Appendix B

Theoretical aspects

B.1 Equivalence of logistic regression and the logistic survival model for current status data

B.1.1 Current status shelf life data

Each consumer is given one sample of a certain product stored during different times under conditions of interest. Consumers have to judge whether they would consume this sample regularly or not. The obtained data are as illustrated in Table B.1.

Table B.1: Current status shelf life data

Consumer	**Storage time**	**Rejection**[a]
1	4	1
2	24	1
3	8	0
4	12	0
⋮	⋮	⋮
n	36	1

[a] 1: rejection; 0: acceptance

B.1.2 Logistic Regression

Let Y be the random variable for the rejection of the product of interest:

$$Y = \begin{cases} 1 & \text{Rejection} \\ 0 & \text{Acceptance} \end{cases} .$$

The logistic regression models the probability that the product is rejected in dependence of one or more covariates. Our covariate of interest is the storage time T, that is, we have the following model, allowing for further covariates resumed in the vector $\boldsymbol{Z}$:

$$p(t, \boldsymbol{z}) = \mathrm{P}(Y = 1 | t, \boldsymbol{z}) = \frac{\exp(\alpha + \beta t + \boldsymbol{\gamma}' \boldsymbol{z})}{1 + \exp(\alpha + \beta t + \boldsymbol{\gamma}' \boldsymbol{z})} , \tag{B.1}$$

which is equivalent to

$$\ln \left(\frac{p(t, \boldsymbol{z})}{1 - p(t, \boldsymbol{z})} \right) = \alpha + \beta t + \boldsymbol{\gamma}' \boldsymbol{z}.$$

Given independent observations $(y_i, t_i, \boldsymbol{z}_i)$, $i = 1, \ldots, n$, the likelihood function for the unknown model parameters is the following:

$$\begin{aligned} L(\alpha, \beta, \boldsymbol{\gamma}) &= \prod_{i=1}^{n} p(t_i, \boldsymbol{z}_i)^{y_i} \big(1 - p(t_i, \boldsymbol{z}_i)\big)^{1-y_i} \\ &= \prod_{i=1}^{n} \left(\frac{\exp(\alpha + \beta t_i + \boldsymbol{\gamma}' \boldsymbol{z}_i)}{1 + \exp(\alpha + \beta t_i + \boldsymbol{\gamma}' \boldsymbol{z}_i)} \right)^{y_i} \left(\frac{1}{1 + \exp(\alpha + \beta t_i + \boldsymbol{\gamma}' \boldsymbol{z}_i)} \right)^{1-y_i} \\ &= \prod_{i=1}^{n} \frac{\exp(\alpha + \beta t_i + \boldsymbol{\gamma}' \boldsymbol{z}_i)^{y_i}}{1 + \exp(\alpha + \beta t_i + \boldsymbol{\gamma}' \boldsymbol{z}_i)} . \end{aligned} \tag{B.2}$$

Maximizing $L(\alpha, \beta, \boldsymbol{\gamma})$ furnishes the maximum likelihood estimators for the model parameters α, β and $\boldsymbol{\gamma}$.

B.1.3 The logistic survival model

The logistic survival model for the shelf life T has the following expression:

$$T = \mu + \boldsymbol{\kappa}' \boldsymbol{z} + \sigma W, \tag{B.3}$$

where $\boldsymbol{z}$ is the vector of covariates to be considered as in model (B.1). The error term distribution W is the standard logistic distribution with distribution function

$$F_W(w) = \frac{e^w}{1 + e^w} .$$

The (independent) observations are summarized by $(u_i, \delta_i, \boldsymbol{z}_i)$, $i = 1, \ldots, n$, being $\delta_i = \mathbf{1}_{\{t_i \le u_i\}}$ the indicator whether the observed time u_i has exceeded or not the shelf life t_i. That is, u_i corresponds to t_i in Section B.1.2, and δ_i to y_i. From model (B.3), we obtain

$$w = \tfrac{1}{\sigma}(t - \mu - \boldsymbol{\kappa}'\boldsymbol{z}),$$

and this leads to the following expression of the likelihood function:

$$\begin{aligned} L(\mu, \boldsymbol{\kappa}, \sigma) &= \prod_{i=1}^{n} F_W(w_i)^{\delta_i} \big(1 - F_W(w_i)\big)^{1-\delta_i} \\ &= \prod_{i=1}^{n} \Big(\frac{\exp\big(\frac{1}{\sigma}(u_i - \mu - \boldsymbol{\kappa}'\boldsymbol{z}_i)\big)}{1 + \exp\big(\frac{1}{\sigma}(u_i - \mu - \boldsymbol{\kappa}'\boldsymbol{z}_i)\big)} \Big)^{\delta_i} \Big(\frac{1}{1 + \exp\big(\frac{1}{\sigma}(u_i - \mu - \boldsymbol{\kappa}'\boldsymbol{z}_i)\big)} \Big)^{1-\delta_i} \\ &= \prod_{i=1}^{n} \frac{\big(\exp\big(\frac{1}{\sigma}(u_i - \mu - \boldsymbol{\kappa}'\boldsymbol{z}_i)\big)\big)^{\delta_i}}{1 + \exp\big(\frac{1}{\sigma}(u_i - \mu - \boldsymbol{\kappa}'\boldsymbol{z}_i)\big)} . \end{aligned} \tag{B.4}$$

We see that the likelihood functions (B.2) and (B.4) are equivalent for $y_i = \delta_i$, $t_i = u_i$, $\alpha = -\mu/\sigma$, $\beta = \frac{1}{\sigma}$, and $\boldsymbol{\gamma} = -\boldsymbol{\kappa}/\sigma$. Therefore, the estimates for the probabilities $p(t, \boldsymbol{z})$ of model (B.1) and $F_T(t) = F_W(\frac{1}{\sigma}(t - \mu - \boldsymbol{\kappa}'\boldsymbol{z}))$ of model (B.3) will amount to the same values.

B.2 Variance estimation of the relative risk

We apply the delta method to obtain a variance estimate of the estimator of the relative risk: $\widehat{\mathrm{RR}} = g(\hat\beta, \hat\sigma) = \exp\big(-\hat\beta/\hat\sigma\big)$. With

$$\begin{aligned} \frac{\partial g}{\partial \beta} g(\beta, \sigma) &= -\frac{e^{-\beta/\sigma}}{\sigma} , \\ \frac{\partial g}{\partial \sigma} g(\beta, \sigma) &= \frac{\beta}{\sigma^2} e^{-\beta/\sigma}, \end{aligned}$$

and denoting the variances of $\hat\beta$ and $\hat\sigma$ by $\sigma^2_{\hat\beta}$ and $\sigma^2_{\hat\sigma}$, as well as, the covariance of $\hat\beta$ and $\hat\sigma$ by $\sigma_{\hat\beta\hat\sigma}$, the formula for the variance of $\widehat{\mathrm{RR}}$ is the following:

$$\begin{aligned} \mathrm{Var}(\widehat{\mathrm{RR}}) &= \Big(-\frac{e^{-\beta/\sigma}}{\sigma} \Big)^2 \sigma^2_{\hat\beta} + \Big(\frac{\beta}{\sigma^2} e^{-\beta/\sigma} \Big)^2 \sigma^2_{\hat\sigma} + 2\Big(-\frac{e^{-\beta/\sigma}}{\sigma} \Big) \frac{\beta}{\sigma^2} e^{-\beta/\sigma} \sigma_{\hat\beta\hat\sigma} \\ &= \frac{\mathrm{RR}^2}{\sigma^2} \Big(\sigma^2_{\hat\beta} - 2\frac{\beta}{\sigma} \sigma_{\hat\beta\hat\sigma} + \frac{\beta^2}{\sigma^2} \sigma^2_{\hat\sigma} \Big). \end{aligned}$$

Thus, the variance of the estimated relative risk can be estimated by:

$$\widehat{\text{Var}}(\widehat{\text{RR}}) = \frac{\widehat{\text{RR}}^2}{\hat{\sigma}^2}\Big(\hat{\sigma}^2_{\hat{\beta}} - 2\frac{\hat{\beta}}{\hat{\sigma}}\hat{\sigma}_{\hat{\beta}\hat{\sigma}} + \frac{\hat{\beta}^2}{\hat{\sigma}^2}\hat{\sigma}^2_{\hat{\sigma}}\Big). \tag{B.5}$$

B.3 Maxima of the likelihood and our proposed reduced version

Consider the likelihood (3.15) on page 46 and our proposed reduced version:

$$L(\boldsymbol{\theta}, \boldsymbol{\omega}) = \prod_{i=1}^{n}\sum_{j=1}^{m}\alpha_{ij} f(u_i(s_j)|s_j)^{\delta_{1i}} S(u_i(s_j)|s_j)^{\delta_{2i}}\big(1 - S(u_i(s_j)|s_j)\big)^{(1-\delta_{1_i})(1-\delta_{2_i})}\omega_j,$$

$$L^*(\boldsymbol{\theta}, \boldsymbol{\omega}^*) = \prod_{i=1}^{n}\sum_{j=1}^{m^*}\alpha^*_{ij} f(u_i(s^*_j)|s^*_j)^{\delta_{1i}} S(u_i(s^*_j)|s^*_j)^{\delta_{2i}}\big(1 - S(u_i(s^*_j)|s^*_j)\big)^{(1-\delta_{1_i})(1-\delta_{2_i})}\omega^*_j,$$

where $\boldsymbol{\omega}^*$ contains all ω_j corresponding to the nonzero elements of the maximum likelihood estimator $\hat{\boldsymbol{\omega}}_n$, obtained by maximizing the likelihood (3.15) simultaneously with respect to $\boldsymbol{\theta}$ and $\boldsymbol{\omega}$. As in Section 3.5.3, we denote by m^* the dimension of $\boldsymbol{\omega}^*$, by s^*_j the values s_j of the support of Z with $\hat{\omega}_j > 0$, and define $\alpha^*_{ij} = \mathbb{1}_{\{s^*_j \in [Z_{l_i}, Z_{r_i}]\}}$, $i = 1, \ldots, n$, $j = 1, \ldots, m^*$. Also, let $\boldsymbol{\omega}^0$ be the subvector of $\boldsymbol{\omega}$ which contains the ω_j corresponding to the zero elements of $\hat{\boldsymbol{\omega}}_n$, and define m^0, s^0_j, and α^0_{ij} analogously to m^*, s^*_j, and α^*_{ij}.

Obviously, the following equality holds:

$$\sum_{j=1}^{m}\alpha_{ij}C(u_i(s_j)|s_j;\boldsymbol{\theta})\omega_j = \sum_{j=1}^{m^*}\alpha^*_{ij}C(u_i(s^*_j)|s^*_j;\boldsymbol{\theta})\omega^*_j + \sum_{j=1}^{m^0}\alpha^0_{ij}C(u_i(s^0_j)|s^0_j;\boldsymbol{\theta})\omega^0_j, \tag{B.6}$$

where $C(\cdot;\boldsymbol{\theta})$ can be replaced by $f(\cdot;\boldsymbol{\theta})$, $S(\cdot;\boldsymbol{\theta})$ or $1-S(\cdot;\boldsymbol{\theta})$. Consequently, the have the equality $L(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\omega}}_n) = L^*(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\omega}}^*_n)$, because $\hat{\omega}^0_{n_j} = 0$, $j = 1, \ldots, m^0$.

Now suppose, there exist $(\boldsymbol{\theta}', \boldsymbol{\omega}^{*\prime})$ such that $L^*(\boldsymbol{\theta}', \boldsymbol{\omega}^{*\prime}) > L^*(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\omega}}^*_n)$. This would imply:

$$\begin{aligned}
L^*(\boldsymbol{\theta}', \boldsymbol{\omega}^{*\prime}) &= \prod_{i=1}^{n} \Big(\sum_{j=1}^{m^*} \alpha^*_{ij} C_i(u_i(s^*_j)|s^*_j; \boldsymbol{\theta}') \omega^{*\prime}_j \Big) \\
&\overset{(B.6)}{=} \underbrace{\prod_{i=1}^{n} \Big[\Big(\sum_{j=1}^{m^*} \alpha^*_{ij} C_i(u_i(s^*_j)|s^*_j; \boldsymbol{\theta}') \omega^{*\prime}_j \Big) + \overbrace{\Big(\sum_{j=1}^{m^0} \alpha^0_{ij} C_i(u_i(s^0_j)|s^0_j; \boldsymbol{\theta}') \hat{\omega}^0_{n_j} \Big)}^{=0} \Big]}_{=L\big(\boldsymbol{\theta}', (\boldsymbol{\omega}^{*\prime}, \hat{\boldsymbol{\omega}}^0_n)\big)} \\
&> \prod_{i=1}^{n} \Big[\Big(\sum_{j=1}^{m^*} \alpha^*_{ij} C_i(u_i(s^*_j)|s^*_j; \hat{\boldsymbol{\theta}}_n) \hat{\omega}^*_{n_j} \Big) + \Big(\sum_{j=1}^{m^0} \alpha^0_{ij} C_i(u_i(s^0_j)|s^0_j; \hat{\boldsymbol{\theta}}_n) \hat{\omega}^0_{n_j} \Big) \Big] \\
&= L(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\omega}}_n) \\
&\implies L\big(\boldsymbol{\theta}', (\boldsymbol{\omega}^{*\prime}, \hat{\boldsymbol{\omega}}^0_n)\big) > L(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\omega}}_n) \quad \lightning
\end{aligned}$$

This contradiction proves, that maximizing the reduced version of the likelihood yields the nonzero elements of the joint maximum likelihood estimator $(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\omega}}_n)$.