Cox Proportional Hazards Model

- Incorporate the effects of covariates
- Parametric survival distributions are not specified
- Semi-parametric models
- Proportional hazards assumption
  \[ h(t; x, \beta) = h_0(t)g(x, \beta) \]
  for some unspecified baseline hazard \( h_0(t) \)
- Hazard ratio:
  \[ \frac{h(t; x_1, \beta)}{h(t; x_0, \beta)} = \frac{g(x_1; \beta)}{g(x_0; \beta)} \]

Since the hazard function is strictly positive, consider the log link,
\[
\log(g(x; \beta)) = x^T \beta = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k
\]
\[
\Rightarrow g(x; \beta) = \exp(x^T \beta)
\]
Then, the hazard ratio is
\[
\exp((x_1^T - x_0^T) \beta)
\]

We have just defined the Cox Proportional Hazards Model
\[ h(t; x, \beta) = h_0(t)\exp(x^T \beta) \]
and
\[ S(t; x, \beta) = [S_0(t)]\exp(x^T \beta) \]
(Cox, 1972 *JRSS B*, 74, 187-220)

Estimation is difficult since \( h_0(t) \) is an infinite dimensional nuisance parameter. What can we do?

The proportional hazards assumption is a rather strong assumption that may not always be reasonable

- Suppose there is a single covariate
  \[ x = \begin{cases} 
  0 & \text{non-surgical intervention} \\
  1 & \text{surgical intervention} 
\end{cases} \]
- Surgical intervention has high early risk, but very good long-term survival prospects
- Non-surgical intervention as low early risk, but poor long-term survival prospects
- The hazard ratio would not be constant; it would initially be high and then decrease over time.

Estimation in Cox Model

- Instead of the full likelihood, use the partial likelihood proposed by Cox (1972, *JRSS B* and 1975, *Biometrika*)
- Consider the conditional probability that an individual dies at time \( t(j) \) given that \( t(j) \) is one of the \( r \) observed death times \( \{t(1), t(2), \ldots, t(r)\} \).
\[
P(\text{individual with values } x(j) \text{ dies at } t(j)) \frac{P(\text{one death at } t(j))}{P(\text{one death at } t(j))}
\]
Since the baseline hazard has an arbitrary form, intervals between successive death times provide no information about the effects of covariates on the hazard function.
• Individuals are assumed to respond independently

\[ P(\text{individual with values } x_{(j)} \text{ dies at } t_{(j)}) = \prod_{k \in R(t_{(j)})} P(\text{individual } k \text{ dies at } t_{(j)}) \]

• Replace probability of death at time \( t_{(j)} \) with probability of death in the interval \([t_{(j)}, t_{(j)} + \Delta] \).

\[ P(\text{individual with } x_{(j)} \text{ dies in } [t_{(j)}, t_{(j)} + \Delta]) = \frac{P(\text{individual } k \text{ dies in } [t_{(j)}, t_{(j)} + \Delta])}{\Delta} \]

• Limit as \( \Delta \to 0 \)

hazard at \( t_{(j)} \) for individual with \( x_{(j)} \)

\( \sum_{k \in R(t_{(j)})} P(\text{individual } k \text{ dies in } [t_{(j)}, t_{(j)} + \Delta]) \)

hazard at \( t_{(j)} \) for individual \( k \)

---

Alternative explanation:

• By definition \( h_i(t) = \frac{f_i(t)}{S_i(t)} \)

• The contribution to the likelihood for an observed failure at time \( t \) is

\[ f_i(t) = h_i(t)S_i(t) = h_0(t)exp(x_i^T \beta)[S_0(t)]^{-exp(x_i^T \beta)} \]

• The contribution to the likelihood for a right censored observation at time \( t \) is

\[ S_i(t) = [S_0(t)]^{-exp(x_i^T \beta)} \]

• Let \( \delta_i = 1 \) if \( t_i \) is a failure time and \( \delta_i = 0 \) if \( t_i \) is a censoring time.

---

Contribution to the likelihood is

\[ h_j(t_{(j)}) = \frac{h_0(t_{(j)}) exp(x_j^T \beta)}{\sum_{k \in R(t_{(j)})} h_k(t_{(j)})} \]

• Cancel out the baseline hazard function to obtain

\[ \exp(x_j^T \beta) \]

The joint partial likelihood is

\[ L_p(\beta) = \prod_{j=1}^{n} \left[ \frac{\exp(x_j^T \beta)}{\sum_{k \in R(t_{(j)})} \exp(x_k^T \beta)} \right]^{\delta_j} \]

where \( \delta_j = 0 \) if \( t_j \) is a censoring time, 1 otherwise.

---

Then, the joint likelihood is

\[ \prod_{i=1}^{n} [f_i(t_{(i)})]^{\delta_i}[S_i(t_{(i)})]^{1-\delta_i} \]

\[ = \prod_{i=1}^{n} [h_i(t_{(i)})S_i(t_{(i)})]^{\delta_i}[S_i(t_{(i)})]^{1-\delta_i} \]

\[ = \prod_{i=1}^{n} h_i(t_{(i)})^{\delta_i}S_i(t_{(i)}) \]

\[ = \prod_{i=1}^{n} \left[ \frac{h_i(t_{(i)})}{\sum_{k \in R(t_{(i)})} h_k(t_{(i)})} \right]^{\delta_i}S_i(t_{(i)}) \]

\[ = \prod_{i=1}^{n} \left[ \frac{\exp(x_i^T \beta)}{\sum_{k \in R(t_{(i)})} \exp(x_k^T \beta)} \right]^{\delta_i}S_i(t_{(i)}) \]

• The partial likelihood is

\[ L(\beta) = \prod_{i=1}^{n} \left[ \frac{\exp(x_i^T \beta)}{\sum_{k \in R(t_{(i)})} \exp(x_k^T \beta)} \right]^{\delta_i} \]
More Details on Estimation in Cox Proportional Hazards Model

- Assuming no ties, the log-partial likelihood is
  \[ \log(L_p(\beta)) = \sum_{i=1}^{n} \delta_i [x_i^T \beta - \log(\sum_{j \in R(t_i)} \exp(x_j^T \beta))] \]

- The partial likelihood depends only on the ordering of the survival times, not the actual values; so it is invariant to monotone transformation of time

- The gradient vector (score function) has elements
  \[ U(\beta) = \left[ \frac{\partial \log(L_p(\beta))}{\partial \beta_k} \right] = \left[ \sum_{i=1}^{n} \delta_i (x_{k,i} - \bar{x}_{k,i}) \right] \]
  where \( \bar{x}_{k,i} = \sum_{j \in R(t_i)} w_{ij} x_{k,j} \)
  and \( w_{ij} = \exp(x_j^T \beta) / \sum_{\ell \in R(t_i)} \exp(x_{\ell}^T \beta) \)

The local information matrix \( I(\beta) \) has elements

\[ I_{k\ell} = -\frac{\partial^2 \log(L_p(\beta))}{\partial \beta_k \partial \beta_\ell} = \sum_{i=1}^{n} \delta_i \sum_{j \in R(t_i)} w_{ij} (x_{k,j} - \bar{x}_{k,i})(x_{\ell,j} - \bar{x}_{\ell,i}) \]

- Large sample distribution
  \( \hat{\beta} \sim N(\beta, I(\beta)^{-1}) \)

- Hypothesis testing
  - Partial likelihood ratio tests
  - Score tests
  - Wald test

Tied Survival Times

Due to the inability to continuously monitor subjects.

- Suppose 5 individuals are at risk and two fail at time \( t \) with risk scores \( r_1 = \exp(x_1^T \beta) \) and \( r_2 = \exp(x_2^T \beta) \)

- We cannot order these two subjects with respect their actual failure times.

- The Breslow approximation: The contribution to the partial likelihood is
  \[ \left( \frac{r_1}{r_1 + r_2} \right) \times \left( \frac{r_1^2}{r_1^2 + r_2^2} \right) \]
  - This is a poor approximation when \( \beta \) is not close to zero
  - \( \hat{\beta} \) is biased toward zero (failed individuals are used too often in the denominator)
  - Default in SAS
• **Efron approximation:** The contribution to the partial likelihood is 
\[
\left( \frac{r_1}{r_1 + r_2 + r_3 + r_4 + r_5} \right) \times \left( \frac{r_2}{r_2 + r_3 + r_4 + r_5} \right)
\]
  - If subjects with tied death times have identical risk scores, this method is exact
  - Biased toward zero for “larger” values of \( \beta \), (e.g., \(|\beta_i| > 2.5\))
  - Smaller bias than the Breslow estimator
  - Default in S-Plus and R

Contribution if three subjects fail at the same time
\[
\left( \frac{r_1}{r_1 + r_2 + r_3 + r_4 + r_5} \right) \times \left( \frac{r_2}{r_2 + r_3 + r_4 + r_5} \right)
\times \left( \frac{r_3}{r_3 + r_4 + r_5} \right)
\]

312

• **Average likelihood method:**
  - “exact” partial likelihood
  - Computationally intensive (TIES=exact option in the MODEL statement in the SAS PHREG procedure)
  - Least amount of bias
  - Contribution to the likelihood for a pair of tied failure times is
  \[
  \frac{1}{2} \left[ \left( \frac{r_1}{r_1 + r_2 + r_3 + r_4 + r_5} \right) \left( \frac{r_2}{r_2 + r_3 + r_4 + r_5} \right) + \left( \frac{r_2}{r_1 + r_2 + r_3 + r_4 + r_5} \right) \left( \frac{r_1}{r_1 + r_3 + r_4 + r_5} \right) \right]
  \]

313

• If there are three individuals with tied failure times, the contribution to the partial likelihood is
\[
\frac{1}{6} \left[ \left( \frac{r_1}{r_1 + r_2 + r_3 + r_4 + r_5} \right) \left( \frac{r_2}{r_2 + r_3 + r_4 + r_5} \right) \left( \frac{r_3}{r_3 + r_4 + r_5} \right) + \left( \frac{r_1}{r_1 + r_2 + r_3 + r_4 + r_5} \right) \left( \frac{r_1}{r_1 + r_3 + r_4 + r_5} \right) \left( \frac{r_3}{r_3 + r_4 + r_5} \right) + \left( \frac{r_2}{r_1 + r_2 + r_3 + r_4 + r_5} \right) \left( \frac{r_3}{r_1 + r_3 + r_4 + r_5} \right) \left( \frac{r_1}{r_1 + r_4 + r_5} \right) + \left( \frac{r_2}{r_1 + r_2 + r_3 + r_4 + r_5} \right) \left( \frac{r_1}{r_1 + r_3 + r_4 + r_5} \right) \left( \frac{r_2}{r_2 + r_4 + r_5} \right) + \left( \frac{r_3}{r_1 + r_2 + r_3 + r_4 + r_5} \right) \left( \frac{r_1}{r_1 + r_3 + r_4 + r_5} \right) \left( \frac{r_2}{r_2 + r_4 + r_5} \right) + \left( \frac{r_3}{r_1 + r_2 + r_3 + r_4 + r_5} \right) \left( \frac{r_2}{r_1 + r_2 + r_4 + r_5} \right) \left( \frac{r_1}{r_1 + r_4 + r_5} \right) \right]
\]

314

• The sum of the \( k! \) terms can be expressed as integral (use integration by parts)
\[
\int_0^\infty \prod_{j=1}^k \left[ 1 - \exp \left( \frac{r_j t}{r_{k+1} + r_{k+2} + \cdots + r_n} \right) \right] e^{-t} dt
\]
(Delong, et al., 1994, *Biometrika*, 81, 607-611)

• Alternatively, you could use a sum of a random sample of the \( k! \) terms.
Interpretation in the Cox Model

- Coefficients are log hazard ratios. For a continuous covariate
  \[ e^{\beta_j} = \frac{\exp(\beta_j(x_j + 1))}{\exp(\beta_j x_j)} \]
  for any value of \( x_j \)

- Also,
  \[ e^{r \beta_j} = \frac{\exp(\beta_j(x_j + r))}{\exp(\beta_j x_j)} \]
  for any value of \( x_j \)

- \( r \beta_j \) is a log-hazard ratio associated with an \( r \) unit increase in \( X_j \), while the other covariates are held constant.

- \( r \beta_j \) is estimated as \( r \hat{\beta}_j \) with standard error \( r \hat{S}_{\beta_j} \), where \( \hat{S}_{\beta_j} \) is the \((j, j)\) element of the inverse of the local estimate of the information matrix for the partial likelihood

- To construct confidence intervals for hazard ratios, exponentiate endpoints of confidence intervals for log hazard ratios
  Compute \( \hat{L} = \hat{\beta}_j - z_{\alpha/2} \hat{S}_{\beta_j} \)
  and \( \hat{U} = \hat{\beta}_j + z_{\alpha/2} \hat{S}_{\beta_j} \)
  Then, compute \((\exp(\hat{L}), \exp(\hat{U}))\) as an approximate \((1 - \alpha) \times 100\%\) confidence interval for \( \exp(\beta) \)

- Estimate risk scores (use the XBETA option in the OUTPUT statement for the PHREG procedure in SAS)
  \[ \hat{r}(x, \hat{\beta}) = \sum \hat{\beta}_k x_k = x^T \hat{\beta} \]

- Estimate of the covariate adjusted survivor function
  \[ \hat{S}(t; x, \hat{\beta}) = [\hat{S}_0(t)]^{\exp(x^T \hat{\beta})} \]

- If the x’s are centered, the baseline survivor function \( S_0(t) \) corresponds to an ‘average’ subject

- We need an estimate of the baseline survivor function \( S_0(t) \)
Estimating the hazard and survivor functions

The Kalbfeisch-Prentice (1973, *Biometrika* 60, 267-278) method:

The probability that a subject with covariates values $x_i$ survives beyond time $t$ is

$$S(t; x_i) = (S_0(t))^{\exp(x_i^T \beta)}$$

For observed failure times $t(1) < t(2) < \ldots < t(r)$ suppose there are $d_j$ deaths and $n_j$ individuals at risk at time $t(j)$.

An approximate joint likelihood is

$$\prod_{j=0}^{r} \left[ \prod_{k \in D_j} \left( 1 - \eta_j \exp(x_i^T \beta) \right) \prod_{k \in R_j - D_j} \eta_j \exp(x_i^T \beta) \right]$$

Estimate $\eta_j$ by maximizing the log of this approximate likelihood. Setting first partial derivatives equal to zero we obtain $\hat{\eta}_j$ as the solution to

$$\sum_{k \in D(t(j))} \frac{\exp(x_i^T \hat{\beta})}{1 - \hat{\eta}_j \exp(x_i^T \beta)} = \sum_{k \in R(t(j))} \exp(x_i^T \hat{\beta})$$

If there are no ties ($d_j = 1$ for all $j=1,2,\ldots,r$), then

$$\hat{\eta}_j = 1 - \frac{\exp(x_i^T \hat{\beta})}{\sum_{k \in R(t(j))} \exp(x_i^T \beta)} \exp(-x_i^T \hat{\beta})$$

Otherwise, an iterative procedure is required to obtain a solution

Let $\eta_j$ denote the conditional probability of failing in $[t(j), t(j+1))$ given that a subject from the baseline population survives to time $t(j)$. Defining $\eta_0 = 1$,

$$S_0(t(j)) = \prod_{i=0}^{j-1} \eta_j$$

and the conditional probability that a subject with covariate values $x_j$ fails in $[t(j), t(j+1))$ given survival to the beginning of the interval is approximately

$$\frac{S_0(t(j)) \exp(x_i^T \beta) - S_0(t(j+1)) \exp(x_i^T \beta)}{S_0(t(j)) \exp(x_i^T \beta)} = 1 - \exp(x_i^T \beta)$$

The conditional probability that a subject with covariate values $x_j$ survives beyond $t(j)$ given survival to the beginning of the interval is approximately

$$\frac{S_0(t(j+1)) \exp(x_i^T \beta)}{S_0(t(j)) \exp(x_i^T \beta)} = \eta_j$$

- $S_0(t)$ is estimated as

$$\hat{S}_0(t) = \prod_{j=1}^{r} \hat{\eta}_j$$

$$= \left[ 1 - \frac{\exp(x_i^T \hat{\beta})}{\sum_{k \in R(t(j))} \exp(x_i^T \beta)} \right] \exp(-x_i^T \hat{\beta})$$

for $t(k) \leq t < t(k+1)$, $k = 1, 2, \ldots, r - 1$ with $\hat{S}_0(t) = 1$ for $t < t(1)$ and $\hat{S}_0(t) = 0$ for $t > t(r)$ unless there are censored times after $t(r)$. In that case, $\hat{S}_0(t) = \hat{S}_0(t(r))$ up to the largest censored time and it is undefined after that time.

- Assume that the baseline hazard is constant between adjacent failure times. Then,

$$\hat{h}_0(t) = \frac{1 - \hat{\eta}_j}{t_{j+1} - t_j}$$

for $t(j) \leq t < t(j+1)$ and $j=1,2,\ldots,r-1$ with $\hat{h}_0(t) = 0$ for $t < t(1)$.
When there are no covariates,
\[ \hat{\eta}_j = \frac{n_j - d_j}{n_j} \]
and
\[ \hat{S}_0(t) = \prod_{j=1}^k \hat{\eta}_j \]
is the Kaplan-Meier estimator.

The baseline cumulative hazard function is estimated as
\[ \hat{H}_0(t) = -\log(\hat{S}_0(t)) = -\sum_{j=1}^k \log(\hat{\eta}_j) \]

For others
\[ \hat{H}_i(t) = \hat{H}_0(t) \exp(x_i^T \hat{\beta}) \]
and
\[ \hat{S}_i(t) = \hat{S}_0(t)^{\exp(x_i^T \hat{\beta})} \]

Breslow (Nelson-Aalen) method:

- Approximate
\[ \hat{\eta}_j^{\exp(x_i^T \hat{\beta})} = \exp\left(\sum_{i \in R(t_j)} \exp(x_i^T \hat{\beta})\right) \]
with \(1 + \exp(x_i^T \hat{\beta})\log(\hat{\eta}_j)\) to obtain
\[ \hat{\eta} = \exp\left(\frac{-d_j}{\sum_{i \in R(t_j)} \exp(x_i^T \hat{\beta})}\right) \]

- Then,
\[ \hat{S}_0(t) = \prod_{j=1}^k \exp\left(\frac{-d_j}{\sum_{i \in R(t_j)} \exp(x_i^T \hat{\beta})}\right) \]
for \(t(k) \leq t < t(k+1), k=1,2,\ldots,r-1\).

- This estimator is not necessarily zero at the longest survival time, when it is uncensored.

Also,
\[ \hat{H}(t) = -\log(\hat{S}_0(t)) = -\sum_{j=1}^k \frac{d_j}{\sum_{i \in R(t_j)} \exp(x_i^T \hat{\beta})} \]

\(\hat{S}_0(t)\) is cheaper to compute than \(\hat{S}_0(t)\).

When there are few ties, \(\hat{S}_0(t)\) will be similar to \(\hat{S}_0(t)\).

Estimates of medians and other percentiles can be obtained from the estimated survival function.

SAS Example: VA data

/* SAS code for fitting a proportional hazards model to data from the VA lung cancer trial of 137 male patients with inoperable lung cancer. This code is posted as vapreg1.sas */

/* Variables */
Treatment: 1=standard, 2=test (chemotherapy)
Celltype: 1=squamous, 2=smallcell, 3=adenocarcinoma, 4=large
Survival in days
Status: 1=dead, 0=censored

<table>
<thead>
<tr>
<th>SCORE FUNCTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 Normal, no evidence of disease</td>
</tr>
<tr>
<td>90 Able to perform normal activity with only minor symptoms</td>
</tr>
<tr>
<td>80 Able to perform normal activity with effort, some symptoms</td>
</tr>
<tr>
<td>70 Able to care for self but unable to do normal activities</td>
</tr>
<tr>
<td>60 Requires occasional assistance</td>
</tr>
<tr>
<td>50 Requires considerable assistance</td>
</tr>
<tr>
<td>40 Disabled, requires special assistance</td>
</tr>
<tr>
<td>30 Severely disabled</td>
</tr>
<tr>
<td>20 Very sick, requires supportive treatment</td>
</tr>
<tr>
<td>10 Moribund</td>
</tr>
</tbody>
</table>

Months from Diagnosis
Age in years
Prior therapy: 0=no, 1=yes
data va;
  infile 'c:\st565\va.dat';
  input rx cellt time status karno months
    age prior_rx;
  prior_rx = prior_rx/10;
  ct2=0; if(cellt=2) then ct2=1;
  ct3=0; if(cellt=3) then ct3=1;
  ct4=0; if(cellt=4) then ct4=1;
run;

proc format; value celltype 1 = 'squamous'
  2 = 'smallcell'
  3 = 'adenocarcinoma'
  4 = 'large';
/* Create a new data file with covariate values at which you want to estimate survival probabilities */
data vanew;
  input rx ct2 ct3 ct4 karno months age prior_rx;
datalines;
  1 0 1 80 64 70 1
  0 0 1 80 64 70 1
RUN;
proc phreg data=va;
  model time*status(0)= rx ct2 ct3 ct4 karno
    months age prior_rx/ties=efron;
  baseline out=va2 covariates=vanew survival=s
  stderr=se lower=l95 upper=u95;
run;
proc print data=va2; run;

proc print data=va2; run;

The PHREG Procedure

Model Information

Data Set WORK.VA
Dependent Variable time
Censoring Variable status
Censoring Value(s) 0
Ties Handling EFRON

Summary of the Number of Event and Censored Values

<table>
<thead>
<tr>
<th>Total</th>
<th>Event</th>
<th>Censored</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>137</td>
<td>128</td>
<td>9</td>
<td>6.57</td>
</tr>
</tbody>
</table>

Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.
Model Fit Statistics

<table>
<thead>
<tr>
<th>Without Covariates</th>
<th>With Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2 LOG L</td>
<td>1010.898</td>
</tr>
<tr>
<td>AIC</td>
<td>1010.898</td>
</tr>
<tr>
<td>BIC</td>
<td>987.610</td>
</tr>
</tbody>
</table>

Testing Global Null Hypothesis: BETA=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio</td>
<td>62.1039</td>
<td>8</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Score</td>
<td>66.7375</td>
<td>8</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Wald</td>
<td>62.3675</td>
<td>8</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Analysis of Maximum Likelihood Estimates

<table>
<thead>
<tr>
<th>Variable</th>
<th>DF</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Chi-Square</th>
<th>Pr&gt;ChiSq</th>
<th>Hazard Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>rx</td>
<td>1</td>
<td>0.29461</td>
<td>0.20755</td>
<td>2.0148</td>
<td>0.1558</td>
<td>1.343</td>
</tr>
<tr>
<td>ct2</td>
<td>1</td>
<td>-0.86156</td>
<td>0.27532</td>
<td>9.7986</td>
<td>&lt;.0001</td>
<td>0.423</td>
</tr>
<tr>
<td>ct3</td>
<td>1</td>
<td>1.19607</td>
<td>0.30092</td>
<td>15.7986</td>
<td>&lt;.0001</td>
<td>3.307</td>
</tr>
<tr>
<td>ct4</td>
<td>1</td>
<td>0.40129</td>
<td>0.28269</td>
<td>2.0151</td>
<td>0.1557</td>
<td>1.494</td>
</tr>
<tr>
<td>karno</td>
<td>1</td>
<td>-0.03282</td>
<td>0.00551</td>
<td>35.4979</td>
<td>&lt;.0001</td>
<td>0.968</td>
</tr>
<tr>
<td>months</td>
<td>1</td>
<td>0.000082</td>
<td>0.000094</td>
<td>0.0001</td>
<td>0.9929</td>
<td>1.000</td>
</tr>
<tr>
<td>age</td>
<td>1</td>
<td>-0.00871</td>
<td>0.07323</td>
<td>0.8763</td>
<td>0.3942</td>
<td>0.991</td>
</tr>
<tr>
<td>prior_rx</td>
<td>1</td>
<td>0.07159</td>
<td>0.02323</td>
<td>0.0950</td>
<td>0.7580</td>
<td>1.074</td>
</tr>
</tbody>
</table>

Fitting the Cox Model in R or Splus

```r
coxph(formula, data=sys.parent(), weights, subset, na.action, eps = 1e-07, init, iter.max=10, method=c("efron","breslow","exact"), singular.ok=T, robust=F, model=F, x=F, y=T)
```
R code for fitting the Cox model to data
from the VA lung cancer trial of 137 male patients with inoperable lung cancer.
This code is posted as vacoxph.ssc

Variables
Treatment: 1=standard, 2=test (chemotherapy)
Celltype: 1=squamous, 2=smallcell, 3=adenoc, 4=large
Survival in days
Status: 1=dead, 0=censored
Karnofsky score
Months from Diagnosis
Age in years
Prior therapy: 0=no, 10=yes

Enter the data into a data frame.

va <- read.table("c:/stat565/va.dat", header=F, col.names=c("rx", "cellt", "time", "status", "karno", "months", "age", "priorrx"))
va$rx <- va$rx-1
va$priorrx <- va$priorrx/10;
va$celltf<-as.factor(va$cellt)
va

library(survival)
options(contrasts=c("contr.treatment", "contr.poly"))
vafit <- coxph(Surv(time, status) ~ rx+celltf+karno + months + age + priorrx, data = va, method="efron", singular.ok=T)
summary(vafit)

Call:
coxph(formula = Surv(time, status) ~ rx + celltf + karno + months + age + priorrx, data = va, method = "efron", singular.ok = T, robust = T, x = T, y = T) n= 137

coef exp(coef) se(coef) robust se z p
rx 0.294605 1.343 0.20755 0.18867 1.562 1.2e-001
celltf2 0.861559 2.367 0.27528 0.31424 2.742 6.1e-003
celltf3 1.196067 3.307 0.30092 0.27596 4.334 1.5e-005
celltf4 0.401290 1.494 0.28269 0.25242 1.590 1.1e-001
karno -0.032815 0.968 0.00551 0.00522 -6.286 3.3e-010
months 0.000082 1.000 0.00914 0.00795 0.010 9.9e-001
age -0.008706 0.991 0.00930 0.01030 -0.845 4.0e-001
priorrx 0.071589 1.074 0.23231 0.22068 0.324 7.5e-001

exp(coef) exp(-coef) lower .95 upper .95
rx 1.343 0.745 0.928 1.943
celltf2 2.367 0.423 1.278 4.382
celltf3 3.307 0.302 1.926 5.680
celltf4 1.494 0.669 0.911 2.450
karno 0.968 1.033 0.958 0.978
months 1.000 1.000 0.985 1.016
age 0.991 1.009 0.972 1.012
priorrx 1.074 0.931 0.697 1.656

Rsquare= 0.364 (max possible= 0.999 )
Likelihood ratio test= 62.1 on 8 df, p=1.8e-010
Wald test = 68.7 on 8 df, p=8.96e-012
Score (logrank) test = 66.7 on 8 df, p=2.19e-011,
Robust = 52.3 p=1.46e-008

Note: The likelihood ratio and score tests assume
independence among observations within a cluster,
the Wald and robust score tests do not.

> names(vafit)
[1] "coefficients" "var" "loglik"
[4] "score" "iter" "linear.predictors"
[7] "residuals" "means" "method"
[10] "n" "terms" "assign"
[13] "naive.var" "rscore" "wald.test"
[16] "n" "formula"
[19] "call"

vafit$x
rx celltf2 celltf3 celltf4 karno months age priorrx
1 0 0 0 0 60 7 69 0
2 0 0 0 0 70 5 64 1
3 0 0 0 0 60 3 38 0
4 0 0 0 0 60 9 63 1
5 0 0 0 0 70 11 65 1
6 0 0 0 0 20 5 49 0
7 0 0 0 0 40 10 69 1
8 0 0 0 0 80 29 68 0
9 0 0 0 0 80 18 43 0
10 0 0 0 0 70 6 70 0
# Model selection

```r
step(vafit, direction="both", trace=1, steps=1000, k=2)
```

Start: AIC = 964.79

```
Surv(time, status) ~ rx + celltf + karno + months + age + priorrx
```

<table>
<thead>
<tr>
<th>Df</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- months 1 962.79
- priorrx 1 962.89
- age 1 963.65
- rx 1 964.79
- celltf 3 977.63
- karno 1 997.78

Step: AIC = 962.79

```
Surv(time, status) ~ rx + celltf + karno + months + age + priorrx
```

<table>
<thead>
<tr>
<th>Df</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- priorrx 1 960.92
- age 1 961.67
- rx 1 962.84
+ months 1 964.79
- celltf 3 975.67
- karno 1 996.82

Step: AIC = 960.92

```
Surv(time, status) ~ rx + celltf + karno + age
```

<table>
<thead>
<tr>
<th>Df</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- age 1 959.83
- rx 1 961.09
+ priorrx 1 962.79
+ months 1 962.89
- celltf 3 973.76
- karno 1 994.86

Step: AIC = 959.83

```
Surv(time, status) ~ rx + celltf + karno
```

<table>
<thead>
<tr>
<th>Df</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- rx 1 959.53
- age 1 960.92
+ priorrx 1 961.67
+ months 1 961.75
- celltf 3 971.93
- karno 1 993.04

Step: AIC = 959.53

```
Surv(time, status) ~ celltf + karno
```

<table>
<thead>
<tr>
<th>Df</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- rx 1 959.83
- age 1 961.09
+ priorrx 1 961.27
+ months 1 961.35
- celltf 3 970.87
- karno 1 992.05

Call:

```
coxph(formula=Surv(time, status) ~ celltf + karno, data = va, method = "efron", singular.ok = T, robust = T, x = T, y = T)
```

<table>
<thead>
<tr>
<th>coef</th>
<th>exp(coef)</th>
<th>se(coef)</th>
<th>robust se</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>celltf2 0.7153</td>
<td>2.04</td>
<td>0.2829</td>
<td>0.3184</td>
<td>2.25</td>
<td>2.5e-02</td>
</tr>
<tr>
<td>celltf3 1.1577</td>
<td>3.29</td>
<td>0.2929</td>
<td>0.2740</td>
<td>4.22</td>
<td>2.4e-05</td>
</tr>
<tr>
<td>celltf4 0.3256</td>
<td>1.39</td>
<td>0.2766</td>
<td>0.2536</td>
<td>1.28</td>
<td>2.0e-01</td>
</tr>
<tr>
<td>karno -0.0311</td>
<td>0.97</td>
<td>0.0051</td>
<td>0.0054</td>
<td>-5.74</td>
<td>9.6e-09</td>
</tr>
</tbody>
</table>

Likelihood ratio test=69.4 on 4 df, p=3.93e-12 n=137
Model Selection with PHREG in SAS

/* This code is posted as vaphreg2.sas */
data va;
  infile 'c:\st565\va.dat';
  input rx cellt time status karno months age prior_rx;
prior_rx = prior_rx/10;
ct2=0; if(cellt=2) then ct2=1;
ct3=0; if(cellt=3) then ct3=1;
ct4=0; if(cellt=4) then ct4=1;
run;
proc phreg data=va;
  model time*status(0)= rx ct2 ct3 ct4 karno
    months age prior_rx/ties=efron
    selection=stepwise sle=.10 sls=.05;
  title "Results from Stepwise Selection";
run;
proc phreg data=va;
  model time*status(0)= rx ct2 ct3 ct4 karno
    months age prior_rx/ties=efron
    selection=backward sls=.10;
  title "Results from Backward Elimination";
run;

Step 1. Variable karno is entered. The model contains the following explanatory variables:
  karno

Model Fit Statistics

<table>
<thead>
<tr>
<th></th>
<th>Without Covariates</th>
<th>With Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2 LOG L</td>
<td>1010.898</td>
<td>968.867</td>
</tr>
<tr>
<td>AIC</td>
<td>1010.898</td>
<td>970.867</td>
</tr>
<tr>
<td>SBC</td>
<td>1010.898</td>
<td>973.719</td>
</tr>
</tbody>
</table>

Testing Global Null Hypothesis: BETA=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio</td>
<td>42.0312</td>
<td>1</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Score</td>
<td>45.3191</td>
<td>1</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Wald</td>
<td>43.3810</td>
<td>1</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Step 2. Variable ct3 is entered. The model contains the following explanatory variables:
  ct3 karno

Model Fit Statistics

<table>
<thead>
<tr>
<th></th>
<th>Without Covariates</th>
<th>With Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2 LOG L</td>
<td>1010.898</td>
<td>959.867</td>
</tr>
<tr>
<td>AIC</td>
<td>1010.898</td>
<td>963.867</td>
</tr>
<tr>
<td>SBC</td>
<td>1010.898</td>
<td>969.571</td>
</tr>
</tbody>
</table>

Testing Global Null Hypothesis: BETA=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio</td>
<td>51.0313</td>
<td>2</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Score</td>
<td>45.3191</td>
<td>2</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Wald</td>
<td>43.3810</td>
<td>2</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Step 3. Variable ct2 is entered. The model contains the following explanatory variables:
  ct2 ct3 karno

Model Fit Statistics

<table>
<thead>
<tr>
<th></th>
<th>Without Covariates</th>
<th>With Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2 LOG L</td>
<td>1010.898</td>
<td>952.902</td>
</tr>
<tr>
<td>AIC</td>
<td>1010.898</td>
<td>958.902</td>
</tr>
<tr>
<td>SBC</td>
<td>1010.898</td>
<td>967.458</td>
</tr>
</tbody>
</table>

Testing Global Null Hypothesis: BETA=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio</td>
<td>57.9965</td>
<td>3</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Score</td>
<td>62.9854</td>
<td>3</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Wald</td>
<td>60.1330</td>
<td>3</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>
NOTE: No (additional) variables met the 0.1 level for entry into the model.

Analysis of Maximum Likelihood Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Standard Error</th>
<th>Estimate</th>
<th>Hazard Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>ct2</td>
<td>0.21562</td>
<td>-0.57401</td>
<td>0.563</td>
</tr>
<tr>
<td>ct3</td>
<td>0.25830</td>
<td>1.00816</td>
<td>2.741</td>
</tr>
<tr>
<td>karno</td>
<td>0.00512</td>
<td>-0.03047</td>
<td>0.970</td>
</tr>
</tbody>
</table>

Summary of Stepwise Selection

<table>
<thead>
<tr>
<th>Step</th>
<th>Entered</th>
<th>Removed</th>
<th>In Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>karno</td>
<td></td>
<td>45.3191</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>2</td>
<td>ct3</td>
<td></td>
<td>10.5638</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>3</td>
<td>ct2</td>
<td></td>
<td>7.2377</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

proc phreg data=va;
   model time*status(0)= rx ct2 ct3 ct4 karno months age prior_rx/ties=efron
   selection=backward sls=.10;
   title "Results from Backward Elimination";
   run;

Step 0. The model contains the following variables:
   rx ct2 ct3 ct4 karno months age prior_rx

Model Fit Statistics

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Without Covariates</th>
<th>With Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2 LOG L</td>
<td>1010.898</td>
<td>948.794</td>
</tr>
<tr>
<td>AIC</td>
<td>1010.898</td>
<td>964.794</td>
</tr>
<tr>
<td>SBC</td>
<td>1010.898</td>
<td>987.610</td>
</tr>
</tbody>
</table>

Testing Global Null Hypothesis: BETA=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio</td>
<td>62.1039</td>
<td>8</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Score</td>
<td>66.7375</td>
<td>8</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Wald</td>
<td>62.3675</td>
<td>8</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Step 1. Variable months is removed. The model contains the following explanatory variables:
   rx ct2 ct3 ct4 karno age prior_rx

Model Fit Statistics

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Without Covariates</th>
<th>With Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2 LOG L</td>
<td>1010.898</td>
<td>948.794</td>
</tr>
<tr>
<td>AIC</td>
<td>1010.898</td>
<td>964.794</td>
</tr>
<tr>
<td>SBC</td>
<td>1010.898</td>
<td>987.610</td>
</tr>
</tbody>
</table>

Testing Global Null Hypothesis: BETA=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio</td>
<td>62.1038</td>
<td>7</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Score</td>
<td>66.6268</td>
<td>7</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Wald</td>
<td>62.3647</td>
<td>7</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Step 5. Variable ct4 is removed. The model contains the following explanatory variables:
   ct2 ct3 karno

Model Fit Statistics

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Without Covariates</th>
<th>With Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2 LOG L</td>
<td>1010.898</td>
<td>952.902</td>
</tr>
<tr>
<td>AIC</td>
<td>1010.898</td>
<td>958.902</td>
</tr>
<tr>
<td>SBC</td>
<td>1010.898</td>
<td>967.458</td>
</tr>
</tbody>
</table>

Testing Global Null Hypothesis: BETA=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio</td>
<td>57.9965</td>
<td>3</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Score</td>
<td>62.9854</td>
<td>3</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Wald</td>
<td>60.1330</td>
<td>3</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>
NOTE: No (additional) variables met the 0.1 level for removal from the model.

Analysis of Maximum Likelihood Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Variable</th>
<th>DF</th>
<th>Standard Estimate</th>
<th>Error</th>
<th>Chi-Square</th>
<th>Pr&gt;ChiSq</th>
<th>Hazard Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>ct2</td>
<td>1</td>
<td>0.57401</td>
<td>0.21562</td>
<td>7.086</td>
<td>0.0078</td>
<td>2.741</td>
<td>0.563</td>
</tr>
<tr>
<td>ct3</td>
<td>1</td>
<td>1.00816</td>
<td>0.25830</td>
<td>15.233</td>
<td>&lt;.0001</td>
<td></td>
<td>2.741</td>
</tr>
<tr>
<td>karno</td>
<td>1</td>
<td>-0.03047</td>
<td>0.00512</td>
<td>35.447</td>
<td>&lt;.0001</td>
<td></td>
<td>0.970</td>
</tr>
</tbody>
</table>

Summary of Backward Elimination

<table>
<thead>
<tr>
<th>Step</th>
<th>Removed</th>
<th>Number</th>
<th>Wald</th>
<th>Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>months</td>
<td>7</td>
<td>0.0001</td>
<td>0.9929</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>prior_rx</td>
<td>6</td>
<td>0.1224</td>
<td>0.7265</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>age</td>
<td>5</td>
<td>0.9315</td>
<td>0.3345</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>rx</td>
<td>4</td>
<td>1.6971</td>
<td>0.1927</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>ct4</td>
<td>3</td>
<td>1.3853</td>
<td>0.2392</td>
<td></td>
</tr>
</tbody>
</table>

proc phreg data=va;
model time*status(0)= rx ct2 ct3 ct4 karno months age prior_rx/ties=efron
  selection=score best=5 start=2 stop=6;
title "Searching Through All Possible Models";
run;

Time-Dependent Covariates

- Covariates whose values change over the course of the study
- Can be used to create models where the proportional hazards property of a constant hazard ratio across time does not hold

- Internal variables: are subject-specific and require periodic measurements on individual subjects (e.g., measure the concentration of a specific substance in blood, or the presence/absence of a certain condition)

- External variables: Factors that apply to all subjects (e.g., average hourly pollen count in an allergy study, or predetermined changes in dosages of a drug)
Let \( x_{ji}(t) \) denote the value for \( j-th \) covariate for the \( i-th \) subject at time \( t \).

The hazard for the \( i-th \) subject is

\[
   h_i(t) = h_0(t) \exp \left( \sum_{j=1}^{k} \beta_j x_{ji}(t) \right)
\]

In this model,

\( h_0(t) \) is the hazard for an individual for whom all covariates are zero at the time origin, and remain zero through time.

The hazard ratio \( h_i(t)/h_0(t) \) can change through time.

At each observed failure time you must know the values of the time dependent covariates for all individuals still at risk.

- Generally no problem for external variables
- No problem when the covariate is a direct function of time
- You could measure internal time dependent covariates using the same schedule of inspection times for each individual (eg., each subject reports to the clinic every six months)
- Otherwise, you may a have huge missing data problem.

The hazard ratio for two subjects is

\[
   \frac{h_i(t)}{h_m(t)} = \exp \left( \beta_1 \left[ x_{1i}(t) - x_{1m}(t) \right] + \cdots + \beta_k \left[ x_{ki}(t) - x_{km}(t) \right] \right)
\]

Then, \( \beta_j \) can be interpreted as the log(hazard ratio) for two individuals whose values of the \( j-th \) explanatory variable differ by one unit at some time point \( t \) and who have the same values for all other covariates at that time point.

Assuming no ties, the log-partial likelihood is

\[
   \log(L_{\hat{p}}(\hat{\beta})) = \sum_{i=1}^{n} \delta_i \left( \sum_{j=1}^{k} \beta_j x_{ji}(t_i) \right) - \log \left( \sum_{t \in R(t_i)} \exp \left( \sum_{j=1}^{k} \beta_j x_{ji}(t_i) \right) \right)
\]

where \( \delta_i = 0 \) for a censored time and is unity otherwise.

Treatment effects can be confounded with effects of time dependent covariates.

Suppose you randomly assign leukaemia patients to either a standard or a new cytotoxic drug.

- White blood cell count is used as a time dependent covariate
- The actual difference in the drugs corresponds to how they affect white blood cell count.
- The difference in the drugs is not significant after adjusting for white blood cell count.
- There is a significant difference in the drugs if white blood cell counts at different times are not included in the model.
Checking the proportional hazards assumption

- Include an interaction with time in the model as a time dependent covariate
  
  \[ x(t) = \text{var} \times \text{time} \]
  
  \[ x(t) = \text{var} \times \log(\text{time}) \]

- To check the proportional hazards assumption for the treatment variable `rx` in the VA lung cancer example, use the following SAS code posted as vacoxph3.sas:

```sas
proc phreg data=va;
  model time*status(0)= rx rxt ct2 ct3 ct4
       karno months age prior_rx /
ties=efron;
  title "Incorrect Check for Proportional Hazards";
run;

proc phreg data=va;
  model time*status(0)= rx rxt ime ct2 ct3 ct4
       karno months age prior_rx /
ties=efron;
  rtime=rx*log(time);
  title "Correct Check for proportional hazards";
run;
```

Incorrect Check for Proportional Hazards

Model Fit Statistics

<table>
<thead>
<tr>
<th>Model Fit Statistics</th>
<th>Without Covariates</th>
<th>Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criterion</td>
<td>-2 LOG L</td>
<td>1010.898</td>
</tr>
<tr>
<td></td>
<td>AIC</td>
<td>1010.898</td>
</tr>
<tr>
<td></td>
<td>SBC</td>
<td>1010.898</td>
</tr>
</tbody>
</table>

Testing Global Null Hypothesis: BETA=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio</td>
<td>143.7263</td>
<td>9</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Score</td>
<td>149.4959</td>
<td>9</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Wald</td>
<td>109.6200</td>
<td>9</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Analysis of Maximum Likelihood Estimates

| Parameter Standard | Hazard Variable DF Estimate Error Chi-Square Pr>ChiSq Ratio |
|--------------------|----------------------|----------------------|----------------------|----------------------|
| rx                 | 1 7.29916 0.87249 69.9883 <.0001 1479.054 |
| rxt                | 1 -1.51682 0.18476 67.4019 <.0001 0.219 |
| ct2                | 1 -0.26993 0.25739 1.0998 0.2943 0.763 |
| ct3                | 1 0.44024 0.30201 2.1249 0.1449 1.553 |
| ct4                | 1 -0.19917 0.28478 0.4891 0.4843 0.819 |
| karno              | 1 -0.01192 0.00610 3.8185 0.0507 0.988 |
| months             | 1 -0.00204 0.00966 0.0446 0.8328 0.996 |
| age                | 1 0.00275 0.00928 0.0879 0.7669 1.003 |
| prior_rx           | 1 0.15749 0.22722 0.4804 0.4882 1.171 |
Correct Check for proportional hazards

Model Fit Statistics

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Without Covariates</th>
<th>With Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2 LOG L</td>
<td>1010.898</td>
<td>948.534</td>
</tr>
<tr>
<td>AIC</td>
<td>1010.898</td>
<td>966.534</td>
</tr>
<tr>
<td>SBC</td>
<td>1010.898</td>
<td>992.202</td>
</tr>
</tbody>
</table>

Testing Global Null Hypothesis: BETA=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr&gt;ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio</td>
<td>62.3642</td>
<td>9</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Score</td>
<td>66.7829</td>
<td>9</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Wald</td>
<td>62.4007</td>
<td>9</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Suppose you wanted to analyze survival times for patients with severe heart disease where some patients receive heart transplants, and you want to allow the hazard to change when a patient receives a transplant.

\[ T = \text{time to death or censoring} \]

\[ x_1 = 1 \text{ if patient had previous surgery} \]
\[ = 0 \text{ otherwise} \]

\[ x_2 = \text{age (in years) when accepted into the study} \]

\[ \text{wait} = \text{time (in days) from acceptance until transplant} \]

\[ x_3(t) = 1 \text{ if patient had transplant by day } t \]
\[ = 0 \text{ otherwise} \]

\[ h_i(t) = h_0(t) \exp(\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3(t)) \]

data set1;
in file ‘c:\stat565\tranplant.dat’;
input Subject T status x1 x2 wait;
run;

proc phreg data=set1;
model T*status(0) = x1 x2 x3 / ties=exact;
if wait>T or wait=. then x3=0;
else x3=1;
run;

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Std</th>
<th>Wald</th>
<th>Pr&gt; Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>x1</td>
<td>0.771</td>
<td>0.360</td>
<td>4.602</td>
</tr>
<tr>
<td>x2</td>
<td>0.031</td>
<td>0.014</td>
<td>4.995</td>
</tr>
<tr>
<td>x3</td>
<td>-0.046</td>
<td>0.303</td>
<td>0.023</td>
</tr>
</tbody>
</table>

Hazard function for the \( i \)-th patient at time \( t \) is

\[ h_i(t) = h_0(t) \exp(\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3(t)) \]
Time Dependent Covariates Measured at Regular Intervals

Example: Recidivism Study of Released Convicts

- 52 weeks of followup
- week = time to arrest (in weeks)
- arrest = 1 if arrest time, otherwise 0
- age = age (in years) at release
- prior = number of prior arrests
- emp1-emp52: coded 1 if employed, 0 otherwise

```
data set1;
  infile 'c:\stat565\recid.dat';
  input subject arrest week age prior emp1-emp52;
run;
```

```
proc phreg data=set1;
  model week*arrest(0) = age prior employed /
ties=efron;
  array emp(*) emp1-emp52;
  employed=emp(week);
run;
```


- You may be interested in effects of employment on the probability of being arrested
- If someone is arrested, that may terminate their employment
- Potential reverse causation is a common problem with time-dependent covariates
- Consider lagged covariate values

```
proc phreg data=set1;
  where week>2;
  model week*arrest(0) = age prior employ1 employ2 / ties=efron;
  array emp(*) emp1-emp52;
  employ1=emp(week-1);
  employ2=emp(week-2);
run;
```

```
  Parameter Std Wald Pr> Risk
  Variable  df Estimate Error Chi-Sq Chi-sq Ratio
  age 1 -0.046 0.022 4.545 0.033 0.955
  prior 1 0.085 0.029 8.644 0.003 1.089
  employed 1 -1.328 0.251 28.070 0.000 0.265
```

- The hazard of arrest may depend on cumulative employment rather than employment status in the previous two weeks

```
data set2; set set1;
  array emp(*) emp1-emp52;
  array cum(*) cum1-cum52;
  cum1=emp1;
  do i=2 to 52;
    cum(i)=cum(i-1) + emp(i);
    end;
  do i = 1 to 52;
    cum(i) = cum(i)/i:
    end;
run;
```

```
proc phreg data=set2;
  where week>1;
  model week*arrest(0) = age prior cumemp / ties=efron;
  array cemp(*) cum1-cum52;
  cumemp=cemp(week-1);
run;
```
Fitting the Cox Model in R with Time-Dependent Covariates

Suppose subject 84 receives a standard treatment and survives 102 days, but subject 94 switches from the new treatment to the standard treatment at day 57 and survives 320 days. Put the data for subject 94 on two lines

<table>
<thead>
<tr>
<th>Subject</th>
<th>Start</th>
<th>Stop</th>
<th>Status</th>
<th>rx</th>
<th>age</th>
<th>conc</th>
</tr>
</thead>
<tbody>
<tr>
<td>84</td>
<td>0</td>
<td>102</td>
<td>1</td>
<td>0</td>
<td>43</td>
<td>31</td>
</tr>
<tr>
<td>94</td>
<td>57</td>
<td>320</td>
<td>0</td>
<td>1</td>
<td>57</td>
<td>22</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subject</th>
<th>Start</th>
<th>Stop</th>
<th>Status</th>
<th>rx</th>
<th>age</th>
<th>conc</th>
</tr>
</thead>
<tbody>
<tr>
<td>43</td>
<td>0</td>
<td>60</td>
<td>0</td>
<td>1</td>
<td>34</td>
<td>45</td>
</tr>
<tr>
<td>43</td>
<td>60</td>
<td>120</td>
<td>0</td>
<td>1</td>
<td>34</td>
<td>36</td>
</tr>
<tr>
<td>43</td>
<td>120</td>
<td>240</td>
<td>0</td>
<td>1</td>
<td>34</td>
<td>28</td>
</tr>
</tbody>
</table>

vafit3 <- coxph(Surv(Start, Stop, Status) ~ rx + age + conc, data = va, method="efron", singular.ok=T)