8. Parametric models in survival analysis

§8.1. General accelerated failure time models for parametric regression

• The accelerated failure time model

Let $T$ be the time to event and $\mathbf{x}$ be a vector of covariates. The accelerated failure time model assumes

$$S(t|\mathbf{x}) = S_0(te^{\mathbf{\theta}^t\mathbf{x}}),$$

where $S(t|\mathbf{x})$ is the survival function of $T$ given the covariates $\mathbf{x}$ and $S_0(t)$ is the baseline survival function.

The survival function with covariate $\mathbf{x}$ can be considered as the survival function with covariate $\mathbf{x} = 0$ accelerated by a factor $e^{\mathbf{\theta}^t\mathbf{x}}$, hence the name of the model.
Let the failure time with the baseline survival function be represented as
\[ \epsilon = e^{\mu + \sigma W}, \]
where \( W \) follows some standard failure time distribution. Then \( T \) can be expressed in two equivalent forms:
\[ T = \exp(\gamma^t \mathbf{x})\epsilon, \]
and
\[ \ln T = \mu + \gamma^t \mathbf{x} + \sigma W. \]
Note: \( \gamma = -\theta \).
For convenience, \( \gamma \) is to be refereed to as the regression coefficients and \( \theta \) is to be referred to as the accelerating coefficients.

The usual distributions specified for \( \epsilon \) in the accelerated model are: \textit{Weibull, log-normal, log-logistic}. 

2
The corresponding distributions for $W$ are: *minimum extreme value, normal, logistic.*

- **Estimation and Inference**

Consider data with right censoring only:

$$\{(T_i, c_i) : i = 1, \cdots, n.\}$$

**Likelihood function:**

$$L(\gamma, \mu, \sigma) = \prod_{i=1}^{n} \left[ f(T_i e^{x_i' \gamma}, \mu, \sigma) e^{x_i' \gamma} c_i [S(T_i e^{-x_i' \gamma}, \mu, \sigma)]^{1-c_i} \right].$$

where $f$ and $S$ are the pdf and survival function of $\epsilon = e^{\mu + \sigma W}$ respectively.

The general theory of MLE applies; that is, the MLEs of the parameters follow an asymptotic multivariate normal distribution, and the variance matrix can be estimated by the inverse of the observed Fisher information matrix.
R functions for parametric regression models

The R function `survreg` fits accelerated failure models that use a log transformation.

```r
survreg.fit = survreg(formula, data=,
                      dist="weibull", ...)  
```

- `formula` is a formula expression for the transformed failure time \( \ln T \).
- `dist` is the assumed distribution for untransformed failure time \( T \). Allowed values include "weibull", "exponential", "gaussian", "logistic", "lognormal" and "loglogistic".

Various estimated quantities can be extracted from the fitted object including:
- Coefficients \( \mu \) and \( \gamma \): `survreg.fit$coef`
- Variance matrix: `survreg.fit$var`
- Parameters \( \mu, \log \sigma \) for baseline model: `survreg.fit$icoef`
- Various residuals: `residuals(survreg.fit, type=)`
• **Univariate analysis** (without covariates)

The function `survreg` can also be used to fit the parametric model without covariates by specifying the formula as

\[
\text{Surv}(\text{time, status}) \sim 1
\]

It returns the estimates and estimated variance matrix for \( \mu \) and \( \ln \sigma \). The parameters \( \mu \) and \( \sigma \) are as in \( \exp(\mu + \sigma W) \). These parameters are related to the conventional parameters of the assumed distribution by certain functional forms. The estimated conventional parameters and their variance can be obtained by the functional forms and the \( \delta \)-method.
§8.2. Accelerated failure time models with proportional hazard rates: Weibull models

- Review on Weibull distribution

Survival function:

\[ S_T(t) = \exp(-\lambda t^\alpha). \]

Hazard rate function:

\[ h_T(t) = \lambda \alpha t^{\alpha-1}. \]

Let \( Y = \ln T \). Then

\[ S_Y(y) = \exp(-\lambda e^{\alpha y}). \]

Redefine \( \lambda = \exp(-\mu/\sigma), \alpha = 1/\sigma \). Then

\[ S_Y(y) = \exp(-e^{\frac{y-\mu}{\sigma}}), \]

which is the survival function of an extreme-value distribution.
Then $T$ can be expressed as a log-linear model as

$$Y = \ln T = \mu + \sigma W,$$

where $W$ is the standard extreme value distribution with survival function

$$S_W(w) = \exp(-e^w).$$

Let $\hat{\mu}, \hat{\sigma}$ be the MLE of $\mu$ and $\sigma$. By the invariance of MLE, the MLE of $\lambda$ and $\alpha$ are obtained as

$$\hat{\lambda} = \exp(-\hat{\mu}/\hat{\sigma}), \quad \hat{\alpha} = 1/\hat{\sigma}.$$  

By the $\delta$ method, the variances and covariance of $\hat{\lambda}$ and $\hat{\alpha}$ are obtained as:

$$\text{Var}(\hat{\lambda}) = \exp(-2\hat{\mu}/\hat{\sigma})[\text{Var}(\hat{\mu})/\hat{\sigma}^2 + \hat{\mu}^2\text{Var}(\hat{\sigma})/\hat{\sigma}^4$$

$$-2\hat{\mu}\text{Cov}(\hat{\mu}, \hat{\sigma})/\hat{\sigma}^3],$$

$$\text{Var}(\hat{\alpha}) = \text{Var}(\hat{\sigma})/\hat{\sigma}^4,$$

$$\text{Cov}(\hat{\lambda}, \hat{\alpha}) = \exp(-\hat{\mu}/\hat{\sigma})[\text{Cov}(\hat{\mu}, \hat{\sigma})/\hat{\sigma}^3 - \hat{\mu}\text{Var}(\hat{\sigma})/\hat{\sigma}^4].$$
**Proportional hazard property of the Weibull accelerated failure time model**

Let 

\[ Y = \ln T = \mu + \gamma^t x + \sigma W. \]

Then the survival function of \( T \) given \( x \) is

\[
S(t|x) = S_0(te^{\theta^t x}) = \exp(-\lambda(te^{\theta^t x})^\alpha),
\]

and the hazard function is

\[
h(t|x) = \lambda \alpha (te^{\theta^t x})^{\alpha-1} e^{\theta^t x} = \lambda \alpha t^{\alpha-1} e^{-\gamma/\sigma} x^{\alpha-1} e^{\theta^t x} = h_0(t)e^{\beta^t x}.
\]

The parameters \( \beta = -\gamma/\sigma \) are referred to as the proportional hazard coefficients.
• Estimation

The R function \texttt{survreg} provides the estimates and the estimated variance and covariances for $\mu$, $\log \sigma$ and $\gamma$. 

The accelerating coefficients $\theta$ and the proportional coefficients $\beta$ can be estimated through the relationship:

$$\theta = -\gamma, \quad \beta = -\gamma/\sigma.$$

The variances and covariances of the estimates are obtained by $\delta$-method as:

$$\text{Cov}(\hat{\theta}_j, \hat{\theta}_k) = \text{Cov}(\hat{\gamma}_j, \hat{\gamma}_k), \quad j, k = 1, \ldots, p.$$ 

$$\text{Cov}(\hat{\beta}_j, \hat{\beta}_k) = \frac{\text{Cov}(\hat{\gamma}_j, \hat{\gamma}_k)}{\hat{\sigma}^2} - \frac{\hat{\gamma}_j \text{Cov}(\hat{\gamma}_j, \hat{\sigma})}{\hat{\sigma}^3} - \frac{\hat{\gamma}_k \text{Cov}(\hat{\gamma}_k, \hat{\sigma})}{\hat{\sigma}^3} + \frac{\hat{\gamma}_k \hat{\gamma}_j \text{Var}(\hat{\sigma})}{\hat{\sigma}^4},$$

$j, k = 1, \ldots, p.$
• **Example** (Laryngeal cancer patients, cont.)

Fit the Laryngeal cancer patient data using Weibull accelerated time model and obtain the estimated proportional hazard coefficients and their variance matrix.

```r
larynx=read.table("larynx.txt")
attach(larynx)

# V1 --- disease stages,
# V2 --- time to death or on study,
# V3 --- patients’ age
# V5 --- censoring indicator

x1=V1
x1[x1!=2]=0
x1[x1==2]=1
x2=V1
x2[x2!=3]=0
x2[x2==3]=1
x3=V1
x3[x3!=4]=0
x3[x3==4]=1
```

10
larynx.weibull = survreg(Surv(V2, V5) ~
    x1 + x2 + x3 + V3, data = larynx,
    dist = "weibull")
summary(larynx.weibull)

hat.sig = 0.885
coeff.gamm = larynx.weibull$coef
var.reg = larynx.weibull$var
m = dim(var.reg)
var.reg[m[1],] = var.reg[m[1],]*hat.sig
var.reg[, m[2]] = var.reg[, m[2]]*hat.sig

var.gamm = var.reg[-m[1], -m[2]]
cov.gamm.sig = var.reg[-m[1], m[2]]
var.sig = var.reg[m[1], m[2]]

coeff.beta = - coeff.gamm / hat.sig
var.beta.1 = var.gamm / hat.sig^2
tmp = coeff.gamm * cov.gamm.sig / hat.sig^3
var.beta.2 = tmp %*% t(rep(1, length(tmp)))
var.beta.3 = t(var.beta.2)
var.beta.4 = coeff.gamm %*% t(coeff.gamm) * var.sig / hat.sig^4
var.beta = var.beta.1 - var.beta.2 - var.beta.3 + var.beta.4
cbind(coef.beta, sqrt(diag(var.beta)))

§8.3. Accelerated failure time models with proportional odds ratio: Log Logistic models

• Review on Log Logistic distribution

Survival function:
\[ S_T(t) = \frac{1}{1 + \lambda t^\alpha}. \]
Hazard rate function:
\[ h_T(t) = \frac{\lambda \alpha t^{\alpha-1}}{1 + \lambda t^\alpha}. \]

Let \( Y = \ln T \). Then
\[ S_Y(y) = \frac{1}{1 + \lambda e^{\alpha y}}. \]
Redefine \( \lambda = \exp(-\mu/\sigma) \), \( \alpha = 1/\sigma \). Then
\[ S_Y(y) = \frac{1}{1 + e^{\frac{y-\mu}{\sigma}}}, \]
which is the survival function of an logistic
distribution.
Then $T$ can be expressed as a log-linear model as

$$Y = \ln T = \mu + \sigma W,$$

where $W$ is the standard logistic distribution with survival function

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

Note: The relationship between the parameters of the log-linear model and the original parameters is the same as that in Weibull models. Hence the estimates of the original parameters can be obtained in the same way as in the Weibull models.

- **Proportional odds ratio property of the Log logistic accelerated failure time model**
Let
\[ Y = \ln T = \mu + \gamma^t \mathbf{x} + \sigma W. \]

Then the survival function of \( T \) given \( \mathbf{x} \) is
\[
S(t|\mathbf{x}) = S_0( t e^{\theta^t \mathbf{x}} ) = \frac{1}{1 + \lambda t^\alpha e^{\theta^t \mathbf{x}}}
\]
and the odds of survival beyond time \( t \) is
\[
\frac{S(t|\mathbf{x})}{1 - S(t|\mathbf{x})} = \frac{1}{\lambda t^\alpha e^{\alpha \theta^t \mathbf{x}}} = \frac{1}{\lambda t^\alpha} e^{-\alpha \theta^t \mathbf{x}} = \frac{S(t|\mathbf{x} = 0)}{1 - S(t|\mathbf{x} = 0)} e^{-\beta^t \mathbf{x}}.
\]

The parameters \( \beta = -\gamma/\sigma \) are referred to as the proportional odds coefficients.

- **Example** (Laryngeal cancer patients, cont.)
Fit the Laryngeal cancer patient data using Log logistic accelerated time model and obtain the estimated proportional odds coefficients and their variance matrix.

```r
larynx.weibull = survreg(Surv(V2, V5) ~ x1 + x2 + x3 + V3, data = larynx, dist = "loglogistic")
summary(larynx.weibull)

hat.sig =
coef.gamm = larynx.weibull$coef
var.reg = larynx.weibull$var
m = dim(var.reg)
var.reg[m[1],] = var.reg[m[1],] * hat.sig
var.reg[, m[2]] = var.reg[, m[2]] * hat.sig

var.gamm = var.reg[-m[1], -m[2]]
cov.gamm.sig = var.reg[-m[1], m[2]]
var.sig = var.reg[m[1], m[2]]

coef.beta = - coef.gamm / hat.sig
var.beta.1 = var.gamm / hat.sig^2
tmp = coef.gamm * cov.gamm.sig / hat.sig^3
var.beta.2 = tmp%*%t(rep(1, length(tmp)))
```
\texttt{var.beta.3=t(var.beta.2)}
\texttt{var.beta.4 = coef.gamm%*%t(coef.gamm)}
\hspace{1em} \ast \texttt{var.sig/hat.sig^4}
\texttt{var.beta = var.beta.1-var.beta.2-var.beta.3}
\hspace{1em} + \texttt{var.beta.4}

\texttt{cbind(coef.beta, sqrt(diag(var.beta)))}

\textbf{§8.4. Diagnostic methods for parametric models}

- \textbf{Checking the distribution assumption of univariate data (without covariates)}

The basic idea is to transform the cumulative hazard function $H$ of the assumed distribution such that the transformed $H$, say $g(H(t))$, is a linear function of certain time scale, say $l(t)$, and then to plot $g(\hat{H}(t))$ against $l(t)$, where $\hat{H}$ is an non-parametric estimate of $H$. 

16
Transformed $H$ and time scale:

<table>
<thead>
<tr>
<th>Model</th>
<th>$H$</th>
<th>$g(H)$</th>
<th>$l(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exponential</td>
<td>$\lambda t$</td>
<td>$\lambda t$</td>
<td>$t$</td>
</tr>
<tr>
<td>Weibull</td>
<td>$\lambda t^\alpha$</td>
<td>$\ln H$</td>
<td>$\ln t$</td>
</tr>
<tr>
<td>Log-logistic</td>
<td>$\ln(1 + \lambda t^\alpha)$</td>
<td>$\ln[\exp(H) - 1]$</td>
<td>$\ln t$</td>
</tr>
</tbody>
</table>

**Procedure:**

1) Obtain the Nelson-Aalen estimator $\hat{H}(t)$.
2) Plot $g(\hat{H}(t))$ against $l(t)$.

Departure from linearity indicates inadequacy of the distribution assumption.

- **Example** (data on allo and auto transplants)

Checking on the distribution assumption of the allo and auto transplant data:

```r
alloauto=read.table("alloauto.txt")
allo = survfit(Surv(V1,V3)~1,data=alloauto,
               subset=(V2==1), type="fleming-harrington")
time.allo = allo$time
```
surv.allo = allo$surv

auto = survfit(Surv(V1,V3)~1,data=alloauto,
                subset=(V2==2), type="fleming-harrington")

time.auto = auto$time
surv.auto = auto$surv

# plot for checking Weibull assumption
par(mfrow=c(1,2))
plot(log(time.allo), log(-log(surv.allo)))
plot(log(time.auto), log(-log(surv.auto)))

# plot for checking Log-logistic assumption
par(mfrow=c(1,2))
plot(log(time.allo), log(exp(-log(surv.allo))-1))
plot(log(time.auto), log(exp(-log(surv.auto))-1))

• Residual plots for accelerated time regression models)

Cox-snell residuals plot:
Cox-snell residuals:

<table>
<thead>
<tr>
<th>Model</th>
<th>( r )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exponential</td>
<td>( \hat{\lambda} t \exp(\hat{\beta}^t \mathbf{x}) )</td>
</tr>
<tr>
<td>Weibull</td>
<td>( \hat{\lambda} t^{\hat{\alpha}} \exp(\hat{\beta}^t \mathbf{x}) )</td>
</tr>
<tr>
<td>Log-logistic</td>
<td>( \ln \left[ \frac{1}{1+\hat{\lambda} t^{\hat{\alpha}} \exp(\hat{\beta}^t \mathbf{x})} \right] )</td>
</tr>
</tbody>
</table>

The plot procedure is the same as in the Cox regression models:

i) Obtain the estimated cumulative hazard function of the Cox-Snell residuals. Denote the estimates by \( \hat{H}_r(r) \).

ii) Plot \( \hat{H}_r(r_j) \) against \( r_j \). Check whether the plot is linear through the origin with a slope 1. If yes, the fitted model is adequate. Any departure from this indicates a lack of fit.
Example (Laryngeal cancer patients, cont.)

Cox-Snell residual plot for Weibull model for the laryngeal cancer data.

```
larynx.weibull = survreg(Surv(V2, V5) ~ x1 + x2 + x3 + V3, data = larynx, dist = "weibull")

hat.sig = larynx.weibull$scale
hat.alpha = 1/hat.sig
reg.linear = larynx.weibull$linear.predictor
reg.linear.mdf = -reg.linear/hat.sig

tt = cbind(Surv(V2, V5))[,1]

cs.resid = exp(reg.linear.mdf)*tt^(hat.alpha)
cs.fit = survfit(Surv(cs.resid, V5) ~ 1, type = "fleming-harrington")
par(mfrow=c(1,1))

plot(cs.fit$time, -log(cs.fit$surv))
lines(c(0,3), c(0,3))
```
Deviance residuals plot

The plot of deviance residuals against either time, observation number or acceleration factor should look like random noise. Otherwise, it is an indication of some inadequacy of the model.

resi=residuals(larynx.weibull, type="deviance")
plot(reg.linear, resi)
plot(tt, resi)