Survival Analysis

Survival Analysis

- Survival Data Characteristics
- Goals of Survival Analysis
- Statistical Quantities
  - Survival function
  - Hazard function
  - Cumulative hazard function
- One-sample Summaries
  - Kaplan-Meier Estimator
  - S.E. Estimation for $\widehat{S}(t)$
  - Life Table Estimation
- Two-sample Summaries
  - Mantel-Haenszel / Log-rank Test
  - Other tests – what? why?
- Regression Methods – Cox Regression
  - Proportional hazards
  - Interpretation of coefficients
  - Estimation & Testing
  - Survival function estimation

Motivation

Example:

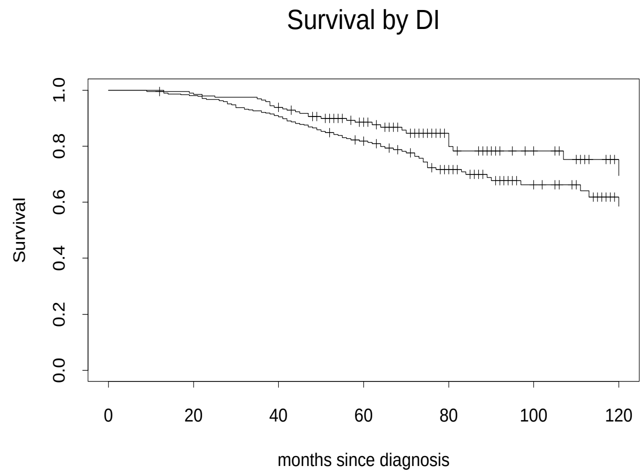
On a subsample of women from a cohort study of breast cancer patients we take new histologic measurements and want to assess the prognostic utility of these measurements.

- Primary Predictor(s): DI, p27 measurement (categorized)
- Other Predictors: stage, lymph nodes, size ...

◦ Outcome(s):

Time-until-death
Death (yes/no)

◦ Issue: most women are not observed until death.



Need a new method?

Q: Why not just use standard linear regression, perhaps taking a log transformation, to analyze the follow-up times?

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

$\boxed{\bullet}$ Useful to have methods that consider (time, status) as the outcome variable.

Survival Data Characteristics

$\boxed{\text{Outcome:}}$ (time, status)

- Time

  Time until an event occurs

  - Define the start time

    diagnosis
    entry into the study
    birth

  - Define the event

    death
    machine failure
    relapse
    discharge

- Event Indicator (status)

  $\delta = 1$ means an event was observed!

  $\delta = 0$ means the time was <u>censored</u>

  study ends before event observed
  patient withdraws / moves
  lost to follow-up

Survival Data

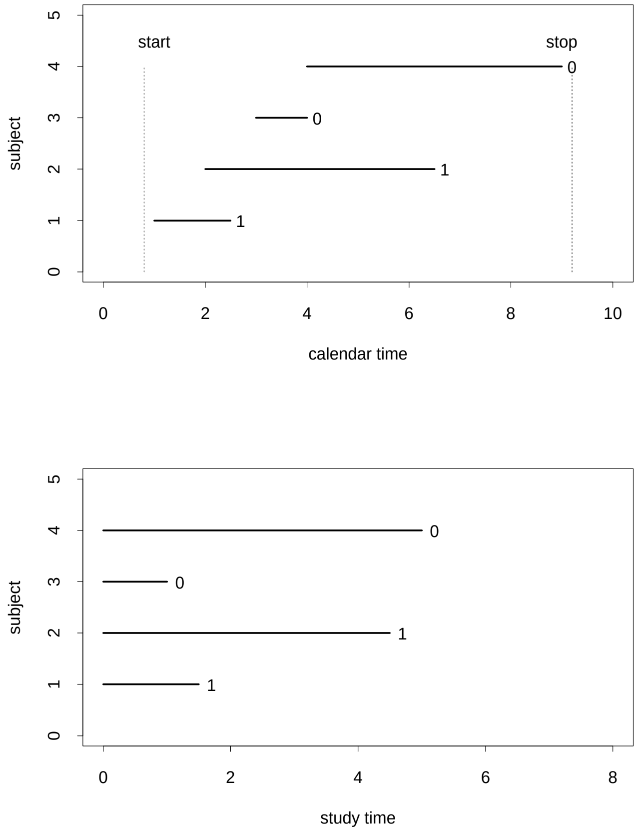
Example: Breast Cancer Histology Data

time	status	aneuploid	s-phase
49	1	1	22.4
73	0	1	6.1
68	0	0	0.8
72	0	0	7.8
80	0	1	4.4
70	0	0	11.1
9	1	0	14.9
72	0	0	14
77	0	0	0.4

(time,status) = (49,1) means:

(time,status) = (73,0) means:

Right Censoring



It's life and death...

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:

$$P[\, T \; < t + \Delta \;\; | \, T \geq t] \quad \approx \quad h(t) \cdot \Delta$$

$$\lim_{\Delta \to 0} \frac{P[\, T \; < t + \Delta \;\; | \, T \geq t]}{\Delta} \quad = \quad h(t)$$

The **hazard function** is the *instantaneous* probability of having an event at time t (per unit time) given that one has survived (ie. not had an event) up to time t.

Estimation of Survival

No Censoring: The job is easy here!

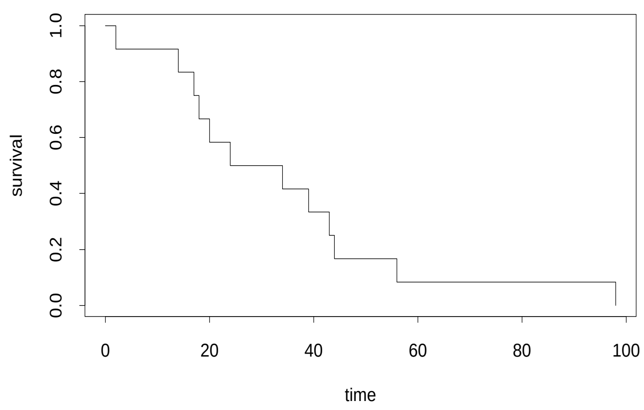
$$\begin{aligned} N &= \text{total number of subjects} \\ n(t) &= \text{number of subjects with } \; T_i > t \end{aligned}$$

$$\widehat{S}(t) \;\; = \;\; \frac{n(t)}{N}$$

<u>Example</u>:

```
N = 12   Median = 29
Quartiles = 17.5, 43.5

Decimal point is 1 place to the right of the colon

   0 : 2
   1 : 478
   2 : 04
   3 : 49
   4 : 34
   5 : 6

High:  98
```

Example w/o censoring



Survival with Censoring

Q: How can we include information from observations like 25+ which we represent as (25,0)?

A: The Kaplan-Meier Estimator.

Before we get to the details of the Kaplan-Meier estimator we'll want to consider an example from *current life tables* that shows us how we can "piece together" survival information.

Example:

Consider information collected in 1989 and 1994 that recorded the age of children in 1989 and then visited them in 1994 to ascertain their survival.

Data:

Age	number	deaths in 5 years	prob. survive 5 years	survive to age
0	200	40	0.800	1.000
5	100	15	0.850	0.800
10	100	10	0.900	0.680
15	100	10	0.900	0.612
20	150	10	0.933	0.558

Conditional Probability

This example shows that we can estimate the probability $P[T > 20]$ by putting together conditional survival probabilities over shorter intervals. Essentially we have

$$\begin{aligned} P[T > 20] &= (1 - P[\text{die by } 20 \mid T > 15]) \cdot P[T > 15] \\ &= (0.900) \cdot P[T > 15] \end{aligned}$$

$$\begin{aligned} P[T > 15] &= (1 - P[\text{die by } 15 \mid T > 10]) \cdot P[T > 10] \\ &= (0.900) \cdot P[T > 10] \end{aligned}$$

$$\begin{aligned} P[T > 10] &= (1 - P[\text{die by } 10 \mid T > 5]) \cdot P[T > 5] \\ &= (0.850) \cdot P[T > 5] \end{aligned}$$

$$\begin{aligned} P[T > 5] &= (1 - P[\text{die by } 5 \mid T > 0]) \\ &= 0.800 \end{aligned}$$

$$\begin{aligned} P[T > 20] &= (0.900) \cdot (0.900) \cdot (0.850) \cdot (0.800) \\ &= 0.5508 \end{aligned}$$

Continuation Probabilities

We can diagram the previous calculations:

Kaplan-Meier Estimator

The **Kaplan-Meier** estimator uses a single sample of 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 Δ. This allows us to estimate $P[\text{die by } t+\Delta \mid T>t]$.

The "at-risk" group declines over time due to subjects that die, and subjects that are lost (censored).

Define:

t_i	:	ith ordered follow-up time
d_i	:	number of deaths at ith ordered time
l_i	:	number of censored observations at ith ordered time
R_i	:	number of subjects at-risk at ith ordered time

$$\begin{aligned}\widehat{S}(t) &= \prod_{t_i \le t}(1-d_i/R_i) \\ &= (1-d_1/R_1)\times(1-d_2/R_2)\times\ldots\times(1-d_j/R_j)\end{aligned}$$

Kaplan-Meier Example

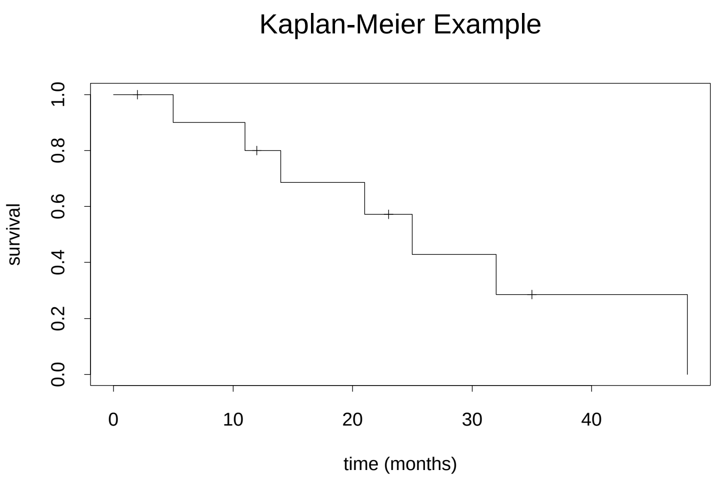
Example:

Observed Death Times : 5, 11, 14, 21, 25, 32, 48

Censored Times : 2, 12, 23, 35

We can record the data in the following table:

time	R_i	d_i	l_i	S_i	d_i/R_i	$(1 - d_i/R_i)$	$\widehat{S}(t)$
2	11	0	1	10	0.000	1.000	1.000
5	10	1	0	9	0.100	0.900	0.900
11	9	1	0	8	0.111	0.889	0.800
12	8	0	1	7	0.000	1.000	0.800
14	7	1	0	6	0.143	0.857	0.686
21	6	1	0	5	0.167	0.833	0.5714



Summary

1. "Time-until" outcomes (survival times) are common in biomedical research.
2. Survival times are often right-skewed.
3. Often a fraction of the times are right-censored.
4. The Kaplan-Meier estimator can be used to estimate and display the distribution of survival times.
5. Life tables are used to combine information across age groups.

Example with STATA

```
********************************************************************
* bc.do                                                            *
*                                                                  *
* PURPOSE:  compute Kaplan-Meier plots                             *
*                                                                  *
* DATE:  01/05/05                                                  *
********************************************************************
infile time status ploidy sphase using bc.dat

label variable time "time (years)"
label variable status "status"
label variable ploidy "ploidy status"
label variable sphase "%S-phase"

label define alab 0 "diploid" 1 "aneuploid"
label values ploidy alab

***
*** variable summaries
***
summarize
table ploidy status

***
*** this defines the failure outcome
***
stset time, failure(status)

***
*** Creates Kaplan-Meier curves
***
sts graph, by(ploidy) saving("bc.plot")
gphpen -deps "bc.plot"

*** show the estimates
sts list, by(ploidy)
```

```
.
. ***
. *** variable summaries
. ***
. summarize

Variable |     Obs        Mean   Std. Dev.       Min        Max
---------+-----------------------------------------------------
    time |     568    65.61092   25.45858          9        120
  status |     568    .2059859   .4047767          0          1
  ploidy |     568    .6478873   .4780499          0          1
  sphase |     568    9.940317   8.841601          0       55.4

. table ploidy status

----------+-----------
ploidy    |   status
status    |    0     1
----------+-----------
  diploid |  169    31
aneuploid |  282    86
----------+-----------


.
. ***
. *** this defines the failure outcome
. ***
. stset time, failure(status)

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

------------------------------------------------------------------------------
      568  total obs.
        0  exclusions
------------------------------------------------------------------------------
      568  obs. remaining, representing
      117  failures in single record/single failure data
    37267  total analysis time at risk, at risk from t =         0
                             earliest observed entry t =         0
                                  last observed exit t =       120
```



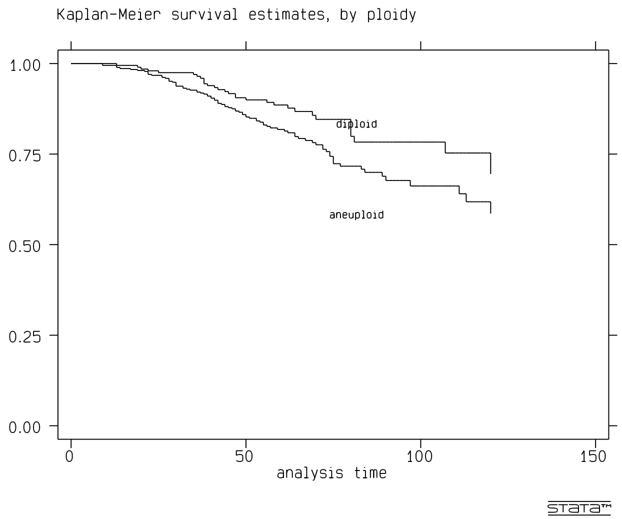
```
.
. *** show the estimates
. sts list, by(ploidy)

         failure _d:  status
   analysis time _t:  time

           Beg.          Net            Survivor      Std.
  Time    Total   Fail   Lost           Function     Error     [95% Conf. Int.]
-------------------------------------------------------------------------------
diploid
     9      200      1      1             0.9950    0.0050     0.9650    0.9993
    12      198      0      1             0.9950    0.0050     0.9650    0.9993
    19      197      1      0             0.9899    0.0071     0.9604    0.9975
    20      196      1      0             0.9849    0.0087     0.9539    0.9951
    22      195      1      0             0.9798    0.0100     0.9472    0.9924
    25      194      1      0             0.9748    0.0111     0.9405    0.9894
    35      193      1      2             0.9697    0.0122     0.9339    0.9863
    36      190      1      0             0.9646    0.0131     0.9273    0.9830
    37      189      1      1             0.9595    0.0140     0.9207    0.9796
    38      187      3      2             0.9441    0.0164     0.9014    0.9687
    39      182      1      2             0.9390    0.0171     0.8950    0.9649
    40      179      0      3             0.9390    0.0171     0.8950    0.9649
    41      176      1      3             0.9336    0.0178     0.8884    0.9609
    42      172      1      1             0.9282    0.0185     0.8817    0.9568
    43      170      0      1             0.9282    0.0185     0.8817    0.9568
    44      169      1      4             0.9227    0.0192     0.8750    0.9527
    45      164      1      1             0.9171    0.0199     0.8681    0.9484
    47      162      2      2             0.9058    0.0212     0.8545    0.9396
    48      158      0      4             0.9058    0.0212     0.8545    0.9396
    49      154      0      5             0.9058    0.0212     0.8545    0.9396
    50      149      1      4             0.8997    0.0219     0.8470    0.9349
    51      144      0      2             0.8997    0.0219     0.8470    0.9349
    52      142      0      3             0.8997    0.0219     0.8470    0.9349
    53      139      0      3             0.8997    0.0219     0.8470    0.9349
    54      136      0      2             0.8997    0.0219     0.8470    0.9349
    55      134      0      3             0.8997    0.0219     0.8470    0.9349
    56      131      1      3             0.8928    0.0228     0.8384    0.9297
    57      127      0      3             0.8928    0.0228     0.8384    0.9297
    58      124      1      3             0.8856    0.0237     0.8294    0.9242
    59      120      0      6             0.8856    0.0237     0.8294    0.9242
    60      114      0      5             0.8856    0.0237     0.8294    0.9242
    61      109      0      5             0.8856    0.0237     0.8294    0.9242
    62      104      1      4             0.8771    0.0250     0.8182    0.9179
    63       99      0      5             0.8771    0.0250     0.8182    0.9179
    64       94      1      3             0.8678    0.0264     0.8058    0.9110
    65       90      0      3             0.8678    0.0264     0.8058    0.9110
    66       87      0      1             0.8678    0.0264     0.8058    0.9110
    67       86      0      1             0.8678    0.0264     0.8058    0.9110
    68       85      0      4             0.8678    0.0264     0.8058    0.9110
    69       81      1      2             0.8570    0.0281     0.7912    0.9034
    70       78      1      3             0.8461    0.0299     0.7766    0.8954
    71       74      0      1             0.8461    0.0299     0.7766    0.8954
    72       73      0      6             0.8461    0.0299     0.7766    0.8954
    73       67      0      2             0.8461    0.0299     0.7766    0.8954
    74       65      0      1             0.8461    0.0299     0.7766    0.8954
    75       64      0      2             0.8461    0.0299     0.7766    0.8954
    76       62      0      2             0.8461    0.0299     0.7766    0.8954
    77       60      0      2             0.8461    0.0299     0.7766    0.8954
    78       58      0      2             0.8461    0.0299     0.7766    0.8954
    79       56      0      2             0.8461    0.0299     0.7766    0.8954
    80       54      3      1             0.7991    0.0386     0.7102    0.8632
```

81	50	1	3	0.7831	0.0410	0.6893	0.8515
82	46	0	2	0.7831	0.0410	0.6893	0.8515
87	44	0	1	0.7831	0.0410	0.6893	0.8515
88	43	0	2	0.7831	0.0410	0.6893	0.8515
89	41	0	1	0.7831	0.0410	0.6893	0.8515
90	40	0	3	0.7831	0.0410	0.6893	0.8515
91	37	0	1	0.7831	0.0410	0.6893	0.8515
92	36	0	1	0.7831	0.0410	0.6893	0.8515
95	35	0	1	0.7831	0.0410	0.6893	0.8515
98	34	0	1	0.7831	0.0410	0.6893	0.8515
100	33	0	3	0.7831	0.0410	0.6893	0.8515
105	30	0	2	0.7831	0.0410	0.6893	0.8515
106	28	0	2	0.7831	0.0410	0.6893	0.8515
107	26	1	1	0.7530	0.0493	0.6403	0.8348
110	24	0	1	0.7530	0.0493	0.6403	0.8348
111	23	0	3	0.7530	0.0493	0.6403	0.8348
112	20	0	1	0.7530	0.0493	0.6403	0.8348
113	19	0	3	0.7530	0.0493	0.6403	0.8348
117	16	0	1	0.7530	0.0493	0.6403	0.8348
118	15	0	1	0.7530	0.0493	0.6403	0.8348
119	14	0	1	0.7530	0.0493	0.6403	0.8348
120	13	1	12	0.6950	0.0719	0.5299	0.8119

aneuploid

13	368	4	0	0.9891	0.0054	0.9713	0.9959
14	364	1	0	0.9864	0.0060	0.9677	0.9943
17	363	1	0	0.9837	0.0066	0.9641	0.9926
19	362	1	0	0.9810	0.0071	0.9605	0.9909
21	361	1	0	0.9783	0.0076	0.9570	0.9891
22	360	3	0	0.9701	0.0089	0.9467	0.9833
23	357	1	0	0.9674	0.0093	0.9433	0.9813
26	356	2	0	0.9620	0.0100	0.9366	0.9773
27	354	1	0	0.9592	0.0103	0.9333	0.9752
28	353	3	0	0.9511	0.0112	0.9235	0.9689
29	350	1	0	0.9484	0.0115	0.9202	0.9668
30	349	4	0	0.9375	0.0126	0.9074	0.9580
32	345	2	2	0.9321	0.0131	0.9011	0.9536
33	341	1	0	0.9293	0.0134	0.8979	0.9513
34	340	1	0	0.9266	0.0136	0.8948	0.9491
36	339	2	3	0.9211	0.0141	0.8885	0.9445
37	334	1	2	0.9184	0.0143	0.8853	0.9422
38	331	1	3	0.9156	0.0145	0.8821	0.9399
39	327	2	1	0.9100	0.0149	0.8757	0.9352
40	324	2	3	0.9044	0.0154	0.8693	0.9304
41	319	2	7	0.8987	0.0158	0.8629	0.9256
42	310	3	3	0.8900	0.0164	0.8531	0.9181
43	304	1	4	0.8871	0.0166	0.8498	0.9156
44	299	2	3	0.8812	0.0170	0.8431	0.9105
45	294	1	5	0.8782	0.0172	0.8397	0.9079
46	288	1	5	0.8751	0.0174	0.8363	0.9053
47	282	2	11	0.8689	0.0179	0.8293	0.8999
48	269	1	4	0.8657	0.0181	0.8256	0.8971
49	264	2	4	0.8591	0.0185	0.8182	0.8914
50	258	2	5	0.8525	0.0190	0.8107	0.8856
51	251	1	6	0.8491	0.0192	0.8069	0.8827
52	244	0	9	0.8491	0.0192	0.8069	0.8827
53	235	2	6	0.8418	0.0197	0.7987	0.8764
54	227	1	6	0.8381	0.0200	0.7945	0.8732
55	220	2	11	0.8305	0.0205	0.7858	0.8666
56	207	1	2	0.8265	0.0208	0.7813	0.8632
57	204	1	4	0.8224	0.0211	0.7766	0.8597
58	199	0	5	0.8224	0.0211	0.7766	0.8597
59	194	1	8	0.8182	0.0214	0.7718	0.8561
60	185	0	7	0.8182	0.0214	0.7718	0.8561
61	178	1	6	0.8136	0.0218	0.7665	0.8522
62	171	1	4	0.8088	0.0221	0.7609	0.8481
63	166	0	10	0.8088	0.0221	0.7609	0.8481
64	156	2	4	0.7985	0.0230	0.7487	0.8394
65	150	1	3	0.7932	0.0235	0.7425	0.8350
66	146	0	1	0.7932	0.0235	0.7425	0.8350
67	145	1	2	0.7877	0.0240	0.7361	0.8304
68	142	0	6	0.7877	0.0240	0.7361	0.8304
69	136	1	5	0.7819	0.0245	0.7293	0.8255
70	130	1	3	0.7759	0.0250	0.7221	0.8205
71	126	0	2	0.7759	0.0250	0.7221	0.8205
72	124	2	2	0.7634	0.0261	0.7074	0.8101
73	120	1	4	0.7570	0.0267	0.6999	0.8048
74	115	2	3	0.7438	0.0278	0.6845	0.7937
75	110	3	2	0.7235	0.0294	0.6611	0.7765
76	105	0	1	0.7235	0.0294	0.6611	0.7765
77	104	1	3	0.7166	0.0299	0.6531	0.7705
78	100	0	4	0.7166	0.0299	0.6531	0.7705
79	96	0	2	0.7166	0.0299	0.6531	0.7705
80	94	0	3	0.7166	0.0299	0.6531	0.7705
81	91	0	3	0.7166	0.0299	0.6531	0.7705

82	88	0	3	0.7166	0.0299	0.6531	0.7705
83	85	1	2	0.7082	0.0307	0.6430	0.7636
84	82	1	3	0.6995	0.0316	0.6328	0.7565
85	78	0	2	0.6995	0.0316	0.6328	0.7565
86	76	0	4	0.6995	0.0316	0.6328	0.7565
87	72	0	3	0.6995	0.0316	0.6328	0.7565
88	69	0	4	0.6995	0.0316	0.6328	0.7565
89	65	1	4	0.6888	0.0329	0.6193	0.7481
90	60	1	2	0.6773	0.0343	0.6050	0.7392
91	57	0	2	0.6773	0.0343	0.6050	0.7392
92	55	0	4	0.6773	0.0343	0.6050	0.7392
93	51	0	1	0.6773	0.0343	0.6050	0.7392
94	50	0	2	0.6773	0.0343	0.6050	0.7392
95	48	0	1	0.6773	0.0343	0.6050	0.7392
96	47	0	3	0.6773	0.0343	0.6050	0.7392
97	44	1	4	0.6619	0.0368	0.5843	0.7284
100	39	0	1	0.6619	0.0368	0.5843	0.7284
102	38	0	1	0.6619	0.0368	0.5843	0.7284
105	37	0	2	0.6619	0.0368	0.5843	0.7284
106	35	0	2	0.6619	0.0368	0.5843	0.7284
109	33	0	1	0.6619	0.0368	0.5843	0.7284
110	32	0	1	0.6619	0.0368	0.5843	0.7284
111	31	1	1	0.6405	0.0413	0.5534	0.7151
113	29	1	1	0.6185	0.0454	0.5229	0.7004
114	27	0	1	0.6185	0.0454	0.5229	0.7004
115	26	0	1	0.6185	0.0454	0.5229	0.7004
116	25	0	1	0.6185	0.0454	0.5229	0.7004
117	24	0	2	0.6185	0.0454	0.5229	0.7004
118	22	0	2	0.6185	0.0454	0.5229	0.7004
119	20	0	1	0.6185	0.0454	0.5229	0.7004
120	19	1	18	0.5859	0.0534	0.4739	0.6820



Survival Analysis

- More on censoring
  - Dependent censoring
  - Independent censoring
  - Interval censoring
  - Left truncation

- Standard errors for KM estimates
  - Greenwood method

- Comparing KM curves: log-rank test
  - Mantel-Haenszel
  - other weighting schemes

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 were censored are representative of the population of responses.

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, $\widehat{S}(t)$.

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

Assumption:

$$
\begin{aligned}
D_i &= \text{the survival time for subject } i \\
C_i &= \text{the censoring time for subject } i
\end{aligned}
$$

$$
\begin{aligned}
T_i &= \min(D_i, C_i) \\
\delta_i &= 1 \quad \text{if } D_i < C_i \text{, and 0 otherwise}
\end{aligned}
$$

$\boxed{\bullet}$ We assume that the censoring time, C_i, is independent of the survival time, D_i.

We observe the pair: (time $= T_i$, status $= \delta_i$).

- Censoring due to the end of study $\Rightarrow$
  - Independent Censoring
- Censoring due to drop-out $\Rightarrow$
  - verify based on reasons for drop-out
- Censoring due to another type of outcome $\Rightarrow$
  - "competing risks", assumed independent

More on Censoring

Interval Censoring:

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

$$a \le T_i < b$$

- HIV vaccine trial with 6 monthly blood testing.

- If everyone shares the same time intervals (ie. 6 month visit schedule) then the outcomes are known as <u>discrete survival</u> times, and logistic regression methods can be used.

Left Truncation:

This occurs when some subjects have a <u>delayed entry</u> 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.

- Breast cancer study where $t = 0$ is the date of diagnosis, but some women are contacted several months (years) after diagnosis and then enter the study.

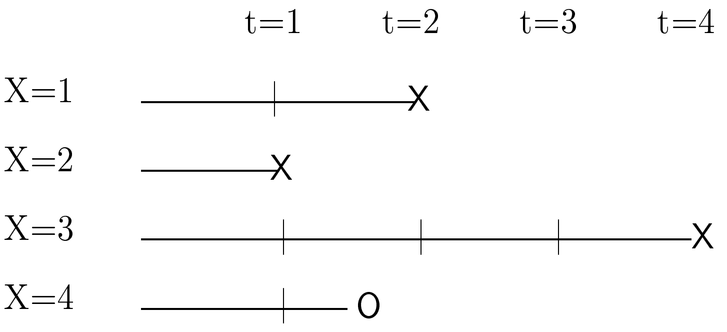
Kaplan-Meier

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?

Efron (1967) gave an intuitive answer: the Kaplan-Meier distributes the "jump" for a censored time to the <u>observed</u> times that are larger than the censored time.

"Distribute to the right"



$\widehat{S}(t)$ Standard Errors

Kaplan-Meier can be used to obtain estimates of survival probabilities such as

$$\widehat{S}(60) \;=\; \text{estimated 60 month survival}$$

Q: Can we obtain a confidence interval for this estimate?

Recall:

t_i : ith ordered follow-up time

d_i : number of deaths at ith ordered time

l_i : number of censored observations at ith ordered time

R_i : number of subjects at-risk at ith ordered time

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

Greenwood's formula:

$$\widehat{V}[\widehat{S}(t)] \;=\; \widehat{S}(t)^2 \sum_{t_i \leq t} \frac{d_i}{R_i(R_i - d_i)}$$

$\widehat{S}(t)$ Standard Errors

Note: Rosner, page 612-613 gives the following:

$$\widehat{V}\{\log[\widehat{S}(t)]\} = \sum_{t_i \leq t} \frac{d_i}{R_i(R_i - d_i)}$$

(where we use R_i in place of S_{i-1}).

In practice, this estimate and the one obtained from Greenwood's formula should be quite similar.

95% Confidence Interval using Greenwood:

$$\text{lower} = \widehat{S}(t) - 1.96 \cdot \widehat{S}(t) \cdot \sqrt{\sum_{t_i \leq t} \frac{d_i}{R_i(R_i - d_i)}}$$

$$\text{upper} = \widehat{S}(t) + 1.96 \cdot \widehat{S}(t) \cdot \sqrt{\sum_{t_i \leq t} \frac{d_i}{R_i(R_i - d_i)}}$$

Computing $\widehat{S}(t)$ Standard Errors

STATA:

- `stset` – to define survival data
- `sts graph` – to create Kaplan-Meier plot
- Can request Greenwood's & easily add to graph!
- Use `sts list` to display.
- `sts test` – for log-rank (+ other) tests

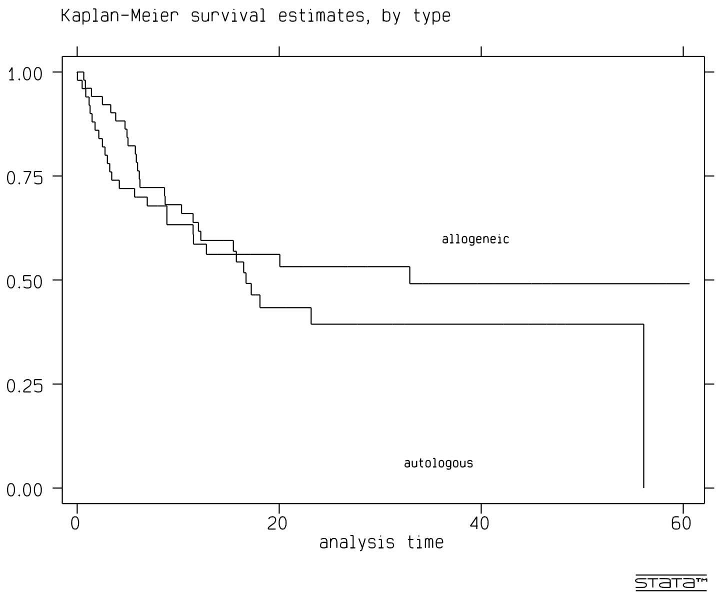
Example:

(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 reinfused 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.

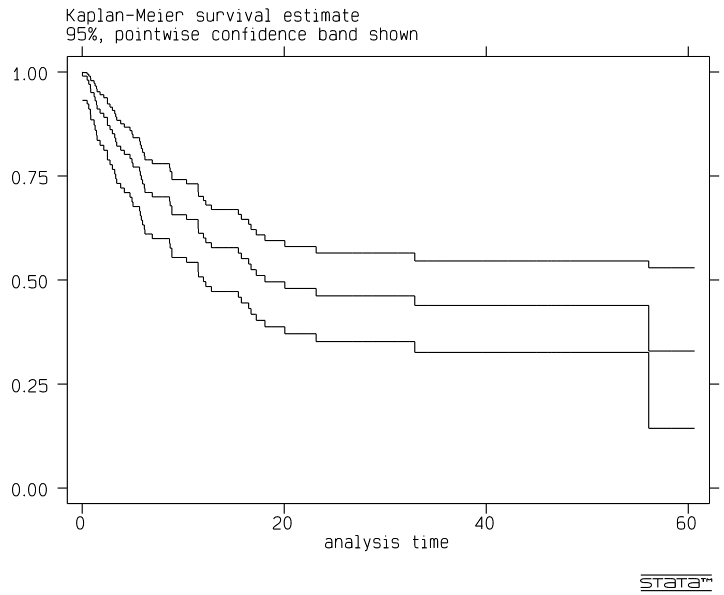
Q: Any difference in survival?

Q: Estimate 5-year survival, with 95% CI.


```
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

***
*** this defines the failure outcome
***
stset time, failure(status)

***
*** this creates Kaplan-Meier curves
***
sts graph, by(type) saving("transplant1.plot")
gphpen -dps "transplant1.plot"

***
*** this computes the log-rank test
***
sts test type, logrank

***
*** combined groups KM with s.e.'s
***
sts graph, gwood level(95) saving("transplant2.plot")
gphpen -dps "transplant2.plot"

***
*** show the S(t) and s.e.'s
***
sts list
sts list, by(type)
```




```
. sts test type, logrank

         failure _d:  status
   analysis time _t:  time


Log-rank test for equality of survivor functions
------------------------------------------------

           |  Events
type       |  observed       expected
-----------+-------------------------
allogeneic |        23          24.82
autologous |        28          26.18
-----------+-------------------------
Total      |        51          51.00

                 chi2(1) =       0.26
                 Pr>chi2 =     0.6077
```




```
. sts list

        failure _d:  status
  analysis time _t:  time
```

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf.	Int.]
.03	101	1	0	0.9901	0.0099	0.9318	0.9986
.493	100	1	0	0.9802	0.0139	0.9231	0.9950
.658	99	1	0	0.9703	0.0169	0.9107	0.9903
.822	98	1	0	0.9604	0.0194	0.8979	0.9849
.855	97	1	0	0.9505	0.0216	0.8852	0.9791
1.184	96	1	0	0.9406	0.0235	0.8725	0.9729
1.283	95	1	0	0.9307	0.0253	0.8601	0.9663
1.414	94	1	0	0.9208	0.0269	0.8479	0.9596
1.48	93	1	0	0.9109	0.0283	0.8358	0.9526
1.776	92	1	0	0.9010	0.0297	0.8238	0.9455
2.138	91	1	0	0.8911	0.0310	0.8120	0.9382
2.5	90	2	0	0.8713	0.0333	0.7887	0.9231
2.763	88	1	0	0.8614	0.0344	0.7772	0.9155
2.993	87	1	0	0.8515	0.0354	0.7658	0.9077
3.224	86	1	0	0.8416	0.0363	0.7545	0.8998
3.322	85	1	0	0.8317	0.0372	0.7433	0.8918
3.421	84	1	0	0.8218	0.0381	0.7322	0.8838
3.816	83	1	0	0.8119	0.0389	0.7211	0.8756
4.178	82	1	0	0.8020	0.0397	0.7101	0.8674
4.441	81	0	1	0.8020	0.0397	0.7101	0.8674
4.737	80	1	0	0.7920	0.0404	0.6990	0.8590
4.836	79	0	1	0.7920	0.0404	0.6990	0.8590
4.934	78	1	0	0.7818	0.0411	0.6878	0.8505
5.033	77	1	0	0.7716	0.0418	0.6767	0.8419
5.691	76	1	0	0.7615	0.0425	0.6656	0.8333
5.757	75	1	0	0.7513	0.0431	0.6546	0.8246
5.855	74	1	1	0.7412	0.0437	0.6436	0.8158
5.987	72	1	0	0.7309	0.0443	0.6326	0.8069
6.151	71	1	0	0.7206	0.0449	0.6215	0.7979
6.217	70	1	0	0.7103	0.0454	0.6106	0.7889
6.447	69	0	1	0.7103	0.0454	0.6106	0.7889
6.941	68	1	1	0.6999	0.0459	0.5995	0.7797
7.993	66	0	1	0.6999	0.0459	0.5995	0.7797
8.651	65	1	0	0.6891	0.0465	0.5880	0.7702
8.711	64	1	0	0.6783	0.0470	0.5766	0.7606
8.882	63	2	0	0.6568	0.0479	0.5540	0.7414
9.145	61	0	1	0.6568	0.0479	0.5540	0.7414
9.441	60	0	1	0.6568	0.0479	0.5540	0.7414
10.33	59	1	0	0.6457	0.0483	0.5424	0.7314

11.48	58	2	0	0.6234	0.0492	0.5192	0.7113
11.51	56	1	0	0.6123	0.0495	0.5077	0.7011
12.01	55	1	1	0.6011	0.0499	0.4963	0.6909
12.1	53	0	1	0.6011	0.0499	0.4963	0.6909
12.24	52	1	0	0.5896	0.0502	0.4844	0.6803
12.4	51	0	1	0.5896	0.0502	0.4844	0.6803
12.8	50	1	0	0.5778	0.0506	0.4723	0.6695
12.99	49	0	1	0.5778	0.0506	0.4723	0.6695
13.06	48	0	1	0.5778	0.0506	0.4723	0.6695
13.85	47	0	1	0.5778	0.0506	0.4723	0.6695
14.47	46	0	1	0.5778	0.0506	0.4723	0.6695
15	45	0	1	0.5778	0.0506	0.4723	0.6695
15.46	44	1	0	0.5646	0.0511	0.4586	0.6577
15.76	43	1	0	0.5515	0.0516	0.4449	0.6458
16.48	42	1	0	0.5384	0.0520	0.4314	0.6338
16.61	41	0	1	0.5384	0.0520	0.4314	0.6338
16.71	40	1	0	0.5249	0.0524	0.4176	0.6214
17.14	39	0	1	0.5249	0.0524	0.4176	0.6214
17.2	38	0	1	0.5249	0.0524	0.4176	0.6214
17.24	37	1	0	0.5107	0.0529	0.4030	0.6085
17.3	36	0	1	0.5107	0.0529	0.4030	0.6085
17.66	35	0	1	0.5107	0.0529	0.4030	0.6085
18.09	34	1	1	0.4957	0.0534	0.3874	0.5949
18.75	32	0	1	0.4957	0.0534	0.3874	0.5949
20.07	31	1	0	0.4797	0.0540	0.3708	0.5805
20.33	30	0	1	0.4797	0.0540	0.3708	0.5805
20.63	29	0	1	0.4797	0.0540	0.3708	0.5805
22.37	28	0	1	0.4797	0.0540	0.3708	0.5805
23.16	27	1	0	0.4620	0.0549	0.3520	0.5648
26.78	26	0	1	0.4620	0.0549	0.3520	0.5648
27.73	25	0	1	0.4620	0.0549	0.3520	0.5648
28.72	24	0	2	0.4620	0.0549	0.3520	0.5648
31.18	22	0	1	0.4620	0.0549	0.3520	0.5648
32.43	21	0	1	0.4620	0.0549	0.3520	0.5648
32.93	20	1	0	0.4389	0.0568	0.3261	0.5459
33.78	19	0	1	0.4389	0.0568	0.3261	0.5459
34.22	18	0	1	0.4389	0.0568	0.3261	0.5459
34.77	17	0	1	0.4389	0.0568	0.3261	0.5459
35.92	16	0	1	0.4389	0.0568	0.3261	0.5459
39.59	15	0	1	0.4389	0.0568	0.3261	0.5459
41.12	14	0	1	0.4389	0.0568	0.3261	0.5459
42.24	13	0	1	0.4389	0.0568	0.3261	0.5459
44.64	12	0	1	0.4389	0.0568	0.3261	0.5459
45	11	0	1	0.4389	0.0568	0.3261	0.5459
46.05	10	0	1	0.4389	0.0568	0.3261	0.5459
46.48	9	0	1	0.4389	0.0568	0.3261	0.5459
46.94	8	0	1	0.4389	0.0568	0.3261	0.5459

47.47	7	0	1	0.4389	0.0568	0.3261	0.5459
48.29	6	0	1	0.4389	0.0568	0.3261	0.5459
48.32	5	0	1	0.4389	0.0568	0.3261	0.5459
56.09	4	1	0	0.3291	0.1041	0.1435	0.5294
57.4	3	0	1	0.3291	0.1041	0.1435	0.5294
58.32	2	0	1	0.3291	0.1041	0.1435	0.5294
60.63	1	0	1	0.3291	0.1041	0.1435	0.5294

Comparing Survival Functions

Q: How can we test (compare) the probability of survival beyond 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.

$$\begin{aligned} H_0 &: S_1(t_0) = S_2(t_0) \\ H_1 &: S_1(t_0) \neq S_2(t_0) \end{aligned}$$

$$\begin{aligned} Z &= \frac{\widehat{S}_1(t_0) - \widehat{S}_2(t_0)}{\sqrt{\widehat{V}[\widehat{S}_1(t_0)] + \widehat{V}[\widehat{S}_2(t_0)]}} \\ Z &\sim N(0,1) \text{ under } H_0 \end{aligned}$$

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.

We have the following estimates from the previous analysis:

$$\begin{aligned} \widehat{S}_1(24) &= 0.5321 \\ \widehat{V}[\widehat{S}_1(24)] &= (0.0746)^2 \end{aligned}$$

$$\begin{aligned} \widehat{S}_2(24) &= 0.3940 \\ \widehat{V}[\widehat{S}_2(24)] &= (0.0790)^2 \end{aligned}$$

$$\begin{aligned} Z &= \frac{0.5321 - 0.3940}{\sqrt{(0.0746)^2 + (0.0790)^2}} \\ &= 1.271 \end{aligned}$$

$$P[\ N(0,1)\ > 1.271] = 0.102 \ \ (\times 2 = 0.204)$$

Comparing Survival Functions

Kaplan-Meier allows a graphical comparison of survival curves for different patient subsets.

Q: What confirmatory tests can we use to compare the entire survival curve for 2 (or more) groups?

A: The log-rank test.

Overview:

$$
\begin{aligned}
H_0 &: S_1(t) = S_2(t) \text{ for all } t \\
H_1 &: S_1(t) \neq S_2(t) \text{ for some } t
\end{aligned}
$$

- 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. Denote the observed failure times as t_j, for $j = 1, 2, \ldots, m$.
2. For each j define:
$$
\begin{aligned}
d_{1j} &= \text{number of deaths in group 1} \\
d_{2j} &= \text{number of deaths in group 2}
\end{aligned}
$$
3. For each j define:
$$
\begin{aligned}
R_{1j} &= \text{number in risk set for group 1} \\
R_{2j} &= \text{number in risk set for group 2}
\end{aligned}
$$
4. $E_{1j} = \left(\frac{R_{1j}}{R_{1j}+R_{2j}}\right)(d_{1j} + d_{2j})$
5. $E_1 = \sum_{j=1}^{m} E_{1j}$, $O_1 = \sum_{j=1}^{m} d_{1j}$
6. The log-rank test statistic is:
$$
\begin{aligned}
X^2 &= (O_1 - E_1)^2 / \widehat{V}_1 \\
\widehat{V}_1 &= \sum_j \frac{R_{1j}R_{2j}(d_{1j}+d_{2j})(R_{1j}+R_{2j}-d_{1j}-d_{2j})}{(R_{1j}+R_{2j})^2(R_{1j}+R_{2j}-1)}
\end{aligned}
$$
7. Under H_0, $X^2 \sim \chi^2(1)$.

Log-rank Test

Note:

For the observed failure time, t_j, we have:

	Dead	Alive	Total
Group 1	d_{1j}	$R_{1j} - d_{1j}$	R_{1j}
Group 2	d_{2j}	$R_{2j} - d_{2j}$	R_{2j}
Total	$d_{1j} + d_{2j}$		$R_{1j} + R_{2j}$

- From this we can see what E_{1j} is (recall 2×2 tables!).
- Mantel-Haenszel $\Rightarrow$ pool across strata: t_j's.

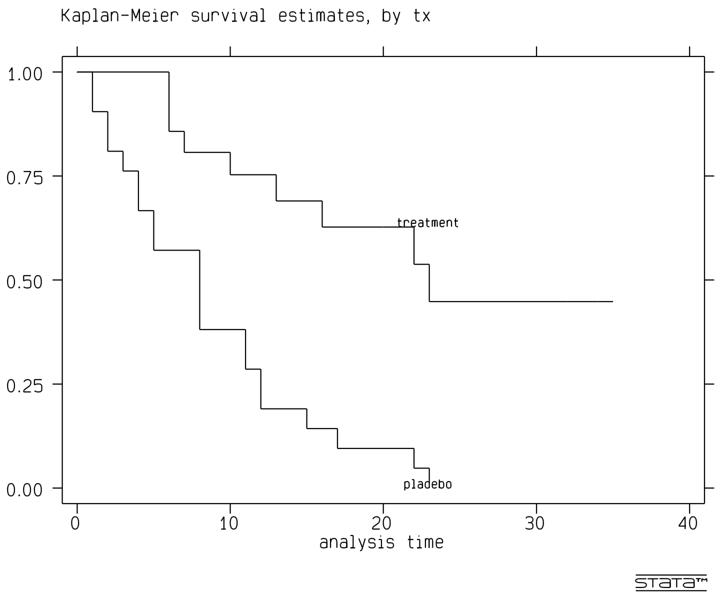
Log-rank Example

Kleinbaum (#2), page 51, Log-rank pages 59-62.

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

Group 1 ($n = 21$) treatment	Group 2 ($n = 21$) placebo
6, 6, 6, 7, 10	1, 1, 2, 2, 3
13, 16, 22, 23	4, 4, 5, 5
6+, 9+, 10+, 11+,	8, 8, 8, 8,
17+, 19+, 20+,	11, 11, 12, 12,
25+, 32+, 32+,	15, 17, 22, 23
34+, 35+	

Note: + denotes censoring



Remission Data

		# failures		# in risk set	
j	t_j	d_{1j}	d_{2j}	R_{1j}	R_{2j}
1	1	0	2	21	21
2	2	0	2	21	19
3	3	0	1	21	17
4	4	0	2	21	16
5	5	0	2	21	14
6	6	3	0	21	12
7	7	1	0	17	12
8	8	0	4	16	12
9	10	1	0	15	8
10	11	0	2	13	8
11	12	0	2	12	6
12	13	1	0	12	4
13	15	0	1	11	4
14	16	1	0	11	3
15	17	0	1	10	3
16	22	1	1	7	2
17	23	1	1	6	1

Leukemia Example

- There are 17 unique failure times ($m = 17$)
- 2×2 table for $t_6 = 6$

	6-MP	Control	Totals
deaths at t_6	3	0	3
survivors past t_6	18	12	30
at risk at t_6	21	12	33

$$O_6 = 3 \qquad E_6 = \frac{21 \times 3}{33} = 1.9$$

$$V_6 = \frac{21 \times 12 \times 3 \times 30}{33^2 \times 32} = 0.651$$

- 2×2 table for $t_{20} = 22$

	6-MP	Control	Totals
deaths at t_{20}	1	1	2
survivors past t_{20}	6	1	7
at risk at t_{20}	7	2	9

$$O_{20} = 1, \qquad E_{20} = \frac{7 \times 2}{9} = 1.56$$

$$V_{20} = \frac{7 \times 2 \times 2 \times 7}{9^2 \times 8} = 0.302$$

	# failures		# in risk set		expected		$O - E$	
j	d_{1j}	d_{2j}	R_{1j}	R_{2j}	E_{1j}	E_{2j}	$(d_{1j} - E_{1j})$	$(d_{2j} - E_{2j})$
1	0	2	21	21	(21/42) 2	(21/42) 2	-1.00	1.00
2	0	2	21	19	(21/40) 2	(19/40) 2	-1.05	1.05
3	0	1	21	17	(21/38) 1	(17/38) 1	-0.55	0.55
4	0	2	21	16	(21/37) 2	(16/37) 2	-1.14	1.14
5	0	2	21	14	(21/35) 2	(14/35) 2	-1.20	1.20
6	3	0	21	12	(21/33) 3	(12/33) 3	1.09	-1.09
7	1	0	17	12	(17/29) 1	(12/29) 1	0.41	-0.41
8	0	4	16	12	(16/28) 4	(12/28) 4	-2.29	2.29
9	1	0	15	8	(15/23) 1	(8/23) 1	0.35	-0.35
10	0	2	13	8	(13/21) 2	(6/18) 2	-1.24	1.24
11	0	2	12	6	(12/18) 2	(6/18) 2	-1.33	1.33
12	1	0	12	4	(12/16) 1	(4/16) 1	0.25	-0.25
13	0	1	11	4	(11/15) 1	(4/15) 1	-0.73	0.73
14	1	0	11	3	(11/14) 1	(3/14) 1	0.21	-0.21
15	0	1	10	3	(10/13) 1	(3/13) 1	-0.77	0.77
16	1	1	7	2	(7/9) 2	(2/9) 2	-0.56	0.56
17	1	1	6	1	(6/7) 2	(1/7) 2	-0.71	0.71
	9	21			19.26	10.74	-10.26	10.26

```
Log-rank test for equality of survivor functions
------------------------------------------------

          |  Events
tx        |  observed       expected
----------+-------------------------
treatment |         9          19.25
placebo   |        21          10.75
----------+-------------------------
Total     |        30          30.00

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

Remission Data:

- In this example we obtain from the variance calculation (not shown)

$$\widehat{V}_1 = 6.270$$

So that the test statistic is:

$$\text{log-rank stat.} \;= (O_1 - E_1)^2/V_1 = (-10.26)^2/6.270 = 16.79$$

We obtain the significance of this statistic by comparison to a $\chi^2(1)$:

$$P[\,\chi^2(1) \;>\; 16.79\,] \;<\; 0.001$$

Note:

- The log-rank statistic is approximately the same as the standard form for "observed versus expected" chi-square statistics:

$$\begin{aligned} X^2 &= \sum_{i=1}^{2}(O_i - E_i)^2/E_i \\ &= (-10.26)^2/19.26 \;+\; (10.26)^2/10.74 \\ &= 15.267 \end{aligned}$$

Generalizations of the log-rank test

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

- When the K groups are formed on the basis on an ordinal variable (ie. are ordered) then a modified version of the log-rank can be used to test for trend (a 1 degree of freedom test). We'll see how we can use Cox regression with a single covariate to obtain an equivalent test.

- **Weighted log-rank tests**

Weighted log-rank statistics

log-rank statistic:

$$O_1 - E_1 = \sum_j (d_{1j} - E_{1j})$$

Q: Should we combine across the failure times (strata, tables) equally or should we give more weight to certain times (earlier, later)?

Proposal:

$$\sum_j w_j (d_{1j} - E_{1j})$$

Define: $R_j = R_{1j} + R_{2j}$.

$$w_j = 1 \quad \Rightarrow \quad \text{log-rank test}$$

$$w_j = R_j \quad \Rightarrow \quad \text{Wilcoxon-Gehan-Breslow test}$$

$$w_j = R_j^{1/2} \quad \Rightarrow \quad \text{Tarone-Ware test}$$

Comments:

- The log-rank test gives equal weight to all times.
  Emphasizes the tail of the survival curve.
- The Wilcoxon-Breslow gives more weight to earlier times.
  Emphasizes beginning of survival curve.

Q: Choice?

○ Which is scientifically more important - early versus late ?

○ The log-rank test is the most powerful for detecting alternatives that correspond to proportional hazards (so related to Cox regression!)

Example: The leukemia remission data

```
. sts test tx, logrank

Log-rank test for equality of survivor functions
------------------------------------------------

          |  Events
tx        |  observed       expected
----------+-------------------------
treatment |         9          19.25
placebo   |        21          10.75
----------+-------------------------
Total     |        30          30.00

                chi2(1) =      16.79
                Pr>chi2 =     0.0000


. sts test tx, wilcoxon

Wilcoxon (Breslow) test for equality of survivor functions
----------------------------------------------------------

          |  Events                          Sum of
tx        |  observed       expected          ranks
----------+--------------------------------------
treatment |         9          19.25          -271
placebo   |        21          10.75           271
----------+--------------------------------------
Total     |        30          30.00             0

                chi2(1) =      13.46
                Pr>chi2 =     0.0002
```



```
 Survival Analysis for TIME

                              Total     Number     Number     Percent
                                        Events     Censored   Censored

  TX           1.00            21          9          12        57.14
  TX           2.00            21         21           0          .00

Overall                        42         30          12        28.57

 Test Statistics for Equality of Survival Distributions for TX

              Statistic        df       Significance

 Log Rank       16.79           1          .0000
 Breslow        13.46           1          .0002
 Tarone-Ware    15.12           1          .0001
```

Example: The breast cancer data

```
. sts test ploidy, logrank

Log-rank test for equality of survivor functions
------------------------------------------------

          |  Events
ploidy    |  observed       expected
----------+-------------------------
diploid   |        31          42.77
aneuploid |        86          74.23
----------+-------------------------
Total     |       117         117.00

                chi2(1) =       5.13
                Pr>chi2 =     0.0235


. sts test ploidy, wilcoxon

Wilcoxon (Breslow) test for equality of survivor functions
----------------------------------------------------------

          |  Events                        Sum of
ploidy    |  observed       expected        ranks
----------+--------------------------------------
diploid   |        31          42.77        -4702
aneuploid |        86          74.23         4702
----------+--------------------------------------
Total     |       117         117.00            0

                chi2(1) =       4.54
                Pr>chi2 =     0.0332
```



```
 Survival Analysis for TIME

                               Total       Number       Number        Percent
                                           Events      Censored      Censored

  PLOIDY            .00          200           31          169          84.50
  PLOIDY           1.00          368           86          282          76.63

Overall                          568          117          451          79.40

 Test Statistics for Equality of Survival Distributions for PLOIDY


               Statistic        df       Significance

 Log Rank           5.13         1            .0235
 Breslow            4.54         1            .0332
 Tarone-Ware        4.96         1            .0259
```

Summary

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 and later times.

Hazard functions and models

- Hazard function
  - Definition
  - Relationship to incidence
  - Cumulative hazard
  - Relationship to survival fnx
- Cox regression
  - Proportional hazards assumption
  - "semi-parametric" model
  - Estimation and Inference
  - Estimation of baseline survival fnx

Hazard 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, Δ."

- Conditional probability divided by Δ, as Δ becomes very small.

- $h(t)$ is a rate between 0 and $+\infty$.

- $h(t)$ depends on the units of time.

- Special cases and synonyms:
  - force of mortality
  - instantaneous incidence rate
  - incidence rate
  - incidence density (where event is disease)

Hazard Rate

Example:

Probability	Δ	Rate=$Prob./\Delta$
$\frac{1}{3}$	$\frac{1}{2}$ day	$\frac{1/3}{1/2} = 0.67/\text{day}$
$\frac{1}{3}$	$\frac{1}{14}$ week	$\frac{1/3}{1/14} = 4.67/\text{week}$

Example: Remission data.

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

Placebo	Treatment
9 events	21 events
359 weeks	182 weeks
Rate=9/359=0.0251	Rate=21/182=0.1154

Note: the (average) hazard ratio is $0.1154/0.0251 = 4.603$.

Cumulative Hazard

Define: Cumulative hazard

$$H(t) = \int_0^t h(s)ds$$

Relationships:

$$h(t) \Longleftrightarrow H(t) \Longleftrightarrow S(t)$$

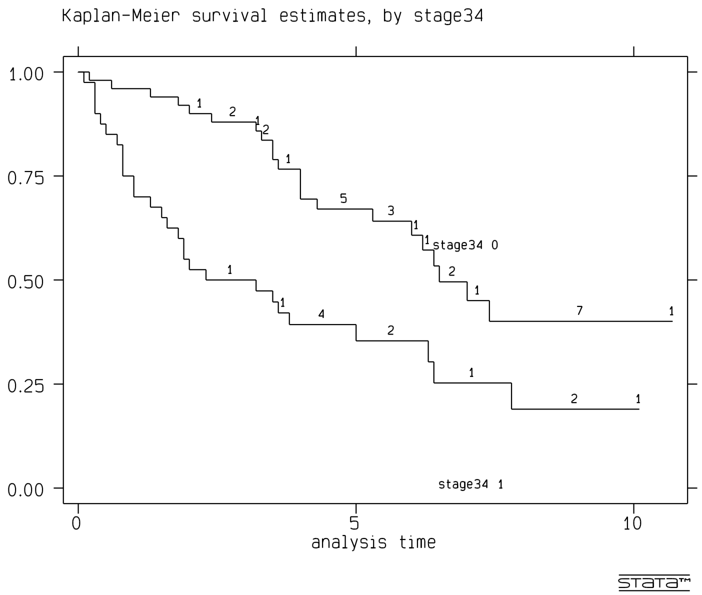
- If we specify the *hazard* then we specify the *cumulative hazard*, and we have specified the *survival function.*

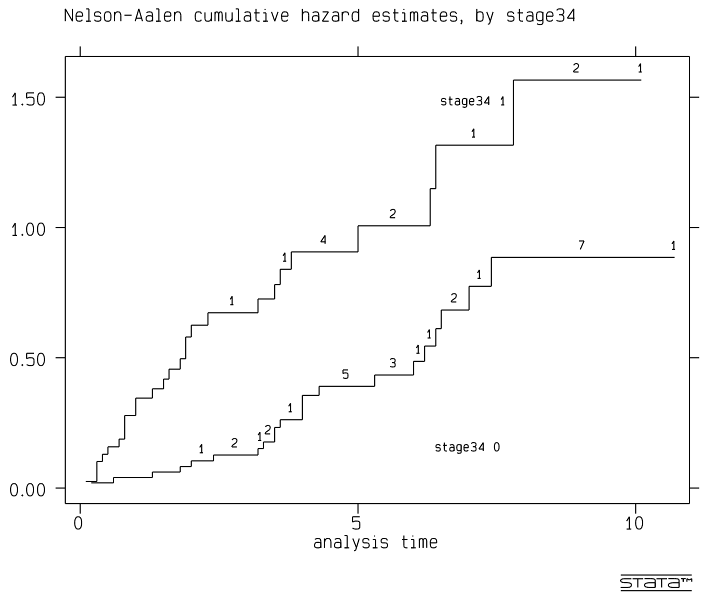
<u>Further Details</u>:

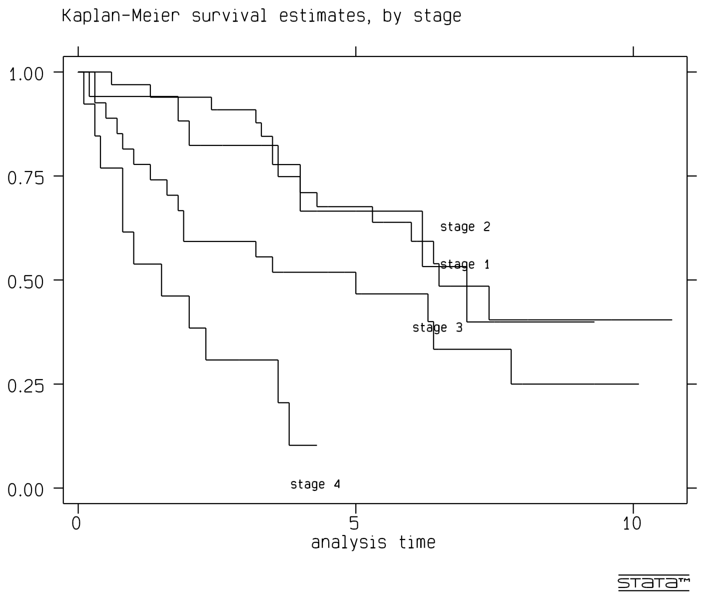
$$\begin{aligned} S(t) &= \exp(\ -H(t)\ \ ) \\[1em] \frac{\partial}{\partial t} S(t) &= -h(t)\ S(t) \\ \frac{\partial}{\partial t} \log S(t) &= -h(t) \end{aligned}$$

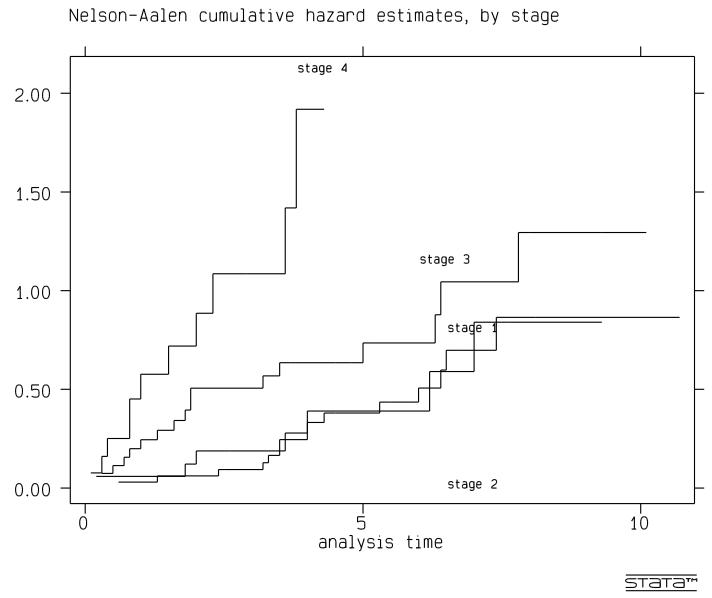
- A direct relationship between the survival function and the cumulative hazard function (see examples that follow).

- The rate-of-change in the survival function (log survival) is given by the hazard function.

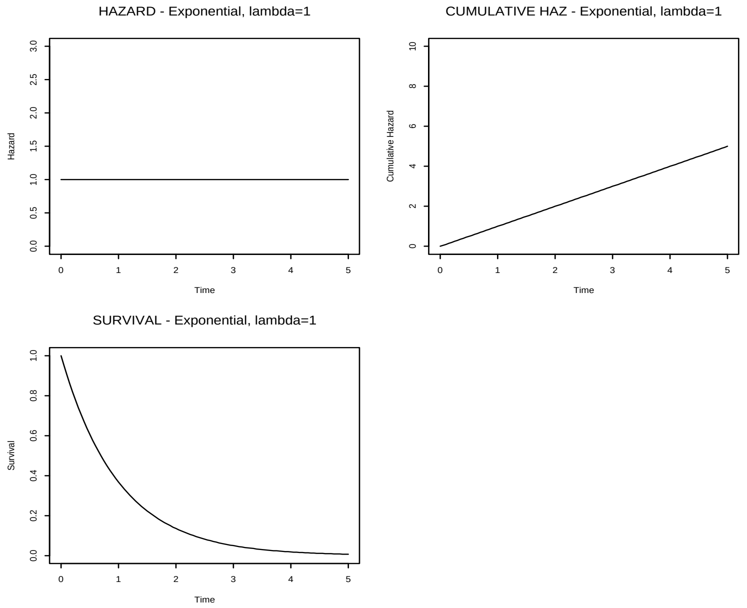



Nelson-Aalen cumulative hazard estimates, by stage34
1.50
1.00
0.50
0.00
0
5
10
analysis time
stage34 1
stage34 0
stata™

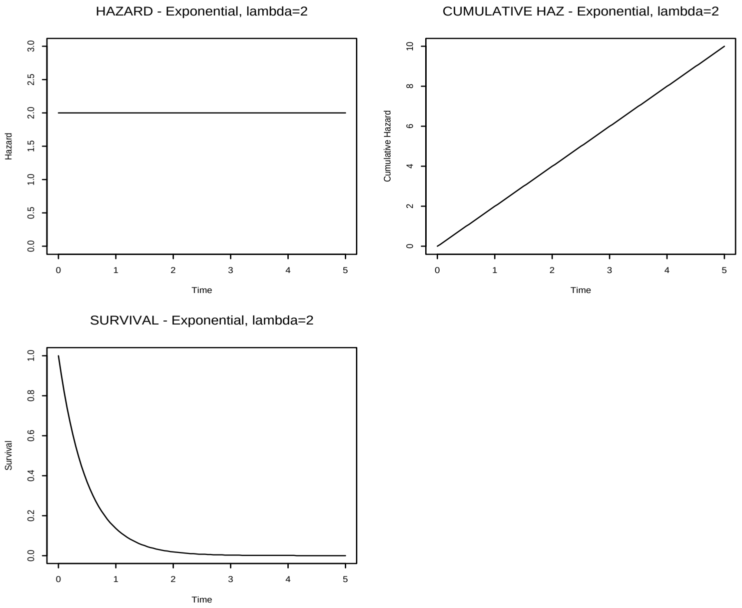

Kaplan-Meier survival estimates, by stage
1.00
0.75
0.50
0.25
0.00
0
5
10
analysis time
stage 2
stage 1
stage 3
stage 4
stata™



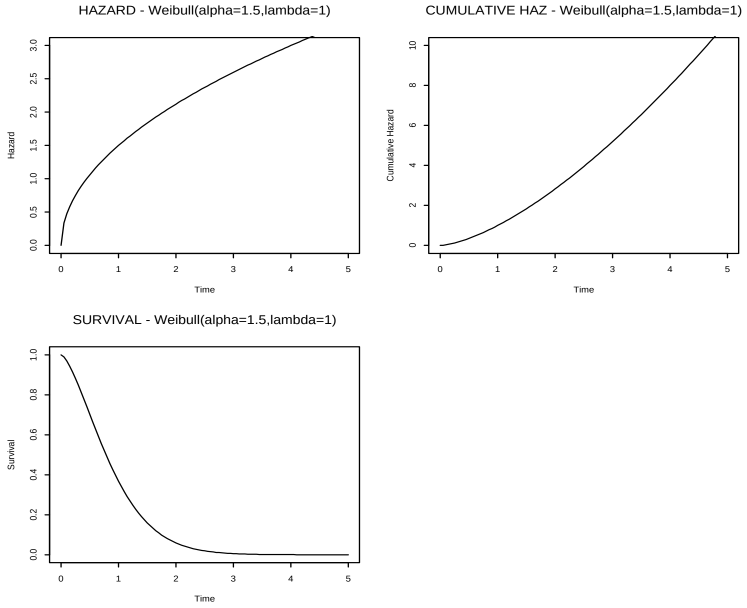
Example: exponential model, $h(t) = 1/\text{year}$



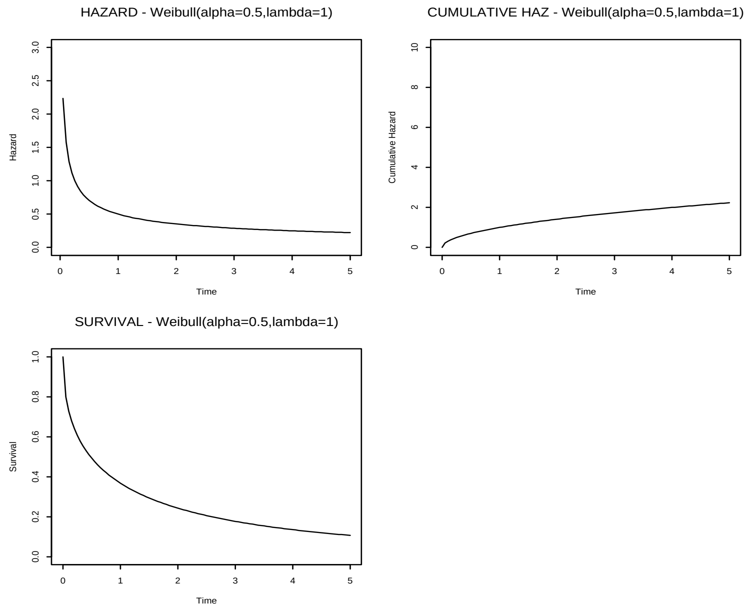
Example: exponential model, $h(t) = 2/\text{year}$



Example: Weibull model, $h(t) = 1.0 \cdot 1.5 \cdot t^{(1.5-1)}/\text{year}$



Example: Weibull model, $h(t) = 1.0 \cdot 0.5 \cdot t^{(0.5-1)}$/year



Motivation

- We can use Kaplan-Meier to characterize survival when there are a few large groups that we want to compare.

- With multiple covariates we can not stratify on all of the predictors at once.

- It is reasonable to expect that many different factors influence survival.

- How to use continuous covariates (without grouping)?.

$\boxed{\text{Proposal:}}$ A regression framework

∘ Cox (1972) proposed modeling the hazard function, $h(t)$, in a seminal paper "Regression Models and Life Tables (with Discussion)".

∘ Cox regression focuses on *hazard ratios*:

$$h(t, X_1)/h(t, X_2).$$

Hazard Models

Additive Model:

$$h(t, \boldsymbol{X}) \;=\; h_0(t) \;+\; \beta_1 \boldsymbol{X}_1 + \beta_2 \boldsymbol{X}_2 + \ldots + \beta_p \boldsymbol{X}_p$$

Multiplicative Model:

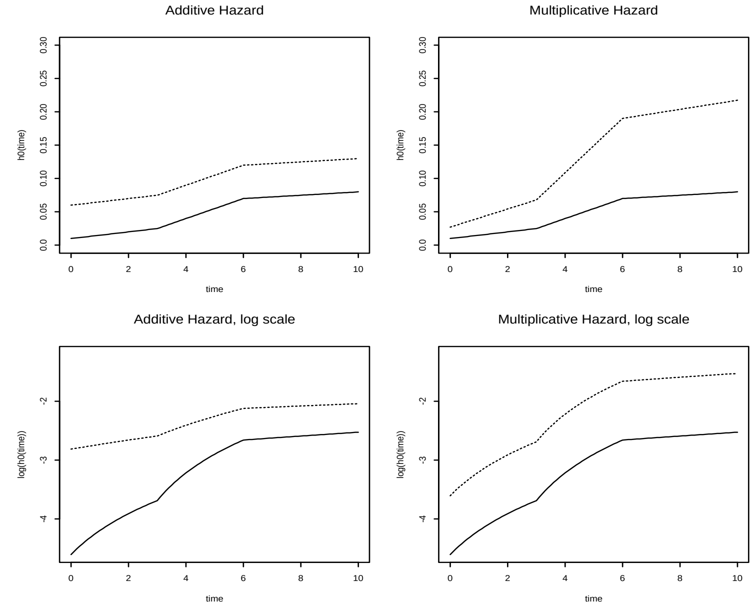
$$\log[h(t, \boldsymbol{X})] \;=\; \log[h_0(t)] \;+\; \beta_1 \boldsymbol{X}_1 + \beta_2 \boldsymbol{X}_2 + \ldots + \beta_p \boldsymbol{X}_p$$

$$h(t, \boldsymbol{X}) \;=\; h_0(t) \exp(\, \beta_1 \boldsymbol{X}_1 + \beta_2 \boldsymbol{X}_2 + \ldots + \beta_p \boldsymbol{X}_p \,)$$

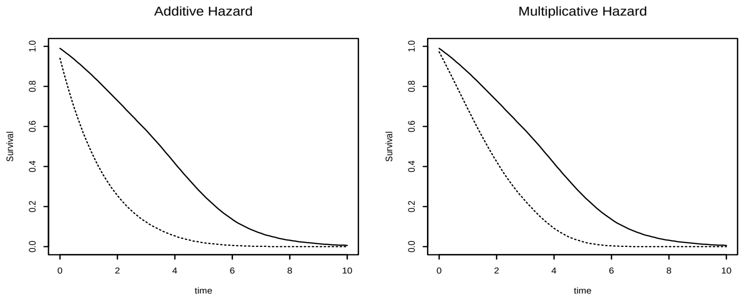
"Proportional Hazards Model"

- $h_0(t)$ is the *baseline hazard.*

Examples



Corresponding Survival Functions



Cox's Proportional Hazards Model

1. With the PH model we can handle several covariates simultaneously.
2. The construction of the model and the interpretation of the terms in the model is just like linear regression and logistic regression, except now we model *hazard ratios*.
3. One amazing contribution of Cox (1972) was an elegant likelihood method that allows estimation of the parameters of interest, $\boldsymbol{\beta}$, without having to estimate the *baseline hazard*, $h_0(t)$. This type of model is known as "semi-parametric" since there is a part of the model that is parametric ($\boldsymbol{\beta}$), and part of the model that is left unspecified (the non-parametric part is $h_0(t)$). The likelihood that Cox constructed is called a "partial likelihood".
4. The main concept is that we are using Cox regression to obtain comparisons between different groups, formed on the basis of covariates, in terms of their instantaneous probability of dying at any point in time. In other words, we model hazard rates.

Cox Regression: Assumptions

Independence:

- Independent observations.
- Independent censoring.

Proportionality:

∘ consider a single binary covariate:

$X = 1$ if treated, and $X = 0$ is control group.

∘ Use of the model

$$h(t, X) = h_0(t) \exp(\ \beta_1 X\ )$$

Implies that the risk of death among subjects in the treated group is $\exp(\beta_1)$ times the risk of death among subjects in the control group *at all times.*

$$\begin{aligned} h(t, X = 1) &= h_0(t) \exp(\beta_1) \\ h(t, X = 0) &= h_0(t) \exp(0) \end{aligned}$$

$$\frac{h(t, X = 1)}{h(t, X = 0)} = \exp(\beta_1)$$

Example: Remission Times

Treatment Group:

	time	status	tx	logwbc
1.	6	1	1	2.31
2.	6	1	1	4.06
3.	6	1	1	3.28
4.	7	1	1	4.43
5.	10	1	1	2.96
6.	13	1	1	2.88
7.	16	1	1	3.6
8.	22	1	1	2.32
9.	23	1	1	2.57
10.	6	0	1	3.2
11.	9	0	1	2.8
12.	10	0	1	2.7
13.	11	0	1	2.6
14.	17	0	1	2.16
15.	19	0	1	2.05
16.	20	0	1	2.01
17.	25	0	1	1.78
18.	32	0	1	2.2
19.	32	0	1	2.53
20.	34	0	1	1.47
21.	35	0	1	1.45

Example: Remission Times

Control Group:

	time	status	tx	logwbc
22.	1	1	2	2.8
23.	1	1	2	5
24.	2	1	2	4.91
25.	2	1	2	4.48
26.	3	1	2	4.01
27.	4	1	2	4.36
28.	4	1	2	2.42
29.	5	1	2	3.49
30.	5	1	2	3.97
31.	8	1	2	3.52
32.	8	1	2	3.05
33.	8	1	2	2.32
34.	8	1	2	3.26
35.	11	1	2	3.49
36.	11	1	2	2.12
37.	12	1	2	1.5
38.	12	1	2	3.06
39.	15	1	2	2.3
40.	17	1	2	2.95
41.	22	1	2	2.73
42.	23	1	2	1.97

STATA Command File:

```
infile time status tx logwbc using leuk2.dat

label variable time "time (weeks)"
label variable status "status"
label variable tx "treatment"
label variable logwbc "log(white blood cell count)"

list

***
*** recode tx
***
recode tx 1=0 2=1

label define tlab 0 "treatment" 1 "placebo"
label values tx tlab


***
*** summarize wbc by tx
***
sort tx
by tx:  summarize logwbc

***
*** center logwbc = important for survival!
***
generate newlwbc = logwbc-3.00


***
*** this defines the failure outcome
***
stset time, failure(status)
stset, noshow


***
*** Univariate analysis with treatment only
***
sts graph, by(tx) saving("leuk2-0.plot")
gphpen -dps "leuk2-0.plot"

***
```

```
*** Cox regression with TX
***
stcox tx, nohr basesurv(shat)
stcox tx

graph shat time, saving("leuk2-0a.plot")
gphpen -dps "leuk2-0a.plot"

stcoxkm, by(tx) saving("leuk2-0b.plot")
gphpen -dps "leuk2-0b.plot"


***
*** let's look at KM curves for levels of WBC
***
generate wbccat = logwbc
recode wbccat 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
label values wbccat wlab

table wbccat

sts graph, by(wbccat) saving("leuk2-1.plot")
gphpen -dps "leuk2-1.plot"

***
*** and log-rank test
***
sts test wbccat, logrank


***
*** Cox regression
***
stcox tx, nohr
stcox tx
lrtest, saving(1)

stcox tx newlwbc, nohr
stcox tx newlwbc
lrtest, saving(2)

xi: stcox i.tx*newlwbc, nohr
xi: stcox i.tx*newlwbc
lrtest, saving(3)

lrtest, using(3) model(2)
```



```
lrtest, using(2) model(1)


***
*** use model 2
***
stcox tx newlwbc, nohr basesurv( s0hat )


***
*** KM and adjusted KM
***
sts graph, by(tx) saving("leuk2-2.plot")
gphpen -dps "leuk2-2.plot"

sts graph, by(tx) adjustfor(newlwbc) saving("leuk2-3.plot")
gphpen -dps "leuk2-3.plot"
```

Cox Regression: Remission Data (tx only)

```
. ***
. *** Cox regression with TX
. ***
. stcox tx, nohr basesurv(shat)

Cox regression -- Breslow method for ties

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

------------------------------------------------------------------------------
      _t |
      _d |      Coef.   Std. Err.       z     P>|z|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
      tx |   1.509191   .4095644      3.685   0.000       .7064599    2.311923
------------------------------------------------------------------------------

. stcox tx

Cox regression -- Breslow method for ties

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

------------------------------------------------------------------------------
      _t |
      _d | Haz. Ratio   Std. Err.       z     P>|z|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
      tx |   4.523072   1.852489      3.685   0.000       2.026804    10.09382
------------------------------------------------------------------------------
```


Summary

1. Interpretation of the hazard.
2. Definition of the cumulative hazard.
3. $S(t) \Longleftrightarrow H(t) \Longleftrightarrow h(t)$
4. Examples using common parametric models (exponential model, weibull model).
5. Cox proportional hazards model:

$$h(t, \boldsymbol{X}) \;=\; h_0(t)\;\exp(\,\beta_1 X_1 + \beta_2 X_2 + \ldots\,)$$

6. Estimation and inference for hazard ratio regression parameters.

(*) Estimation of $\widehat{S}(t, \boldsymbol{X})$

Note:

$$
\begin{aligned}
H(t, X) &= \int_0^t h(s, X) ds \\
&= \int_0^t h_0(s) \exp(X\beta) ds \\
&= H_0(t) \exp(X\beta)
\end{aligned}
$$

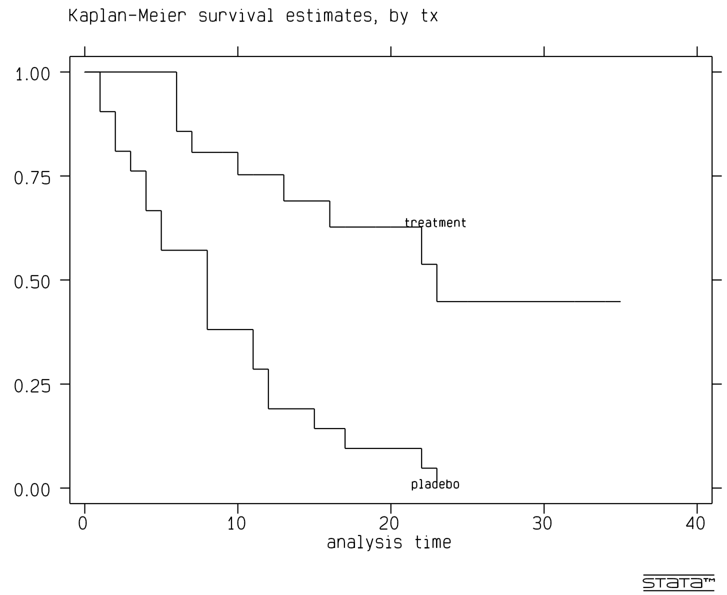
$$
\begin{aligned}
S(t, X) &= \exp(\; -H(t, X) \;) \\
&= \exp(-H_0(t) \cdot \exp(X\beta) \;)
\end{aligned}
$$

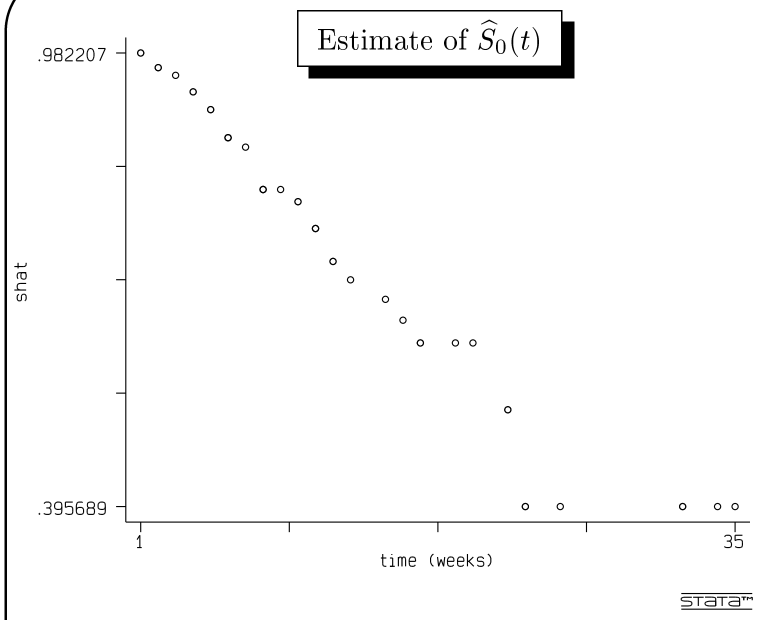
$$
S(t, X) = [\exp(-H_0(t) \;)]^{\exp(X\beta)}
$$

$$
S(t, X) = [S_0(t)]^{\exp(X\beta)}
$$

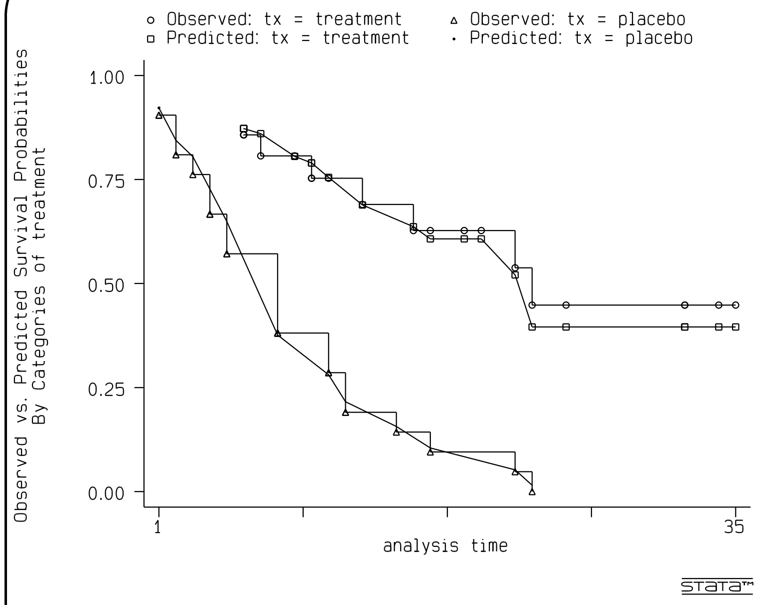
- An estimate of $S_0(t)$ is provided by STATA.

- From $\widehat{S}_0(t)$ and $\widehat{\boldsymbol{\beta}}$ we can obtain fitted survival functions for any covariate value(s).





Estimate of $\widehat{S}_0(t)$
.982207
.395689
shat
1
35
time (weeks)
STATA™



Observed: tx = treatment
Predicted: tx = treatment
Observed: tx = placebo
Predicted: tx = placebo
Observed vs. Predicted Survival Probabilities
By Categories of treatment
1.00
0.75
0.50
0.25
0.00
1
35
analysis time
STATA™

Recap on PH Model

1. We assume that the hazard ratio comparing $X = 1$ to $X = 0$ is 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, $\boldsymbol{\beta}$, we can obtain an estimate of the baseline survival function, $\widehat{S}_0(t)$, and fitted survival functions for any value of X.

Multiple Predictors

Example: Remission data

Response = time until death or relapse.

Covariates = treatment group, WBC count.

Models:

model 0

$$\log[h(t, \boldsymbol{X})] = \log[h_0(t)] + \beta_2 \log(\text{wbc})$$

model 1

$$\log[h(t, \boldsymbol{X})] = \log[h_0(t)] + \beta_1 \text{ Tx}$$

model 2

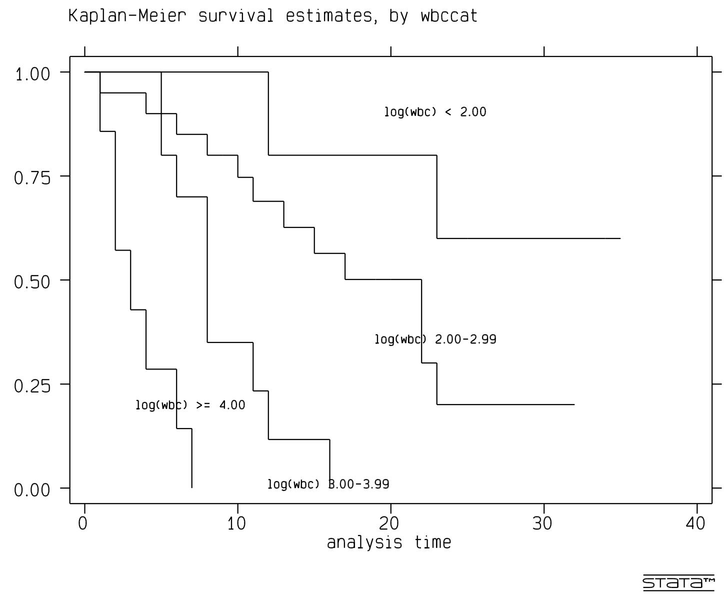
$$\log[h(t, \boldsymbol{X})] = \log[h_0(t)] + \beta_1 \text{ Tx} + \beta_2 \log(\text{wbc})$$

model 3

$$\log[h(t, \boldsymbol{X})] = \log[h_0(t)] + \beta_1 \text{ Tx} + \beta_2 \log(\text{wbc}) + \beta_3 \text{ Tx} \times \log(\text{wbc})$$

Remission Data:

```
. ***
. *** summarize wbc by tx
. ***
. sort tx
. by tx:  summarize logwbc

-> tx=treatment
Variable |     Obs        Mean   Std. Dev.       Min        Max
---------+-----------------------------------------------------
  logwbc |      21     2.63619   .7738764       1.45       4.43

-> tx=  placebo
Variable |     Obs        Mean   Std. Dev.       Min        Max
---------+-----------------------------------------------------
  logwbc |      21    3.224286   .9722786        1.5          5
.
. ***
. *** center logwbc = important for survival!
. ***
. generate newlwbc = logwbc-3.00

. ***
. *** let's look at KM curves for levels of WBC
. ***
. generate wbccat = logwbc
. recode wbccat 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 wbccat wlab
.
. table wbccat
-------------------+-----------
            wbccat |      Freq.
-------------------+-----------
   log(wbc) < 2.00 |          5
log(wbc) 2.00-2.99 |         20
log(wbc) 3.00-3.99 |         10
  log(wbc) >= 4.00 |          7
-------------------+-----------
.
. ***
. *** KM plots for wbc
. ***
. sts graph, by(wbccat) saving("leuk2-1.plot")
```

Model 1:

```
                                                  LR chi2(1)      =     15.21
Log likelihood  =   -86.379622                    Prob > chi2     =    0.0001

------------------------------------------------------------------------------
      _t |
      _d |      Coef.   Std. Err.       z     P>|z|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
      tx |   1.509191    .4095644      3.685   0.000       .7064599    2.311923
------------------------------------------------------------------------------
```

Model 2:

```
                                                  LR chi2(2)      =     43.41
Log likelihood  =    -72.27926                    Prob > chi2     =    0.0000

------------------------------------------------------------------------------
      _t |
      _d |      Coef.   Std. Err.       z     P>|z|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
      tx |   1.294067     .422104      3.066   0.002       .4667586    2.121376
 newlwbc |   1.604343    .3293283      4.872   0.000       .9588716    2.249815
------------------------------------------------------------------------------
```

Likelihood Ratio Test: H_0 : coefficient of log(wbc) = 0

```
. lrtest, using(2) model(1)
Cox:  likelihood-ratio test                          chi2(1)     =      28.20
                                                     Prob > chi2 =     0.0000
```


Model 3:

```
                                                  LR chi2(3)      =     43.84
Log likelihood  =   -72.065719                    Prob > chi2     =    0.0000

------------------------------------------------------------------------------
      _t |
      _d |      Coef.   Std. Err.       z     P>|z|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
   Itx_1 |   1.328354    .4252541      3.124   0.002       .4948713    2.161837
 newlwbc |   1.802788     .446717      4.036   0.000       .9272389    2.678337
ItXnew_1 |  -.3421952    .5197407     -0.658   0.510      -1.360868    .6764778
------------------------------------------------------------------------------
```

Likelihood Ratio Test:

H_0 : coefficient of (Tx)x(log(wbc)) = 0

```
. lrtest, using(2) model(1)
. lrtest, using(3) model(2)
Cox:  likelihood-ratio test                          chi2(1)     =       0.43
                                                     Prob > chi2 =      0.513
```

Model Summary

model	terms	$\exp(\widehat{\beta}_1)$	$-2\log L$
1	Tx	4.523	172.76
2	Tx + log(wbc)	3.648	144.56
3	Tx + log(wbc) + Tx · log(wbc)	3.774*	144.13

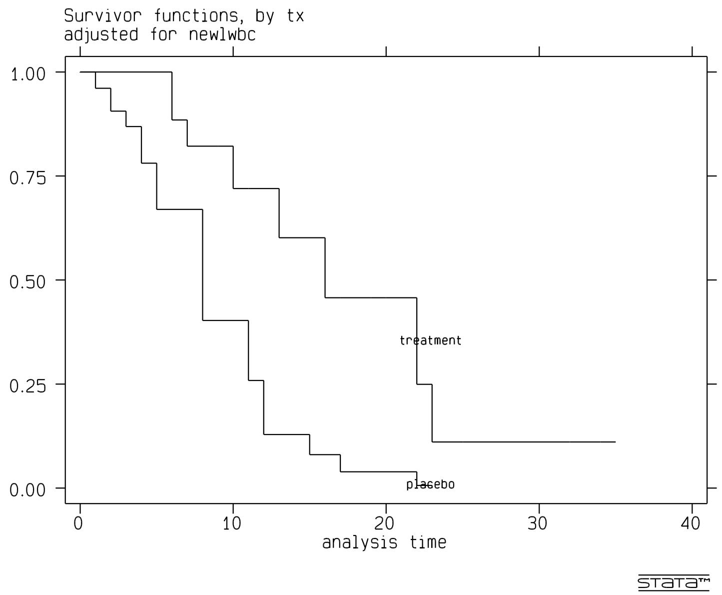
* for log(wbc)=3.0

test	LR stat	df	p-val
model 1 verus null	15.21	1	< 0.001
model 2 versus model 1	28.20	1	< 0.001
model 3 versus model 2	0.43	1	0.513

Survival for Tx groups – adjusted for log(WBC)

$$\widehat{S}(t, \mathrm{Tx}=1, \log(wbc)=3) = \left[\widehat{S}_0(t)\right]^{\exp(1.294)}$$

$$\widehat{S}(t, \mathrm{Tx}=0, \log(wbc)=3) = \left[\widehat{S}_0(t)\right]^{\exp(0.0)}$$



Estimating Hazard Ratios

Consider two values for the covariates

$$\boldsymbol{X}^{(0)} = (\boldsymbol{X}_1^{(0)}, \boldsymbol{X}_2^{(0)}, \ldots, \boldsymbol{X}_p^{(0)})$$
$$\boldsymbol{X}^{(1)} = (\boldsymbol{X}_1^{(1)}, \boldsymbol{X}_2^{(1)}, \ldots, \boldsymbol{X}_p^{(1)})$$

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

Model:

$$\begin{aligned} h(t, \boldsymbol{X}) &= h_0(t)\exp(\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p) \\ &= h_0(t)\exp(\sum_{j=1}^{p} X_j \beta_j) \end{aligned}$$

Hazard Ratio (HR):

$$\begin{aligned} h(t, \boldsymbol{X}^{(0)}) &= h_0(t)\exp(\sum_{j=1}^{p} \boldsymbol{X}_j^{(0)} \beta_j) \\ h(t, \boldsymbol{X}^{(1)}) &= h_0(t)\exp(\sum_{j=1}^{p} \boldsymbol{X}_j^{(1)} \beta_j) \\ HR &= \exp(\sum_{j=1}^{p} \boldsymbol{X}_j^{(1)} \beta_j - \sum_{j=1}^{p} \boldsymbol{X}_j^{(0)} \beta_j) \\ &= \exp(\sum_{j=1}^{p} \beta_j (\boldsymbol{X}_j^{(1)} - \boldsymbol{X}_j^{(0)})\ ) \end{aligned}$$

Example: Remission Data, Model 3

$$\begin{aligned} \boldsymbol{X}^{(1)} &= (\text{Tx} = 1, \text{newlwbc} = 0.5) \\ \boldsymbol{X}^{(0)} &= (\text{Tx} = 0, \text{newlwbc} = 0.5) \end{aligned}$$

$$\begin{aligned} \widehat{HR} &= \frac{\exp(\ 1.328(1.0) + 1.803(0.5) - 0.342(1.0)(0.5)\ )}{\exp(\ 1.328(0.0) + 1.803(0.5) - 0.342(0.0)(0.5)\ )} \\ &= \exp(\ 1.328 - 0.342(0.5)\ ) \ = \ 3.180 \end{aligned}$$

Summary

1. We evaluate **confounding** similar to other regression models – is there a meaningful change in the summary of interest (hazard ratio) after controlling for the potential confounder?
2. We use **Wald** and **Likelihood ratio** statistics to test whether certain coefficients are zero.
3. We can use the estimated PH regression coefficients to obtain risk comparisons in terms of hazard ratios.
4. We can use the estimated PH regression coefficients and the estimate of the baseline survival, $\widehat{S}_0(t)$, to obtain an estimate of the survival function for any covariate value, $\boldsymbol{X}$.
5. We assume that the hazards are **proportional** across the values of each covariate.
6. We assume that the comparison of hazards for $X = 1$ versus $X = 0$ does not depend on the time, t.
7. **Q**: How can we check the PH assumption?

Checking for proportionality

- Graphical approaches
  - $-\log\{-\log[S(t, \boldsymbol{X})]\}$ plots
  - Observed and fitted $S(t, \boldsymbol{X})$
  - Residual plots
- Confirmatory approaches
  - Test of goodness-of-fit
  - (Creating time-dependent variables)
- Correction
  - Stratification
  - Add covariate $\times$ (log) time to the model (Biostat 537!)

-log-log Plots

Recall: Under a PH assumption

$$S(t,X) = [S_0(t)]^{\exp(\beta X)}$$

$$\log[S(t,X)] = \exp(\beta X) \cdot \log[S_0(t)]$$

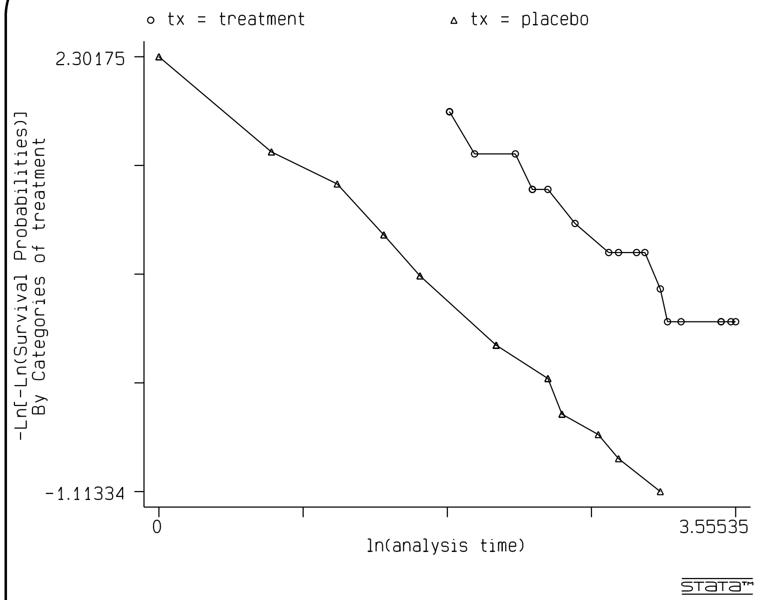
$$\log\{-\log[S(t,X)]\} = \beta X + \log\{-\log[S_0(t)]\}$$

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

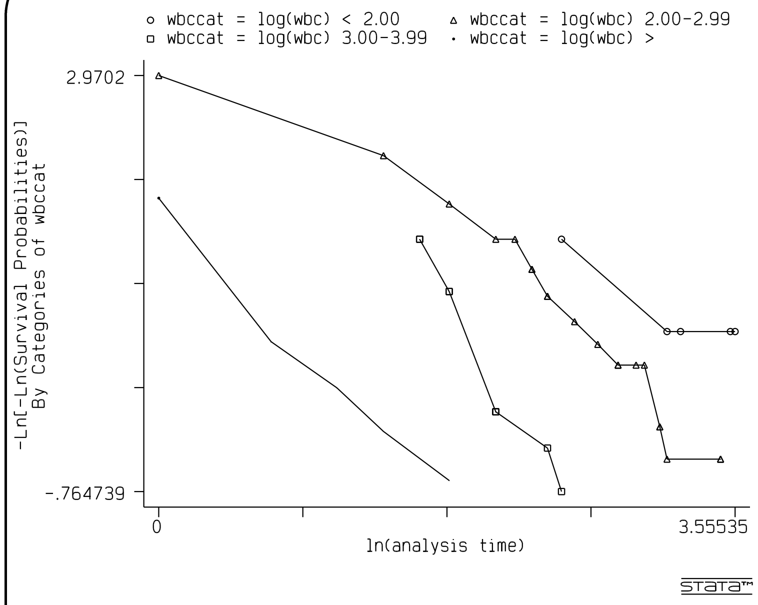
$$\beta = \log\{-\log[S(t,X=1)]\} - \log\{-\log[S(t,X=0)]\}$$

Idea:

- Plot $\log\{-\log[\widehat{S}(t,X)]\}$ versus time and look for "parallel" curves.



Remission data



Remission data

-log-log Plots

Comments:

- $-\log\{-\log[\widehat{S}(t,X)]\}$ or $\log\{-\log[\widehat{S}(t,X)]\}$
- Plot against time, or log(time).
- Use Kaplan-Meier for $\widehat{S}(t,X)$ (either unadjusted or adjusted).
- Crossing (in the middle) is an indication of trouble.
- Interpret plots recognizing that there is *variation* since these are *estimates* of the survival functions.

Issues:

- How parallel is parallel?
    - subjective decision
    - conservative strategy: assume PH is OK.
- Categorization of continuous predictors.
- Adjusted versus unadjusted $\widehat{S}(t,X)$.

Observed and Expected Survival

Idea:

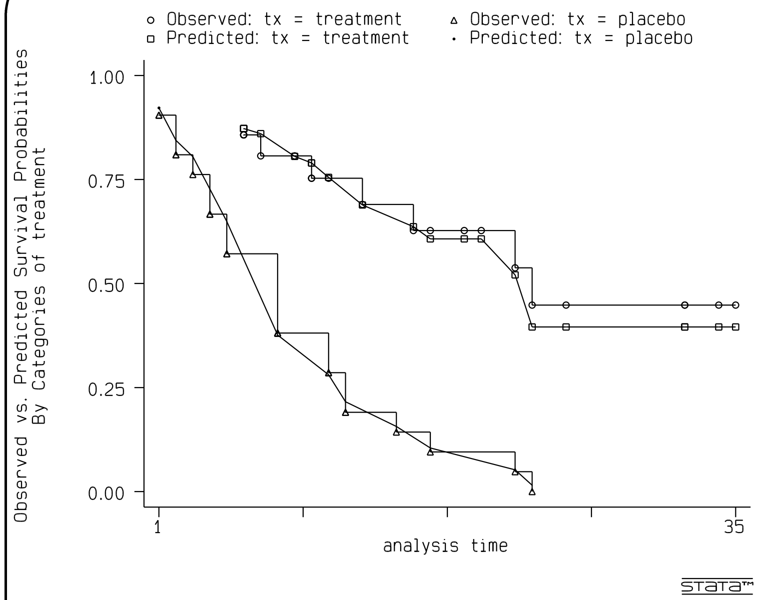
- Compare Kaplan-Meier estimates to fitted survival curves obtained from Cox regression.

Issues:

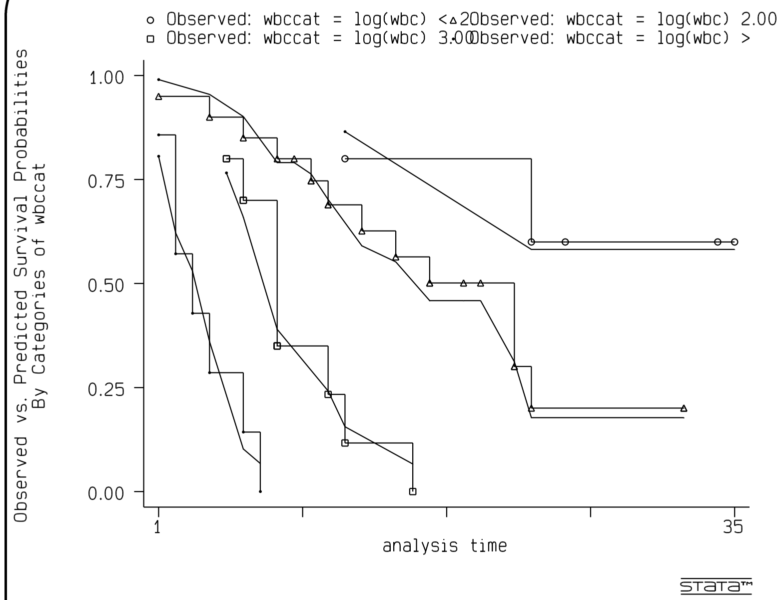
- If we adjust for other predictors in the Cox regression then we may impact the fitted survival. This can make comparison to KM estimates difficult (unless we can adjust those as well).

- How close is close?
    - Subjective decision

- Continuous covariates



Remission data



Remission data

Goodness-of-fit Tests

$\boxed{\star}$ Several packages (STATA - yes!) now include hypothesis tests for proportionality of hazards.

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

$$
\begin{aligned}
H_0 &: \beta_j(t) = \beta_j \\
H_1 &: \beta_j(t) \quad \text{has a trend in time}
\end{aligned}
$$

- Here $\exp(\ \beta_j(t)\ )$ represents the hazard ratio comparing $X_j = 1$ to $X_j = 0$ at time t, controlling for other predictors:

$$
\begin{aligned}
\frac{h(t, X_1 = 1, X_2 = x_2)}{h(t, X_1 = 0, X_2 = x_2)} &= \frac{h_0(t) \exp(\ \beta_1(t) \cdot (1) + \ \beta_2 x_2\ )}{h_0(t) \exp(\ \beta_1(t) \cdot (0) + \ \beta_2 x_2\ )} \\
&= \exp(\ \beta_1(t)\ ) \\
&? \ \exp(\ \beta_1\ )
\end{aligned}
$$

- These tests use a certain residual (Schoenfeld residual) that can also be used to check the PH assumption.

Cox regression: Remission data

```
. stcox tx newlwbc, nohr scaledsch(resid0*)

Cox regression -- Breslow method for ties

No. of subjects =           42                   Number of obs   =        42
No. of failures =           30
Time at risk    =          541
                                                 LR chi2(2)      =     43.41
Log likelihood  =    -72.27926                   Prob > chi2     =    0.0000

------------------------------------------------------------------------------
      _t |
      _d |      Coef.   Std. Err.       z     P>|z|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
      tx |   1.294067    .422104      3.066   0.002       .4667586    2.121376
 newlwbc |   1.604343   .3293283      4.872   0.000       .9588716    2.249815
------------------------------------------------------------------------------
```

Model Checking: Remission data

```
. ***
. *** Model checks
. ***
.
. stphtest, detail
note: cannot perform global test because schoenfeld(newvars) option was
      not specified when stcox was estimated

      Test of proportional hazards assumption

      Time:  Time
      ----------------------------------------------------------------
                  |       rho            chi2       df       Prob>chi2
      ------------+---------------------------------------------------
      tx          |      0.01159         0.00        1          0.9536
      newlwbc     |      0.03915         0.07        1          0.7960
      ------------+---------------------------------------------------
```


Residual Analysis

- For Cox regression there are several types of residuals!
  - Cox-Snell: overall model fit
  - Martingale: functional form for X's
  - Schoenfeld: checking the PH assumption
  - Score, Deviance: leverage, outliers

Schoenfeld:

Let $\boldsymbol{X}_i = (X_{i1}, X_{i2}, \ldots)$ be the covariate associated with the observed failure time, t_i. Let $R(i)$ represent the subjects that are at-risk for this failure time.

Define:

$$r_{ij} \quad = \quad X_{ij} - [\text{weighted average of the } X_j\text{'s for R(i) }]$$

$$r_{ij} \quad = \quad \text{“observed”} \quad - \quad [\text{“expected” under PH model}]$$

- There is a residual for each predictor variable.
- The residuals are only for the *observed* failure times.

Residual Analysis

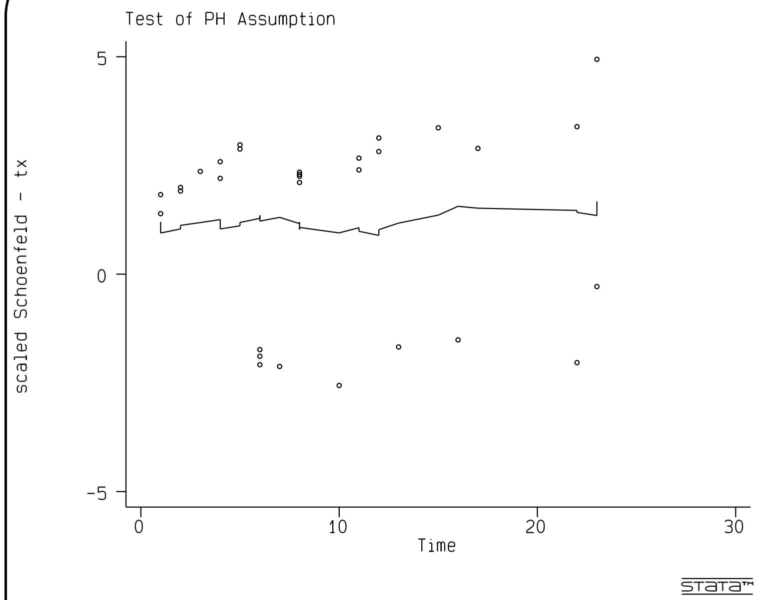
Use: Plot residual versus time.

• Interpretation:

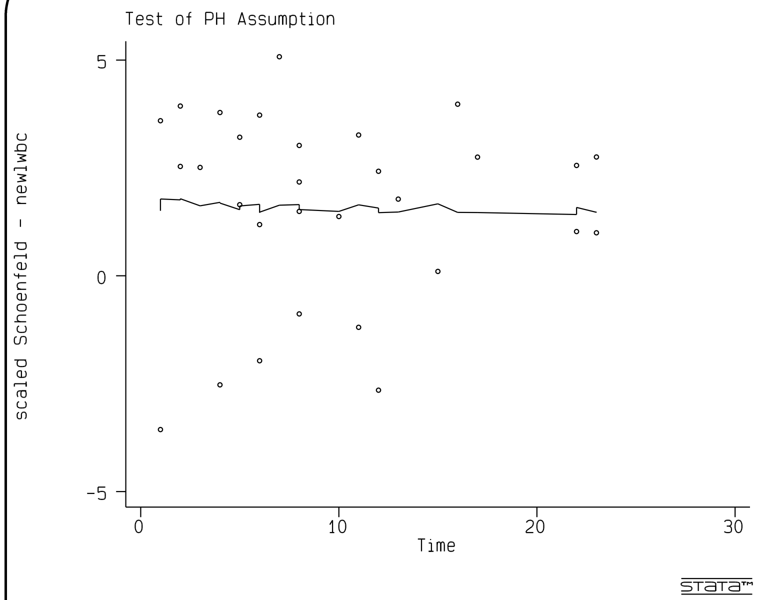
◦ If a smooth through the residuals is **constant** over time, then the agreement between the observed covariate (for the person who failed) and the prediction assuming a PH model is good.
⇒ PH assumption looks fine.

◦ If an **increasing** trend is observed, then the observed failures are occurring more often than expected among subjects with **high** values at later follow-up times.
⇒ Hazard ratio is **increasing** over time. PH violated.

◦ If a **decreasing** trend is observed, then the observed failures are occurring more often than expected among subjects with **low** values at the later follow-up times.
⇒ Hazard ratio is **decreasing** over time. PH violated.



Remission data



Remission data

Example: Methadone Treatment

The following analysis considers a dataset from a study by Caplehorn et al. ("Methadone Dosage and Retention of Patients in Maintenance Treatment", *Med. J. Aust.*, 1991). These data record the time in days spent by heroin addicts from entry to departure from one of two methadone clinics. There are two additional covariates, namely, *prison record* and *maximum methadone dose*, both believed to correlate with the time spent in the clinic.

Objectives:

- Describe the relationship between the covariates and time until clinic discharge.
- Is *prison* an important predictor?
- Is *dose* an important predictor?

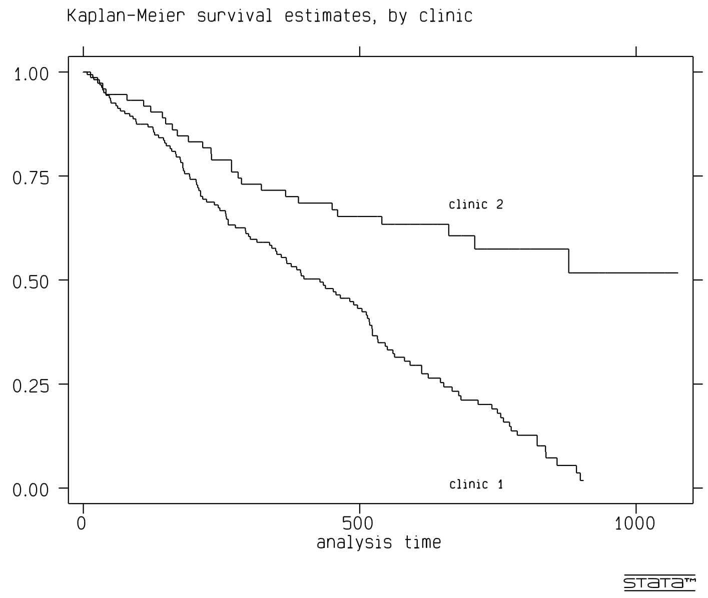
Exploratory Data Analysis :

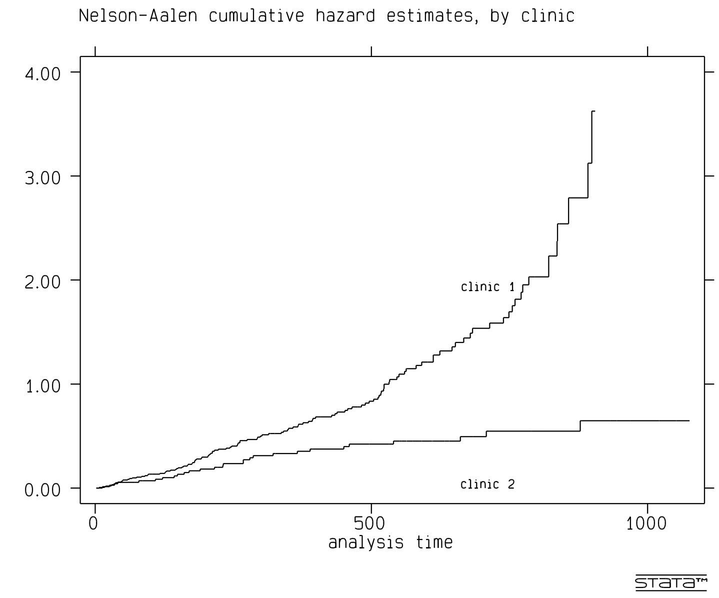
```
.
. ***
. *** EDA for predictors
. ***
. summarize dose

Variable |     Obs        Mean   Std. Dev.       Min        Max
---------+-----------------------------------------------------
    dose |     238    60.39916   14.45013         20        110

. centile dose, centile( 10 25 50 75 90 )

                                                   -- Binom. Interp. --
Variable |     Obs  Percentile      Centile        [95% Conf. Interval]
---------+-------------------------------------------------------------
    dose |     238         10            40              40          40
         |                 25            50              50          55
         |                 50            60              60          60
         |                 75            70              65     74.2803
         |                 90            80              80          80

.
. generate dosecat = dose

. recode dosecat min/49=1 50/59=2 60/69=3 70/max=4
(238 changes made)

. label define dlab 1 "dose <= 49" 2 "dose 50-59" 3 "dose 60-69" 4 "70 <= dose"

. label values dosecat dlab

.
. tabulate clinic prison, row chi

     study |     prison record
    clinic |        no        yes |     Total
-----------+----------------------+----------
  clinic 1 |        88         75 |       163
           |     53.99      46.01 |    100.00
-----------+----------------------+----------
  clinic 2 |        39         36 |        75
           |     52.00      48.00 |    100.00
-----------+----------------------+----------
```



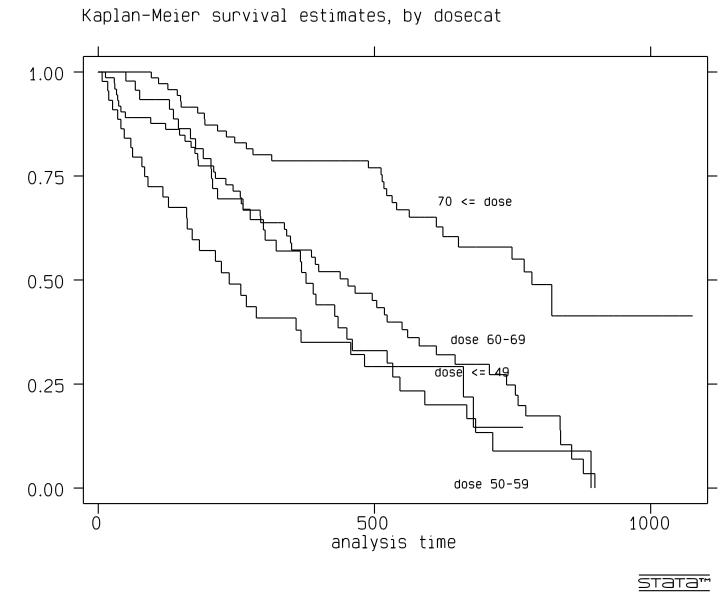
```
     Total |       127        111 |       238
           |     53.36      46.64 |    100.00

          Pearson chi2(1) =   0.0815   Pr = 0.775

.
. tabulate clinic dosecat, row chi

     study |                    dosecat
    clinic | dose <= 4  dose 50-5  dose 60-6  70 <= dos |     Total
-----------+--------------------------------------------+----------
  clinic 1 |        27         38         62         36 |       163
           |     16.56      23.31      38.04      22.09 |    100.00
-----------+--------------------------------------------+----------
  clinic 2 |        18         10         12         35 |        75
           |     24.00      13.33      16.00      46.67 |    100.00
-----------+--------------------------------------------+----------
     Total |        45         48         74         71 |       238
           |     18.91      20.17      31.09      29.83 |    100.00

          Pearson chi2(3) =  22.4646   Pr = 0.000

.
. tabulate prison dosecat, row chi

    prison |                    dosecat
    record | dose <= 4  dose 50-5  dose 60-6  70 <= dos |     Total
-----------+--------------------------------------------+----------
        no |        27         29         32         39 |       127
           |     21.26      22.83      25.20      30.71 |    100.00
-----------+--------------------------------------------+----------
       yes |        18         19         42         32 |       111
           |     16.22      17.12      37.84      28.83 |    100.00
-----------+--------------------------------------------+----------
     Total |        45         48         74         71 |       238
           |     18.91      20.17      31.09      29.83 |    100.00

          Pearson chi2(3) =   4.8712   Pr = 0.181

.
. sort clinic

. by clinic:  summarize dose

-> clinic= clinic 1
Variable |     Obs        Mean   Std. Dev.       Min        Max
---------+-----------------------------------------------------
```

```
    dose |      163    58.95706   12.40338         20         80

-> clinic= clinic 2
Variable |      Obs        Mean   Std. Dev.       Min        Max
---------+-----------------------------------------------------
    dose |       75    63.53333   17.81613         40        110


.
. sort prison

. by prison:  summarize dose

-> prison=       no
Variable |      Obs        Mean   Std. Dev.       Min        Max
---------+-----------------------------------------------------
    dose |      127    60.07874   15.73572         20        110

-> prison=      yes
Variable |      Obs        Mean   Std. Dev.       Min        Max
---------+-----------------------------------------------------
    dose |      111    60.76577   12.88407         40        100
```

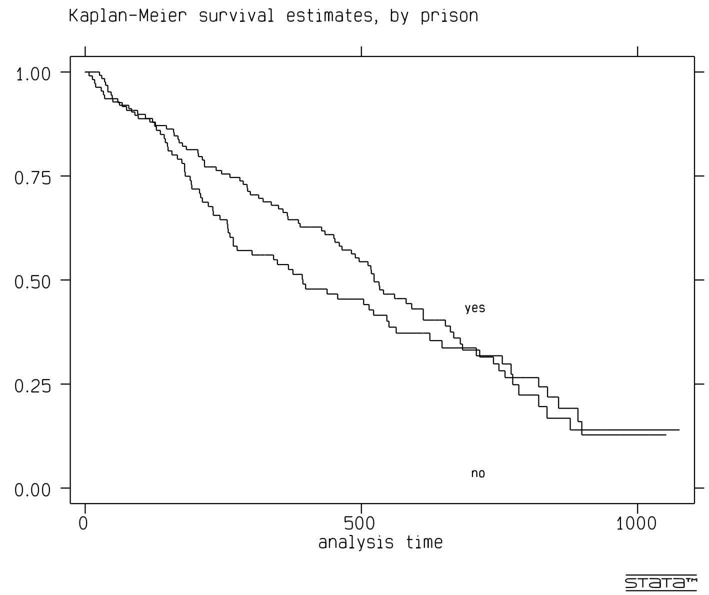





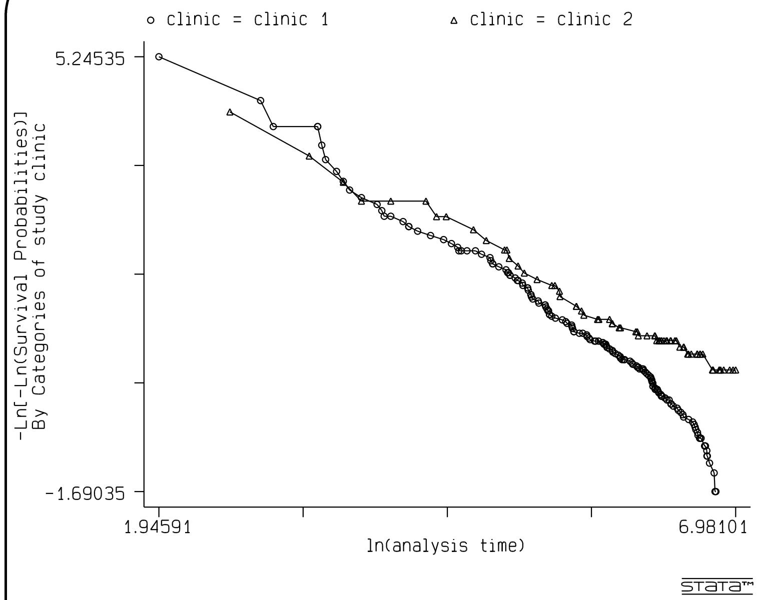
Nelson-Aalen cumulative hazard estimates, by clinic
4.00
3.00
2.00
1.00
0.00
0
500
1000
analysis time
clinic 1
clinic 2



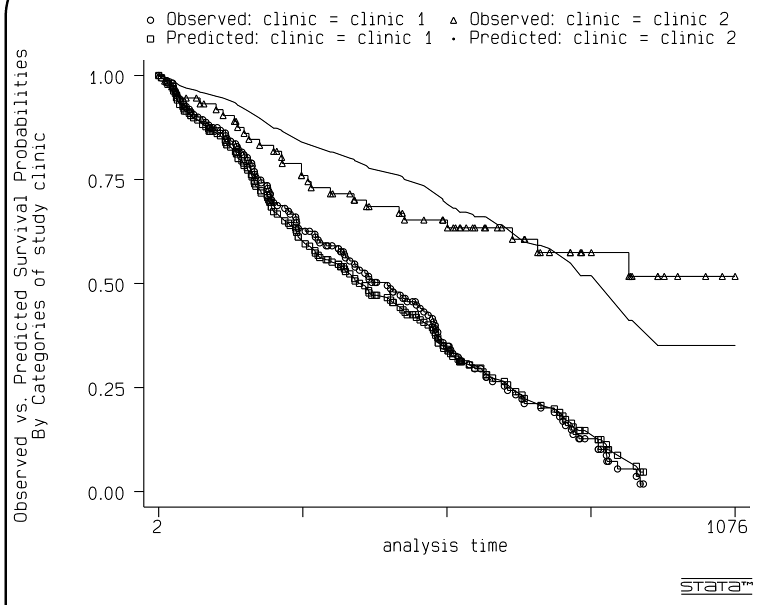
Kaplan-Meier survival estimates, by dosecat
1.00
0.75
0.50
0.25
0.00
0
500
1000
analysis time
70 <= dose
dose 60-69
dose <= 49
dose 50-59



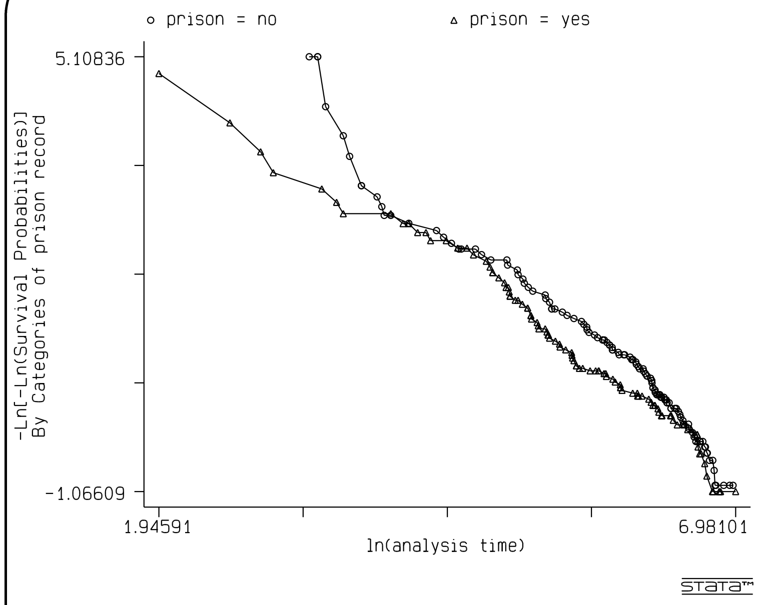
Kaplan-Meier survival estimates, by prison
1.00
0.75
0.50
0.25
0.00
0
500
1000
analysis time
yes
no



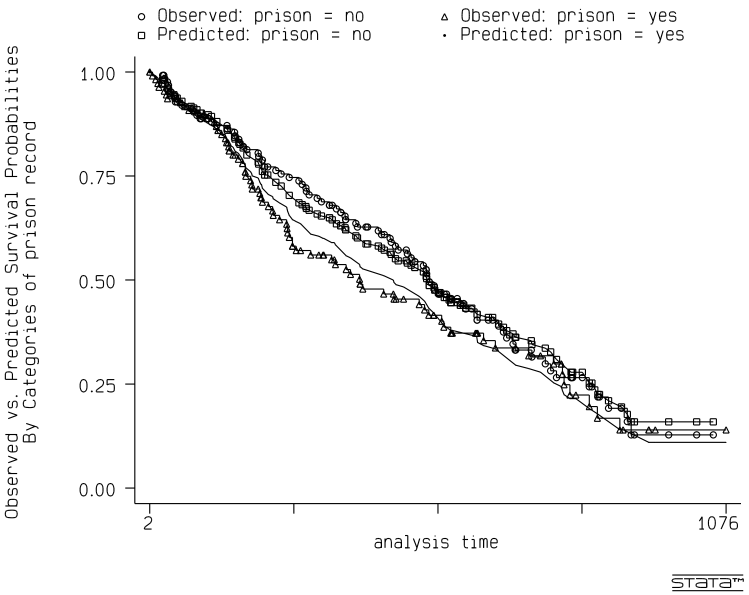
clinic = clinic 1
clinic = clinic 2
5.24535
-1.69035
1.94591
6.98101
ln(analysis time)
-Ln[-Ln(Survival Probabilities)]
By Categories of study clinic



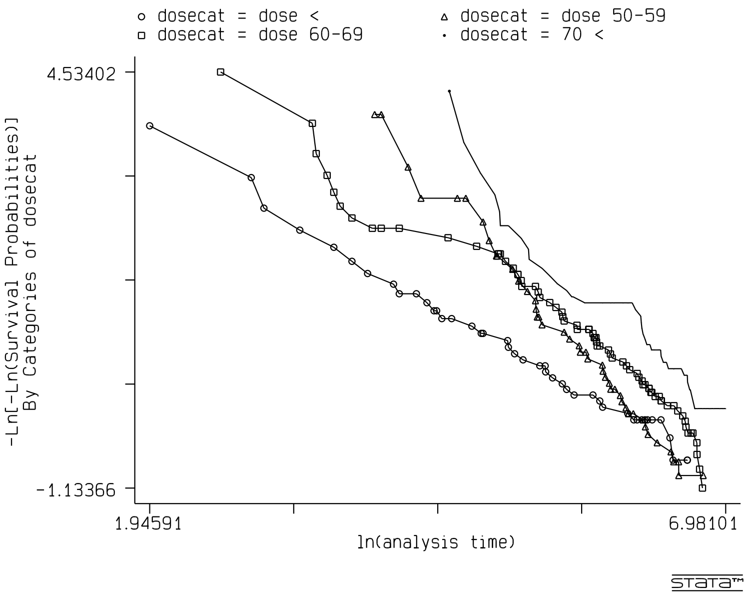
Observed: clinic = clinic 1
Observed: clinic = clinic 2
Predicted: clinic = clinic 1
Predicted: clinic = clinic 2
Observed vs. Predicted Survival Probabilities
By Categories of study clinic
1.00
0.75
0.50
0.25
0.00
2
1076
analysis time



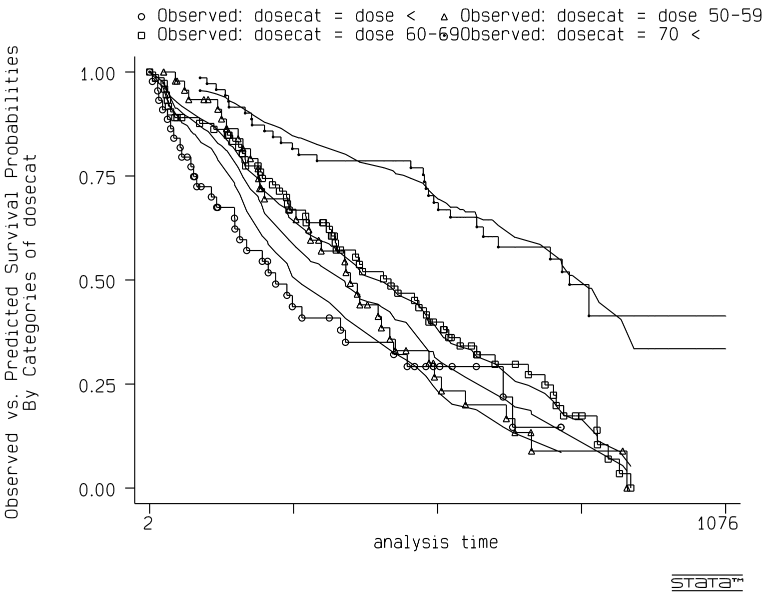
prison = no
prison = yes
-Ln[-Ln(Survival Probabilities)]
By Categories of prison record
5.10836
-1.06609
1.94591
6.98101
ln(analysis time)



Observed: prison = no
Observed: prison = yes
Predicted: prison = no
Predicted: prison = yes
Observed vs. Predicted Survival Probabilities
By Categories of prison record
1.00
0.75
0.50
0.25
0.00
2
1076
analysis time

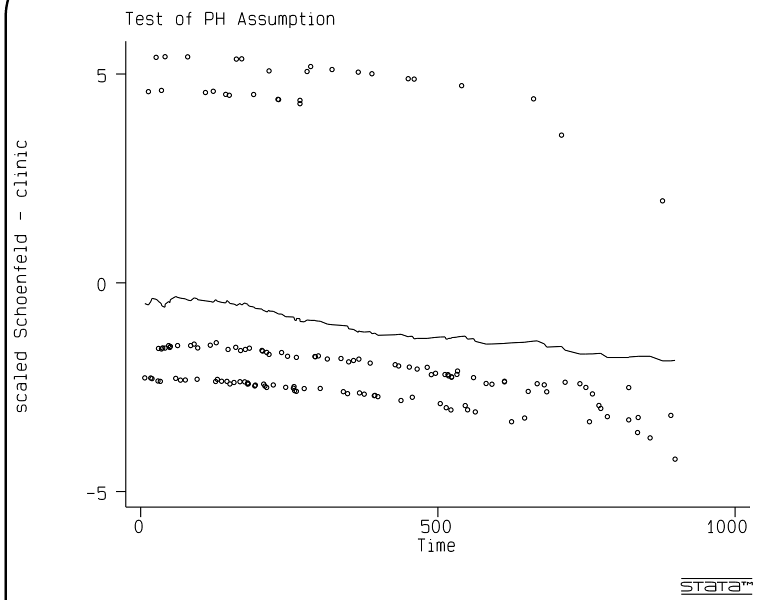


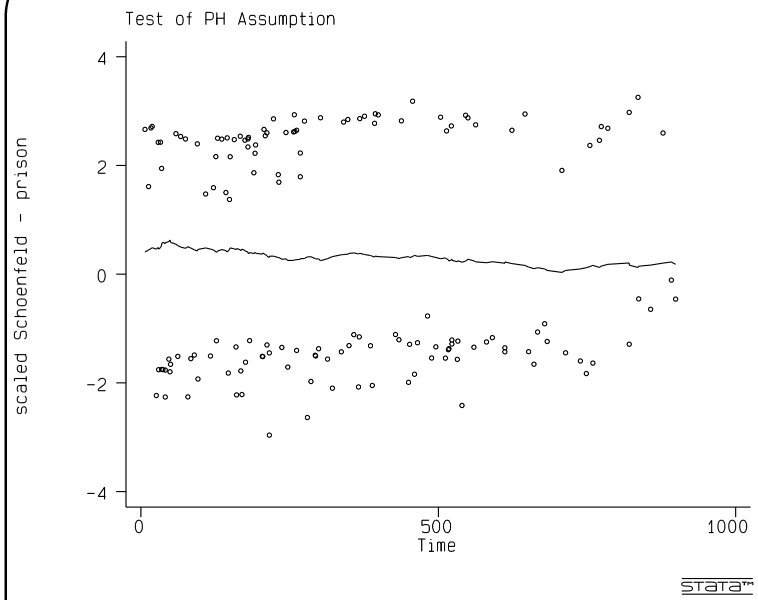
dosecat = dose <
dosecat = dose 50-59
dosecat = dose 60-69
dosecat = 70 <
-Ln[-Ln(Survival Probabilities)]
By Categories of dosecat
4.53402
-1.13366
1.94591
6.98101
ln(analysis time)

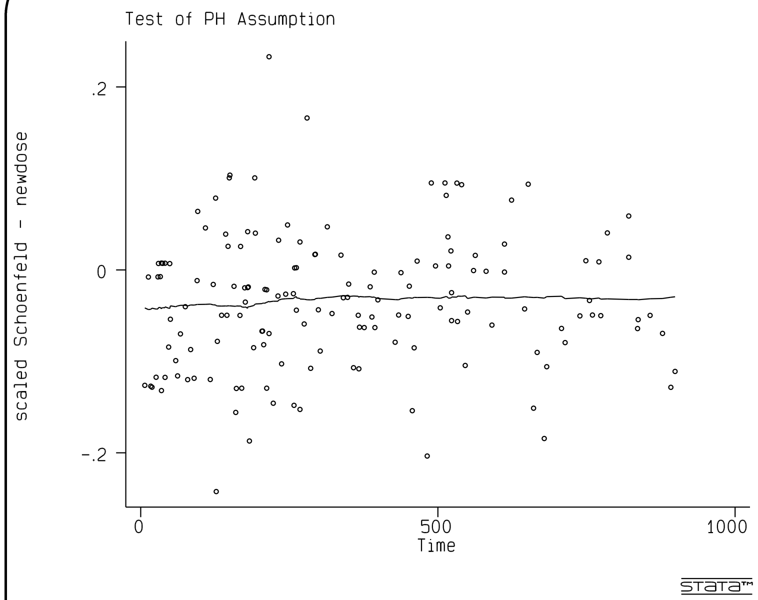


Confirmatory Analysis:

```
. ***
. *** Cox regression
. ***
. stcox clinic prison newdose, nohr basesurv(s0hat) scaledsch(resid0*)

Cox regression -- Breslow method for ties

No. of subjects =          238                  Number of obs   =       238
No. of failures =          150
Time at risk    =        95812
                                                LR chi2(3)      =     64.52
Log likelihood  =   -673.40242                  Prob > chi2     =    0.0000

------------------------------------------------------------------------------
      _t |
      _d |      Coef.   Std. Err.       z     P>|z|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
  clinic |   -1.00887   .2148709     -4.695   0.000      -1.430009   -.5877304
  prison |   .3265108   .1672211      1.953   0.051      -.0012366    .6542581
 newdose |  -.0353962   .0063795     -5.548   0.000      -.0478997   -.0228926
------------------------------------------------------------------------------
.
.
. *** Model checks
.
. stphtest, detail
note: cannot perform global test because schoenfeld(newvars) option was
      not specified when stcox was estimated

      Test of proportional hazards assumption

      Time:  Time
      ----------------------------------------------------------------
                  |       rho            chi2       df       Prob>chi2
      ------------+---------------------------------------------------
      clinic      |    -0.26344         11.66        1         0.0006
      prison      |    -0.03654          0.20        1         0.6541
      newdose     |     0.06184          0.51        1         0.4748
      ------------+---------------------------------------------------
```


Test of PH Assumption
scaled Schoenfeld - clinic
5
0
-5
0
500
1000
Time
STATA™


Test of PH Assumption
scaled Schoenfeld - prison
4
2
0
-2
-4
0
500
1000
Time
STATA™



What to do about Clinic?

Q: Can we still make PH inference about *prison* and *dose* even though *clinic* does not satisfy the PH assumption?

A: Yes. In order to do this we can perform a "stratified" analysis. This is different than using dummy variables, and is different than using separate analyses by clinic.

Idea:

We can use a model where within each clinic we have the same PH model, but we allow clinics to have different baseline hazards:

$$\begin{aligned} \text{clinic 1} \quad &: \quad h(t, \boldsymbol{X}) \;=\; h_{0,1}(t) \exp(\, \beta_1 \cdot \text{prison} \;+\; \beta_2 \cdot \text{dose}\,) \\ \text{clinic 2} \quad &: \quad h(t, \boldsymbol{X}) \;=\; h_{0,2}(t) \exp(\, \beta_1 \cdot \text{prison} \;+\; \beta_2 \cdot \text{dose}\,) \end{aligned}$$

Such a model is said to be "stratified on clinic". It is analogous to allowing interactions between clinic and time, but we still have additive terms for the other covariates (no dependence on time).

Stratified Analysis:

```
. ***
. *** Revised Cox regression
. ***
. stcox prison newdose, strata(clinic) nohr basesurv(s1hat) scaledsch(resid1*)

Stratified Cox regr. -- Breslow method for ties

No. of subjects =          238                   Number of obs   =       238
No. of failures =          150
Time at risk    =        95812
                                                 LR chi2(2)      =     33.94
Log likelihood  =     -597.714                   Prob > chi2     =    0.0000

------------------------------------------------------------------------------
      _t |
      _d |      Coef.   Std. Err.       z     P>|z|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
  prison |   .3887882   .1689154      2.302   0.021       .0577201    .7198563
 newdose |  -.0351449    .006465     -5.436   0.000      -.0478162   -.0224737
------------------------------------------------------------------------------
                                                           Stratified by clinic

.
. stphtest, detail
note: cannot perform global test because schoenfeld(newvars) option was
      not specified when stcox was estimated

      Test of proportional hazards assumption

      Time:  Time
      ----------------------------------------------------------------
                  |       rho            chi2       df       Prob>chi2
      ------------+---------------------------------------------------
      prison      |   -0.01671           0.04        1         0.8380
      newdose     |    0.07592           0.77        1         0.3788
      ------------+---------------------------------------------------
```


513 - Summary (Logistic Regression and Survival)

- **Logistic Regression**
  - Multivariate Problem
    - Binary response
    - General predictors
  - Logistic function
  - Logistic regression model
    - Linear predictor
  - Applying the logistic model
    - Probabilities
    - Odds ratios
  - Some special cases
    - Single binary exposure
    - Binary exposure / binary stratifying
      - Additive model
      - Interaction
      - Interpretations
  - Estimation
    - Random assumptions
      - Binomial with $n = 1$
      - Independence
    - Maximum likelihood
  - Inference

Wald Statistics
Likelihood Ratio Statistics

- More complicated models
  - Nominal predictors
  - Ordinal predictors
  - Continuous (quantitative) predictors
- Case-control analysis
  - Key assumptions
  - Ignore the intercept ⇒ only ORs
- Models for prediction
  - Assessment of accuracy
    - sensitivity
    - specificity
    - ROC curves
  - Validation data
- Model building strategies
  - Confirmatory analysis
    - Variable specification
    - Backward elimination – Interactions
    - Backward elimination – Confounders
  - Multiple comparisons

- **Survival Analysis**
  - Survival data characteristics
    - (time, status)
    - Right censoring
  - Survival function
  - Hazard function
  - Estimation of Survival
    - Life table method
    - Kaplan-Meier
      - Greenwood's standard errors
  - More on censoring
    - Independent censoring
  - Comparing survival curves
    - At a single time, t_0
    - Log-rank test
      - Weighted log-rank tests
  - Hazard, Cumulative Hazard, Survival
    - Definitions
    - Relationships
    - Examples
  - Cox proportional hazards model
    - Baseline hazard
    - Proportionality assumption
      - Examples
    - Estimation of $S(t, \boldsymbol{X})$ using PH model

Multiple predictors

Inference

- Wald
- Likelihood ratios

Estimating hazard ratios

Predictive model building (Lumley)

Checking the PH assumption

- log-minus-log plots
- Goodness-of-fit tests
- Residual plots

Stratified Cox regression