Survival Analysis
Part III
Survival Analysis

- Survival Data Characteristics
- Goals of Survival Analysis
- Statistical Quantities
  - Survival function
  - Hazard function
- One-sample Summaries
  - Kaplan-Meier Estimator
  - S.E. Estimation for KM estimator
  - Life Table Estimation
- Two-sample Summaries
  - Mantel-Haenszel / Log-rank test
- Regression Methods - Cox Regression
  - Proportional hazards
  - Interpretation of coefficients
  - Estimation & Testing
  - Survival function estimation
  - Diagnostics
Motivation

On a sample of women from a cohort study of breast cancer patients we take histologic measurements with the goal of assessing the prognostic utility of these measurements.

- **Primary Predictor(s):**
  - ploidy (diploid, aneuploid)
  - % time S-phase
- **Other predictors:** stage, lymph nodes, size …
- **Outcome(s):**
  - Time-until-death (or end of follow-up)
  - Death (yes/no)
- **Question:** Do women with diploid/aneuploid cells survive longer?
- **Complication:** Time to death is not observed on all women.
### Survival Analysis – Example

**Example: Breast Cancer Histology Data**

<table>
<thead>
<tr>
<th>time</th>
<th>status</th>
<th>aneuploid</th>
<th>s-phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>49</td>
<td>1</td>
<td>1</td>
<td>22.4</td>
</tr>
<tr>
<td>73</td>
<td>0</td>
<td>1</td>
<td>6.1</td>
</tr>
<tr>
<td>68</td>
<td>0</td>
<td>0</td>
<td>0.8</td>
</tr>
<tr>
<td>72</td>
<td>0</td>
<td>0</td>
<td>7.8</td>
</tr>
<tr>
<td>80</td>
<td>0</td>
<td>1</td>
<td>4.4</td>
</tr>
<tr>
<td>70</td>
<td>0</td>
<td>0</td>
<td>11.1</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>0</td>
<td>14.9</td>
</tr>
<tr>
<td>72</td>
<td>0</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>77</td>
<td>0</td>
<td>0</td>
<td>0.4</td>
</tr>
</tbody>
</table>

(time, status) = (49, 1) means: Subject died at month 49

(time, status) = (72, 0) means: Subject lost to follow-up/censored at month 72, i.e. time to death > 72 months
Need a New Method?

Q: Why not just use standard linear regression, perhaps taking a log transformation, to analyze the follow-up times (i.e. time to death)?

Q: Why not just use logistic regression to analyze dead/alive status as the outcome variable?

Conclusion: Useful to have methods that consider (time, status) as the outcome variable.
Outcome: (time, status)

- Time - time until an event occurs, t
  - Define the start time
    - e.g. diagnosis, entry into the study, birth, time of randomization
  - Define the event
    - e.g. death, machine failure, relapse, discharge

- Event Indicator (status), $\delta$
  - $\delta = 1$ means an event was observed at t
  - $\delta = 0$ means the time to event was censored at t

Censored = study ends, patient withdraws/moves or is lost to follow-up before event is observed.
Right Censoring

The diagram illustrates right censoring in survival analysis. The top graph represents calendar time, with start and stop times indicated for different subjects. The bottom graph shows study time, also with start and stop times for subjects. The censoring is indicated by the horizontal lines extending to the right, marking the end of the observation period for each subject.
One way of summarizing data like this is the *life table*:

### Life Table – Breast Cancer Example

<table>
<thead>
<tr>
<th>Beg. Interval</th>
<th>Total</th>
<th>Deaths</th>
<th>Censored</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 20</td>
<td>568</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>20 40</td>
<td>557</td>
<td>36</td>
<td>18</td>
</tr>
<tr>
<td>40 60</td>
<td>503</td>
<td>37</td>
<td>167</td>
</tr>
<tr>
<td>60 80</td>
<td>299</td>
<td>21</td>
<td>130</td>
</tr>
<tr>
<td>80 100</td>
<td>148</td>
<td>9</td>
<td>67</td>
</tr>
<tr>
<td>100 120</td>
<td>72</td>
<td>3</td>
<td>37</td>
</tr>
<tr>
<td>120 140</td>
<td>32</td>
<td>2</td>
<td>30</td>
</tr>
</tbody>
</table>
Some Definitions: Survival Distributions

Survival function:

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

The survival function is the probability that the survival time, \( T \), is greater than the specific time \( t \).

Hazard function:

\[ P[T < t + \Delta \mid T \geq t] \approx h(t) \cdot \Delta \]

\[ \lim_{\Delta \to 0} \frac{P[T < t + \Delta \mid T \geq t]}{\Delta} = h(t) \]

The hazard function is the conditional “failure” rate or the instantaneous probability of having an event at time \( t \) (per unit time) given that one has survived (i.e. not had an event) up to time \( t \).
Estimation of Survival

No Censoring:

\[ N = \text{total number of subjects} \]
\[ n(t) = \text{number of subjects with } T_i > t \]

\[ \hat{S}(t) = \frac{n(t)}{N} \]

Example: \( N = 12 \)

Ordered observations:
2, 14, 17, 18, 20, 24, 34, 39, 43, 44, 56, 98

Median = 29
Quartiles = 17.5, 43.5
Survival without Censoring

. input time
time
  1. 2
  2. 14
  :
  13. end
. stset time
. sts graph

Note: Step size is $1/N=1/12=0.08333$ for each “failure”
Survival with Censoring

Q: How can we include information from observations like “25+” (survived more than 25 weeks) which we represent as (25,0)?

A: The Kaplan-Meier estimator of S(t).

But, before we get to the details of the Kaplan-Meier estimator we'll consider how we can use the life table information to “piece together” survival information.
**Life Table Construction**

*First*, approximate the conditional probability of death in each interval, given survival to the start of the interval:

\[
P(\text{Death in interval} | \text{Alive at start of interval}) \approx \frac{\# \text{ died in interval}}{\# \text{ alive at start} - \frac{1}{2} (\# \text{ lost in interval})}
\]

**Breast Cancer data**

<table>
<thead>
<tr>
<th>Interval</th>
<th>Beg. Total</th>
<th>Deaths</th>
<th>Lost</th>
<th>Conditional Probability of Death</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>20</td>
<td>568</td>
<td>9</td>
<td>(0.0159=9/(568-2/2))</td>
</tr>
<tr>
<td>20</td>
<td>40</td>
<td>557</td>
<td>36</td>
<td>(0.0657=36/(557-18/2))</td>
</tr>
<tr>
<td>40</td>
<td>60</td>
<td>503</td>
<td>37</td>
<td>(0.0882)</td>
</tr>
<tr>
<td>60</td>
<td>80</td>
<td>299</td>
<td>21</td>
<td>(0.0897)</td>
</tr>
<tr>
<td>80</td>
<td>100</td>
<td>148</td>
<td>9</td>
<td>(0.0786)</td>
</tr>
<tr>
<td>100</td>
<td>120</td>
<td>72</td>
<td>3</td>
<td>(0.0561)</td>
</tr>
<tr>
<td>120</td>
<td>140</td>
<td>32</td>
<td>2</td>
<td>(0.1176)</td>
</tr>
</tbody>
</table>
Life Table Construction

Second, multiply conditional probabilities of survival over each interval to determine the unconditional probability of survival:

\[
P[T > 80] = (P[T > 80 \mid T > 60]) \cdot P(T > 60) \\
= (1 - 0.0897) \cdot P(T > 60) \\
P[T > 60] = (P[T > 60 \mid T > 40]) \cdot P(T > 40) \\
= (1 - 0.0882) \cdot P(T > 40) \\
P[T > 40] = (P[T > 40 \mid T > 20]) \cdot P(T > 20) \\
= (1 - 0.0657) \cdot P(T > 20) \\
P[T > 20] = (P[T > 20 \mid T > 0]) \cdot P(T > 0) \\
= (1 - 0.0159) \cdot P(T > 0) \\
P[T > 0] = ?
\]

\[
P[T > 80] = 0.9103 \cdot 0.9118 \cdot 0.9343 \cdot 0.9841 = 0.7631
\]
## Life Table Construction Using STATA

**. ltable time status, intervals(20)**

<table>
<thead>
<tr>
<th>Interval</th>
<th>Beg. Total</th>
<th>Deaths</th>
<th>Lost</th>
<th>Survival</th>
<th>Std. Error</th>
<th>[95% Conf. Int.]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>20</td>
<td>568</td>
<td>9</td>
<td>2</td>
<td>0.9841</td>
<td>0.0052</td>
</tr>
<tr>
<td>20</td>
<td>40</td>
<td>557</td>
<td>36</td>
<td>18</td>
<td>0.9195</td>
<td>0.0115</td>
</tr>
<tr>
<td>40</td>
<td>60</td>
<td>503</td>
<td>37</td>
<td>167</td>
<td>0.8384</td>
<td>0.0165</td>
</tr>
<tr>
<td>60</td>
<td>80</td>
<td>299</td>
<td>21</td>
<td>130</td>
<td>0.7631</td>
<td>0.0217</td>
</tr>
<tr>
<td>80</td>
<td>100</td>
<td>148</td>
<td>9</td>
<td>67</td>
<td>0.7032</td>
<td>0.0277</td>
</tr>
<tr>
<td>100</td>
<td>120</td>
<td>72</td>
<td>3</td>
<td>37</td>
<td>0.6637</td>
<td>0.0343</td>
</tr>
<tr>
<td>120</td>
<td>140</td>
<td>32</td>
<td>2</td>
<td>30</td>
<td>0.5856</td>
<td>0.0600</td>
</tr>
</tbody>
</table>

---

![Graph showing proportion surviving over time](image)

**Proportion Surviving**

**Time since diagnosis (months)**
A disadvantage of the life table approach is that the intervals are arbitrary – you can get different estimates of survival depending on what set of intervals you choose.

The actuarial approximation (subtracting off half the number lost during the interval) is just that: an approximation

Q: Is there a “right” set of intervals?

A: Yes, the shorter the intervals, the better (less bias). It turns out that the shortest (useful) intervals result from using the observed death (event) times as the interval endpoints. Reducing the interval lengths beyond that has no effect on the estimated survival.

The estimate of survival based on this set of shortest intervals is called the Kaplan-Meier or product limit estimate.
The **Kaplan-Meier** estimator uses the data in a way similar to the life table. At any given time, $t$, we can count the number of subjects that are **at-risk**, that is, known to be alive, and then see how many deaths occur in the next (small) time interval of length $\Delta$. This allows us to estimate $P[\text{die by } t + \Delta \mid T > t]$.

The “at-risk” group gets smaller over time due to subjects that die, and subjects that are lost (censored).

Define:

- $t_i$: $i$’th ordered follow-up time
- $d_i$: number of deaths at $i$’th ordered time
- $l_i$: number of censored observations at $i$’th ordered time
- $R_i$: number of subjects at-risk at $i$’th ordered time

$$\hat{S}(t) = \prod_{t_i \leq t} (1 - d_i / R_i)$$

$$= (1 - d_1 / R_1) \times (1 - d_2 / R_2) \times \ldots \times (1 - d_j / R_j)$$

**Kaplan-Meier Estimator**
**Kaplan-Meier – Simple Example**

Observed “Death” Times: 5, 11, 14, 21, 25, 32, 48
Censored Times: 2, 12, 25, 35

<table>
<thead>
<tr>
<th>time</th>
<th>$R_i$</th>
<th>$d_i$</th>
<th>$l_i$</th>
<th>$d_i/R_i$</th>
<th>$(1 - d_i/R_i)$</th>
<th>$\hat{S}(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>11</td>
<td>0</td>
<td>1</td>
<td>0.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>0.100</td>
<td>0.900</td>
<td>0.900</td>
</tr>
<tr>
<td>11</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>0.111</td>
<td>0.889</td>
<td>0.800</td>
</tr>
<tr>
<td>12</td>
<td>8</td>
<td>0</td>
<td>1</td>
<td>0.000</td>
<td>1.000</td>
<td>0.800</td>
</tr>
<tr>
<td>14</td>
<td>7</td>
<td>1</td>
<td>0</td>
<td>0.143</td>
<td>0.857</td>
<td>0.686</td>
</tr>
<tr>
<td>21</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>0.167</td>
<td>0.833</td>
<td>0.571</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>0.200</td>
<td>0.800</td>
<td>0.457</td>
</tr>
<tr>
<td>32</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0.333</td>
<td>0.667</td>
<td>0.305</td>
</tr>
<tr>
<td>35</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0.000</td>
<td>1.000</td>
<td>0.305</td>
</tr>
<tr>
<td>48</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

We really only needed to list the failure (“death”) times. Why?
Kaplan-Meier Using STATA

```
. stset time, failure(status=1) /*time and status are the key variables*/
  
    failure event:  status == 1
  obs. time interval:  (0, time]
  exit on or before:  failure

  +--------------------------------------------------+
<table>
<thead>
<tr>
<th>time   status   _st   _d   _t   _t0</th>
</tr>
</thead>
<tbody>
<tr>
<td>5        1     1    1    5     0</td>
</tr>
<tr>
<td>11        1     1    1   11     0</td>
</tr>
<tr>
<td>...  0</td>
</tr>
<tr>
<td>25        0     1    0   25     0</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>32        1     1    1   32     0</td>
</tr>
<tr>
<td>48        1     1    1   48     0</td>
</tr>
<tr>
<td>...  0</td>
</tr>
<tr>
<td>25        0     1    0   25     0</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>35        0     1    0   35     0</td>
</tr>
</tbody>
</table>
```

11 total obs.
0 exclusions

11 obs. remaining, representing
7 failures in single record/single failure data
230 total analysis time at risk, at risk from t = 0
  earliest observed entry t = 0
  last observed exit t = 48

. list, noobs

<table>
<thead>
<tr>
<th>time</th>
<th>status</th>
<th>_st</th>
<th>_d</th>
<th>_t</th>
<th>_t0</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>21</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>21</td>
</tr>
<tr>
<td>25</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>32</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>32</td>
</tr>
<tr>
<td>48</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>48</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>25</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>35</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>35</td>
<td>0</td>
</tr>
</tbody>
</table>
**Kaplan-Meier Using STATA**

`. sts list`

<table>
<thead>
<tr>
<th>Time</th>
<th>Beg. Total</th>
<th>Fail</th>
<th>Lost</th>
<th>Survivor Function</th>
<th>Std. Error</th>
<th>[95% Conf. Int.]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>11</td>
<td>0</td>
<td>1</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>0.9000</td>
<td>0.0949</td>
<td>0.9853</td>
</tr>
<tr>
<td>11</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>0.8000</td>
<td>0.1265</td>
<td>0.9459</td>
</tr>
<tr>
<td>12</td>
<td>8</td>
<td>0</td>
<td>1</td>
<td>0.8000</td>
<td>0.1265</td>
<td>0.9459</td>
</tr>
<tr>
<td>14</td>
<td>7</td>
<td>1</td>
<td>0</td>
<td>0.6857</td>
<td>0.1515</td>
<td>0.8871</td>
</tr>
<tr>
<td>21</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>0.5714</td>
<td>0.1638</td>
<td>0.8146</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>0.4571</td>
<td>0.1662</td>
<td>0.7298</td>
</tr>
<tr>
<td>32</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0.3048</td>
<td>0.1666</td>
<td>0.6174</td>
</tr>
<tr>
<td>35</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0.3048</td>
<td>0.1666</td>
<td>0.6174</td>
</tr>
<tr>
<td>48</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0.0000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

`. sts graph, censored(single)`

**Kaplan-Meier survival estimate**

---

Spring 2013  Biostat 513  304
Censoring

We saw earlier that if we have N uncensored times then the Kaplan-Meier curve simply takes “steps” of $1/N$ for every observed failure time.

**Q:** What happens to the “steps” for censored observations?

**A:** Efron (1967) gave an intuitive answer: the Kaplan-Meier distributes the “jump” for a censored time to the observed times that are larger than the censored time.

“Distribute to the right”

\[
\begin{array}{cccc}
 t=1 & t=2 & t=3 & t=4 \\
 X = 1 & \overbrace{\phantom{X}} & & \\
 X = 2 & & X & \\
 X = 3 & & & X \\
 X = 4 & & & O \\
\end{array}
\]
Breast Cancer dataset

- Time to death
- “ploidy” is possible predictor (aneuploid vs diploid)

```plaintext
summarize

Variable | Obs   | Mean   | Std. Dev. | Min | Max
---------|-------|--------|-----------|-----|-----
time     |  568  | 65.61092| 25.45858  |  9  | 120
status   |  568  | .20598 | .40477    |  0  | 1
ploidy   |  568  | .64788 | .47804    |  0  | 1
sphase   |  568  | 9.94031| 8.84160   |  0  | 55.4

.table ploidy status

---------+-----------
ploidy   | status
---------+-----------
diploid  | 169 1
aneuploid| 282 86
---------+-----------
```
### Kaplan-Meier – Example 2

```
. list
+------------------------------------+
<table>
<thead>
<tr>
<th>time   status     ploidy   sphase</th>
</tr>
</thead>
</table>
1. |   49        1   aneuploid 22.4 |
2. |   73        0   aneuploid 6.1  |
3. |   68        0      dipoid  .8  |
4. |   72        0      dipoid  7.8  |
5. |   80        0   aneuploid 4.4  |
6. |   70        0      dipoid 11.1 |
+------------------------------------|

. stset time, failure(status=1)
. sts graph, by(ploidy)
```

Kaplan-Meier survival estimates, by ploidy

![Kaplan-Meier survival estimates, by ploidy](image-url)
1. “Time-until” outcomes (survival times) are common in biomedical research.

2. Survival times are often right-skewed.

3. Often some of the times are right-censored.

4. The Kaplan-Meier estimator can be used to estimate and display the distribution of survival times.

Coming up …

• More on censoring
• Standard error estimates for KM plots
• Comparing KM curves
Censoring

Censoring is a form of missing data, or a data selection process. As such, censoring may lead to selection bias unless we can assume that the observations that are censored at time $t$ are representative of all subjects who “survive” to $t$.

Example:
Suppose that in a clinical trial we remove subjects from the study when they are still alive but appear to be particularly ill (or particularly well). If we treat these as censored and then assume that they were representative we would obtain biased estimates of survival probabilities, $S(t)$.

This is an example of dependent censoring. All of the procedures that we'll discuss assume that the censoring is independent of the survival times, $T_i$. 
Censoring

• Consider two extreme assumptions about censored observations:
  
  1. All censored subjects die immediately after censorship
  2. All censored subjects outlive everyone else

• The Kaplan-Meier estimate lies between these extremes.

• Intuitively, the KM estimate takes each censored observation and distributes it over all failure times that are greater than the censored observation ("distribute to the right")

Consider an illustrative dataset, constructed from our earlier simple example
Censoring

Observed “Death” Times : 5, 11, 14, 21, 25, 32, 48
Censored Times : 2, 12, 25, 35

. preserve
. * Worst case scenario: death follows censorship
. replace status=1 if status==0
. gen grp=3
. save "worst.dta"
. restore

. preserve
. * Best case scenario: censoreds outlive everyone . . * else
. replace time=50 if status==0
. gen grp=1
. save "best.dta"
. restore

. gen grp=2
. append using "best.dta"
. append using "worst.dta"
. lab def grpname 1 "Best" 2 "KM" 3 "Worst"
. lab val grp grpname

. stset time, failure(status=1)

. sts graph, by(grp) cens(s) title("Bounds on Kaplan-Meier") plot1(lpat(dash) lw(medthick)) plot2(lpat(solid) lw(medthick)) plot3(lpat(dash) lw(medthick))
Bounds on Kaplan-Meier

- grp = Best
- grp = KM
- grp = Worst
Censoring

Notation:

- $D_i =$ the survival time for subject $i$
- $C_i =$ the censoring time for subject $i$

We observe:

- $T_i = \min(D_i, C_i)$
- $\delta_i = 1$ if $D_i < C_i$, and 0 otherwise

We assume that $C_i$ is independent of $D_i$.

Examples …

- Censoring due to the end of study ("administrative censoring") $\Rightarrow$ independent censoring
- Censoring due to drop-out $\Rightarrow$ “verify” independence based on reasons for drop-out
- Censoring due to another type of outcome ("competing risks") $\Rightarrow$ often assumed independent
**Interval Censoring & Truncation**

**Interval Censoring:**
This occurs when we do not observe the exact time of failure, but rather two time points between which the event occurred:

\[ a \leq T_i < b \]

E.g. HIV vaccine trial with 6 monthly blood testing. If everyone shares the same time intervals (i.e. 6 month visit schedule) then the outcomes are known as discrete survival times, and logistic regression methods can be used.

**Left Truncation:**
This occurs when some subjects have a *delayed entry* into the study. This can lead to bias since the subject must have lived long enough to enter at a later time. Kaplan-Meier and Cox regression can accommodate this aspect.

E.g. Study of post-menopausal hormone use where \( t = 0 \) is the *date of starting hormones* but women enter the study years later.
Kaplan-Meier can be used to obtain estimates of survival probabilities such as

\[ \hat{S}(60) = \text{estimated 60 month survival} \]

**Q:** How can we obtain a confidence interval for this estimate?

**Recall:**

- \( t_i \): i’th ordered follow-up time
- \( d_i \): number of deaths at i’th ordered time
- \( l_i \): number of censored observations at i’th ordered time
- \( R_i \): number of subjects at-risk at i’th ordered time

**Greenwood’s formula:**

\[
\hat{S}(t) = \prod_{t_i \leq t} \left(1 - \frac{d_i}{R_i}\right)
\]

\[
\hat{V}[\hat{S}(t)] = S(t)^2 \sum_{t_i \leq t} \frac{d_i}{R_i(R_i - d_i)}
\]
**Standard Errors and CIs for \( \hat{S}(t) \)**

**95% Confidence Interval using Greenwood:**

\[
\text{lower : } \hat{S}(t) - 1.96 \cdot \hat{S}(t) \sqrt{\sum_{t_i \leq t} \frac{d_i}{R_i (R_i - d_i)}} \\
\text{upper : } \hat{S}(t) + 1.96 \cdot \hat{S}(t) \sqrt{\sum_{t_i \leq t} \frac{d_i}{R_i (R_i - d_i)}}
\]

Because \( S(t) \) must lie in the interval \([0,1]\), an alternative approach is to determine confidence intervals for a transformation of \( S(t) \) and then transform back, for instance using:

\[
\hat{V} \{ \log[\log \hat{S}(t)] \} = \left[ \sum_{t_i \leq t} \frac{d_i}{R_i (R_i - d_i)} \right] / [\log \hat{S}(t)]^2
\]

In practice, confidence intervals based on either approach should be similar except, possibly, when \( S(t) \) is near 0 or 1.
Standard Errors and CIs for $\hat{S}(t)$

**STATA:**

- `stset` - to define survival data
- `sts graph` - to create Kaplan-Meier plot
- `sts graph, gwood` - request Greenwood's CI added to graph
- `sts gen surv=s survse=se(s) survlse=se(lls)`
  - saves estimate of KM survival curve, standard error and se of $\log(-\log(S(t)))$
- `sts list` - to display
- `sts test` - for log-rank (& other) tests
Standard Errors and CIs for $\hat{S}(t)$

```
.sts graph, gw cens(s)
```

Kaplan-Meier survival estimate
Comparison of Survival Curves

(Klein and Moeschberger, 1997): Data from 101 patients with advanced acute myelogenous leukemia were reported to the International Bone Marrow Transplant Registry. Fifty-one patients had received an autologous (auto) bone marrow transplant in which, after high doses of chemotherapy, their own bone marrow was used to replace their destroyed immune system. Fifty patients had an allogeneic (allo) bone marrow transplant where marrow from an HLA matched sibling was used to replenish their immune systems.

Q: Estimate 2-year survival in each group, with 95% CI.
Q: Any difference in survival between the groups?

```
.infile time type status using transplant.dat
.label variable time "time (months)"
.label variable status "status"
.label variable type "transplant type"
.label define tlab 1 "allogeneic" 2 "autologous"
.label values type tlab
.stset time, failure(status)
```
Kaplan-Meier allows a graphical comparison of survival curves for different patient subsets.

Comparison at a single time point
- How to choose the time point?

Comparison of entire curves
- Log rank test
- Generalized Wilcoxon tests
Comparing Survival at Fixed Times

Q: How can we test (compare) the probability of survival at a certain time, $t_0$, for two groups of subjects?

A: Given the Kaplan-Meier survival estimator and Greenwood's variance estimator we can use a $Z$ statistic.

$$H_0 : S_1(t_0) = S_2(t_0)$$

$$H_1 : S_1(t_0) \neq S_2(t_0)$$

$$Z = \frac{\hat{S}_1(t_0) - \hat{S}_2(t_0)}{\sqrt{\hat{V}[\hat{S}_1(t_0)] + \hat{V}[\hat{S}_2(t_0)]}}$$

$Z \sim N(0,1)$ under $H_0$

Note: this reduces to a two sample test of binomial proportions if there is no censoring
Where do we put $t_0$?

- Test result may depend on the choice of $t_0$

- Sometimes there is a clear scientific reason for a particular choice
Comparing Survival at Fixed Times

Example: Using the 50 allogeneic patients and the 51 autologous patients we can test whether the two groups differ with respect to two year survival.

```
  . sts list, by(type)

         Time  Beg. Total  Fail  Net  Lost   Survivor      Std.            [95% Conf. Int.]
          Time    Total     Fail     Lost     Function     Error
-------------------------------------------------------------------------------
    allogeneic
      .03       50      1      0            0.9800    0.0198     0.8664    0.9972
      .493      49      1      0            0.9600    0.0277     0.8494    0.9898
      .855      48      1      0            0.9400    0.0336     0.8254    0.9802
     1.184      47      1      0            0.9200    0.0384     0.8007    0.9692
     1.283      46      1      0            0.9000    0.0424     0.7763    0.9571
     1.48       45      1      0            0.8800    0.0460     0.7522    0.9442
     1.776      44      1      0            0.8600    0.0491     0.7286    0.9307
     2.138      43      1      0            0.8400    0.0518     0.7054    0.9166
     2.5        42      1      0            0.8200    0.0543     0.6826    0.9020
     2.763      41      1      0            0.8000    0.0566     0.6602    0.8870
     2.993      40      1      0            0.7800    0.0586     0.6381    0.8716
      ::
     11.51      26      1      0            0.5861    0.0716     0.4334    0.7108
     12.1       25      0      1            0.5861    0.0716     0.4334    0.7108
     12.8       24      1      0            0.5617    0.0726     0.4086    0.6896
     12.99      23      0      1            0.5617    0.0726     0.4086    0.6896
     13.85      22      0      1            0.5617    0.0726     0.4086    0.6896
     16.61      21      0      1            0.5617    0.0726     0.4086    0.6896
     17.14      20      0      1            0.5617    0.0726     0.4086    0.6896
     20.07      19      1      0            0.5321    0.0746     0.3772    0.6649
     20.33      18      0      1            0.5321    0.0746     0.3772    0.6649
     22.37      17      0      1            0.5321    0.0746     0.3772    0.6649
     26.78      16      0      1            0.5321    0.0746     0.3772    0.6649
```
### Comparing Survival at Fixed Times

<table>
<thead>
<tr>
<th>Time</th>
<th>Beg. Total</th>
<th>Net Fail</th>
<th>Lost</th>
<th>Survivor Function</th>
<th>Std. Error</th>
<th>[95% Conf. Int.]</th>
</tr>
</thead>
<tbody>
<tr>
<td>.658</td>
<td>51</td>
<td>1</td>
<td>0</td>
<td>0.9804</td>
<td>0.0194</td>
<td>0.8689 0.9972</td>
</tr>
<tr>
<td>.822</td>
<td>50</td>
<td>1</td>
<td>0</td>
<td>0.9608</td>
<td>0.0272</td>
<td>0.8522 0.9900</td>
</tr>
<tr>
<td>1.414</td>
<td>49</td>
<td>1</td>
<td>0</td>
<td>0.9412</td>
<td>0.0329</td>
<td>0.8286 0.9806</td>
</tr>
<tr>
<td>2.5</td>
<td>48</td>
<td>1</td>
<td>0</td>
<td>0.9216</td>
<td>0.0376</td>
<td>0.8044 0.9698</td>
</tr>
<tr>
<td>3.322</td>
<td>47</td>
<td>1</td>
<td>0</td>
<td>0.9020</td>
<td>0.0416</td>
<td>0.7804 0.9580</td>
</tr>
<tr>
<td>11.48</td>
<td>31</td>
<td>1</td>
<td>0</td>
<td>0.6385</td>
<td>0.0683</td>
<td>0.4886 0.7549</td>
</tr>
<tr>
<td>12.01</td>
<td>30</td>
<td>1</td>
<td>1</td>
<td>0.6172</td>
<td>0.0693</td>
<td>0.4670 0.7365</td>
</tr>
<tr>
<td>12.24</td>
<td>28</td>
<td>1</td>
<td>0</td>
<td>0.5951</td>
<td>0.0702</td>
<td>0.4447 0.7173</td>
</tr>
<tr>
<td>12.4</td>
<td>27</td>
<td>0</td>
<td>1</td>
<td>0.5951</td>
<td>0.0702</td>
<td>0.4447 0.7173</td>
</tr>
<tr>
<td>13.06</td>
<td>26</td>
<td>0</td>
<td>1</td>
<td>0.5951</td>
<td>0.0702</td>
<td>0.4447 0.7173</td>
</tr>
<tr>
<td>14.47</td>
<td>25</td>
<td>0</td>
<td>1</td>
<td>0.5951</td>
<td>0.0702</td>
<td>0.4447 0.7173</td>
</tr>
<tr>
<td>15</td>
<td>24</td>
<td>0</td>
<td>1</td>
<td>0.5951</td>
<td>0.0702</td>
<td>0.4447 0.7173</td>
</tr>
<tr>
<td>15.46</td>
<td>23</td>
<td>1</td>
<td>0</td>
<td>0.5693</td>
<td>0.0718</td>
<td>0.4174 0.6954</td>
</tr>
<tr>
<td>15.76</td>
<td>22</td>
<td>1</td>
<td>0</td>
<td>0.5434</td>
<td>0.0730</td>
<td>0.3908 0.6730</td>
</tr>
<tr>
<td>16.48</td>
<td>21</td>
<td>1</td>
<td>0</td>
<td>0.5175</td>
<td>0.0740</td>
<td>0.3649 0.6502</td>
</tr>
<tr>
<td>16.71</td>
<td>20</td>
<td>1</td>
<td>0</td>
<td>0.4916</td>
<td>0.0747</td>
<td>0.3396 0.6270</td>
</tr>
<tr>
<td>17.2</td>
<td>19</td>
<td>0</td>
<td>1</td>
<td>0.4916</td>
<td>0.0747</td>
<td>0.3396 0.6270</td>
</tr>
<tr>
<td>17.24</td>
<td>18</td>
<td>1</td>
<td>0</td>
<td>0.4643</td>
<td>0.0754</td>
<td>0.3130 0.6025</td>
</tr>
<tr>
<td>17.3</td>
<td>17</td>
<td>0</td>
<td>1</td>
<td>0.4643</td>
<td>0.0754</td>
<td>0.3130 0.6025</td>
</tr>
<tr>
<td>17.66</td>
<td>16</td>
<td>0</td>
<td>1</td>
<td>0.4643</td>
<td>0.0754</td>
<td>0.3130 0.6025</td>
</tr>
<tr>
<td>18.09</td>
<td>15</td>
<td>1</td>
<td>1</td>
<td>0.4334</td>
<td>0.0764</td>
<td>0.2824 0.5752</td>
</tr>
<tr>
<td>18.75</td>
<td>13</td>
<td>0</td>
<td>1</td>
<td>0.4334</td>
<td>0.0764</td>
<td>0.2824 0.5752</td>
</tr>
<tr>
<td>20.63</td>
<td>12</td>
<td>0</td>
<td>1</td>
<td>0.4334</td>
<td>0.0764</td>
<td>0.2824 0.5752</td>
</tr>
<tr>
<td>23.16</td>
<td>11</td>
<td>1</td>
<td>0</td>
<td>0.3940</td>
<td>0.0790</td>
<td>0.2416 0.5429</td>
</tr>
<tr>
<td>27.73</td>
<td>10</td>
<td>0</td>
<td>1</td>
<td>0.3940</td>
<td>0.0790</td>
<td>0.2416 0.5429</td>
</tr>
</tbody>
</table>
Comparing Survival at Fixed Times

We have the following estimates from the previous pages:

\[ \hat{S}_1(24) = 0.5321 \]
\[ \hat{V}[\hat{S}_1(24)] = (0.0746)^2 \]
\[ \hat{S}_2(24) = 0.3940 \]
\[ \hat{V}[\hat{S}_2(24)] = (0.0790)^2 \]

\[ Z = \frac{0.5321 - 0.3940}{\sqrt{(0.0746)^2 + (0.0790)^2}} \]

\[ = 1.271 \]

p = 0.204 (2-sided)
Comparing Survival Curves

• The log-rank test provides a means of comparing the entire survival curve for 2 (or more) groups

Overview:

\[ H_0 : S_1(t) = S_2(t) \text{ for all } t \]
\[ H_1 : S_1(t) \neq S_2(t) \text{ for some } t \]

• For each observed failure time calculate the expected number of failures in each group if \( S_1(t) = S_2(t) \).
• Compare the total expected failures in each group, \( E_j \), to the total observed failures, \( O_j \).
• A large-sample \( \chi^2(1) \) test.
• Mantel-Haenszel test with strata formed by observed failure times.
**Log-rank Test**

1. Notation: \( t_1 < t_2 < \ldots t_J \) are ordered failure times in the pooled sample (both groups combined).

2. For each \( j \) define:
   - \( d_{1j} = \) number of deaths in group 1 at time \( t_j \)
   - \( d_{2j} = \) number of deaths in group 2 at time \( t_j \)

3. For each \( j \) define
   - \( R_{1j} = \) number at risk in group 1 at time \( t_j \)
   - \( R_{2j} = \) number at risk in group 2 at time \( t_j \)

4. For each time \( t_j \) create a 2x2 table:

<table>
<thead>
<tr>
<th></th>
<th>Group 1</th>
<th>Group 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deaths at ( t_j )</td>
<td>( d_{1j} )</td>
<td>( d_{2j} )</td>
<td>( d_{1j}+d_{2j} )</td>
</tr>
<tr>
<td>Survivors past ( t_j )</td>
<td>( R_{1j}-d_{1j} )</td>
<td>( R_{2j}-d_{2j} )</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>( R_{1j} )</td>
<td>( R_{2j} )</td>
<td>( R_{1j}+R_{2j} )</td>
</tr>
</tbody>
</table>
5. In group 1 we observe \( O_{1j} = d_{1j} \) at time \( t_j \). Under \( H_0 \) we would expect

\[
E_{1j} = \left( \frac{R_{1j}}{R_{1j} + R_{2j}} \right) (d_{1j} + d_{2j})
\]

6. Define \( E_1 = \sum_{j=1}^{J} E_{1j} \); \( O_1 = \sum_{j=1}^{J} d_{1j} \)

7. The log-rank test statistic is:

\[
X^2 = \frac{(O_1 - E_1)^2}{\hat{V}_1}
\]

where

\[
\hat{V}_1 = \sum_{j} \frac{R_{1j} R_{2j} (d_{1j} + d_{2j})(R_{1j} + R_{2j} - d_{1j} - d_{2j})}{(R_{1j} + R_{2j})^2 (R_{1j} + R_{2j} - 1)}
\]

8. Under \( H_0 \): \( S_1(t) = S_2(t) \), \( X^2 \sim \chi^2(1) \)
Log-rank Example


Remission times (in weeks) for two groups of leukemia patients.

**Group 1 (n = 21) treatment**

6, 6, 6, 7, 10, 13, 16, 22, 23, 6+, 9+, 10+, 11+, 17+, 19+, 20+, 25+, 32+, 32+, 34+, 35+

**Group 0 (n = 21) placebo**

1, 1, 2, 2, 3, 4, 4, 5, 5, 8, 8, 8, 8, 11, 11, 12, 12, 15, 17, 22, 23

Note: + denotes censoring
### Log-rank Example

<table>
<thead>
<tr>
<th>j</th>
<th>( t_j )</th>
<th>( d_{1j} )</th>
<th>( d_{2j} )</th>
<th>( R_{1j} )</th>
<th>( R_{2j} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>21</td>
<td>19</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>21</td>
<td>17</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>21</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>21</td>
<td>14</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>21</td>
<td>12</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>1</td>
<td>0</td>
<td>17</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>0</td>
<td>4</td>
<td>16</td>
<td>12</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>15</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td>11</td>
<td>0</td>
<td>2</td>
<td>13</td>
<td>8</td>
</tr>
<tr>
<td>11</td>
<td>12</td>
<td>0</td>
<td>2</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>12</td>
<td>13</td>
<td>1</td>
<td>0</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>13</td>
<td>15</td>
<td>0</td>
<td>1</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>14</td>
<td>16</td>
<td>1</td>
<td>0</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td>15</td>
<td>17</td>
<td>0</td>
<td>1</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>16</td>
<td>22</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>17</td>
<td>23</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>1</td>
</tr>
</tbody>
</table>
Log-rank Example

.stset time, failure(status=1)
.sts graph, by(group)
Log-rank Example

```
. sts test group

Log-rank test for equality of survivor functions
-----------------------------------------------
<table>
<thead>
<tr>
<th>Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>group</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

chi2(1) = 16.79
Pr>chi2 = 0.0000
```
What can we detect with the Log-rank Test?

We will be able to detect this ..

but not this
Generalizations of the Log-rank Test

• The observed and expected calculations can be extended naturally to more than two groups. The resulting log-rank test will have a $\chi^2$ random variable with $K - 1$ degrees of freedom (where $K$ is the number of groups) under $H_0$.

• When the $K$ groups are formed on the basis of an ordinal variable (i.e. are ordered) then a modified version of the log-rank can be used to test for trend (a 1 degree of freedom test). Cox regression (coming up) with a single covariate can be used to obtain an equivalent test.
**Weighted Log-rank Tests**

*Weighted* tests are available when *a priori* one has reason to emphasize some time intervals (i.e. earlier or later) over others.

**Idea** ($d_{1j}$ is obs. # deaths in group 1 at time $j$; $E_{1j}$ is exp. # deaths):

Log-Rank statistic based on: $\Sigma_j (d_{1j} - E_{1j})$

Weighted log rank statistic based on: $\Sigma_j w_j(d_{1j} - E_{1j})$

Possible weights:

- $w_j = 1 \implies$ Log-Rank test
- $w_j = R_j \implies$ Wilcoxon-Gehan-Breslow test
- $w_j = R_j^{1/2} \implies$ Tarone-Ware test
- $w_j = \hat{S}(t_j) \implies$ Peto-Prentice test

where

$$R_j = R_{1j} + R_{2j}, \text{ the total number at risk at time } t_j$$
Weighted Log-rank Tests

Comments:

• The log-rank test gives equal weight to all times, which emphasizes the tail of the survival curve (relatively).

• The Wilcoxon-Breslow and Peto-Prentice give more weight to earlier times, which emphasizes beginning of survival curve.

• Wilcoxon-Breslow and Tarone-Ware depend on censoring pattern. Log-rank and Peto-Prentice do not.

How to decide?

• Which is scientifically more important - early versus late?

• The log-rank test is the most powerful for detecting alternatives that correspond to proportional hazards (a common assumption when comparing survival curves)
Weighted Log-rank Example
Leukemia Data

```
. sts test group, logrank
Log-rank test for equality of survivor functions

<table>
<thead>
<tr>
<th>Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>group</td>
</tr>
<tr>
<td>observed</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

chi2(1) = 16.79
Pr>chi2 = 0.0000
```

```
. sts test group, wilcoxon
Wilcoxon (Breslow) test for equality of survivor functions

<table>
<thead>
<tr>
<th>Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>group</td>
</tr>
<tr>
<td>observed</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

chi2(1) = 13.46
Pr>chi2 = 0.0002
**Weighted Log-rank Example**

**Leukemia Data**

```stata
. sts test group, tw
Tarone-Ware test for equality of survivor functions

<table>
<thead>
<tr>
<th>tx</th>
<th>Events observed</th>
<th>Events expected</th>
<th>Sum of ranks</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>21</td>
<td>10.75</td>
<td>51.162748</td>
</tr>
<tr>
<td>1</td>
<td>9</td>
<td>19.25</td>
<td>-51.162748</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>30</td>
<td>30.00</td>
<td>0</td>
</tr>
</tbody>
</table>

chi2(1) = 15.12  
Pr>chi2 = 0.0001

. sts test group, p
Peto-Peto test for equality of survivor functions

<table>
<thead>
<tr>
<th>tx</th>
<th>Events observed</th>
<th>Events expected</th>
<th>Sum of ranks</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>21</td>
<td>10.75</td>
<td>6.3622095</td>
</tr>
<tr>
<td>1</td>
<td>9</td>
<td>19.25</td>
<td>-6.3622095</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>30</td>
<td>30.00</td>
<td>0</td>
</tr>
</tbody>
</table>

chi2(1) = 14.08  
Pr>chi2 = 0.0002
```
Test Statistics for Equality of Survival Distributions

<table>
<thead>
<tr>
<th>Statistic</th>
<th>df</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Rank</td>
<td>1</td>
<td>0.0000</td>
</tr>
<tr>
<td>Wilcoxon-Breslow</td>
<td>1</td>
<td>0.0002</td>
</tr>
<tr>
<td>Tarone-Ware</td>
<td>1</td>
<td>0.0001</td>
</tr>
<tr>
<td>Peto-Prentice</td>
<td>1</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

Weighted Log-rank Example

Leukemia Data
SUMMARY
Comparison of Survival Curves

1. We can compare survival probabilities at any single time, $t_0$, with a familiar 2-sample statistic.

2. We can compare the entire survival function for 2 groups using the log-rank test.

3. The log-rank test can easily be extended to $K$ groups ($K \geq 2$).

4. Alternative tests have been proposed that allow different weight to be given to earlier or later times.
Comparison of Survival Curves
Gotcha’s

1. Immortal time bias (survivor treatment selection bias)
   • Popes live longer than artists
2. Survivorship bias (left truncation bias)
   • WHI clinical trial vs observational study
3. Choice of time scale can be important
   • Study time vs calendar time vs time since disease onset vs time since diagnosis
   • Assumption that time “t” means the same for everyone
Hazard Functions and Models: Outline

- Hazard function
  - Definition
  - Relationship to incidence
  - Cumulative hazard
  - Relationship to survival function
- Cox regression
  - Proportional hazards assumption
  - “Semi-parametric” model
  - Estimation and inference
  - Estimation of baseline hazard/survival function
Recall:

\[ h(t) = \lim_{\Delta \to 0} \frac{P(t \leq T < t + \Delta \mid T \geq t)}{\Delta} \]

- Probability of an event in the next small time interval \((t; t + \Delta)\),
given survival until time \(t\), divided by the length of the time
interval, \(\Delta\).
- Conditional probability divided by \(\Delta\), as \(\Delta\) becomes very small.
- \(h(t)\) is a rate between 0 and \(+\infty\).
- \(h(t)\) depends on the unit of time.
- Special cases and synonyms:
  - hazard rate
  - force of mortality
  - instantaneous incidence rate
  - incidence rate
  - incidence density (where event is disease)
Hazard Functions

Definitions

- Survival function: \( S(t) = P(T > t) \)
- Density function: \( f(t) = -S'(t) = -dS(t)/dt \)
- Hazard function: \( h(t) = f(t)/S(t) \)
- Cumulative hazard function: \( H(t) = \int_0^t h(s) ds \)

- There is a one-to-one relationship between \( h(t) \) and \( S(t) \) (\( f(t) \) too!) - if you know any one then you know the others.
- High hazard (high risk of failure) is associated with rapidly declining portion survivor curve. Low hazard is associated with flat portion of survivor curve.

\[
S(t) = \exp(-\int_0^t h(s) ds) \]
\[
\frac{\partial}{\partial t} S(t) = -h(t) S(t) \]
\[
\frac{\partial}{\partial t} \log S(t) = -h(t) \]
Hazard Rate

- Hazard is a rate - probability of “failure” per unit time e.g.

<table>
<thead>
<tr>
<th>P(fail)</th>
<th>Time</th>
<th>Hazard Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/3</td>
<td>1/2 day</td>
<td>(1/3)/(1/2) = .67/day</td>
</tr>
<tr>
<td>1/3</td>
<td>1/14 week</td>
<td>(1/3)/(1/14) = 4.67/wk</td>
</tr>
</tbody>
</table>

Ave. Hazard Rate = number of events divided by the total exposure time

Example: Remission duration in acute leukemia

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Events</td>
<td>9</td>
</tr>
<tr>
<td>Time</td>
<td>359 weeks</td>
</tr>
</tbody>
</table>

Ave Hazard rate: 9/359 = .0251  21/182 = .1154

Note: The ratio of the average hazard rates is .0251/.1154 = .2175
**Hazard Rate**

- Average hazard rate is only part of the story, however. We also want to know how the hazard varies over time.
- We can estimate the hazard within subintervals.
- A simple estimate of the hazard is

\[
\hat{h}(t_i) = \frac{d_i}{R_i \Delta t_i}
\]

- Since this estimate typically is quite rough, a “smoothed” estimate of the hazard is often presented.
Hazard Rate: Example

Observed Death Times: 5, 11, 14, 21, 25, 32, 48
Censored Times: 2, 12, 25, 35

<table>
<thead>
<tr>
<th>time</th>
<th>$R_i$</th>
<th>$d_i$</th>
<th>$d_i/R_i$</th>
<th>$\Delta t_i$</th>
<th>h(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0.000</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>1</td>
<td>0.100</td>
<td>6</td>
<td>0.0167</td>
</tr>
<tr>
<td>11</td>
<td>9</td>
<td>1</td>
<td>0.111</td>
<td>3</td>
<td>0.037</td>
</tr>
<tr>
<td>14</td>
<td>7</td>
<td>1</td>
<td>0.143</td>
<td>7</td>
<td>0.020</td>
</tr>
<tr>
<td>21</td>
<td>6</td>
<td>1</td>
<td>0.167</td>
<td>4</td>
<td>0.042</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>1</td>
<td>0.200</td>
<td>7</td>
<td>0.028</td>
</tr>
<tr>
<td>32</td>
<td>3</td>
<td>1</td>
<td>0.333</td>
<td>16</td>
<td>0.021</td>
</tr>
<tr>
<td>48</td>
<td>1</td>
<td>1</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Hazard Rate: Example
Hazard Example
Leukemia Data

```
.sts graph, by(group) hazard
```

Smoothed hazard estimates, by group

Kaplan-Meier survival estimates, by group
Hazard Function – Parametric Models

• Just as we specified models for probability functions and probability densities, we can specify models for hazard functions.

• In fact, when we choose a probability model for a set of survival times we are implicitly choosing a hazard function, and vice versa (1-1 relationship between hazard and probability).

• The simplest hazard (probability) model for failure time data is the exponential model.

• Exponential model:
  - Constant hazard
  - “No memory” \( P(T \leq t + \Delta t \mid T \geq t) = P(0 \leq T \leq \Delta t) \)

\[
\begin{align*}
h(t) &= \lambda \\
S(t) &= \exp(-\lambda t) \\
f(t) &= \lambda \exp(-\lambda t)
\end{align*}
\]

Note: Sometimes written with parameter \( \mu=1/\lambda \)
Hazard Function – Exponential Model

- Hazard
- Cumulative Hazard
- Density
- Survival
**Hazard Function – Weibull Model**

Weibull model: \[ h(t) = \lambda \beta t^{\beta - 1} \]
\[ S(t) = \exp(-\lambda t^\beta) \]
\[ f(t) = \lambda \beta t^{\beta - 1} \exp(-\lambda t^\beta) \]

Key point: \( \beta > 1 \), hazard increases; \( \beta < 1 \), hazard decreases

\( \beta < 1 \)
Hazard Function – Weibull Model

$\beta = 2$

<table>
<thead>
<tr>
<th>Hazard</th>
<th>Density</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cumulative Hazard</td>
<td>Survival</td>
</tr>
</tbody>
</table>

Spring 2013  Biostat 513  353
Why Model the Hazard Function?

• Reasonable to expect that many factors may influence survival (and therefore, hazard).

• Kaplan-Meier can be used to characterize survival when there are a few large groups, but with multiple covariates we cannot stratify on all at once.

• How to consider effect of continuous covariates (without grouping)?

• How to incorporate time-dependent covariates?

Approach: A regression model for h(t)

• Cox (1972) proposed modeling the hazard function, h(t), in a seminal paper, “Regression Models and Life Tables (with Discussion)”

• Key parameter is the hazard ratio: h(t,X₁)/h(t,X₂)
Hazard Models

Additive Model:

\[ h(t, X) = h_0(t) + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p \]

- \( h_0(t) \) is the baseline hazard (like the intercept - but it is a curve, not a constant)
- Effect of covariates is additive on baseline rates
- Makes sense to think, e.g., that with diagnosis of breast cancer certain quantitative characteristics add to underlying “force of mortality”

Multiplicative Model (“proportional hazards”; “Cox model”):

\[
\begin{align*}
\log[h(t, X)] &= \log h_0(t) + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p \\
h(t, X) &= h_0(t) \exp(\beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p)
\end{align*}
\]

- Effect of covariates is multiplicative on baseline rates
- Multiplicative model guarantees positive hazard
Hazard Models – Additive vs Multiplicative

Additive Hazard

Multiplicative Hazard

Additive Hazard, log scale

Multiplicative Hazard, log scale

Additive Hazard

Multiplicative Hazard

Spring 2013  Biostat 513  356
Cox’s Proportional Hazards Model

1. With the PH model we can accommodate several covariates simultaneously.

2. The construction of the model and the interpretation of the terms in the model is much like linear regression and logistic regression, except now we model hazard ratios.

3. This type of model is known as “semi-parametric” since there is a part of the model that is parametric (the $\beta$ part), and part of the model that is left unspecified (nonparametric) (the $h_0(t)$ part).

4. Cox (1972) introduced the elegant “partial likelihood” method that allows estimation of the parameters of interest, $\beta$, without having to estimate the baseline hazard, $h_0(t)$.

5. Cox regression is used to compare different groups, formed on the basis of covariates, in terms of their instantaneous probability of failing at any time, $t$. 
Independence:
- Independent observations.
- Independent censoring.

Proportionality of hazards:
- Consider a single binary covariate:
  \( X = 1 \) if treated, and \( X = 0 \) is control group.
- The model
  \[
  h(t;X) = h_0(t) \exp(X\beta)
  \]
  implies that the risk of death among subjects in the treated group is \( \exp(\beta) \) times the risk of death among subjects in the control group at all times.
- Hazard ratio = \( h(t;X = 1)/h(t;X = 0) = \exp(\beta) \)
Example: Leukemia Remission Times

. use leukemia.dta
. stset time, failure(status)

  failure event:  status != 0 & status < .
obs. time interval:  (0, time]
exit on or before:  failure

------------------------------------------------------------------------------
42  total obs.
  0  exclusions
------------------------------------------------------------------------------
42  obs. remaining, representing
30  failures in single record/single failure data
541  total analysis time at risk, at risk from t =         0
    earliest observed entry t =         0
    last observed exit t =        35

. sts list

<table>
<thead>
<tr>
<th>Time</th>
<th>Beg. Total</th>
<th>Net Fail</th>
<th>Lost</th>
<th>Survivor</th>
<th>Std. Function</th>
<th>Error</th>
<th>Std. [95% Conf. Int.]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>42</td>
<td>2</td>
<td>0</td>
<td>0.9524</td>
<td>0.0329</td>
<td>0.8227</td>
<td>0.9879</td>
</tr>
<tr>
<td>2</td>
<td>40</td>
<td>2</td>
<td>0</td>
<td>0.9048</td>
<td>0.0453</td>
<td>0.7658</td>
<td>0.9631</td>
</tr>
<tr>
<td>3</td>
<td>38</td>
<td>1</td>
<td>0</td>
<td>0.8810</td>
<td>0.0500</td>
<td>0.7373</td>
<td>0.9486</td>
</tr>
<tr>
<td>4</td>
<td>37</td>
<td>2</td>
<td>0</td>
<td>0.8333</td>
<td>0.0575</td>
<td>0.6819</td>
<td>0.9168</td>
</tr>
<tr>
<td>5</td>
<td>35</td>
<td>2</td>
<td>0</td>
<td>0.7857</td>
<td>0.0633</td>
<td>0.6286</td>
<td>0.8822</td>
</tr>
<tr>
<td>6</td>
<td>33</td>
<td>3</td>
<td>1</td>
<td>0.7143</td>
<td>0.0697</td>
<td>0.5521</td>
<td>0.8265</td>
</tr>
<tr>
<td>20</td>
<td>10</td>
<td>0</td>
<td>1</td>
<td>0.3411</td>
<td>0.0774</td>
<td>0.1966</td>
<td>0.4909</td>
</tr>
<tr>
<td>22</td>
<td>9</td>
<td>2</td>
<td>0</td>
<td>0.2653</td>
<td>0.0765</td>
<td>0.1311</td>
<td>0.4204</td>
</tr>
<tr>
<td>23</td>
<td>7</td>
<td>2</td>
<td>0</td>
<td>0.1895</td>
<td>0.0710</td>
<td>0.0753</td>
<td>0.3431</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>0.1895</td>
<td>0.0710</td>
<td>0.0753</td>
<td>0.3431</td>
</tr>
<tr>
<td>32</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>0.1895</td>
<td>0.0710</td>
<td>0.0753</td>
<td>0.3431</td>
</tr>
<tr>
<td>34</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0.1895</td>
<td>0.0710</td>
<td>0.0753</td>
<td>0.3431</td>
</tr>
<tr>
<td>35</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.1895</td>
<td>0.0710</td>
<td>0.0753</td>
<td>0.3431</td>
</tr>
</tbody>
</table>
Example: Leukemia Remission Times

```
.sls graph, by(group)
```
Example: Leukemia Remission Times

. stcox group, nohr

Cox regression -- Breslow method for ties

No. of subjects = 42          Number of obs = 42
No. of failures = 30          Number of failures = 30
Time at risk = 541            Number of obs = 42
LR chi2(1) = 15.21
Log likelihood = -86.379622   Prob > chi2 = 0.0001

------------------------------------------------------------------------------
  _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  group |  -1.509191   .4095644    -3.68   0.000    -2.311923   -.7064599
------------------------------------------------------------------------------

. stcox

Cox regression -- Breslow method for ties

No. of subjects = 42          Number of obs = 42
No. of failures = 30          Number of obs = 42
Time at risk = 541            Number of obs = 42
LR chi2(1) = 15.21
Log likelihood = -86.379622   Prob > chi2 = 0.0001

------------------------------------------------------------------------------
  _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  group |   .2210887   .0905501    -3.68   0.000     .0990706    .4933877
------------------------------------------------------------------------------
Example: Leukemia Remission Times

. stcox group, nohr exactp

Cox regression -- exact partial likelihood

No. of subjects = 42  Number of obs = 42
No. of failures = 30
Time at risk = 541
Log likelihood = -74.543101  LR chi2(1) = 16.25
                  Prob > chi2 = 0.0001

                  _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  group |  -1.628244   .4331313    -3.76   0.000    -2.477166   -.7793222
-------------

. stcox group, nohr efron

Cox regression -- Efron method for ties

No. of subjects = 42  Number of obs = 42
No. of failures = 30
Time at risk = 541
Log likelihood = -85.008425  LR chi2(1) = 16.35
                  Prob > chi2 = 0.0001

                  _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  group |  -1.572125   .4123967    -3.81   0.000    -2.380408   -.7638424
-------------

Spring 2013  Biostat 513  362
Estimation of Hazard Ratio

Following Cox regression we can calculate an estimated hazard ratio (HR), comparing hazards at different covariate levels.

\[
    HR = \exp(\hat{\beta}(X_1 - X_0))
\]

- \(\hat{\beta}\) : estimated regression coefficient(s)
- \(X_j\) : covariates

We assume that the hazard ratio comparing \(X_1\) to \(X_0\) is constant over time.

Example: Leukemia remission

```
   . stcox group, exactp
```

Cox regression -- exact partial likelihood

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of subjects</td>
<td>42</td>
<td>Number of obs</td>
</tr>
<tr>
<td>No. of failures</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Time at risk</td>
<td>541</td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-74.543101</td>
<td>LR chi2(1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Prob &gt; chi2</td>
</tr>
</tbody>
</table>

| _t | Haz. Ratio   | Std. Err. | z    | P>|z| | [95% Conf. Interval] |
|----|--------------|-----------|------|-------|----------------------|
|    | group        |           |      |       |                      |
|    | .1962739     | .0850124  | -3.76| 0.000 | .0839809 .4587168    |

```
Estimation of Survivor Function $S(t,X)$

- The baseline survivor function, $S_0(t)$, can be estimated using a generalization of the K-M estimate to the case where the hazard function depends on covariates.
- The survivor function can then be estimated using the fitted PH model.
- The `stcox` option `basesurv(var)` saves an estimate of the baseline survivor function, $S_0(t)$, in the variable `var`.
- The `stcox` option `basech(var)` saves an estimate of the baseline cumulative hazard function, $H_0(t)$, in the variable `var`.
- Using $S_0(t)$ it is possible to estimate $S(t,X)$ for any $X$ since it can be shown that:

$$
\hat{S}(t,X) = \left[ \hat{S}_0(t) \right]^{\exp(X\hat{\beta})}
$$
Example: Leukemia Remission Times

```stata
   . stcox group, nohr basesurv(s0hat)
   . scatter s0hat time
```

![Graph showing baseline survivorship over time](image-url)
Example: Leukemia Remission Times

. stcox group, nohr basesurv(s0hat)
. generate shat = s0hat^exp(-1.509191*group)
. sts graph, by(group) addplot( (line shat time if group==0, sort lcol(red)) (line shat time if group==1, sort lcol(blue))) legend(off)
Example: Leukemia Remission Times

```
. stcoxkm, by(group) pred1opts(s(i)) pred2opts(s(i))
```
Recap on Cox PH Model

1. We assume that the hazard ratio comparing $X=1$ to $X=0$ is \textit{constant} over time.

2. There is no intercept in the PH model – the “intercept” is really the unspecified baseline hazard, $h_0(t)$

3. Given an estimate of the regression parameter, $\hat{\beta}$, and an estimate of the baseline survival function, $\hat{S}_0(t)$, we can obtain fitted survival functions for any value of $X$. 
1. Interpretation of the hazard.
2. $S(t) \Leftrightarrow h(t)$
3. Examples using common parametric models (exponential model, weibull model).
4. Cox proportional hazards model:
   \[ h(t;X) = h_0(t) \exp(\beta_1 X_1 + \beta_2 X_2 + \ldots) \]
5. Estimation and inference for hazard ratio regression parameters.
**Multiple Predictors**

**Example:** Remission duration in acute leukemia

Response = time until death or relapse.

Covariates = treatment group, WBC count, (sex).

**Models:**

Model 0 \[ \log[h(t,X)] = \log[h_0(t)] + \beta_2 \log(wbc) \]

Model 1 \[ \log[h(t,X)] = \log[h_0(t)] + \beta_1 Tx \]

Model 2 \[ \log[h(t,X)] = \log[h_0(t)] + \beta_1 Tx + \beta_2 \log(wbc) \]

Model 3 \[ \log[h(t,X)] = \log[h_0(t)] + \beta_1 Tx + \beta_2 \log(wbc) + \beta_3 Tx \times \log(wbc) \]
Example: Remission Duration

. generate wbc = log(wbc)
. recode wbc min/1.99=1 2.00/2.99=2 3.00/3.99=3 4.00/max=4
. label define wlab 1 "log(wbc) < 2.00" 2 "log(wbc) 2.00 - 2.99" 3 "log(wbc) 3.00 - 3.99" 4 "log(wbc) >= 4.00"
. label values wbc wlab
. table wbc

<table>
<thead>
<tr>
<th>wbc</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>log(wbc) &lt; 2.00</td>
<td>5</td>
</tr>
<tr>
<td>log(wbc) 2.00 - 2.99</td>
<td>20</td>
</tr>
<tr>
<td>log(wbc) 3.00 - 3.99</td>
<td>10</td>
</tr>
<tr>
<td>log(wbc) &gt;= 4.00</td>
<td>7</td>
</tr>
</tbody>
</table>

. sts graph, by(wbc)

Kaplan-Meier survival estimates, by wbc
### Example: Remission Duration

```stata
. sort group
. by group: summ logwbc

-> group = 0

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>logwbc</td>
<td>21</td>
<td>3.224286</td>
<td>.9722786</td>
<td>1.5</td>
<td>5</td>
</tr>
</tbody>
</table>

-> group = 1

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>logwbc</td>
<td>21</td>
<td>2.63619</td>
<td>.7738764</td>
<td>1.45</td>
<td>4.43</td>
</tr>
</tbody>
</table>

. *** Center log(WBC) at 3 before analysis
. gen logwbc3=logwbc-3
```
Example: Remission Duration

Model 0:

. stcox logwbc3, nohr exactp

Cox regression -- exact partial likelihood

LR chi2(1) = 35.23
Log likelihood = -65.056775                     Prob > chi2 = 0.0000

| t | Coef.  | Std. Err. | z    | P>|z| | [95% Conf. Interval] |
|---|--------|-----------|------|------|----------------------|
|   | logwbc3 | 1.737357  | 0.3232761 | 5.37 | 0.000 | 1.103748 - 2.370967 |

. est store model0

Model 1:

. stcox group, nohr exactp

LR chi2(1) = 16.25
Log likelihood = -74.543101                     Prob > chi2 = 0.0001

| t | Coef.  | Std. Err. | z    | P>|z| | [95% Conf. Interval] |
|---|--------|-----------|------|------|----------------------|
|   | group  | -1.628244 | 0.4331313 | -3.76 | 0.000 | -2.477166 - 0.7793222 |

. est store model1
### Example: Remission Duration

**Model 2:**

```stata
. stcox group logwbc3, nohr exactp
```

Log likelihood = -59.38471                         LR chi2(2)    =  46.57
Prob > chi2    =  0.0000

|         | Coef.  | Std. Err. |      z | P>|z| | [95% Conf. Interval] |
|---------|--------|-----------|--------|------|----------------------|
| group   | -1.444 | 0.45485   | -3.18  | 0.001| -2.3358    -0.5528   |
| logwbc3 | 1.763  | 0.35923   | 4.91   | 0.000| 1.0594    2.4675    |

. est store model2

**Likelihood Ratio Tests:**

```stata
. lrtest model1 model2
```

(Assumption: model1 nested in model2)

LR chi2(1) =  30.32  Prob > chi2 =  0.0000  H₀: ?

```stata
. lrtest model0 model2
```

(Assumption: model0 nested in model2)

LR chi2(1) =  11.34  Prob > chi2 =  0.0008  H₀: ?
Example: Remission Duration

Model 3:
\[ \text{. xi: stcox i.group*logwbc3, exactp nohr} \]

LR chi2(3) = 47.01
Log likelihood = -59.164688  Prob > chi2 = 0.0000

| t    | Coef.  | Std. Err. | z     | P>|z|   | [95% Conf. Interval] |
|------|--------|-----------|-------|-------|----------------------|
| _Igroup_1 | -1.48818 | 0.4646956 | -3.20 | 0.001 | -2.398967 to -0.5773934 |
| logwbc3 | 1.601659 | 0.4254097 | 3.76  | 0.000 | 0.7678715 to 2.435447  |
| _IgroXlogw-1 | 0.3801314 | 0.5709466 | 0.67  | 0.506 | -0.7389034 to 1.499166 |

. est store model3

Likelihood Ratio Test:

\[ H_0 : \]

. lrtest model2 model3

Likelihood-ratio test
(Assumption: model2 nested in model3)
LR chi2(1) = 0.44
Prob > chi2 = 0.5071
### Model Summary

<table>
<thead>
<tr>
<th>Model</th>
<th>$\exp(\beta_1)$</th>
<th>Log $L$</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.196</td>
<td>-74.54</td>
<td>151.09</td>
</tr>
<tr>
<td>2</td>
<td>0.236</td>
<td>-59.38</td>
<td>122.76</td>
</tr>
<tr>
<td>3</td>
<td>0.226*</td>
<td>-59.16</td>
<td>124.32</td>
</tr>
</tbody>
</table>

* for log(wbc) = 3.0

<table>
<thead>
<tr>
<th>Test</th>
<th>LR stat</th>
<th>df</th>
<th>p-val</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1 vs. null</td>
<td>16.25</td>
<td>1</td>
<td>0.0001</td>
</tr>
<tr>
<td>Model 2 vs. Model 0</td>
<td>11.34</td>
<td>1</td>
<td>0.0008</td>
</tr>
<tr>
<td>Model 3 vs. Model 2</td>
<td>0.44</td>
<td>1</td>
<td>0.5071</td>
</tr>
</tbody>
</table>

**Survival for Tx groups – adjusted for log(WBC):**

\[
\hat{S}(t, \text{Tx}=0, \log(\text{wbc})=3) = \hat{S}_0(t)
\]

\[
\hat{S}(t, \text{Tx}=1, \log(\text{wbc})=3) = [\hat{S}_0(t)]^{\exp(-1.444)}
\]
Model 2:
. stcox group logwbc3, nohr bases(s0hat) basech(H0hat) basehc(h0hat)
. stcurve, survival at1(group=0 logwbc3=0) at2(group=1 logwbc3=0)

Estimated survival for log(wbc) = 3 by treatment group  (0 = placebo; 1 = treated)
**Example: Remission Duration**

```
. stcurve, survival at1(group=0 logwbc3=1) at2(group=1 logwbc3=1)
```

Estimated survival for log(wbc) = 4 by treatment group (0 = placebo; 1 = treated)
Example: Remission Duration

```
. stcurve, survival at1(group=0 logwbc3=1) at2(group=1 logwbc3=0)
```

Compare survival for placebo patient with log(wbc) = 4 and treated patient with log(wbc)=3.
Consider two values for the covariates

\[ X^{(0)} = (X_1^{(0)}, X_2^{(0)}, \ldots, X_p^{(0)}) \]

\[ X^{(1)} = (X_1^{(1)}, X_2^{(1)}, \ldots, X_p^{(1)}) \]

Q: What is the hazard ratio comparing \( X^{(1)} \) to \( X^{(0)} \) if we use a PH model?

Model:

\[
\begin{align*}
    h(t, X) &= h_0(t) \exp(\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p) \\
    &= h_0(t) \exp(\sum_{j=1}^{p} X_j \beta_j)
\end{align*}
\]
**Estimating Hazard Ratios**

Hazard Ratio (HR): \( h(t, X^{(0)}) = h_0(t) \exp\left( \sum_{j=1}^{p} X_j^{(0)} \beta_j \right) \)

\( h(t, X^{(1)}) = h_0(t) \exp\left( \sum_{j=1}^{p} X_j^{(1)} \beta_j \right) \)

\[
HR = \exp\left( \sum_{j=1}^{p} X_j^{(1)} \beta_j - \sum_{j=1}^{p} X_j^{(0)} \beta_j \right) 
= \exp\left( \sum_{j=1}^{p} \beta_j \left( X_j^{(1)} - X_j^{(0)} \right) \right)
\]

Example: Remission Data, Model 3 (Interaction)

\( X^{(1)} = (\text{Group} = 1, \log\text{wbc} = 3) \)

\( X^{(0)} = (\text{Group} = 0, \log\text{wbc} = 4) \)

\[
\hat{HR} = \frac{\exp(-1.44(1) + 1.76(0))}{\exp(-1.44(0) + 1.76(1))} = .0407
\]
Example: Remission Duration

What is the estimated hazard ratio comparing these two patients: treated patient with log(wbc)=3 and placebo patient with log(wbc) = 4?

```
. lincom group-logwbc3
( 1)  group - logwbc3 = 0

                  _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
(1) |  -3.207747   .6051697    -5.30   0.000    -4.393858   -2.021636
-------------
```

```
. lincom group-logwbc3, hr
( 1)  group - logwbc3 = 0

                  _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
(1) |   .0404476   .0244777    -5.30   0.000      .012353    .1324386
-------------
```
SUMMARY

1. Adjust for **confounding** as in other regression models – is there a meaningful change in the summary of interest (hazard ratio) after controlling for the potential confounder(s)?

2. Use **Wald** and **Likelihood ratio** statistics to test whether certain coefficients are zero (including tests for EM).

3. Preferable to center covariates to enhance interpretation and alleviate collinearity

4. Use estimated PH regression coefficients to obtain risk comparisons in terms of hazard ratios.

5. Use the estimated PH regression coefficients and the estimate of the baseline survival, \( \hat{S}_0(t) \), to obtain an estimate of the survival function for any covariate value, \( X \).

6. We assume that the hazards are **proportional** in \( t \) across the values of each covariate.
Assessing Model Adequacy

• Proportional hazards
  - Graphical methods
  - Analytical methods

• Residuals (Biost 537)
  - Schoenfeld residuals
    • proportional hazards
  - Martingale and deviance residuals
    • functional form for covariates
    • leverage and outliers
Checking for Proportionality

- Graphical approaches
  - plots of $-\log\{-\log[S(t,X)]\}$
  - Observed and fitted $S(t,X)$
  - Residual plots (Biost 537)
- Confirmatory approaches
  - Test for proportionality
- Correction for failure of PH assumption
  - Stratification
  - Add \textit{covariate} \times (\log) time to the model (See Biostat 537)
-Log-log Plots

Recall: Under a PH assumption

\[ S(t, X) = [S_0(t)]^{\exp(\beta X)} \]
\[ \log[S(t, X)] = \exp(\beta X) \cdot \log[S_0(t)] \]
\[ \log\{-\log[S(t, X)]\} = \beta X + \log\{-\log[S_0(t)]\} \]

This implies that the separation between log-log plots should be constant over time:

\[ \beta = \log\{-\log[S(t, X=1)]\} - \log\{-\log[S(t, X=0)]\} \]

Idea:

• Plot \(-\log\{-\log[\hat{S}(t, X)]\}\) versus time or ln(time) and assess whether the curves are “parallel”.
Example: Remission Duration

- Check PH assumption for Group

```
.stphplot, by(group)
```

![Graph showing the PH assumption for remission duration](image)
Example: Remission Duration

- Check PH assumption for logwbc (using categories)

```
stphplot, by(wbccat)
```

![Graph showing survival analysis with PH assumption check for logwbc using categories.](image)

- Limitation: 

```
wbccat = log(wbc) < 2.00
wbccat = log(wbc) 2.00 - 2.99
wbccat = log(wbc) 3.00 - 3.99
wbccat = log(wbc) >
```
Example: Remission Duration

- Check PH assumption for sex (also in dataset, not “significant” in PH model, but …)

```
stphplot, by(sex)
```

- PH assumption not okay
Example: Remission Duration

.sts graph, by(sex) haz

Smoothed hazard estimates, by sex

analysis time

sex = 0  sex = 1
-Log-log Plots

Comments:

- \( -\log(-\log[\hat{S}(t,X)]) \) or \( \log(-\log[\hat{S}(t,X)]) \)
- Plot against time (or \( \ln(\text{time}) \))
- Use Kaplan-Meier for \( \hat{S}(t,X) \)
- Crossing (in middle) or convergence is an indication of violation of PH assumption
- Interpret plots recognizing that there is variation since these are estimates of the survival functions.

Issues:

- How parallel is parallel?
  - subjective decision
- Categorization of continuous predictors
Idea:

- Compare Kaplan-Meier estimates to fitted survival curves obtained from Cox regression.

Issues:

- If we adjust for other predictors in the Cox regression then we may impact the fitted survival. This can make comparison to KM estimates difficult (unless we can adjust those as well).
- How close is close?
  - subjective decision
- Continuous covariates
Example: Remission Duration

```
stcoxkm, by(group)
```

![Survival curve for remission duration](image-url)
Example: Remission Duration

```
.stcoxkm, by(wbccat)
```

- Plots with multiple categories are often less useful unless well-separated, as here
Example: Remission Duration

```
.stcoxkm, by(sex)
```

```
Observed: sex = female
Predicted: sex = female
Observed: sex = male
Predicted: sex = male
```

- PH not okay?
Tests for Proportional Hazards

• STATA (and other packages) now include hypothesis tests for proportionality of hazards.

• Such tests are obtained from a fitted Cox regression and test the proportional hazards assumption:

\[ H_0 : \beta_j(t) = \beta_j \]
\[ H_1 : \beta_j(t) \text{ has a trend in time} \]

• Here \( \exp(\beta_j(t)) \) represents the hazard ratio comparing \( X_j = 1 \) to \( X_j = 0 \) at time \( t \), controlling for other predictors in the model.

\[
\frac{h(t, X_1 = 1, X_2 = x_2)}{h(t, X_1 = 0, X_2 = x_2)} = \frac{h_0(t) \exp(\beta_1(t) \cdot 1 + \beta_2 x_2)}{h_0(t) \exp(\beta_1(t) \cdot 0 + \beta_2 x_2)}
\]

\[ = \exp(\beta_1(t)) \]

• These tests use “Schoenfeld” residuals
Example: Remission Duration

```
. stcox group logwbc3, nohr efron schoenfeld(SCH*) scaledsch(SCA*)
Cox regression -- Efron method for ties
No. of subjects =           42                     Number of obs   =        42
No. of failures =           30Time at risk    =          541
LR chi2(2)      =     46.71
Log likelihood  =   -69.828101                     Prob > chi2     =    0.0000
------------------------------------------------------------------------------
     _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+---------------------------------------------------------------
    group  |  -1.386075   .4247984    -3.26   0.001    -2.218665   -.5534859
  logwbc3  |    1.69089   .3358976     5.03   0.000     1.032543    2.349238
------------------------------------------------------------------------------
. estat phtest, detail
Test of proportional-hazards assumption
Time:  Time
         |        rho   chi2 df    Prob>chi2
-------------+------------------------------------------
        group |   -0.03047   0.02  1      0.8751
      logwbc3 |    0.03923   0.07  1      0.7922
 global test |        0.09   2       0.9569
```
Example: Remission Duration

.stcox sex, nohr efron

  failure _d:  status
  analysis time _t:  time

Cox regression -- Efron method for ties

No. of subjects = 42                     Number of obs = 42
No. of failures = 30                      Time at risk = 541
LR chi2(1) = 0.60                         Prob > chi2 = 0.4396
Log likelihood = -92.885605               Prob > chi2 = 0.4396

------------------------------------------------------------------------------
   _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
------------------------------------------------------------------------------
    sex |  -.3101396   .4042745    -0.77   0.443    -1.102503    .4822238
------------------------------------------------------------------------------
.
.estat phtest

  Test of proportional-hazards assumption

  Time:  Time

  +-------------------------------+----------------+----------------+----------------+
  |                     chi2  df  Prob>chi2 |
  +-------------------------------+----------------+----------------+
  global test |       10.90  1    0.0010
  +-------------------------------+----------------+----------------+
Example: Remission Duration

```
estat phtest, plot(sex)
```
Example: Remission Duration

```
. stcox tx logwbc3 sex, nohr efron

  failure _d:  status  
analysis time _t:  time

Cox regression -- Efron method for ties

No. of subjects =         42                     Number of obs =        42
No. of failures =         30                     LR chi2(3)      =     47.19
Time at risk    =          541                     Prob > chi2     =    0.0000

Log likelihood  =   -69.590483                     LR chi2(3)      =     47.19
                     Prob > chi2     =    0.0000

------------------------------------------------------------------------------
                   _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       tx |  -1.503591   .4615127    -3.26   0.001    -2.408139   -.5990429
       logwbc3 |   1.681942   .3365836     5.00   0.000     1.022251    2.341634
       sex |    .314678   .4545115     0.69   0.489    -.5761482    1.205504
------------------------------------------------------------------------------

. estat phtest, detail

Test of proportional-hazards assumption

Time:  Time

<table>
<thead>
<tr>
<th>rho</th>
<th>chi2</th>
<th>df</th>
<th>Prob&gt;chi2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tx</td>
<td>0.05700</td>
<td>0.11</td>
<td>1</td>
</tr>
<tr>
<td>logwbc3</td>
<td>0.06864</td>
<td>0.21</td>
<td>1</td>
</tr>
<tr>
<td>sex</td>
<td>-0.29209</td>
<td>2.56</td>
<td>1</td>
</tr>
<tr>
<td>global test</td>
<td>2.82</td>
<td>3</td>
<td>0.4206</td>
</tr>
</tbody>
</table>
```

Example: Remission Duration

.estat phtest, plot(sex)

Test of PH Assumption

scaled Schoenfeld - sex

0 2 4 6

0 5 10 15 20 25

Time

bandwidth = .8
Example: Remission Duration

Is the relationship with logwbc linear?

. gen lwbc2=logwbc3^2

. stcox group logwbc3 lwbc2, nohr

LR chi2(3)      =     44.50
Log likelihood  =   -71.73582                     Prob > chi2     =    0.0000

------------------------------------------------------------------------------
     _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      group |  -1.366605   .4303963    -3.18   0.001    -2.210167   -.5230442
     logwbc3 |   1.510339   .3221063     4.69   0.000     .8790224    2.141656
       lwbc2 |   .2710913   .2558792     1.06   0.289    -.2304227    .7726052
------------------------------------------------------------------------------

Example: Remission Duration
## Model Checking Summary
### Cox PH Model

<table>
<thead>
<tr>
<th>Assumption</th>
<th>Technique</th>
<th>Stata</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>PH</td>
<td>-log-log plot</td>
<td><code>stphplot</code></td>
<td>graphical, cat. covars only</td>
</tr>
<tr>
<td></td>
<td>obs v pred KM</td>
<td><code>stcoxkm</code></td>
<td>graphical, cat. covars only</td>
</tr>
<tr>
<td></td>
<td>$H_0: \beta(t) = \beta$</td>
<td><code>estat phtest</code></td>
<td>test each cov., any kind of cov.</td>
</tr>
</tbody>
</table>
What to do if PH fails?

Time dependent covariates:
- Interaction between covariate and t or log t
- Smoothed estimates of $\beta(t)$
- See Biost 537 and other advanced courses

Stratification:
- Discrete (grouped) covariates only
- Separate baseline hazard for each covariate class
- Graphical analysis – no quantitative estimate of HR
Stratified Cox Model

Remission duration study also includes information on gender

```
. tab sex

<table>
<thead>
<tr>
<th>sex</th>
<th>Freq.</th>
<th>Percent</th>
<th>Cum.</th>
</tr>
</thead>
<tbody>
<tr>
<td>female</td>
<td>22</td>
<td>52.38</td>
<td>52.38</td>
</tr>
<tr>
<td>male</td>
<td>20</td>
<td>47.62</td>
<td>100.00</td>
</tr>
<tr>
<td>Total</td>
<td>42</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>
```

```
. stcox group logwbc3 sex, nohr efron scaledsch(SCA*)
LR chi2(3) = 47.19
Log likelihood = -69.590483
Prob > chi2 = 0.0000

| _t | Coef.  | Std. Err. | z     | P>|z| | [95% Conf. Interval] |
|----|--------|-----------|-------|-----|----------------------|
| group | -1.5036 | 0.4615 | -3.26 | 0.001 | -2.408139 to -0.5990429 |
| logwbc3 | 1.6819 | 0.3366 | 5.00 | 0.000 | 1.022251 to 2.341634 |
| sex | 0.3147 | 0.4545 | 0.69 | 0.489 | -0.5761482 to 1.205504 |
```

```
. estat phtest, detail log

<table>
<thead>
<tr>
<th>Test of proportional-hazards assumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time: Log(t)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>rho</th>
<th>chi2</th>
<th>df</th>
<th>Prob&gt;chi2</th>
</tr>
</thead>
<tbody>
<tr>
<td>group</td>
<td>0.1331</td>
<td>0.59</td>
<td>1</td>
<td>0.4431</td>
</tr>
<tr>
<td>logwbc3</td>
<td>0.0807</td>
<td>0.30</td>
<td>1</td>
<td>0.5867</td>
</tr>
<tr>
<td>sex</td>
<td>-0.3620</td>
<td>3.94</td>
<td>1</td>
<td>0.0472</td>
</tr>
</tbody>
</table>
```
Stratified Cox Model

```
.stphplot, by(sex)
```

![Graph showing stratified Cox model analysis time vs. -ln(-ln(Survival Probability)) for different sexes.](image)
**Stratified Cox Model**

**Q:** Evidence is against PH assumption for sex (though less strong in multivariate model). If we don’t accept PH assumption on sex, can we still make PH inference about group and logwbc even if sex does not satisfy the PH assumption?

**A:** Yes. In order to do this we can perform a “stratified” analysis. This is different from using dummy variables, and is different from using separate analyses by gender.
Stratified Cox Model

Idea:

We can use a model where, within each gender, we have the same PH model, but we allow men and women to have different baseline hazards:

- women: \( h(t,X) = h_{0,0}(t) \exp(\beta_1 \cdot \text{group} + \beta_2 \cdot \text{logwbc3}) \)
- men: \( h(t,X) = h_{0,1}(t) \exp(\beta_1 \cdot \text{group} + \beta_2 \cdot \text{logwbc3}) \)

Such a model is said to be “stratified on sex”. It is analogous to allowing interactions between sex and time, but we still have additive terms for the other covariates (no dependence on time).
**Stratified Cox Model**

**Proportional Hazards Model:**

\[ \log[h(t, X)] = \log[h_0(t)] + \beta_1 \text{group} + \beta_2 \log \text{wbc3} + \beta_3 \text{sex} \]

**Stratified Cox Model:**

\[ \log[h(t, X)] = \log[h_{0,sex}(t)] + \beta_1 \text{group} + \beta_2 \log \text{wbc3} \]

**Q:** What’s the interpretation of \( \beta_1 \) in each model?
Stratified Cox Model

Proportional Hazards:

F: \( h(t, X) = h_0(t) \exp(\beta_1\text{group} + \beta_2\text{logwbc3}) \)

M: \( h(t, X) = h_0(t) \exp(\beta_1\text{group} + \beta_2\text{logwbc3} + \beta_3) \)

Separate Models:

F: \( h(t, X) = h_{0,1}(t) \exp(\beta_1^{(1)}\text{group} + \beta_2^{(1)}\text{logwbc3}) \)

M: \( h(t, X) = h_{0,2}(t) \exp(\beta_1^{(2)}\text{group} + \beta_2^{(2)}\text{logwbc3}) \)

Stratified Model #1:

F: \( h(t, X) = h_{0,F}(t) \exp(\beta_1\text{group} + \beta_2\text{logwbc3}) \)

M: \( h(t, X) = h_{0,M}(t) \exp(\beta_1\text{group} + \beta_2\text{logwbc3}) \)
Stratified Cox Model

Stratified Model #2:

\[ h(t, X) = h_{0,sex}(t) \exp(\beta_{1\text{group}} + \beta_{2\text{logwbc3}} + \beta_{3\text{group} \times \text{sex}}) \]

F: \[ h(t, X) = h_{0,F}(t) \exp(\beta_{1\text{group}} + \beta_{2\text{logwbc3}}) \]

M: \[ h(t, X) = h_{0,M}(t) \exp(\beta_{1\text{group}} + \beta_{2\text{logwbc3}} + \beta_{3\text{group}}) \]
### Example: Remission Duration

**Proportional hazards model**

```
. stcox group logwbc3 sex, efron nohr

No. of subjects =           42                     Number of obs =        42
No. of failures =           30Time at risk    =          541
LR chi2(3)      =     47.19
Log likelihood  =   -69.590483                     Prob > chi2     =    0.0000

------------------------------------------------------------------------------
     _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      group |  -1.503591   .4615127    -3.26   0.001    -2.408139   -.5990429
     logwbc3 |   1.681942   .3365836     5.00   0.000     1.022251    2.341634
         sex |    .314678   .4545115     0.69   0.489    -.5761482    1.205504
------------------------------------------------------------------------------
```
Example: Remission Duration

Separate Models

Females

. stcox group logwbc3 if sex==0, nohr efron

Log likelihood = -33.090979
LR chi2(2) = 6.65
Prob > chi2 = 0.0361

|     _t  |      Coef.   Std. Err.   z   P>|z|   [95% Conf. Interval] |
|-------|--------------|-----------------|-----|----------|-----------------|
|  group |  -.3112706   .5635539  -0.55  0.581    -1.415816    .7932747 |
| logwbc3|  1.206146    .5034893   2.40  0.017     .2193255    2.192967 |

Males

. stcox group logwbc3 if sex==1, nohr efron

Log likelihood = -20.760908
LR chi2(2) = 29.18
Prob > chi2 = 0.0000

|     _t  |      Coef.   Std. Err.   z   P>|z|   [95% Conf. Interval] |
|-------|--------------|-----------------|-----|----------|-----------------|
|  group |  -1.977887   .739202  -2.68  0.007    -3.426697   -.5290782 |
| logwbc3|  1.742777    .5357723   3.25  0.001     .6926825    2.792871 |
Stratified Cox Model – Remission Duration

Stratified #1

```
  . stcox group logwbc3, strata(sex) efron nohr bases(S) basech(H)
  LR chi2(2)  =  32.06
  Log likelihood  = -55.734815                     Prob > chi2     = 0.0000

  ------------------------------------------------------------------------------
   _t    |     Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
  -------------+----------------------------------------------------------------
    group |  -0.998104   .4735546    -2.11   0.035    -1.926254   -.0699538
   logwbc3 |   1.453654   .3440687     4.22   0.000     .7792919    2.128017
  ------------------------------------------------------------------------------
```

Stratified by sex

```
  . estimates store model1
  Stratified #1

  Stratified #2

  . gen txsex=group*sex
  . stcox group logwbc3 txsex, strata(sex) nohr efron
  LR chi2(3)  =  35.28
  Log likelihood  = -54.126889                     Prob > chi2     = 0.0000

  ------------------------------------------------------------------------------
   _t    |     Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
  -------------+----------------------------------------------------------------
    group |  -0.286573   .5685327    -0.50   0.614    -1.400877    .8277308
   logwbc3 |   1.472627   .3517843     4.19   0.000     .7831426    2.162112
   txsex  |  -1.642102   .9140899    -1.80   0.072    -3.433685    .1494813
  ------------------------------------------------------------------------------
```

Stratified by sex

```
  . estimates store model2
  Stratified #2

  . lrtest model1 model2
  Likelihood-ratio test                                  LR chi2(1)  =  3.22
  (Assumption: model1 nested in model2)                  Prob > chi2 = 0.0729
```
Time Varying Covariates

- Stratification allows us to make inferences about everything except the stratification variable. In effect, we have adjusted away the effect of sex.
- Suppose we wanted to make inferences on sex. Is it possible to model the time dependent HR?
- Yes! Include an interaction between sex and time.

Note: A different, but related, situation arises when we have a covariate for which we have measured changes over time e.g. sexual behavior, CD4 count
**Time Varying Covariates**

Time-independent model:

\[ \log h(t; X) = \log h_0(t) + \beta_1 \text{sex} \]

Time-dependent model:

\[ \log h(t; X) = \log h_0(t) + \beta_1 \text{sex} + \beta_2 \text{sex} \ast t \]

(note: no main effect for \( t \) – it is part of \( \log h_0(t) \))

What is the HR for sex?

\[
\frac{h(t; \text{sex} = 1)}{h(t; \text{sex} = 0)} = \frac{h_0(t) \exp(\beta_1 + \beta_2 t)}{h_0(t) \exp(0)} = \exp(\beta_1 + \beta_2 t)
\]

\( \beta_2 \) positive \( \Rightarrow \) HR increases over time

\( \beta_2 \) negative \( \Rightarrow \) HR decreases over time
Stratified Cox Model – Remission Data

```
. stcox tx logwbc sex, nohr efron tvc(sex) texp(_t)
```

Cox regression -- Efron method for ties

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of subjects</td>
<td>42</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of failures</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time at risk</td>
<td>541</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-67.641012</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LR chi2(4)</td>
<td>51.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prob &gt; chi2</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| _t | Coef. | Std. Err. | s   | D>|s| | [95% Conf. Interval] |
|----|-------|-----------|-----|--------|---------------------|
| main |       |           |     |         |                     |
| tx  | -1.186389 | .4828406 | -2.46 | 0.014 | -2.132739 | -2.2400388 |
| logwbc | 1.535217 | .3405832 | 4.51 | 0.000 | .8676864 | 2.202748 |
| sex | 1.727716 | .8966549 | 1.93 | 0.054 | -.8296948 | 3.485128 |
| tvc |       |           |     |         |                     |
| sex | -.1655309 | .0904991 | -1.83 | 0.067 | -.3429058 | .011844 |

Note: variables in tvc equation interacted with _t
Stratified Cox Model – Remission Data

\[ H_{\text{sex}}(t = 5) = \]
lincom [main]sex + 5*[tv{c}]sex, hr

|       | Haz. Ratio | Std. Err. | z     | P>|z|   | [95% Conf. Interval] |
|-------|------------|-----------|-------|-------|---------------------|
| (1)   | 2.459755   | 1.397205  | 1.58  | 0.113 | .8079549 – 7.488532 |

\[ H_{\text{sex}}(t = 25) = \]
lincom [main]sex + 25*[tv{c}]sex, hr

|       | Haz. Ratio | Std. Err. | z     | P>|z|   | [95% Conf. Interval] |
|-------|------------|-----------|-------|-------|---------------------|
| (1)   | .0897653   | .1417142  | -1.53 | 0.127 | .0040673 – 1.981143 |

Overall LR test for sex:
\text{stcox tx logwbc, nohr efron}
estimates store model0
\text{stcox tx logwbc sex, nohr efron tv{c}sex}
estimates store model1
\text{lrtest model0 model1}

Likelihood-ratio test
(Assumption: model0 nested in model1) LR chi2(2) = 4.37 Prob > chi2 = 0.1122
**Survival Analysis and Sample Size**

**Q:** What are the considerations for determining the sample size necessary when the study endpoint is a time-until-event?

**Planned Analysis:**

- Assessment of percent surviving beyond $t^*$.
  - Need to know $\alpha$, power, $S(t^*, X=0)$, $S(t^*, X=1)$
  - Comparison of proportions (see STATA sampsi!)
- Assessment of survival function and/or hazard ratio.
  - Log-rank / Cox regression.
- Power depends on events only! $L = \left(\frac{Z_{1-\alpha/2} + Z_{1-\beta}}{\log(HR)/2}\right)^2$
  - $L$ is number of events (i.e. deaths)
  - $HR$ is relative hazard.
- $N/arm$ based on expected event rate, duration of follow-up and $HR$. 
Sample Size Example

HPTN 039

Assume:
- 2 treatment arms with $N$ subjects each
- $h_0 = 0.035/yr$ with $HR = 0.5$
- $f = 1$ year follow-up
- $\alpha = 0.05$, Power $= 1 - \beta = 0.90$

Compute:
- $L = \left[\frac{(1.96+1.28)}{(-0.3466)}\right]^2 = 87.4$
- $N = \frac{L}{h_0 \times f \times (1 + HR)} = 87.4/\left[(0.035)(1)(1.5)\right] = 1665$
- Inflate $N$ further for expected lost to follow-up
- Compare to 1865/arm (sampsi) for $p_0 = .035$ to $p_1 = 0.0175$ at 1 year.
- Less censoring >> greater savings in sample size
Specifics

Categorical Data and Stratification

Logistic Regression:

• Fitting and testing the model
  • Maximum likelihoods, likelihood ratio tests, Wald tests
  • Interpreting the model results (odds ratios)
  • Confounding, effect modification
• Prediction
  • Models give predicted probabilities
  • Evaluation of a “good fitting” models
    • Sensitivity, specificity, ROC curve, AUC
Survival (time-to-event) Analysis:

- Outcome consists of (time, status), subjects at risk (at time, \( t \))
- Censoring (informative vs independent) and truncation
- Survival summaries
  - Hazard, \( h(t) \) = rate, instantaneous risk of failure of those at risk
  - Survival, \( S(t) \) = probability of remaining “alive” at time \( t \)
    - Estimate using Kaplan-Meier estimate
      - At a given point in time (2 yr) or a given \( S(t) \) (median)
      - Confidence intervals (Greenwood)
  - \( S(t) \leftrightarrow h(t) \leftrightarrow f(t) \)
  - Comparing survival curves (log-rank, weighted log-rank tests)
Survival (time-to-event) Analysis:

- Cox proportional hazards regression, \( h(t \mid x) = h_0(t) \exp(x \beta) \)
- \( h_0(t) \) is an arbitrary “baseline hazard” curve
- Proportional hazards assumption – HR is constant in time
- Fitting and testing the model
  - Partial likelihood: likelihood ratio tests, Wald tests
- Interpreting the model results
  - HR = \( e^{(X1-X0)\beta} \) comparing groups with covariates \( X1 \) and \( X0 \)
  - Survival curves – \( S(t \mid x) = S_0(t)^{\exp(HR)} \)
  - Confounding and effect modification
Biostat 513 Review

Survival (time-to-event) Analysis:

- Cox proportional hazards regression, $h(t \mid x) = h_0(t) \exp(x\beta)$
- Checking model assumptions:
  - Log-log plots of $S(t)$
  - Observed (KM) vs predicted (Cox) survival curves
  - Use of Schoenfeld residuals (to test PH assumption)
- Remedies
  - Stratification (for PH assumption)
    - Analogous to time by predictor interactions
  - Time varying covariates, time dependent predictors
Final Exam

When: Wednesday, June 12, 2013

Where: HSB T-439

- Exam will be closed notes, closed book
- You can bring
  - one 8.5” by 11” sheet of notes
  - Midterm review sheet
  - Final review sheet
  - Hand calculator
  - *Devices with internet access capability are not permitted*