3_24;
print hazA3S1_12 hazA3S1_24 hazA3S2_12 hazA3S2_24 hazA3S3_12 hazA3S3_24;

print survA1S1_12 survA1S1_24 survA1S2_12 survA1S2_24 survA1S3_12 survA1S3_24;
print survA2S1_12 survA2S1_24 survA2S2_12 survA2S2_24 survA2S3_12 survA2S3_24;
print survA3S1_12 survA3S1_24 survA3S2_12 survA3S2_24 survA3S3_12 survA3S3_24;

print oddsA1S1_12 oddsA1S1_24 oddsA1S2_12 oddsA1S2_24 oddsA1S3_12 oddsA1S3_24;
print oddsA2S1_12 oddsA2S1_24 oddsA2S2_12 oddsA2S2_24 oddsA2S3_12 oddsA2S3_24;
print oddsA3S1_12 oddsA3S1_24 oddsA3S2_12 oddsA3S2_24 oddsA3S3_12 oddsA3S3_24;

print hazratioA1S1_12 hazratioA1S1_24 hazratioA1S2_12 hazratioA1S2_24 hazratioA1S3_12 hazratioA1S3_24;
print hazratioA2S1_12 hazratioA2S1_24 hazratioA2S2_12 hazratioA2S2_24 hazratioA2S3_12 hazratioA2S3_24;
print hazratioA3S1_12 hazratioA3S1_24 hazratioA3S2_12 hazratioA3S2_24 hazratioA3S3_12 hazratioA3S3_24;

print oddsratioA1S1_12 oddsratioA1S1_24 oddsratioA1S2_12 oddsratioA1S2_24 oddsratioA1S3_12 oddsratioA1S3_24;
print oddsratioA2S1_12 oddsratioA2S1_24 oddsratioA2S2_12 oddsratioA2S2_24 oddsratioA2S3_12 oddsratioA2S3_24;
print oddsratioA3S1_12 oddsratioA3S1_24 oddsratioA3S2_12 oddsratioA3S2_24 oddsratioA3S3_12 oddsratioA3S3_24;

**********Median Lifetime*********************************************************;
medianBASELINE=((1/lambdaBASELINE)#log(2))##(1/alpha);
medianA1S1=((1/lambdaA1S1)#log(2))##(1/alpha);
medianA1S2=((1/lambdaA1S2)#log(2))##(1/alpha);
medianA1S3=((1/lambdaA1S3)#log(2))##(1/alpha);
medianA2S1=((1/lambdaA2S1)#log(2))##(1/alpha);
medianA2S2=((1/lambdaA2S2)#log(2))##(1/alpha);
medianA2S3=((1/lambdaA2S3)#log(2))##(1/alpha);
medianA3S1=((1/lambdaA3S1)#log(2))##(1/alpha);
medianA3S2=((1/lambdaA3S2)#log(2))##(1/alpha);
medianA3S3=((1/lambdaA3S3)#log(2))##(1/alpha);

print 'medianlifetimeBASELINE=' medianBASELINE;
print 'medianlifetime(age=A1,score=S1)=' medianA1S1
      'medianlifetime(age=A1,score=S2)=' medianA1S2
      'medianlifetime(age=A1,score=S3)=' medianA1S3;
print 'medianlifetime(age=A2,score=S1)=' medianA2S1
      'medianlifetime(age=A2,score=S2)=' medianA2S2
      'medianlifetime(age=A2,score=S3)=' medianA2S3;
print 'medianlifetime(age=A3,score=S1)=' medianA3S1
      'medianlifetime(age=A3,score=S2)=' medianA3S2
      'medianlifetime(age=A3,score=S3)=' medianA3S3;
```

Bibliography

[1] AGRESTI, A. (1990). *Categorical data analysis.* New York: Wiley.

[2] BAIN, L.J. and ENGELHARDT, M. (1991). *Statistical analysis of reliability and life-testing models:theory and methods.* New York: Marcel Dekker.

[3] BENNETT, S. (1983). Analysis of survival data by the proportional odds model. Ph.D. Thesis, University of Reading, United Kingdom.

[4] CANTOR, A. (1998). *Extending Sas survival analysis techniques for medical research.* USA: SAS Institute Inc.

[5] COLLETT, D. (1994). *Modelling survival data in medical research.* London: Chapman Hall.

[6] COX, D.R. (1972). Regression models and life-tables. *Journal of the royal statistical society,* **34***: 187-220.*

[7] COX, D.R. and OAKES, D. (1984). *Analysis of survival data.* London: Chapman Hall.

[8] COX, D.R. (1995). Some remarks on the analysis of survival data. *Proceedings of the First Seattle Symposium in Biostatistics: Survival Analysis 123-169.*

[9] CROWDER, M.J., KIMBER, A.C., SMITH, R.L. and SWEETING, T.J. (1991) *Statistical analysis of reliability data.* London: Chapman Hall.

[10] CROWTHER, N.A.S. and MATTHEWS, G.B. (unpublished paper). Fitting distributions to grouped data.

[11] CROWTHER, N.A.S. and MATTHEWS, G.B. (1995). A Maximum likelihood estimation procedure when modelling in terms of constraints. *S.A. Statistical Journal,* **29(1)**, *29-51.*

[12] CROWTHER, N.A.S. and SHAW, T.M. (1989). An estimation procedure in categorical data analysis. *S.A. Statistical Journal,* **23(1)***: 63-83.*

[13] ELANDT-JOHNSON, R.C. and JOHNSON, N.L. (1990). *Survival models and data analysis.* New York: Wiley.

[14] FARRINGTON, C.P. (1996). Interval censored survival data: a generalized linear modeling approach. *Statistics in Medicine*, **15**: *283-292.*

[15] FINKELSTEIN, D.M. (1986). A proportional hazards model for interval censored failure time data. *Biometrics*, **42**: *845-854.*

[16] GROSS, A.J. and CLARK, V.A. (1975). *Survival distributions: reliability applications in the Biomedical Sciences.* New York: Wiley.

[17] HEITJAN, D.F. (1989). Inference from grouped continuous data: a review. *Ststistical Science*, **4**: *164-183.*

[18] HOSMER, D.W. and LEMESHOW, S. (1999). *Applied survival analysis: regression modeling of time to event data.* New York: Wiley.

[19] HUANG, J. and WELLNER, J.A. (1995). Interval censored survival data: a review of recent progress. *Proceedings of the First Seattle Symposium in Biostatistics: Survival Analysis 123-169.*

[20] JOUBERT, H.M. (1991). Fitting of distributions in the case of survey data. M.Sc. dissertation, University of South Africa, South Africa.

[21] KALBFLEISCH, J.D. and PRENTICE, R.L. (1980). *The statistical analysis of failure time data.* New York: Wiley.

[22] KLEIN, J.P. and MOESCHBERGER, M.L. (1997). *Survival analysis techniques for censored and truncated data.* New York: Springer-Verlag.

[23] KLEINBAUM, D.G. (1997). *Survival analysis: a self-learning text.* New York: Springer-Verlag.

[24] LAWLESS, J.F. (1982). *Statistical models and methods for lifetime data.* New York: Wiley.

[25] LE, C.T. (1997). *Applied survival analysis.* New York: Wiley.

[26] LEE, E.T. (1992). *Statistical methods for survival data analysis.* New York: Wiley.

[27] LINDSEY, J.K. (1998). A study of interval censoring in parametric regression models. *Lifetime Data Analysis*, **4**: *329-354.*

[28] LOUZADA-NETO, F. (1997). Extended hazard regression model for reliability and survival analysis. *Lifetime Data Analysis*, **3**: *367-381.*

[29] MALLER, R. and ZHOU, X. (1996). *Survival analysis with long-term survivors.* New York: Wiley.

[30] MATTHEWS, G.B. (1995). Maximum likelihood estimation when modelling in terms of constraints. Ph.D. Thesis, University of Pretoria, South Africa.

[31] MATTHEWS, G.B. and CROWTHER, N.A.S. (1998). A Maximum likelihood estimation procedure for the generalized linear model with restrictions. *S.A. Statistical Journal,* **32(2)**, *119-144.*

[32] McCullagh, P. (1980). Regression models for ordinal data (with discussion). *Journal of the Royal Statistical Society,* **42**: *109-142.*

[33] McKEAGUE, I.W. and ZHANG, M. (1996). Fitting Cox's proportional hazards model using grouped survival data. *Lifetime Data: Models in Reliability and Survival Analysis, 227-232.*

[34] MILLER, R.G. (1998). *Survival analysis.* New York: Wiley.

[35] NELSON, W. (1990). *Accelerated testing, statistical models, test plans, and data analysis.* New York: Wiley.

[36] ODELL, P.M., ANDERSON, K.M. and D'AGOSTINO, R.B. (1992). Maximum likelihood estimation for interval-censored data using a Weibull-based accelerated failure time model. *Biometrics,* **48**: *951-959.*

[37] PANJER, H.H. and WILLMOT, G.E. (1992). *Insurance risk models.* USA: Society of Actuaries.

[38] PARMAR, M.K.B. and MACHIN, D. (1995). *Survival analysis: a practical approach.* Chichester: Wiley.

[39] RAFTERY, A.E., TANNER, M.A. and WELLS, M.T. (2001). *Statistics in the 21st Century.* London: Chapman and Hall.

[40] SAS Institute Inc. (2000). *SAS/IML*^{TM} *user's guide for personal computers, Version 8.1* Cary, NC: SAS Institute Inc.

[41] WEI, L.J. (1992). The accelerated failure time model: a useful alternative to the Cox regression model in survival analysis. *Statistics in Medicine,* **11**: *1871-1879.*

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Abstract

Fitting of survival functions for grouped data

on insurance policies

by

Elizabeth Magrietha Louw

Supervisor: **Professor N.A.S. Crowther**

Degree: **Ph D**

The aim of the research is the statistical modelling of parametric survival distributions of grouped survival data of long- and shortterm policies in the insurance industry, by means of a method of maximum likelihood estimation subject to constraints.

This methodology leads to explicit expressions for the estimates of the parameters, as well as for approximated variances and covariances of the estimates, which gives exact maximum likelihood estimates of the parameters. This makes direct extension to more complex designs feasible.

The statistical modelling offers parametric models for survival distributions, in contrast with non-parametric models that are used commonly in the actuarial profession. When the parametric models provide a good fit to data, they tend to give more precise estimates of the quantities of interest such as odds ratios, hazard ratios or median lifetimes. These estimates form the statistical foundation for scientific decisionmaking with respect to actuarial design, maintenance and marketing of insurance policies.

Although the methodology in this thesis is developed specifically for the insurance industry, it may be applied in the normal context of research and scientific decisionmaking, that includes for example survival distributions for the medical, biological, engineering, econometric and sociological sciences.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Ekserp

Fitting of survival functions for grouped data on insurance policies

deur

Elizabeth Magrietha Louw

Promotor: **Professor N.A.S. Crowther**
Graad: **Ph D**

Die doelwit van die navorsing is die statistiese modellering van parametriese oorlewingsverdelings van gegroepeerde oorlewingsdata van lang- en korttermyn polisse in die versekeringsbedryf, deur middel van 'n metode van maksimum aanneemlikheidsberaming onderworpe aan beperkings.

Hierdie metode lei tot eksplisiete uitdrukkings vir die beramings van die parameters, asook vir benaderde variansies en kovariansies van die beramers, wat eksakte maksimum aanneemlikheidsberamings van die parameters gee. Dit maak direkte uitbreiding na meer komplekse ontwerpe moontlik.

Die statistiese modellering bied parametriese modelle vir oorlewingsverdelings, in teenstelling met nie-parametriese modelle wat algemeen in die aktuariële professie gebruik word. Indien die parametriese modelle 'n goeie passing by die data gee, sal hulle meer akkurate beramings van die hoeveelhede van belang, soos kruisprodukverhoudings, risikoverhoudings of mediaan leeftye gee. Hierdie beramings vorm die statistiese grondslag vir wetenskaplike besluitneming met betrekking tot aktuariële ontwerp, onderhoud en bemarking van versekeringspolisse.

Alhoewel die metodologie in hierdie tesis spesifiek vir die versekeringsbedryf ontwikkel is, kan dit in die normale konteks van navorsing en wetenskaplike besluitneming, wat byvoorbeeld oorlewingsverdelings vir die mediese, biologiese, ingenieurswese, ekonometriese en sosiologiese wetenskappe insluit, toegepas word.