6. Proportional Hazard Regression II

§6.1. Hazard regression model with time-dependent covariates

• Data:

\[(T_i, \delta_i, (X_i(t), 0 \leq t \leq T_i)) : i = 1, \ldots, n,\]

where \(X_i(t)\) is an array of values of the covariates including both time-independent and time-dependent covariates. For example, \(X_i(t)\) could be of the form:

\[x_{1i},\]
\[x_{2i}(t_1), \ldots, x_{2i}(T_i)\]
\[\ldots\]

• Partial likelihoods

\[t_1 < \cdots < t_D : \text{ordered event times.}\]
\[x_{(i)k}(t_i): \text{kth covariate observed at } t_i \text{ of the individual whose failure time is } t_i.\]
\( \mathcal{R}(t_i) \): risk set of individuals at time \( t_i \).

The partial likelihoods with time-dependent covariates have the same form as those with time-independent covariates.

For example, in the case of no ties,

\[
L(\beta) = \prod_{i=1}^{D} \frac{\exp\left\{\sum_{k=1}^{p} \beta_k x_{(i)k}(t_i)\right\}}{\sum_{j \in \mathcal{R}(t_i)} \exp\left\{\sum_{k=1}^{p} \beta_k x_{jk}(t_i)\right\}}.
\]

Note, an individual could be in the risk sets at different times, say, \( \mathcal{R}(t_i) \) and \( \mathcal{R}(t_l) \), the covariate values associated with this individual are generally different at different times.
• **R function coxph with time-dependent covariates**

The hazard regression model with time-dependent covariates can still be fitted using the function `coxph` with a special treatment on the data.

The observation for an individual is represented in multiple copies of records, each record consisting of `start.time`, `stop.time`, `status` and covariate values.

The response variable is generated by

\[
\text{Surv}(\text{start.time}, \text{stop.time}, \text{status})
\]
Data management

Case 1. Time-dependent covariate is continuous and can be observed or predicted at any time of events.

Original data:

<table>
<thead>
<tr>
<th>Subject</th>
<th>time</th>
<th>status</th>
<th>Z1</th>
<th>Z2</th>
<th>Z1</th>
<th>Z2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4.01(1)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>55</td>
<td>1</td>
<td>0</td>
<td>5.2(1)</td>
<td>5.45(55)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>60</td>
<td>0</td>
<td>1</td>
<td>3.8(1)</td>
<td>4(55)</td>
<td>4.1(60)</td>
</tr>
</tbody>
</table>

The representation of data appropriate for coxph:

<table>
<thead>
<tr>
<th>Subject</th>
<th>start.time</th>
<th>stop.time</th>
<th>status</th>
<th>Z1</th>
<th>Z2</th>
<th>Z1</th>
<th>Z2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4.01</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5.2</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>55</td>
<td>1</td>
<td>0</td>
<td>5.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>3.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>55</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>55</td>
<td>60</td>
<td>0</td>
<td>1</td>
<td>4.1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Case 2. Time-dependent covariate taking discrete values over time.

Original data:

<table>
<thead>
<tr>
<th>Subject</th>
<th>time</th>
<th>status</th>
<th>Z1</th>
<th>Z2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>55</td>
<td>1</td>
<td>0</td>
<td>0 (20)</td>
</tr>
<tr>
<td>3</td>
<td>60</td>
<td>0</td>
<td>1</td>
<td>0 (40)</td>
</tr>
</tbody>
</table>

where 0(t₀) means: Z2 = 0 for t < t₀; Z2 = 1 for t ≥ t₀.

The representation of data appropriate for \texttt{coxph}:

<table>
<thead>
<tr>
<th>Subject</th>
<th>start.time</th>
<th>stop.time</th>
<th>status</th>
<th>Z1</th>
<th>Z2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>55</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>40</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>60</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
The complete data set bmt.txt contains the following variables:

- **g**—Disease Group
  - 1—ALL
  - 2—AML Low Risk
  - 3—AML High Risk
- **T1**—Time To Death Or On Study Time
- **T2**—Disease Free Survival Time (Time To Relapse, Death Or End Of Study)
- **d1**—Death Indicator
  - 1—Dead 0—Alive.
- **d2**—Relapse Indicator
  - 1—Relapsed, O—Disease Free.
- **d3**—Disease Free Survival Indicator
  - 1—Dead Or Relapsed, O—Alive Disease Free.
- **TA**—Time To Acute Graft-Versus-Host Disease.
- **dA**—Acute GVHD Indicator.
  - 1—Developed Acute GVHD
  - 0—Never Developed Acute GVHD.
- **TC**—Time To Chronic Graft-Versus-Host Disease.
- **dC**—Chronic GVHD Indicator
  - 1—Developed Chronic GVHD
0—Never Developed Chronic GVHD.

TP—Time To Return of Platelets to Normal Levels
dP—Platelet Recovery Indicator
   1—Platelets Returned To Normal,
   0—Platelets Never Returned to Normal

Z1—Patient Age In Years
Z2—Donor Age In Years
Z3—Patient Sex
   1—Male, 0—Female
Z4—Donor Sex
   1—Male, 0—Female
Z5—Patient CMV Status
   1—CMV Positive, 0—CMV Negative
Z6—Donor CMV Status
   1—CMV Positive, 0—CMV Negative
Z7—Waiting Time to Transplant In Days
Z8—FAB
   1—FAB Grade 4 Or 5 and AML, 0—Otherwise
Z9—Hospital
   1—The Ohio State University,
   2—Alferd, 3—St. Vincent, 4—Hahnemann
Z10—MTX Used as a Graft-Versus-Host- Prophylactic
   1—Yes 0—No

The following three intermediate events are to be considered in the analysis:
(i) Development of acute graft-versus-host disease (aGVHD),
(ii) Development of chromic graft-versus-host disease (cGVHD),
(iii) Return of platelet count to normal level.

These three events are represented by the variables: \( TA, dA, TC, dC, TP, dP \). Alternatively, they can be represented by three time-dependent covariates as follows:

\[
x_A(t) = \begin{cases} 
0 & \text{if } t < TA, \\
1 & \text{otherwise.}
\end{cases}
\]

\[
x_P(t) = \begin{cases} 
0 & \text{if } t < TP, \\
1 & \text{otherwise.}
\end{cases}
\]

\[
x_C(t) = \begin{cases} 
0 & \text{if } t < TC, \\
1 & \text{otherwise.}
\end{cases}
\]
Types of individuals and the corresponding records:

There are 8 types of individuals as listed below:

<table>
<thead>
<tr>
<th>Type</th>
<th>aGVHD</th>
<th>cGVHD</th>
<th>pRecover</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>2</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>3</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>4</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>5</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>6</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>7</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>8</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>

For an individual of Type 1, there is only one record:

```
start stop status xA xC xP
0   T2   d3   0   0   0
```
For an individual of Type 2, 3 or 4, there are two records. Say, for a type 2 individual:

\[
\begin{array}{cccccc}
\text{start} & \text{stop} & \text{status} & xA & xC & xP \\
0 & TA & 0 & 0 & 0 & 0 \\
TA & T2 & d3 & 1 & 0 & 0 \\
\end{array}
\]

For an individual of Type 5, 6 or 7, there are three records. Say, for a type 5 individual, suppose TA < TC:

\[
\begin{array}{cccccc}
\text{start} & \text{stop} & \text{status} & xA & xC & xP \\
0 & TA & 0 & 0 & 0 & 0 \\
TA & TC & 0 & 1 & 0 & 0 \\
TC & T2 & d3 & 1 & 1 & 0 \\
\end{array}
\]
For an individual of Type 8, there are four records. Suppose \( TA < TC < TP \):

```
<table>
<thead>
<tr>
<th>start</th>
<th>stop</th>
<th>status</th>
<th>xA</th>
<th>xC</th>
<th>xP</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>TA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TA</td>
<td>TC</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TC</td>
<td>TP</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>TP</td>
<td>T2</td>
<td>d3</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
```

**R codes for data management:**

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

# Group indicator
    g1=bmt$g
    g1[g1!=2]=0
    g1[g1==2]=1

    g2=bmt$g
    g2[g2!=3]=0
    g2[g2==3]=1
```

11
detach(bmt)
bmt=data.frame(bmt,g1,g2)
attach(bmt)

# Data management on time-dependent covariate

bmt000=bmt [dA==0&dC==0&dP==0,]
bmt100=bmt [dA==1&dC==0&dP==0,]
bmt010=bmt [dA==0&dC==1&dP==0,]
bmt001=bmt [dA==0&dC==0&dP==1,]
bmt110=bmt [dA==1&dC==1&dP==0,]
bmt101=bmt [dA==1&dC==0&dP==1,]
bmt011=bmt [dA==0&dC==1&dP==1,]
bmt111=bmt [dA==1&dC==1&dP==1,]

n000=dim(bmt000)[1]
n100=dim(bmt100)[1]
n010=dim(bmt010)[1]
n001=dim(bmt001)[1]
n110=dim(bmt110)[1]
n101=dim(bmt101)[1]
n011=dim(bmt011)[1]
n111=dim(bmt111)[1]
# Management of Type 000
start = rep(0, n000)
stop = bmt000$T2
status = bmt000$d3
xP = xC = xA = rep(0, n000)
bmt000 = data.frame(bmt000, start, stop, status, xA, xC, xP)

# Management of Type 100
stt1 = rep(0, n100)
stt2 = stp1 = apply(cbind(bmt100$TA, bmt100$T2), 1, min)
stp2 = bmt100$T2
ss1 = rep(0, n100)
ss1[bmt100$T2 == stp1] = bmt100$d3[bmt100$T2 == stp1]
ss2 = bmt100$d3
xA1 = rep(0, n100)
xA2 = rep(1, n100)

start = c(stt1, stt2)
stop = c(stp1, stp2)
status = c(ss1, ss2)
xA = c(xA1, xA2)
xP = xC = rep(0, 2 * n100)

bmt100 = rbind(bmt100, bmt100)
bmt100 = data.frame(bmt100, start, stop, status, xA, xC, xP)
bmt100 = bmt100[start < stop,]
# Management of Type 110

\[stt1 = \text{rep}(0, n110)\]
\[stt2 = \text{stp1} = \text{apply}(\text{cbind}(\text{bmt110}$TA, \text{bmt110}$TC, \text{bmt110}$T2), 1, \text{min})\]
\[stt3 = \text{stp2} = \text{apply}(\text{cbind}(\text{apply}(\text{cbind}(\text{bmt110}$TA, \text{bmt110}$TC), 1, \text{max}), \text{bmt110}$T2), 1, \text{min})\]
\[stp3 = \text{bmt110}$T2\]

\[ss1 = \text{rep}(0, n110)\]
\[ss1[\text{bmt110}$T2 == \text{stp1}] = \text{bmt110}$d3[\text{bmt110}$T2 == \text{stp1}]\]
\[ss2 = \text{rep}(0, n110)\]
\[ss2[\text{bmt110}$T2 == \text{stp2}] = \text{bmt110}$d3[\text{bmt110}$T2 == \text{stp2}]\]
\[ss3 = \text{bmt110}$d3\]

\[xC2 = xA2 = xC1 = xA1 = \text{rep}(0, n110)\]
\[xA2[\text{bmt110}$TA < \text{bmt110}$TC] = 1\]
\[xC2[\text{bmt110}$TC < \text{bmt110}$TA] = 1\]
\[xC3 = xA3 = \text{rep}(1, n110)\]

\[\text{start} = c(stt1, stt2, stt3)\]
\[\text{stop} = c(stp1, stp2, stp3)\]
\[\text{status} = c(ss1, ss2, ss3)\]
\[xA = c(xA1, xA2, xA3)\]
\[xC = c(xC1, xC2, xC3)\]
\[xP = \text{rep}(0, 3*n110)\]
bmt110=rbind(bmt110,bmt110,bmt110)
bmt110=data.frame(bmt110,start,stop,status,xA,xC,xP)
bmt110=bmt110[start<stop,]

# Management of Type 111
stt1=rep(0,n111)
stt2=stp1=apply(cbind(bmt111$TA,bmt111$TC,bmt111$TP,
bmt111$T2),1,min)
stt3=stp2=apply(cbind(apply(cbind(bmt111$TA,bmt111$TC,
bmt111$TP),1,median),bmt111$T2),1,min)
stt4=stp3=apply(cbind(apply(cbind(bmt111$TA,bmt111$TC,
bmt111$TP),1,max),bmt111$T2),1,min)
stp4=bmt111$T2

ss3=ss2=ss1=rep(0,n111)
ss1[bmt111$T2==stp1]=bmt111$d3[bmt111$T2==stp1]
ss2[bmt111$T2==stp2]=bmt111$d3[bmt111$T2==stp2]
ss3[bmt111$T2==stp3]=bmt111$d3[bmt111$T2==stp3]
ss4=bmt111$d3

xA1=xC1=xP1=rep(0,n111)

xA2=xC2=xP2=rep(0,n111)
xA2[bmt111$TA==apply(cbind(bmt111$TA,bmt111$TC,
bmt111$TP),1,min)]=1
xC2[\texttt{bmt111$TC==apply(cbind(bmt111$TA,bmt111$TC, bmt111$TP),1,min)}]=1

xP2[\texttt{bmt111$TP==apply(cbind(bmt111$TA,bmt111$TC, bmt111$TP),1,min)}]=1

xA3=xC3=xP3=\texttt{rep(0,n111)}

xA3[\texttt{bmt111$TA==apply(cbind(bmt111$TA,bmt111$TC, bmt111$TP),1,median)}]=1

xA3[\texttt{xA2==1}]=1

xC3[\texttt{bmt111$TC==apply(cbind(bmt111$TA,bmt111$TC, bmt111$TP),1,median)}]=1

xC3[\texttt{xC2==1}]=1

xP3[\texttt{bmt111$TP==apply(cbind(bmt111$TA,bmt111$TC, bmt111$TP),1,median)}]=1

xP3[\texttt{xP2==1}]=1

xA4=xC4=xP4=\texttt{rep(1,n111)}

\texttt{start=c(stt1, stt2, stt3, stt4)}
\texttt{stop=c(stp1, stp2, stp3, stp4)}
\texttt{status=c(ss1, ss2, ss3, ss4)}
\texttt{xA=c(xA1, xA2, xA3, xA4)}
\texttt{xC=c(xC1, xC2, xC3, xC4)}
\texttt{xP=c(xP1, xP2, xP3, xP4)}
bmt111=rbind(bmt111,bmt111,bmt111,bmt111)
bmt111=data.frame(bmt111,start,stop,status,xA,xC,xP)
bmt111=bmt111[start<stop,]

# Combine all portions together
bmt.tdc=rbind(bmt000,bmt100,bmt010,bmt001,bmt110,
bmt101,bmt011,bmt111)

Fit the model with the additional time-dependent covariates:

```r
fit.tdc=coxph(Surv(start,stop,status)~g1+g2+Z1+Z2+Z8+xA+xC+xP,
data=bmt.tdc,method="breslow")
print(fit.tdc)
```

- **Testing of proportional hazard rate assumption by using time-dependent covariates**

To test whether the hazard rates with different values of a fixed-time covariate $Z_1$ are proportional, introduce an artificial covariate:

$$Z_2(t) = Z_1 \times g(t),$$
where $g(t)$ is a known function. Usually, $g(t)$ is taken as $\ln(t)$.

For different values of the fixed-time covariate, say $Z_1$ and $Z_1^*$,

$$
\frac{h(t|Z_1)}{h(t|Z_1^*)} = \exp\{\beta_1(Z_1-Z_1^*)+\beta_2g(t)(Z_1-Z_1^*)\}.
$$

Testing for the proportional hazard rate for $Z_1$ is equivalent to testing $\beta_2 = 0$.

- **Modeling non-proportional hazard rate through proportional hazard model with time-dependent covariates**

  **Step 1:** Fit a model with proportional hazard rates separately in two regions divided by a constant $\tau$. 

Introduce the artificial covariates as

\[ Z_2(t) = \begin{cases} 
Z_1 & \text{if } t < \tau, \\
0 & \text{if } t \geq \tau,
\end{cases} \]

\[ Z_3(t) = \begin{cases} 
Z_1 & \text{if } t \geq \tau, \\
0 & \text{if } t < \tau,
\end{cases} \]

where \( \tau \) is a constant.

The hazard becomes:

\[ h(t|Z(t)) = \begin{cases} 
h_0(t) \exp(\theta Z_1) & \text{if } t \leq \tau, \\
h_0(t) \exp[\theta Z_1] & \text{if } t > \tau.
\end{cases} \]

Equivalently,

\[ h(t|Z(t)) = h_0(t) \exp(\beta_1 Z_1 + \beta_2 Z_2 + \beta_3 Z_3). \]

The choice of \( \tau \):

Let \( t_i, i = 1, \ldots, D \) be the distinct event times. For each \( t_i \), fit the model with the artificial time-dependent covariates \( Z_2(t) \) and \( Z_3(t) \), and compute the likelihood of the model. Then
choose the $t_i$ corresspondiong to the largest likelihood as the value of $\tau$.

**Step 2:** Test proportional hazard rates in each region. If in any region, the proportional hazard rate is rejected, further divide the region and repeat the procedure in step 1. For example, to test proportional hazard in the region $t < \tau$, let

$$Z_2(t) = \begin{cases} 
Z_1 g(t) & \text{if } t < \tau, \\
0 & \text{if } t \geq \tau,
\end{cases}$$

$$Z_3(t) = \begin{cases} 
Z_1 & \text{if } t \geq \tau, \\
0 & \text{if } t < \tau,
\end{cases}$$

The hazard becomes:

$$h(t|Z(t)) = \begin{cases} 
 h_0(t) \exp[\beta_1 Z_1 + \beta_2 g(t) Z_2] & \text{if } t < \tau, \\
h_0(t) \exp(\beta_3 Z_1) & \text{if } t \geq \tau.
\end{cases}$$
Equivalently
\[ h(t | Z(t)) = h_0(t) \exp(\beta_1 Z_1 + \beta_2 Z_2 + \beta_3 Z_3). \]

Testing proportional hazard on \( t \leq \tau \) is equivalent to test \( \beta_2 = 0 \).

If the null hypothesis is rejected, for some \( \tau_1 \) in between 0 and \( \tau \), define
\[
Z_2(t) = \begin{cases} 
Z_1 & \text{if } t < \tau_1, \\
0 & \text{otherwise}, 
\end{cases}
\]
\[
Z_4(t) = \begin{cases} 
Z_1 & \text{if } \tau_1 \leq t \leq \tau, \\
0 & \text{otherwise}. 
\end{cases}
\]

Repeat Step 1 with \( \tau_1 \) playing the role of \( \tau \).

**Step 3:** Repeat step 2 until there is no region on which proportional hazard rate can be rejected.
§6.2. Stratified proportional hazard models

- Stratification

When the proportional hazard assumption is not valid for certain covariates, it is sometimes proper to stratify on those covariates so that within each stratum the proportional hazard assumption holds.

After stratification, each stratum will have an arbitrary baseline hazard rate. For example, for stratum $j$:

$$h_j(t|Z) = h_{j0}(t) \exp(\beta^t Z).$$

For individuals from different strata, say stratum $j$ and $k$:

$$\frac{h_j(t|Z_j)}{h_k(t|Z_k)} = \frac{h_{j0}(t)}{h_{k0}(t)} \exp[\beta(Z_j - Z_k)].$$
Partial Likelihood and Estimation

The covariates which are used to stratify the population are no longer considered as covariates in the linear predictor of the proportional hazard model.

It is assumed that other covariates have the same effects across over all strata. Let $\beta$ be the coefficient vector corresponding to these covariates in the linear predictor. The above assumption means that $\beta$ is the same for all strata.

Suppose the population is stratified into $K$ strata. The log partial likelihood function of $\beta$ is given by

$$L(\beta) = L_1(\beta) + \cdots + L_K(\beta),$$

where $L_j(\beta)$ is the log partial likelihood function based on data from stratum $j$ only.
The MLE of $\beta$ is obtained by maximizing $L(\beta)$.

Syntax of `coxph` for fitting a stratified proportional hazard model:

```r
coxph(Surv(...)~ ...+strata(strata.variable), ...)
```

- **A remark on the inference of $\beta$**

  Under the assumption that $\beta$ is the same for all strata, the global and local tests on $\beta$ are carried out by the same methods as in the un-stratified models such as the Likelihood Ratio, Wald and Score tests.

- **Test for homogeneity of $\beta$**

  In the discussion above, it is assumed that the coefficient vector $\beta$ is the same across all strata. In general, the coefficient vector might be different from stratum to stratum. Let $\beta_j$ denote the coefficient vector for stratum $j$. It
is of interest to test the hypothesis of homogeneity:

\[ H_0 : \beta_1 = \cdots = \beta_K. \]

**Likelihood Ratio test:**

\[ X_{LR} = 2 \sum_{j=1}^{K} [L_j(\hat{\beta}_j) - L_j(\hat{\beta})], \]

where \( \hat{\beta} \) is the overall MLE of \( \beta \), and \( \hat{\beta}_j \) is the MLE of \( \beta \) using the data in stratum \( j \) only.

**Wald test:**

Let

\[ C = \begin{pmatrix} 
I_p & -I_p & 0 & \cdots & 0 & 0 \\
0 & I_p & -I_p & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & I_p & -I_p 
\end{pmatrix}, \]

\[ B = (\beta_1^t, \cdots, \beta_K^t)^t. \]
Then
\[ CB = 0 \iff \beta_1 = \cdots = \beta_K. \]

Let
\[
\hat{B} = (\hat{\beta}_1^t, \ldots, \hat{\beta}_K^t)^t, \\
\hat{\Sigma} = \text{Diag}(\hat{\Sigma}_1, \ldots, \hat{\Sigma}_s),
\]
where $\hat{\Sigma}_j$ is the estimated variance matrix of $\hat{\beta}_j$. The Wald statistic is given by
\[
X^2_W = [C \hat{B}]^t[C \hat{\Sigma} C^t]^{-1}[C \hat{B}].
\]

Under the hypothesis of homogeneity of $\beta$, both $X_{LR}$ and $X^2_W$ have an asymptotic $\chi^2$-distribution with d.f. $(K - 1)p$.

**Remark:** Equality of components of $\beta$ in different strata can be tested similarly.
Example: Bone Marrow Transplant (cont.)

Fit of model stratified on MTX ($Z_{10}$):

```r
fit.stra=coxph(Surv(start,stop,status)~g1+g2++Z1+Z2+Z1Z2+Z8+xP+strata(Z10),
data=bmt.tdc,method="breslow")
print(fit.stra)
```

Test for homogeneity of $\beta$

Likelihood Ratio test:

```r
L=fit.stra$loglik
stra1=coxph(Surv(start,stop,status)~g1+g2++Z1+Z2+Z1Z2+Z8+xP,data=bmt.tdc,
subset=(Z10==0),method="breslow")
L1=stra1$loglik

stra2=coxph(Surv(start,stop,status)~g1+g2++Z1+Z2+Z1Z2+Z8+xP,data=bmt.tdc,
subset=(Z10==1),method="breslow")
L2=stra2$loglik
```
XLR=2*(L1^2+L2^2-L)^2
p.value=1-pchisq(XLR,7)
c(XLR, p.value)

Wald Test:

A1=stra1$var
A2=stra2$var
b1=stra1$coef
b2=stra2$coef

b=c(b1,b2)
C=cbind(diag(rep(1,7)), -diag(rep(1,7)))
A=rbind(cbind(A1, diag(rep(0,7))),
        cbind(diag(rep(0,7)), A2))

XW=t(b)*t(C)*solve(C*A*t(C))*C*b
p.value=1-pchisq(XW,7)
c(XW, p.value)
§6.3. Proportional hazard regression with left truncation data

- Partial likelihood

Data:

\[(V_i, T_i, \delta_i, (X_i(t), 0 \leq t \leq T_i)) : i = 1, \ldots, n,\]

where \(V_i\) is the left truncation time of individual \(i\).

The data has the same structure as in the case of existence of time-dependent covariates except there is an additional observation on the left truncation.

The partial likelihood:

\[
L(\beta) = \prod_{i=1}^{D} \frac{\exp\{\sum_{k=1}^{p} \beta_k x(i)_k(t_i)\}}{\sum_{j \in \mathcal{R}(t_i)} \exp\{\sum_{k=1}^{p} \beta_k x(j)_k(t_i)\}},
\]

where the risk set \(\mathcal{R}(t_i)\) at time \(t_i\) is modified
as
\[ \mathcal{R}(t_i) = \{ j \mid V_j < t_i < T_j \}. \]

- **coxph for left truncated data**

  \( V \): vector of left truncation times,

  \( T \): vector of event times or times on study,

  \( d \): vector of indicators of censoring,

  \( X = (X_1, \ldots, X_p) \): vector of covariates.

  The function `coxph` is used in the form

  \[
  \text{coxph(Surv}(V,T,d) \sim X_1+\ldots+X_p,\ldots)\]

- **Examples**

  **Channing House data:**

  The data set contains the following variables:

  - **entry**: age of entry into retirement home, months.
**time**: age of death or left retirement home, months.

**time.diff**: difference between time and entry.

**gender**: 1 male, 2 female.

**status**: 1 dead, 0 alive.

The death time of an elderly is left truncated at his or her entry age.

```r
canning=read.table("channing.txt")
x=canning$gender
x[x==2]=0
canning1=data.frame(canning,x)
canning1=canning1[canning1$time.diff>0,]
canning.fit=coxph(Surv(entry,time,status)~x,data=canning1)
print(canning.fit)

H_0: no difference between male and female w.r.t. survival.
Likelihood Ratio test: p-value = 0.0749.
```
Bone Marrow Transplant (cont.):
The study is interested in the disease-free survival of patients who have had their platelets returned to normal level. Only the patients who have had their platelets returned to normal can be sampled for this study. The disease-free time is thus truncated at the time the platelets returned to normal level.

```r
bmtTP=bmt[bmt$dP==1,]
bmtTP.fit=coxph(Surv(TP,T2,d3)~g1+g2++Z1+Z2+
  Z1Z2+Z8+strata(Z10),
  data=bmtTP,method="efron")
print(bmtTP.fit)
```