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 × 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

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

⇒ 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

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

⇒ 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 $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_p)'$

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

- 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 $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_p)'$

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

- 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

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

⇒ 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

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

⇒ 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

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

- 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

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

  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

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

- 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 $\dfrac{h_j(t)}{h_0(t)}$, converges from the value $\exp(-\sum\limits_{k=1}^{p} \beta_k Z_k)$ 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

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

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.

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

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 $\boldsymbol{Z} = \mathbf{0}$, 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

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

  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

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

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

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

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

$$i = 1, 2, 3, 4 \quad \text{and} \quad l = 1, 2, 3.$$

$\boldsymbol{p}_{1l} = (p_{1l,1}, p_{1l,2}, p_{1l,3}, p_{1l,4}, p_{1l,5}, p_{1l,6}, p_{1l,7})'$ 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 $\quad l =$ 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 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	Risk Factor A			Total
Period	Level A_1	Level A_2	Level A_3	
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 $\boldsymbol{S}_{A_l} =$ 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

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

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:

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

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

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

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}\\[1em]\mathbf{1}\\\mathbf{1}\\\mathbf{1}\\\mathbf{1}\\[1em]\mathbf{1}\\\mathbf{1}\\\mathbf{1}\\\mathbf{1}\end{pmatrix}+\beta_{A_1}\cdot\begin{pmatrix}\mathbf{1}\\\mathbf{1}\\\mathbf{1}\\\mathbf{1}\\[1em]\mathbf{0}\\\mathbf{0}\\\mathbf{0}\\\mathbf{0}\\[1em]-\mathbf{1}\\-\mathbf{1}\\-\mathbf{1}\\-\mathbf{1}\end{pmatrix}+\beta_{A_2}\cdot\begin{pmatrix}\mathbf{0}\\\mathbf{0}\\\mathbf{0}\\\mathbf{0}\\[1em]\mathbf{1}\\\mathbf{1}\\\mathbf{1}\\\mathbf{1}\\[1em]-\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\\[1em]\ln\boldsymbol{x}_1\\\ln\boldsymbol{x}_2\\\ln\boldsymbol{x}_3\\\ln\boldsymbol{x}_4\\[1em]\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\\[1em]\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\\[1em]\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}$$

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

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

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

$\boldsymbol{C}$ 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

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

with

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

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

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

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

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

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

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

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}

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

Take the summation over the risk factor levels, that gives

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

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

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

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

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

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

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

$$
\begin{aligned}
&= \underbrace{\begin{pmatrix}
1 & 1 & 0 & \ln \boldsymbol{x}_1 & 0 & 0 \\
1 & 1 & 0 & \ln \boldsymbol{x}_2 & 0 & 0 \\
1 & 1 & 0 & \ln \boldsymbol{x}_3 & 0 & 0 \\
1 & 1 & 0 & \ln \boldsymbol{x}_4 & 0 & 0 \\
 & & & & & \\
1 & 0 & 1 & 0 & \ln \boldsymbol{x}_1 & 0 \\
1 & 0 & 1 & 0 & \ln \boldsymbol{x}_2 & 0 \\
1 & 0 & 1 & 0 & \ln \boldsymbol{x}_3 & 0 \\
1 & 0 & 1 & 0 & \ln \boldsymbol{x}_4 & 0 \\
 & & & & & \\
1 & -1 & -1 & 0 & 0 & \ln \boldsymbol{x}_1 \\
1 & -1 & -1 & 0 & 0 & \ln \boldsymbol{x}_2 \\
1 & -1 & -1 & 0 & 0 & \ln \boldsymbol{x}_3 \\
1 & -1 & -1 & 0 & 0 & \ln \boldsymbol{x}_4
\end{pmatrix}} \cdot \begin{pmatrix} \ln \lambda \\ \beta_{A_1} \\ \beta_{A_2} \\ \alpha_{A_1} \\ \alpha_{A_2} \\ \alpha_{A_3} \end{pmatrix} \\
\Rightarrow \ln\left(\frac{\boldsymbol{\pi}_S}{\boldsymbol{1} - \boldsymbol{\pi}_S}\right) &= \quad \boldsymbol{X}_1 \qquad\qquad \cdot \begin{pmatrix} \ln \lambda \\ \beta_{A_1} \\ \beta_{A_2} \\ \alpha_{A_1} \\ \alpha_{A_2} \\ \alpha_{A_3} \end{pmatrix}
\end{aligned}
$$

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

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

$\boldsymbol{C}$ 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 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

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

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

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

where

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

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

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

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

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

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

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

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

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

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

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

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} \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\[1em] \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\[1em] \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \end{pmatrix} + \beta_{A_1} \cdot \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\[1em] \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\[1em] -\mathbf{1} \\ -\mathbf{1} \\ -\mathbf{1} \\ -\mathbf{1} \end{pmatrix} + \beta_{A_2} \cdot \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\[1em] \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\[1em] -\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 \\[1em] \ln \boldsymbol{x}_1 \\ \ln \boldsymbol{x}_2 \\ \ln \boldsymbol{x}_3 \\ \ln \boldsymbol{x}_4 \\[1em] \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 \\[1em] \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 \\[1em] \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}$$

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

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

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

$\boldsymbol{C}$ 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 Weibull distributions for the three levels of the risk factor A.

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

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

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

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

$$= \ln\lambda \cdot \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \end{pmatrix} + \beta_{A_1} \cdot \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ -\mathbf{1} \\ -\mathbf{1} \\ -\mathbf{1} \\ -\mathbf{1} \end{pmatrix} + \beta_{A_2} \cdot \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ -\mathbf{1} \\ -\mathbf{1} \\ -\mathbf{1} \\ -\mathbf{1} \end{pmatrix} + \alpha_{A_1} \cdot \begin{pmatrix} \ln \boldsymbol{x}_1 \\ \ln \boldsymbol{x}_2 \\ \ln \boldsymbol{x}_3 \\ \ln \boldsymbol{x}_4 \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} + \alpha_{A_2} \cdot \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \ln \boldsymbol{x}_1 \\ \ln \boldsymbol{x}_2 \\ \ln \boldsymbol{x}_3 \\ \ln \boldsymbol{x}_4 \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} + \alpha_{A_3} \cdot \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \ln \boldsymbol{x}_1 \\ \ln \boldsymbol{x}_2 \\ \ln \boldsymbol{x}_3 \\ \ln \boldsymbol{x}_4 \end{pmatrix}$$

$$= \underbrace{\begin{pmatrix} \mathbf{1} & \mathbf{1} & \mathbf{0} & \ln \boldsymbol{x}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{1} & \mathbf{1} & \mathbf{0} & \ln \boldsymbol{x}_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{1} & \mathbf{1} & \mathbf{0} & \ln \boldsymbol{x}_3 & \mathbf{0} & \mathbf{0} \\ \mathbf{1} & \mathbf{1} & \mathbf{0} & \ln \boldsymbol{x}_4 & \mathbf{0} & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \ln \boldsymbol{x}_1 & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \ln \boldsymbol{x}_2 & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \ln \boldsymbol{x}_3 & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \ln \boldsymbol{x}_4 & \mathbf{0} \\ \mathbf{1} & -\mathbf{1} & -\mathbf{1} & \mathbf{0} & \mathbf{0} & \ln \boldsymbol{x}_1 \\ \mathbf{1} & -\mathbf{1} & -\mathbf{1} & \mathbf{0} & \mathbf{0} & \ln \boldsymbol{x}_2 \\ \mathbf{1} & -\mathbf{1} & -\mathbf{1} & \mathbf{0} & \mathbf{0} & \ln \boldsymbol{x}_3 \\ \mathbf{1} & -\mathbf{1} & -\mathbf{1} & \mathbf{0} & \mathbf{0} & \ln \boldsymbol{x}_4 \end{pmatrix}} \cdot \begin{pmatrix} \ln\lambda \\ \beta_{A_1} \\ \beta_{A_2} \\ \alpha_{A_1} \\ \alpha_{A_2} \\ \alpha_{A_3} \end{pmatrix}$$

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

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

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

$\boldsymbol{C}$ 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

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

with

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

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

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

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

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

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

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

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

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

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

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

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

or

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

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

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	Risk Factor Level									Total
Period	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

$$\boldsymbol{S}_{lm} = \left(\begin{array}{ccccccccccccccccccccccc}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0
\end{array}\right).$$

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

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

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

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

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 \\[1em] 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\[1em] 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 \\[1em] 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\[1em] -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 \\[1em] 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\[1em] -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 \\[1em] 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ -1 \\ -1 \\ -1 \\ -1 \\[1em] 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 \\[1em] 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \\ -1 \\ -1 \\ -1 \\ -1 \\[1em] 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 \\[1em] \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 \\[1em] \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}
$$

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

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

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

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

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

$$\ln\left(-\ln S(\boldsymbol{x})\right) = \begin{pmatrix}
1 & 1 & 0 & 1 & 0 & \ln \boldsymbol{x}_1 \\
1 & 1 & 0 & 1 & 0 & \ln \boldsymbol{x}_2 \\
1 & 1 & 0 & 1 & 0 & \ln \boldsymbol{x}_3 \\
1 & 1 & 0 & 1 & 0 & \ln \boldsymbol{x}_4 \\
1 & 1 & 0 & 0 & 1 & \ln \boldsymbol{x}_1 \\
1 & 1 & 0 & 0 & 1 & \ln \boldsymbol{x}_2 \\
1 & 1 & 0 & 0 & 1 & \ln \boldsymbol{x}_3 \\
1 & 1 & 0 & 0 & 1 & \ln \boldsymbol{x}_4 \\
1 & 1 & 0 & -1 & -1 & \ln \boldsymbol{x}_1 \\
1 & 1 & 0 & -1 & -1 & \ln \boldsymbol{x}_2 \\
1 & 1 & 0 & -1 & -1 & \ln \boldsymbol{x}_3 \\
1 & 1 & 0 & -1 & -1 & \ln \boldsymbol{x}_4 \\
 & & & & & \\
1 & 0 & 1 & 1 & 0 & \ln \boldsymbol{x}_1 \\
1 & 0 & 1 & 1 & 0 & \ln \boldsymbol{x}_2 \\
1 & 0 & 1 & 1 & 0 & \ln \boldsymbol{x}_3 \\
1 & 0 & 1 & 1 & 0 & \ln \boldsymbol{x}_4 \\
1 & 0 & 1 & 0 & 1 & \ln \boldsymbol{x}_1 \\
1 & 0 & 1 & 0 & 1 & \ln \boldsymbol{x}_2 \\
1 & 0 & 1 & 0 & 1 & \ln \boldsymbol{x}_3 \\
1 & 0 & 1 & 0 & 1 & \ln \boldsymbol{x}_4 \\
1 & 0 & 1 & -1 & -1 & \ln \boldsymbol{x}_1 \\
1 & 0 & 1 & -1 & -1 & \ln \boldsymbol{x}_2 \\
1 & 0 & 1 & -1 & -1 & \ln \boldsymbol{x}_3 \\
1 & 0 & 1 & -1 & -1 & \ln \boldsymbol{x}_4 \\
 & & & & & \\
1 & -1 & -1 & 1 & 0 & \ln \boldsymbol{x}_1 \\
1 & -1 & -1 & 1 & 0 & \ln \boldsymbol{x}_2 \\
1 & -1 & -1 & 1 & 0 & \ln \boldsymbol{x}_3 \\
1 & -1 & -1 & 1 & 0 & \ln \boldsymbol{x}_4 \\
1 & -1 & -1 & 0 & 1 & \ln \boldsymbol{x}_1 \\
1 & -1 & -1 & 0 & 1 & \ln \boldsymbol{x}_2 \\
1 & -1 & -1 & 0 & 1 & \ln \boldsymbol{x}_3 \\
1 & -1 & -1 & 0 & 1 & \ln \boldsymbol{x}_4 \\
1 & -1 & -1 & -1 & -1 & \ln \boldsymbol{x}_1 \\
1 & -1 & -1 & -1 & -1 & \ln \boldsymbol{x}_2 \\
1 & -1 & -1 & -1 & -1 & \ln \boldsymbol{x}_3 \\
1 & -1 & -1 & -1 & -1 & \ln \boldsymbol{x}_4
\end{pmatrix} \cdot \begin{pmatrix} \ln\lambda \\ \beta_{A_1} \\ \beta_{A_2} \\ \beta_{B_1} \\ \beta_{B_2} \\ \alpha \end{pmatrix}$$

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

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

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

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.