Analysis of Failure Time Data

2a: Survival Times of women with tumors which were negatively or positively stained w/ HPA

Histochernical marker (Helix pomatia agglutinin, HPA) binds to those breast cancer cells assoc. w/ metastasis to local lymph nodes. HPA stained cells identified microscopically.

<table>
<thead>
<tr>
<th>Neg. Stain</th>
<th>23</th>
<th>181</th>
</tr>
</thead>
<tbody>
<tr>
<td>47</td>
<td>198*</td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>208*</td>
<td></td>
</tr>
<tr>
<td>70*</td>
<td>212*</td>
<td></td>
</tr>
<tr>
<td>71*</td>
<td>224*</td>
<td></td>
</tr>
<tr>
<td>100*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>101*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>148</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N1 = 13

<table>
<thead>
<tr>
<th>Pos. Stain</th>
<th>5</th>
<th>68</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>76*</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>105*</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>107*</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>109*</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>113</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>116</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>118</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>143</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>154*</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>162*</td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>188*</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>212*</td>
<td></td>
</tr>
<tr>
<td>59</td>
<td>217*</td>
<td></td>
</tr>
<tr>
<td>61</td>
<td>225*</td>
<td></td>
</tr>
</tbody>
</table>

N2 =
We are interested in making inferences about:

- Median time to failure for each population sampled.
- Proportions surviving 5 years, 10 years, 15 years, etc. for each group.
- Are the proportions statistically different?

Confidence intervals for differences, etc.

Complications:

- Data is definitely not normally distributed so usual theory fails.

Censoring also creates a new difficulty to overcome.
Survival Curve:

Let $T =$ time to failure of a randomly selected individual.

Fix a time of interest, $t$.

Then the survival probability is defined as

$$\text{Probability} \left( T > t \right)$$

which is the proportion of individuals in the entire population sampled from that survive at least $t$ units of time.

Define

$$S(t)$$

Example:

- Median survival time = 10 years
- Most people are dead within 20 years
We use data to estimate survival curves, e.g., Breast Cancer data.

Ex. Consider BC data (Pos stain)

Then

\[
\Pr (T > 61 \text{ months}) = \frac{16}{32} = 0.5
\]

\[
\Pr (T > 26 \text{ months}) = \frac{24}{32} = 0.75
\]

\[
\Pr (T > 116 \text{ months}) = \frac{8+\text{something}}{32} \approx 0.25
\]

The problem with the last example is that some of the previously censored individuals are probably still alive at 116 months and this fact needs to be accounted for.

Kaplan-Meier Estimate of S(t)

*Kaplan-Meier (JASA, 1958)* is the most cited statistics paper since WWII.

BMDP1L calculates it.
BMDP1L Program File

Data Input

<table>
<thead>
<tr>
<th>Time</th>
<th>Censor</th>
<th>Stain</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>147</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2241</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>225</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Censor = 0  => alive
Censor = 1  => dead
Stain = -1  => negative
Stain = 1   => positive

/ Problem Title is 'Breast Cancer'.
/ Input File is 'BC.data'.
Variables are 3.
Format = Free.
/ Variable Names = survival, censor, stain.
/ Form Time = survival,
Unit = months,
Status = censor,
Response = 1.
/ Estimate Method = product,
Plots = surv.
Group is stain.
/ Group Codes (3) are -1, 1.
   (1) pos
   (2) neg
Estimated median survival time w/ negative stain is roughly 180 months and for positive stain is approximately 70 months.

\[ Pr(T > 10 \text{ years} \mid \text{neg stain}) = .35 \]
\[ Pr(T > 10 \text{ years} \mid \text{pos stain}) = .75 \]
\[ Pr(T > 5 \text{ years} \mid \text{neg stain}) = .55 \]
\[ Pr(T > 5 \text{ years} \mid \text{pos stain}) = .85 \]

Are these numbers "really different," e.g., are they indicative of a real difference in the "true" or population proportions?

The log rank test pools the information from sample differences like those above corresponding to all possible "death times." A statistic is obtained that is approximately (standard normal) \( \chi^2 \)-distributed. Under the hypothesis that the population survival curves for
The two groups are identical. The combined test is usually called the Mantel-Haenszel test. The value of the statistic for these data is

\[ Z_{obs} = (1.88)^2 \]

with corresponding \( P\text{-value} = .061 \).

So if the theory holds, in large samples, we have seen a relatively rare event. Were the null hypothesis to be true,

Reject hypothesis of identical survival curves at .06 level of significance.
Life Tables: (BMDP12)

Here, data arrives in grouped form.

Ex: 48 myeloma patients.

<table>
<thead>
<tr>
<th>Year</th>
<th>Deaths</th>
<th>Censored</th>
<th>at risk at beginning of interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 1</td>
<td>16</td>
<td>4</td>
<td>48</td>
</tr>
<tr>
<td>Year 2</td>
<td>10</td>
<td>4</td>
<td>28</td>
</tr>
<tr>
<td>Year 3</td>
<td>1</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>Year 4</td>
<td>3</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>Year 5</td>
<td>2</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>Over 5 years</td>
<td>4</td>
<td>1</td>
<td>5</td>
</tr>
</tbody>
</table>

\[ S(t) \]

\[ \frac{1}{0.65} \]

\[ \frac{1}{0.40} \]

\[ \frac{1}{0.37} \]

\[ \frac{1}{0.28} \]

\[ \frac{1}{0.21} \]

No information about how curve should drop in between grouping end points.
Accelerated Failure Time Model
Cox model

\[ T = \text{Time to failure after diagnosis with lung cancer} \]

\[ X_1 = \text{Karnofsky Score} \]
(measured how well off you are at beginning of the study)
\[ 20 = \text{Bad}, 50 = \text{Good} \]

\[ X_2 = \text{Age in years} \]

\[ X_3 = \text{Prior Therapy} \]
\[ \text{yes} = 10, \text{no} = 0 \]

\[ X_4 = \text{Type of tumor} \]
\[ \text{squamous, small cell, adeno, large cell} \]

\[ X_5 = \text{Treatment} \]

\[ X_6 = \text{Time from diagnosis to randomization} \]
Let \( X = (x_1, x_2, x_3, x_4, x_5, x_6) \) be the vector of covariate values.

Let

\[
S(t \mid X) = P(T > t \mid X)
\]

The survivor function for given covariate information.

The ultimate goal is to characterize the effects of covariates on survival.
Model 1: (AFT)

\[ \ln T = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_6 X_6 + \text{error} \]

Looks like usual regression model (log stat 108) in log failure time.

Two problems:

1. Errors may not be normal
2. Censored data

If we let error have what is called the "extreme value" distribution, the model we get is called Weibull regression.

If we let error have the "logistic" distribution, we get log logistic regression.

If we let error have a normal distribution, we get log normal regression.
We are interested in testing hypotheses like

\[ H_0 : B_1 = 0 \quad \text{vs} \quad H_A : B_1 \neq 0 \]

eg. is Karnofsky score useful for predicting survival?

BMDP2L output gives

\[
\begin{align*}
\hat{B}_1 & \quad , \quad \hat{B}_2 & \quad , \quad \text{etc} \\
\text{S}e(\hat{B}_1) & \quad , \quad \text{S}e(\hat{B}_2) & \quad , \quad \text{etc}
\end{align*}
\]

Under \( H_0 : B_1 = 0 \) and with large sample sizes, these statistics may be compared with standard normal percentage points.

Note: Cell type is categorical and here we are interested in testing whether the effect of tumor types are the same for all 4 types, or not.

Results in a chi-square test based on 3 degrees of freedom under \( H_0 \).
Weibull Model Results

Veterans Admin. Lung Cancer Trial
n = 137; 5128 dead; 9 lost

Variable        coef/se

Karnofsky       6.24 *
Age             +0.81
Prior Therapy   -0.2
Treatment       +1.14
Time from diag. 0

Cell Type (22.3) **

The value 22.3 is referred to a chi square (3) distribution
and the p-value is much smaller than .01.

The other values are referred to standard normal and only Karnofsky is important.

Stepwise methods of model selection result in the model...
<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Err</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kaminsky</td>
<td>0.34</td>
<td>0.13</td>
<td>Highly Sig.</td>
</tr>
<tr>
<td>Prior Histrogy</td>
<td>1.88</td>
<td>0.45</td>
<td>Just Sig.</td>
</tr>
<tr>
<td>Cell Type</td>
<td>(14.52)*</td>
<td></td>
<td>Highly Sig.</td>
</tr>
</tbody>
</table>

(* *) Again refer to chi square 3.

We get similar results for log logistic.

Plots: P.15, 17, 18, 19.
log logistic

\[ K = 90 \]
\[ \rho = 0 \]
Hazard Function:

\[
\text{Hazard of death in 1st interval} = \frac{\text{10 deaths}}{\text{20 possible}} = \frac{1}{2}
\]

\[
\text{Hazard in 2nd interval} = \frac{\text{10 deaths}}{(20 \text{ possible}) \times 2 \text{ years}} = \frac{1}{4}
\]

\[
\text{Hazard in 3rd interval} = \frac{\text{10 deaths}}{(10) \times 1 \text{ year}} = 1
\]
Theoretical Hazard

\[ h(t) = \frac{\# \text{ deaths in } (t, t+\Delta t)}{\text{ population}} \]
\[ \approx \frac{\# \text{ at risk just prior to } t}{\text{ at risk just prior to } t} \]

Let \( X \) = covariate information for a particular individual.

Let \( \tilde{X} \) = covariate information for a second individual.

Cox model: (Cox 72)

(2nd most cited paper in statistics)

\[ h(t | X) = \frac{\# \text{ free from } t}{h(t | \tilde{X})} \]
\[ = (\text{constant}) \cdot h(t | \tilde{X}) \]

\( \text{ eg. my hazard or risk is some # times yours, no matter what t we are looking at.} \)
Suppose \( \text{const} = 2 \).

So my chances of dying in the next time period are double your chances, regardless of whether we are thinking about 1 year after the beginning of the study, or 5 years, or 10 years, etc.

\[
h(t) = e^{x_1b_1 + x_2b_2 + \cdots + x_6b_6}
\]

Your risk = \( h(t \mid X) \)

When \( h(t) \) is called the "Baseline hazard".
So if, for given $B_1, B_2, \ldots$, your $X_1 B_1 + X_2 B_2 \ldots$ is large, your risk is larger than that for someone whose covariate values gives a small value for this quantity. Baseline corresponds to a value of zero.

Survival Functions:

$S(t)$

Your Survival Function
Lung Cancer Data

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Coef. SE</th>
<th>Wald Coef. SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Karnofsky</td>
<td>-5.97</td>
<td>6.24</td>
</tr>
<tr>
<td>Age</td>
<td>-0.71</td>
<td>-1.44</td>
</tr>
<tr>
<td>Prior Therapy</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disease Duration</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Cell Type</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(‡‡) Refer to Table 5.4.1 (3).

Note the difference in signs. This is because risk goes up for large \( \chi^2 \_1 + \chi^2 \_2 + \cdots \) in the Cox model while it goes down in the accelerated failure time model. So interpretations are consistent. It is the magnitude that matters and we see that we get the same qualitative results from both analyses e.g. Karnofsky and cell type are very important.

Survival curve plots based on Cox model would be qualitatively the same as those we looked at.
Important Issues

1. Model checking / model fit
   Need to see if Cox model or AFT models fit well, or not.

2. Confidence intervals
   for relative risk, survival probabilities, median survival time.

3. Time dependent covariates.