MODULE 4: INTRODUCTION TO SURVIVAL ANALYSIS

Summer Institute in Statistics for Clinical Research
University of Washington
July, 2016

Barbara McKnight, Ph.D.
Professor
Department of Biostatistics
University of Washington

SESSION 1: SURVIVAL DATA: EXAMPLES

Module 4: Introduction to Survival Analysis
Summer Institute in Statistics for Clinical Research
University of Washington
July, 2016

Barbara McKnight, Ph.D.
Professor
Department of Biostatistics
University of Washington

OVERVIEW

- Session 1
  - Introductory examples
  - The survival function
  - Survival Distributions
  - Mean and Median survival time

- Session 2
  - Censored data
  - Risk sets
  - Censoring Assumptions
  - Kaplan-Meier Estimator and CI
  - Median and CI

- Session 3
  - Two-group comparisons: logrank test
  - Trend and heterogeneity tests for more than two groups

- Session 4
  - Introduction to Cox regression

OVERVIEW – MODULE 8

Module 8: Survival analysis for Observational Data

- More complicated Cox models
  - Adjustment
  - Interaction
- Hazard function Estimation
- Competing Risks
- Choice of time variable
- Left Entry
- Time-dependent covariates

OVERVIEW – MODULE 12

Module 12: Survival analysis in Clinical Trials

- Estimating survival after Cox model fit
- More two-sample tests
  - Weighted logrank
  - Additional tests
- Adjustment, precision and post-randomization variables
- Power
- Choice of outcome
- Information accrual in sequential monitoring

PRELIMINARIES

- No prior knowledge of survival analysis techniques assumed
- Familiarity with standard one- and two-sample statistical methods (estimation and testing) is assumed
- Emphasis on application rather than mathematical details
- Examples

SESSIONS/BREAKS

- 8:30 – 10:00
  - Break until 10:30
- 10:30 – 12:00
  - Break until 1:30
- 1:30 – 3:00
  - Break until 3:30
- 3:30 – 5:00

WHAT IS SURVIVAL ANALYSIS ABOUT?

- Studies the occurrence of an event over time
  - Time from randomization to death (cancer RCT)
  - Time from acceptance into a heart transplant program to death
  - Time from randomization to diagnosis of Alzheimer's Disease
  - Time from birth to removal of supplementary oxygen therapy
  - Time from marriage until separation or divorce
  - Time until failure of light bulb
- Explores factors that are thought to influence the chance that the event occurs
  - Treatment
  - Age
  - Gender
  - Body Mass Index
  - Depression
  - ….. others

YOUR EXAMPLES

EXAMPLE 1

- Levamisole and Fluorouracil for adjuvant therapy of resected colon carcinoma
  Moertel et al, 1990, 1995
- 1296 patients
- Stage B_2 or C
- 3 unblinded treatment groups
  - Observation only
  - Levamisole (oral, 1yr)
  - Levamisole (oral, 1yr) + fluorouracil (intravenous 1yr)

EXAMPLE 1

- Randomization
  - Adaptive
  - B_2, extent of invasion, time since surgery
  - C, extent of invasion, time since surgery, number of lymph nodes involved

EXAMPLE 1

- Statistical analysis
  - Kaplan-Meier survival curves
  - Log-rank statistic
  - Cox proportional-hazards model for all multivariable analysis
  - Backward regression, maximal partial-likelihood estimate statistic
  - O'Brien-Fleming boundary for sequential monitoring

EXAMPLE 1

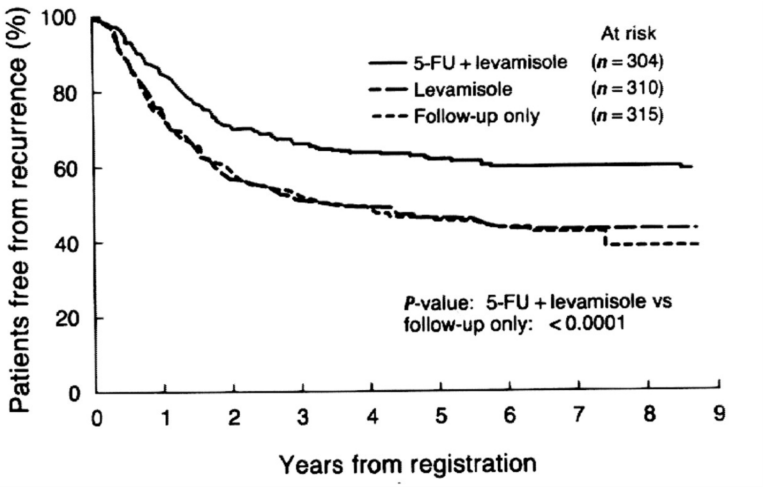


Figure 1: Recurrence-free interval according to treatment arm. Patients who died without recurrence have been censored. 5-FU = fluorouracil.

EXAMPLE 1

- Results (stage C) after 2nd interim analysis
- Fluorouracil + Levamisole reduced the
  - Recurrence rate by 40% (p<0.0001)
  - Death rate by 33% (p<0.0007)
- Levamisole reduced the
  - Recurrence rate by 2%
  - Death rate by 6%
- Toxicity was mild (with few exceptions)
- Patient compliance excellent

EXAMPLE 1

- R survival package data "colon"
  - 929 eligible patients (971 randomized – 42 ineligible)
  - Treatment groups (rx)
  - Sex, age
  - Obstruction of colon by tumor (obstruct)
  - Perforation of colon (perfor)
  - Adherence to nearby organs (adhere)
  - Number of lymph nodes with detectable cancer (nodes)
  - Days until event or censoring (time)
  - Censoring status (status)

EXAMPLE 1

- Multivariable analysis:
  - Proportional hazards model
  - "we kept the variable of treatment in the model and used backward regression for other covariates"
  - Other covariates ($P < 0.01$)
    - Depth of primary tumor invasion,
    - Invasion of adjacent structures
    - Regional implants
    - Number of metastatic lymph nodes
    - Histological differentiation
    - Preoperative carcinoembryonic antigen level

EXAMPLE 1

- Multivariable results:
  - "After adjustment for minor imbalances in prognostic variables among treatment arms, therapy with fluorouracil plus levamisole was again found to have an advantage over observation (40% reduction in recurrence rate; P < 0.0001)."
  - "Levamisole alone had no detectable advantage (2% reduction in recurrence rate; P = 0.86)."

EXAMPLE 2 – ALZHEIMER'S

- Petersen et al. 2005, NEJM
- Mild cognitive impairment
- Vitamin E and Donepezil and Placebo
- Time from randomization to AD diagnosis
- Length of treatment 3 years
- Double blind
- Outcome: Possible or probable AD

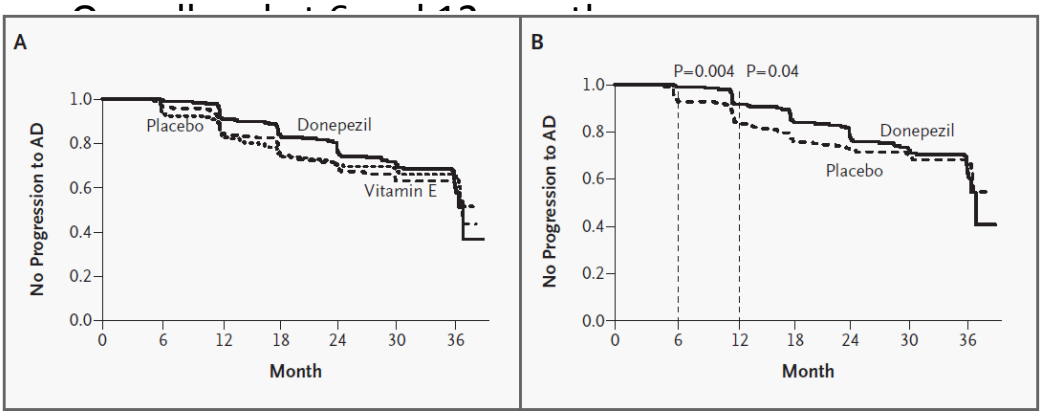
EXAMPLE 2 – ALZHEIMER'S

- 769 enrolled
- 212 developed possible or probable AD
- "There were no significant differences … during the three years of treatment"
- Vitamin E vs Placebo
  - Hazard Ratio 1.02 (95% CI, 0.74, 1.41), p-value 0.91
- Donepezil vs Placebo
  - Hazard Ratio 0.80 (95% CI, 0.57, 1.13), p-value 0.42

EXAMPLE 2 – ALZHEIMER'S

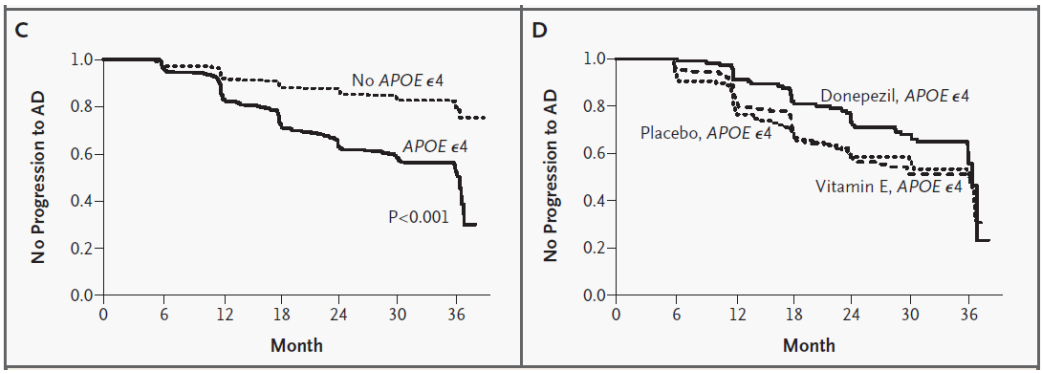
- Prespecified analyses
- At 6 months intervals
  - Donepezil vs Placebo significantly reduced likelihood of progression to AD during the first 12 months (p-value 0.04)
  - Finding supported by secondary outcome measures
  - Subgroup ≥ 1 apolipoprotein E ϵ4 alleles significantly reduced likelihood of progression to AD over 3 years
  - Vitamin E vs Placebo: no significant differences
  - Vitamin E vs Placebo: also no significance for above subgroup

EXAMPLE 2 – RESULTS



EXAMPLE 2 – RESULTS

- APOE ϵ4 results



EDITORIAL

- "long-awaited results"
- Donepezil standard therapy for AD
- "Implications …. Enormous"
  - Clear-cut negative findings for Vitamin E
  - Especially noteworthy
  - Despite dearth of evidence of its efficacy

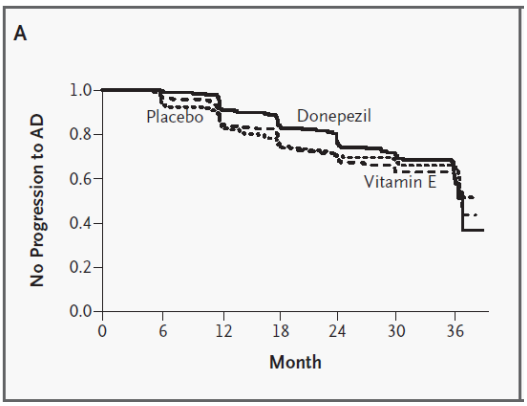
  - Findings for donepezil "much less clear"
  - "not quite as disappointing"

EDITORIAL COMMENTS

- "rate of progression … somewhat lower in the treatment group during the first year of the study"
- "by two years, even this small effect had worn off"
- Possible explanation: "Reduced statistical power later in the study as the number of subjects at risk declined owing to death, withdrawal and development of AD
- Secondary analyses suggest… benefits wore off

EXAMPLE 2 – RESULTS

- Interesting steps…..



"COUNTER" EXAMPLE

- Resuscitation Outcomes Consortium
  - Out-of-hospital cardiac arrest
  - Traumatic injury
- Prehospital interventions
- Exception from informed consent
- 10 Regional Centers
  - 7 US
  - 3 Canada

"COUNTER" EXAMPLE

- Times
  - Event (cardiac arrest, traumatic injury)
  - 911 call
  - Arrival of EMS
  - Treatment start
  - Potential outcomes
    - Return of spontaneous circulation (Cardiac arrest)
    - ED admission
    - Survival to hospital discharge
    - Neurologically intact survival
    - 28-day survival
    - 6-month neurological outcomes

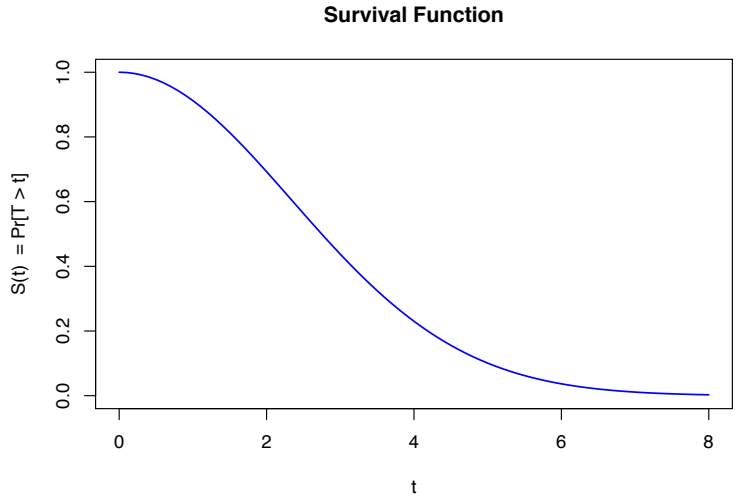
"COUNTER" EXAMPLE

- Time of injury/cardiac arrest
  (ordinarily unknown)
- 911 call
- Cardiac arrest: Many deaths before admission to hospital
- Trauma: Many deaths within the first 24 – 48 hours

SURVIVAL DATA AND FUNCTION

- Original applications in biometry were to survival times in cancer clinical trials
- Many other applications in biometry: eg. disease onset ages
- Interest centers not only on average or median survival time but also on probability of surviving beyond 2 years, 5 years, 10 years, etc.
- Best described with the entire survival function S(t).
  - For T = a subject's survival time, S(t) = P[T> t].
  - Characterizes the entire distribution of survival times T.
  - Gives useful information for each t.

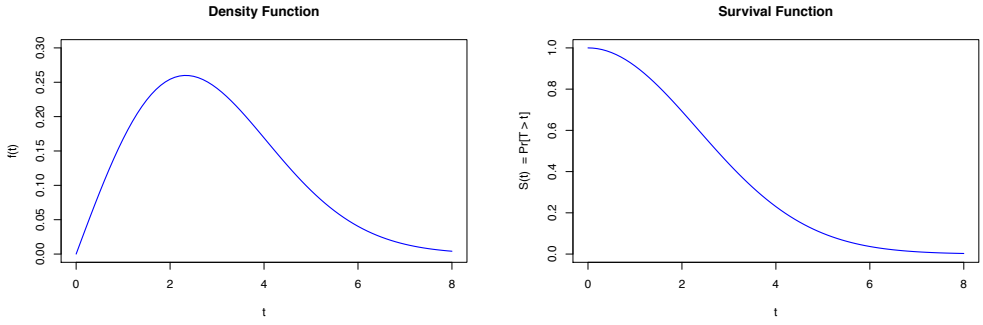
SURVIVAL FUNCTION



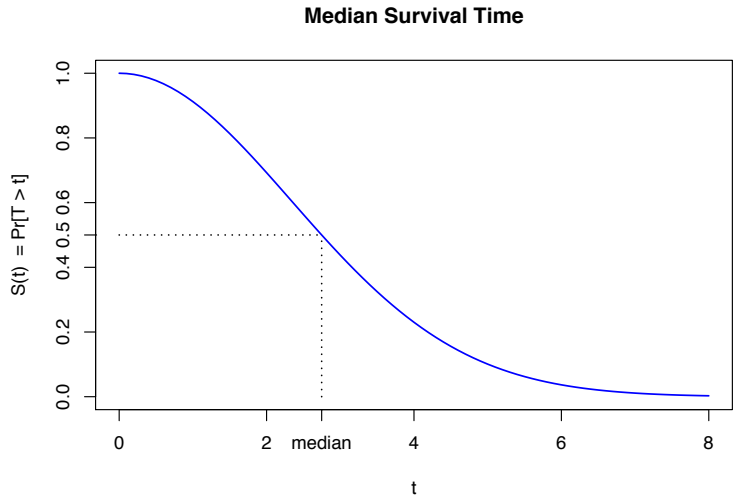
SURVIVAL DISTRIBUTION

- Continuous probability distribution of times *T*
- Only non-negative *T's* are possible: Pr(*T*<0)=0
- Density function $f(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \Pr(t \le T < t + \Delta t)$
- Area under the $f(t)$ curve between two points is the probability *T* is between the two points.

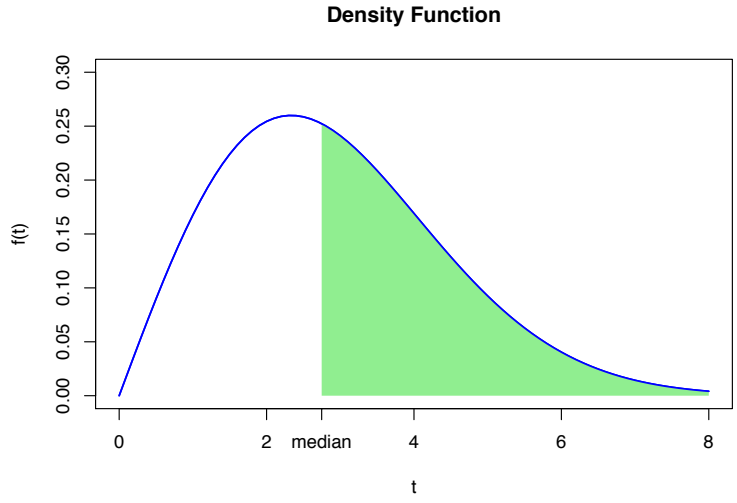
DENSITY AND SURVIVAL FUNCTIONS



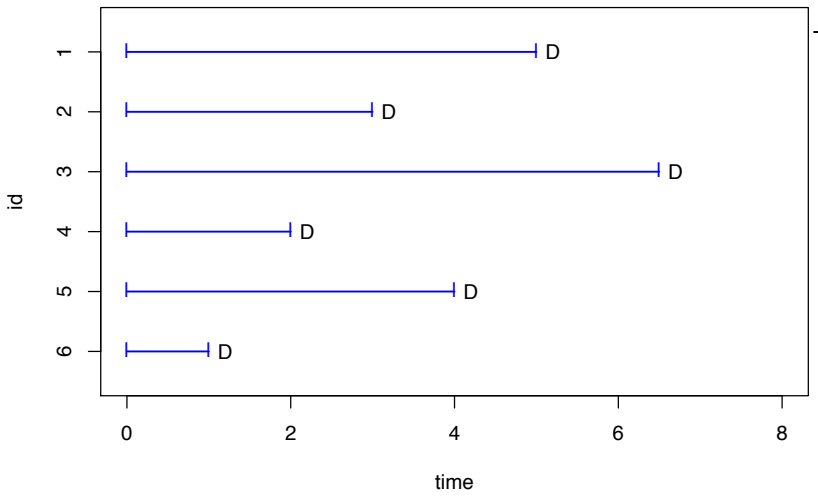
MEDIAN SURVIVAL TIME



MEDIAN SURVIVAL TIME



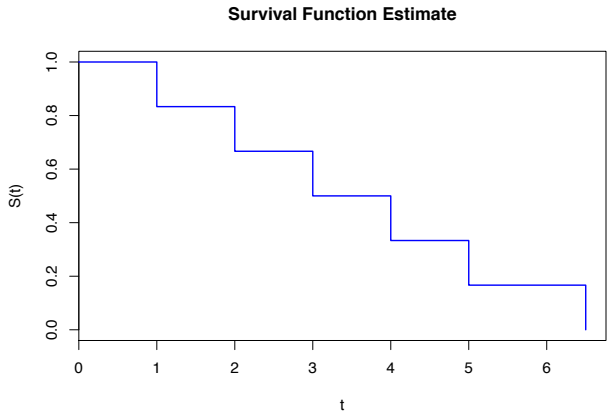
ILLUSTRATIVE DATA



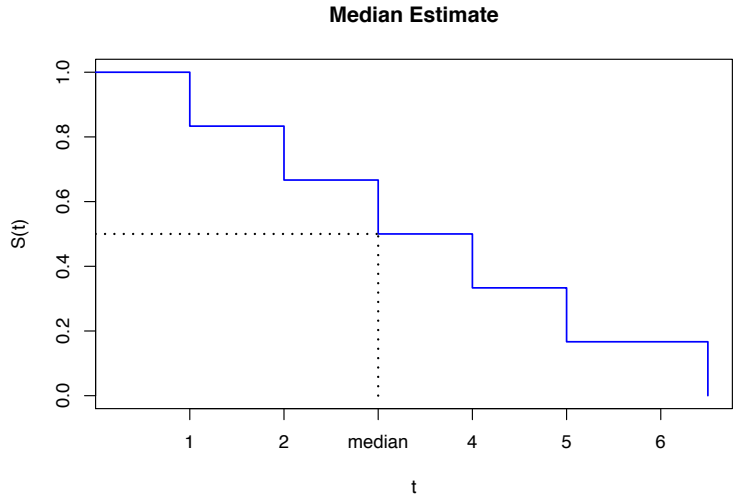
id	Y
1	5
2	3
3	6.5
4	2
5	4
6	1

SURVIVAL FUNCTION ESTIMATE

- Nonparametric Estimate: reduce estimate by 1/n every time there is an event (death): Empirical survival function estimate



MEDIAN ESTIMATE



By convention: median is earliest time where survival estimate ≤ .5

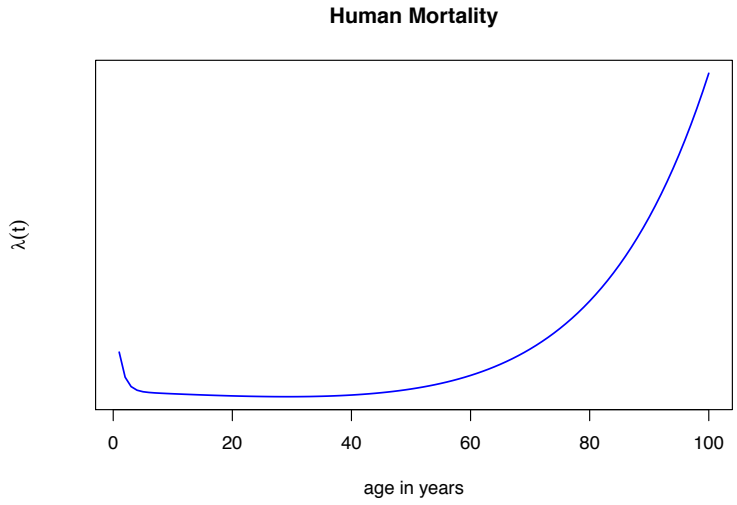
OTHER WAYS TO DESCRIBE A SURVIVAL DISTRIBUTION

- So far we have looked at the density function and survival function S(t).
- Also of interest: "hazard" function λ(t)

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \Pr[t \leq T < t + \Delta t | T \geq t]$$

- Instantaneous rate at which death occurs at t in those who are alive at t
- Examples:
  - Age-specific death rate
  - Age-specific disease incidence rate

HAZARD FUNCTION FOR HUMANS

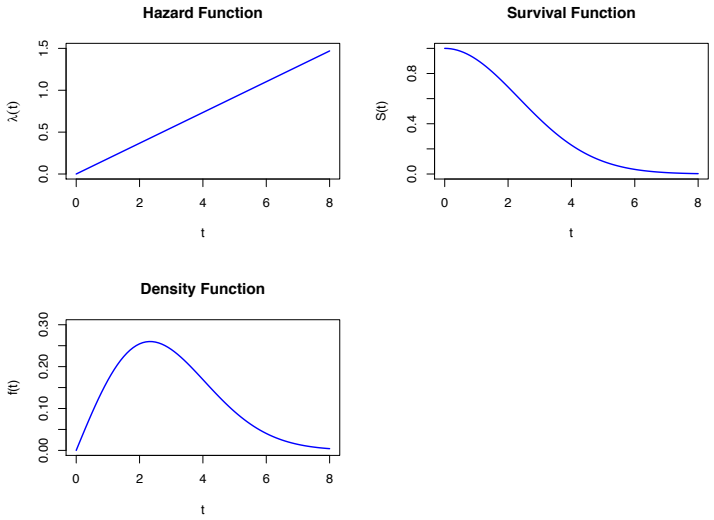


EQUIVALENT CHARACTERIZATIONS

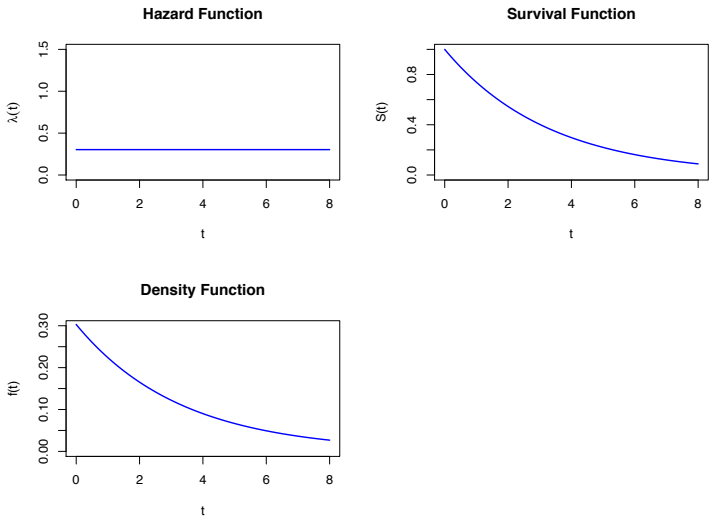
- Any <u>one</u> of the density function(f(t)), the survival function(S(t)) or the hazard function(λ(t)) is enough to determine the survival distribution.
- They are each functions of each other:

  - $S(t) = \int_t^\infty f(s)ds = e^{-\int_0^t \lambda(s)ds}$
  - $f(t) = -\frac{d}{dt}S(t) = \lambda(t)e^{-\int_0^t \lambda(s)ds}$
  - $\lambda(t) = \frac{f(t)}{S(t)}$

EQUIVALENT CHARACTERIZATIONS



EQUIVALENT CHARACTERIZATIONS



In R

Call up packages we will use (assumes installed)

```
library(survival)
library(ggplot2)
library(ggfortify)
library(rms)
```

Get data (in survival package)

```
data(veteran)
```


Look at data.

```
head(veteran)
```

```
##   trt celltype time status karno diagtime age prior
## 1   1 squamous   72      1    60        7  69     0
## 2   1 squamous  411      1    70        5  64    10
## 3   1 squamous  228      1    60        3  38     0
## 4   1 squamous  126      1    60        9  63    10
## 5   1 squamous  118      1    70       11  65    10
## 6   1 squamous   10      1    20        5  49     0
```

Survival Curve

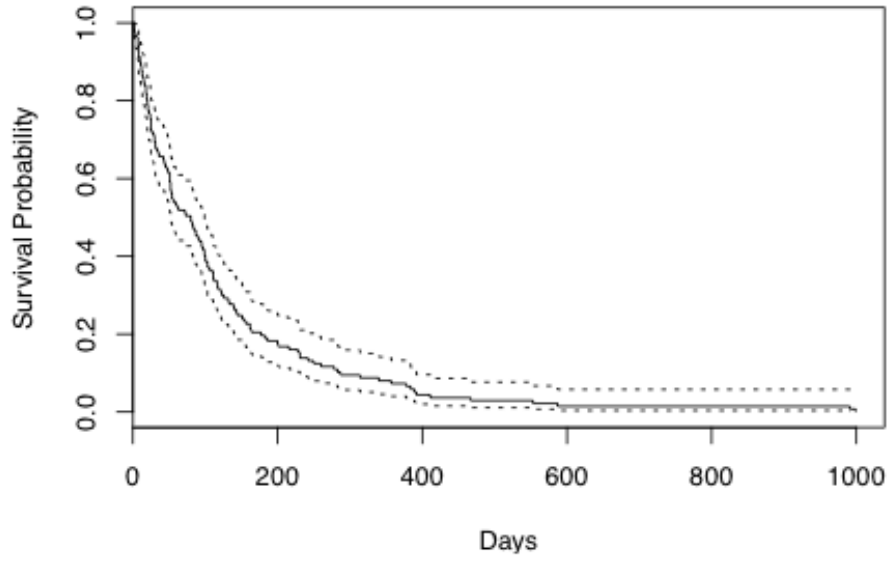
Survival time variable, make survival object and get descriptives

```
Y <- Surv(veteran$time)
Shat <- survfit(Y ~ 1)
Shat
```

```
## Call: survfit(formula = Y ~ 1)
##
##       n  events  median 0.95LCL 0.95UCL
##     137     137      80      52      99
```

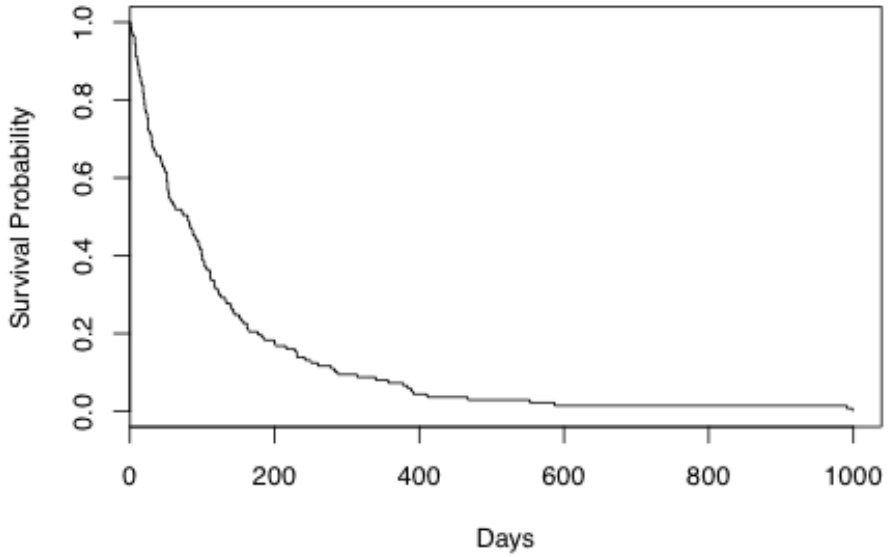

Plot Survival Curve

```
plot(Shat, xlab = "Days", ylab = "Survival Probability")
```



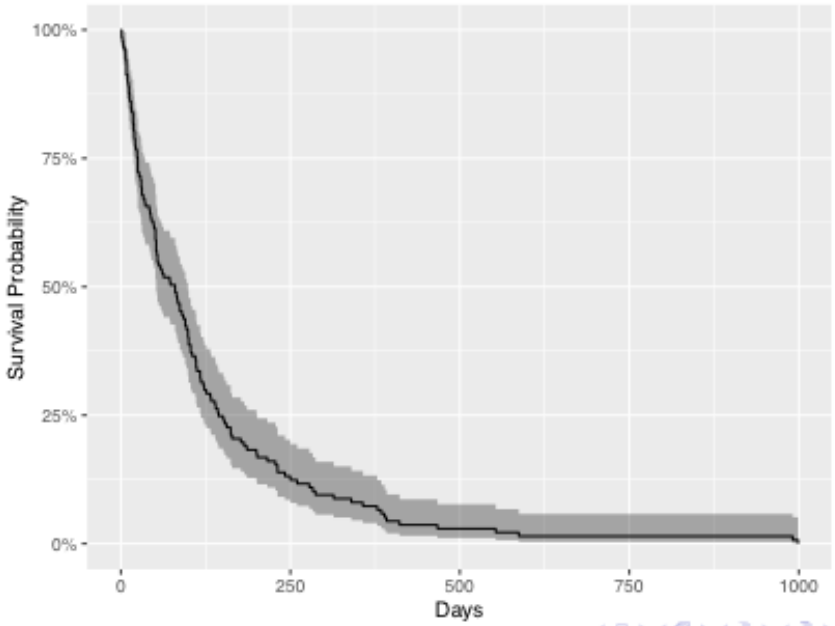
Plot Survival Curve: Other Options

```
plot(Shat, conf.int = FALSE, xlab = "Days",
     ylab = "Survival Probability")
```



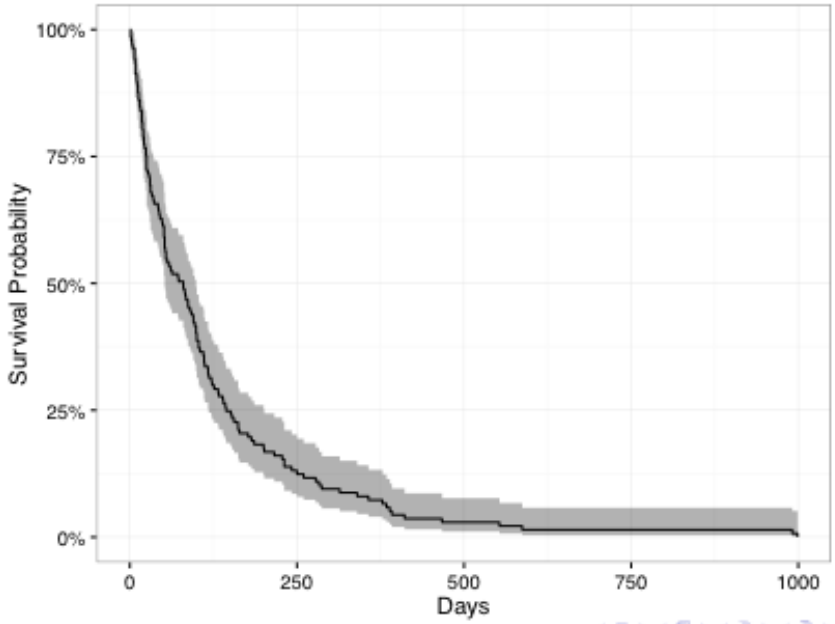
Using ggplot2 and ggfortify

```
autoplot(Shat) + labs(x = "Days", y = "Survival Probability")
```



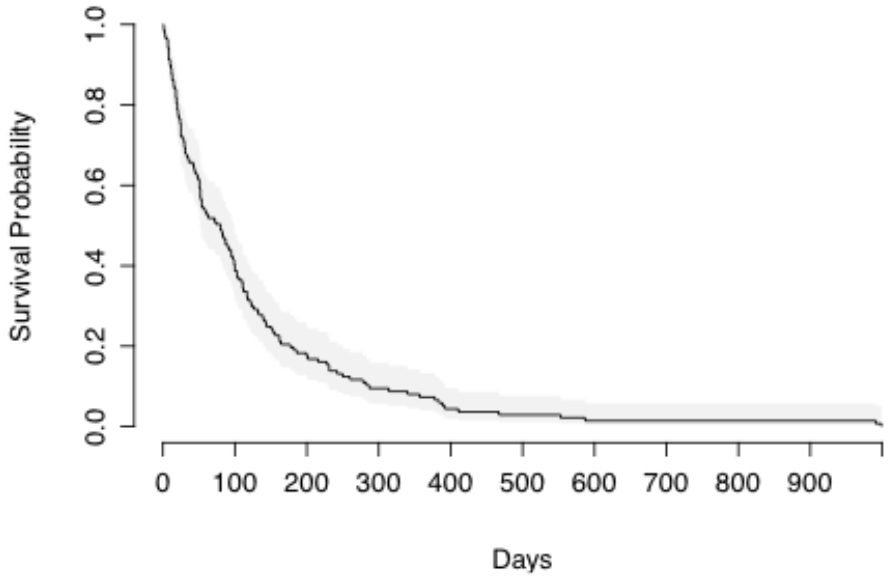
Adding black and white theme.

```
autoplot(Shat) +  theme_bw() + labs(x = "Days", y = "Survival Probability")
```



Using rms

```
Shat2 <- npsurv(Y ~ 1)
survplot(Shat2, xlab = "Days")
```



Subset of the data: squamous tumors

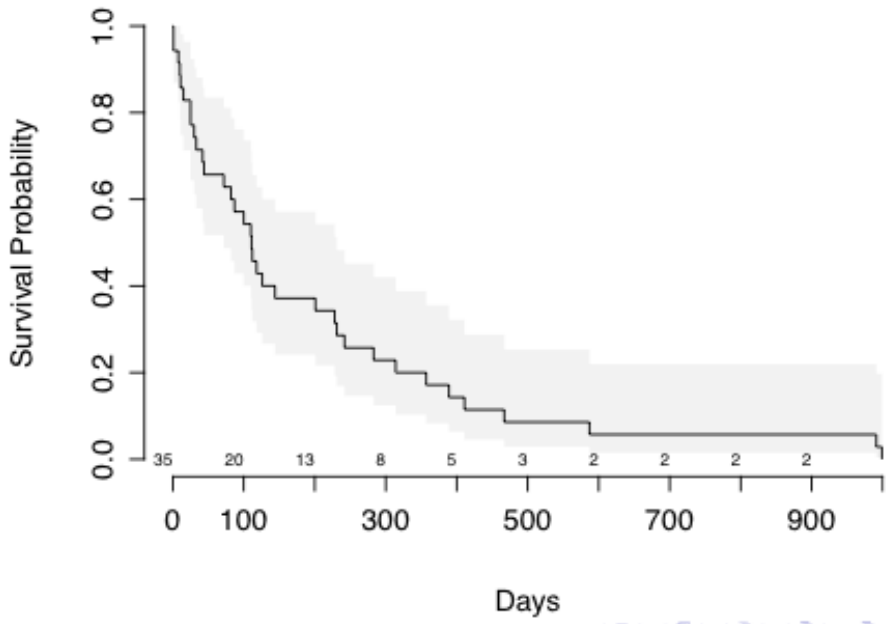
```
with(veteran, table(celltype))
```

```
## celltype
##  squamous smallcell     adeno     large
##        35        48        27        27
```

```
sqdata <- veteran[veteran$celltype == "squamous",]
Ysq <- Surv(sqdata$time)
Shatsq <- npsurv(Ysq~ 1)
```


Plot for Subset of the data: squamous tumors

```
survplot(Shatsq, xlab = "Days", n.risk = TRUE)
```



Your turn

Plot the survival curve for small cell tumors.

In R

Call up packages we will use (assumes installed)

```
library(survival)
library(ggplot2)
library(ggfortify)
library(rms)
```

Get data (in survival package)

```
data(veteran)
```

Look at data.

```
head(veteran)
```

```
##   trt celltype time status karno diagtime age prior
## 1   1 squamous   72      1    60        7  69     0
## 2   1 squamous  411      1    70        5  64    10
## 3   1 squamous  228      1    60        3  38     0
## 4   1 squamous  126      1    60        9  63    10
## 5   1 squamous  118      1    70       11  65    10
## 6   1 squamous   10      1    20        5  49     0
```

Survival Curve

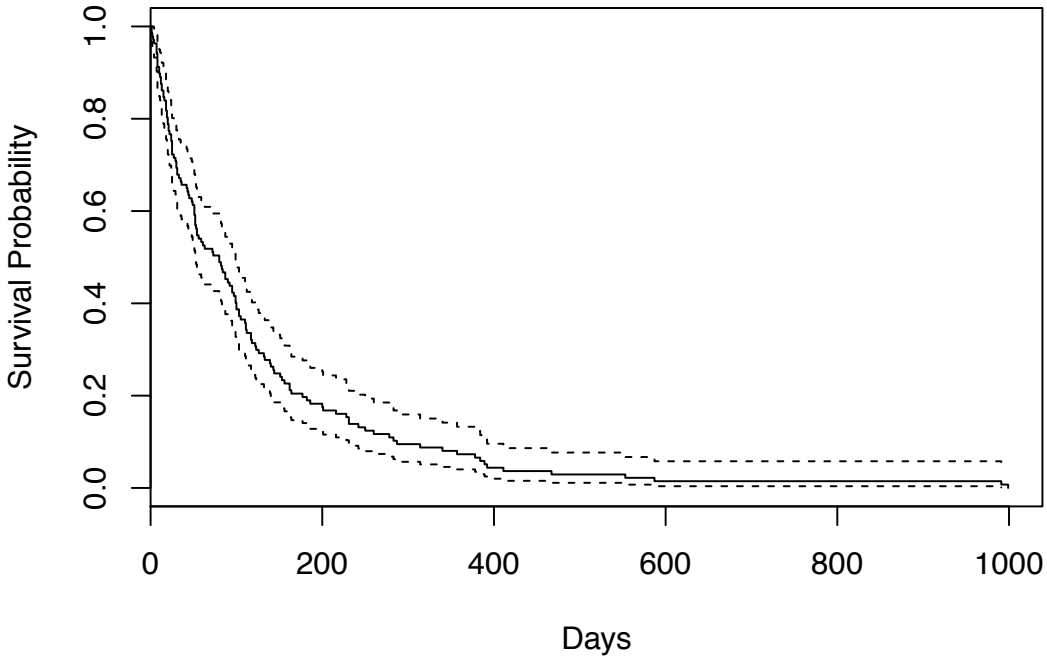
Survival time variable, make survival object and get descriptives

```
Y <- Surv(veteran$time)
Shat <- survfit(Y ~ 1)
Shat
```

```
## Call: survfit(formula = Y ~ 1)
##
##       n  events  median 0.95LCL 0.95UCL
##     137     137      80      52      99
```

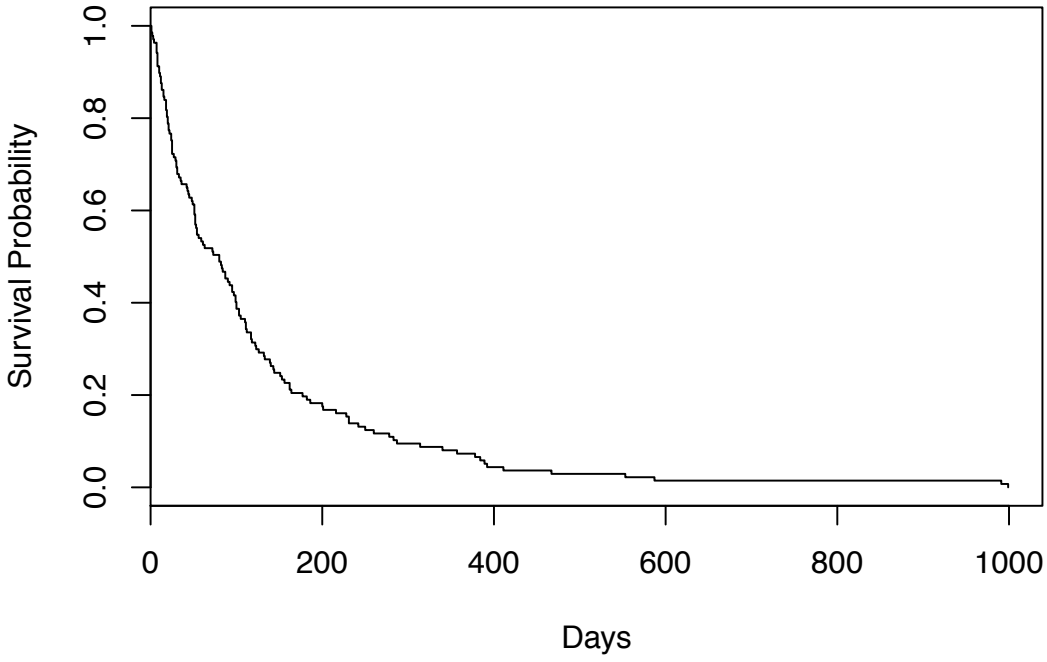
Plot Survival Curve

```
plot(Shat, xlab = "Days", ylab = "Survival Probability")
```



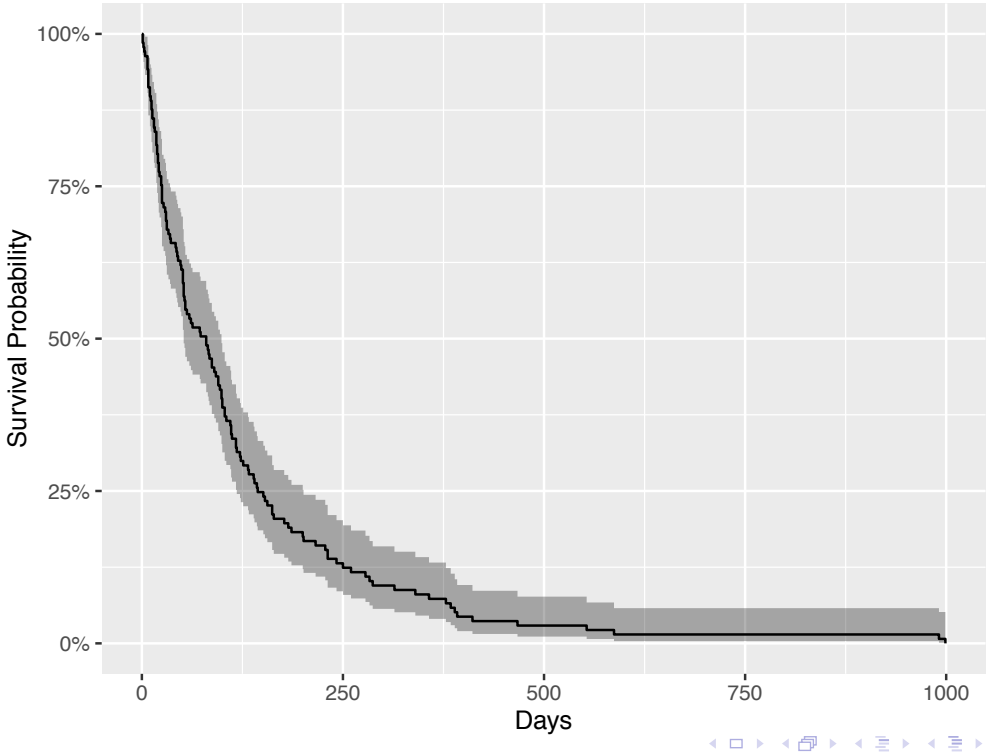
Plot Survival Curve: Other Options

```
plot(Shat, conf.int = FALSE, xlab = "Days",
     ylab = "Survival Probability")
```



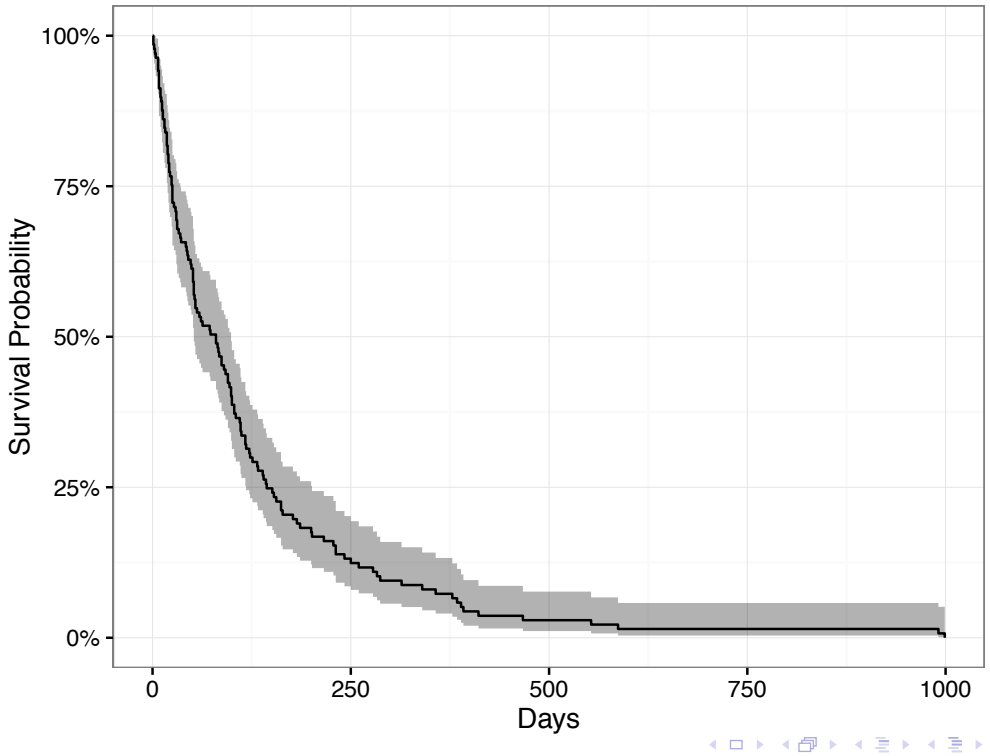
Using ggplot2 and ggfortify

```
autoplot(Shat) + labs(x = "Days", y = "Survival Probability")
```



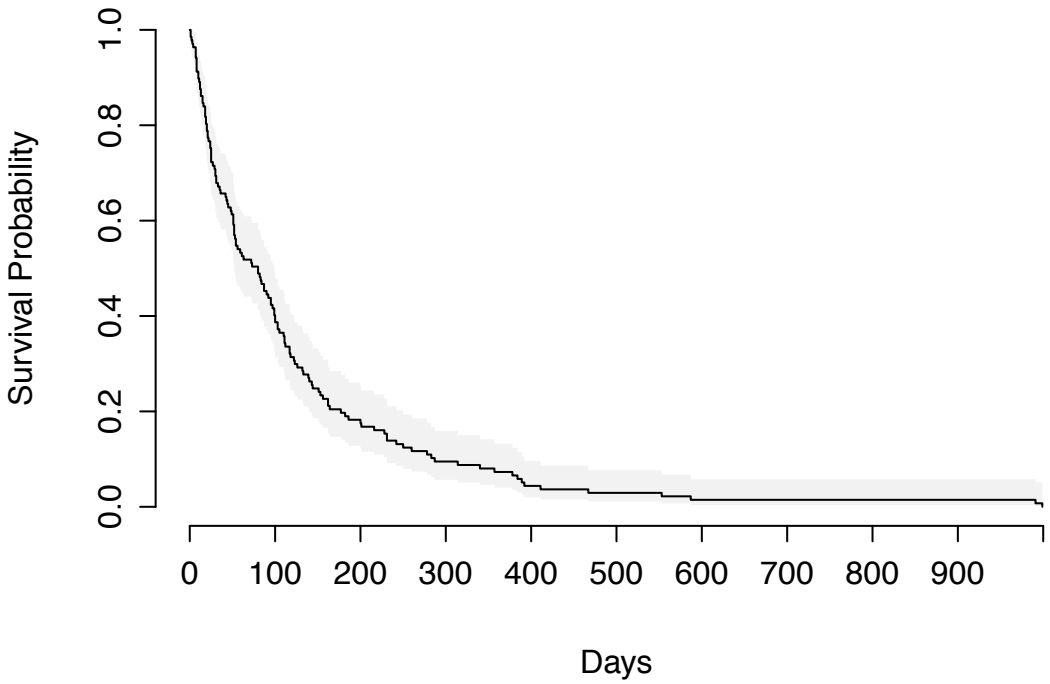
Adding black and white theme.

```
autoplot(Shat) +  theme_bw() + labs(x = "Days", y = "Survival Probability")
```



Using rms

```
Shat2 <- npsurv(Y ~ 1)
survplot(Shat2, xlab = "Days")
```



Subset of the data: squamous tumors

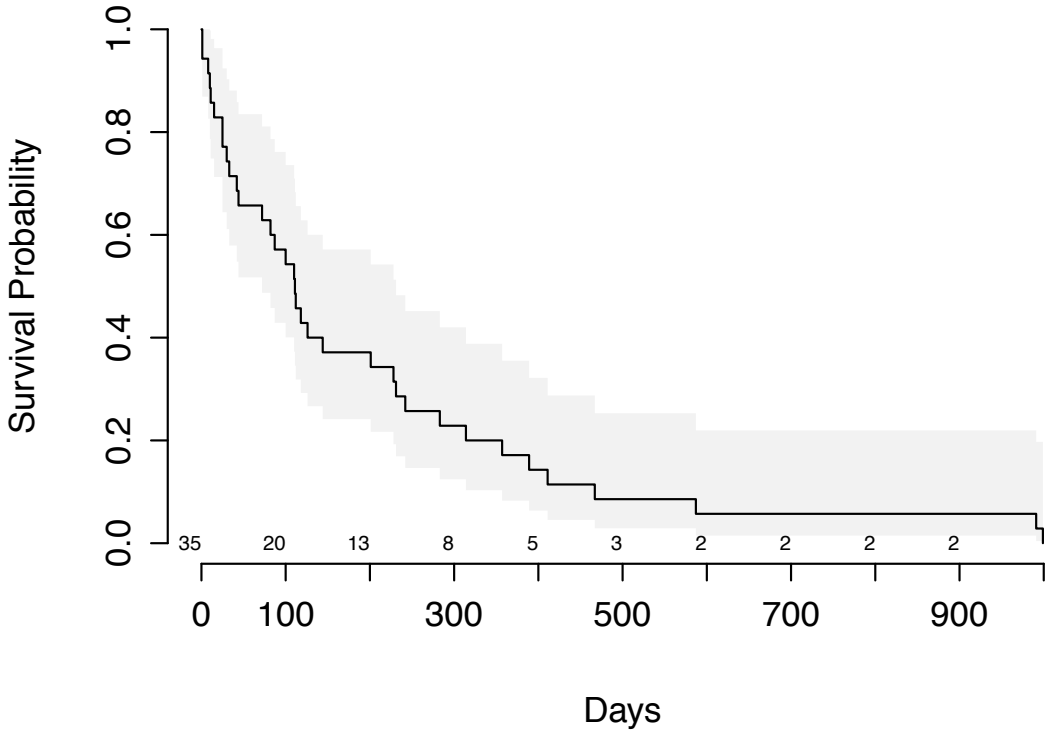
```
with(veteran, table(celltype))
```

```
## celltype
##  squamous smallcell     adeno     large
##        35        48        27        27
```

```
sqdata <- veteran[veteran$celltype == "squamous",]
Ysq <- Surv(sqdata$time)
Shatsq <- npsurv(Ysq~ 1)
```

Plot for Subset of the data: squamous tumors

```
survplot(Shatsq, xlab = "Days", n.risk = TRUE)
```



Your turn

Plot the survival curve for small cell tumors.

SESSION 2:
ONE-SAMPLE METHODS

Module 4: Introduction to Survival Analysis
Summer Institute in Statistics for Clinical Research
University of Washington
June, 2016

Barbara McKnight, Ph.D.
Professor
Department of Biostatistics
University of Washington

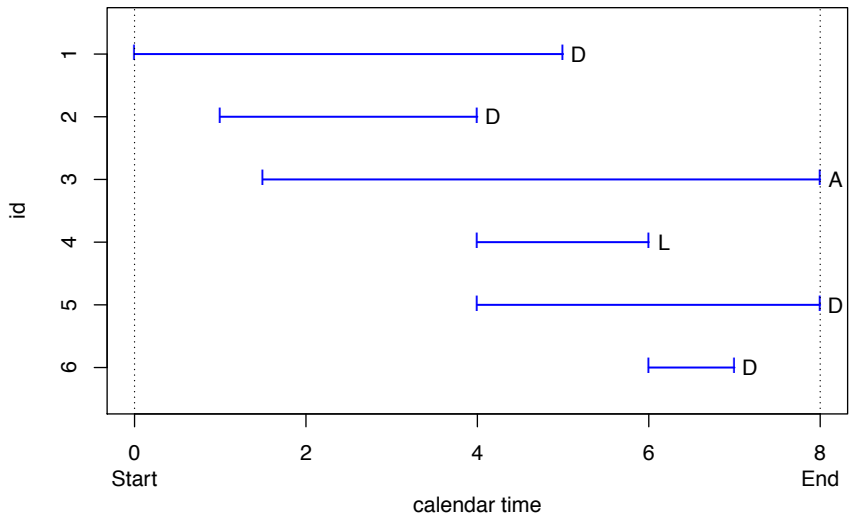
OUTLINE

- Session 2:
  - Censored data
  - Risk sets
  - Censoring assumptions
  - Kaplan-Meier Estimator
  - Median estimator
  - Standard errors and CIs

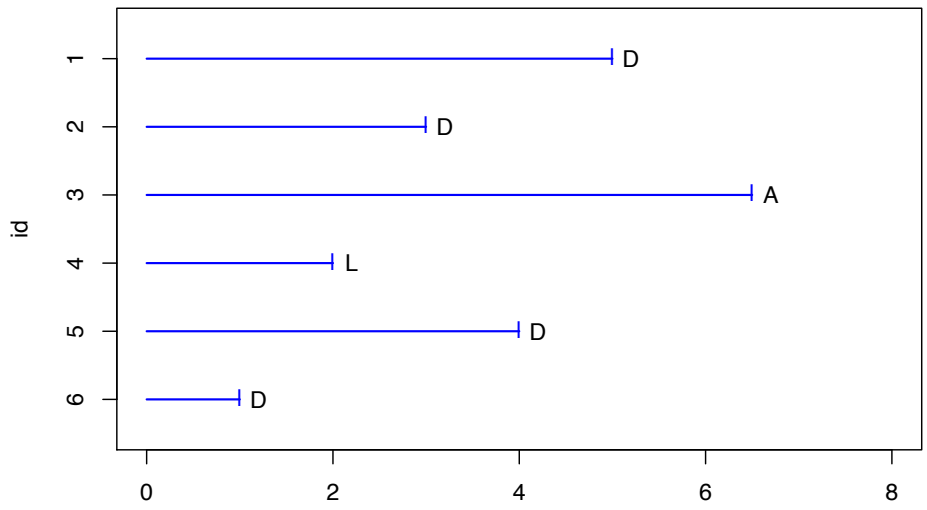
OUTLINE

- Session 2:
  - Censored data
  - Risk sets
  - Censoring assumptions
  - Kaplan-Meier Estimator
  - Median estimator
  - Standard errors and CIs

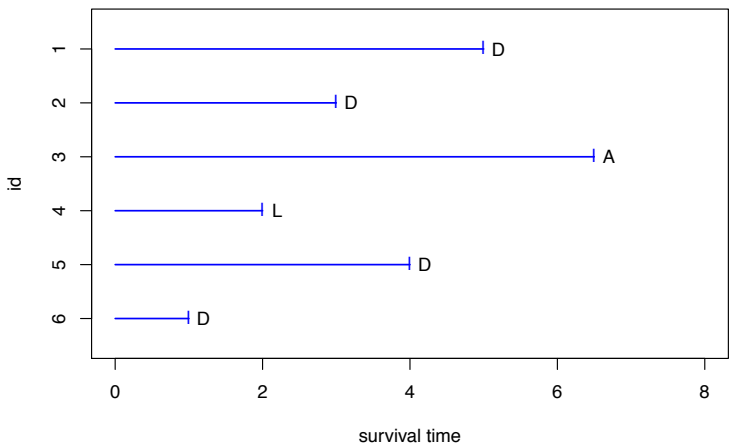
CLINICAL TRIAL



CENSORED DATA



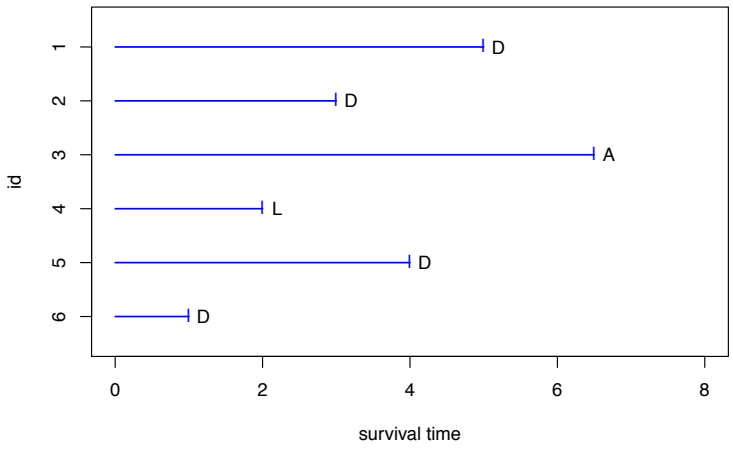
CENSORED DATA



id	Y	δ
1	5	1
2	3	1
3	6.5	0
4	2	0
5	4	1
6	1	1

"Censored" observations give some information about their survival time.

CENSORED DATA



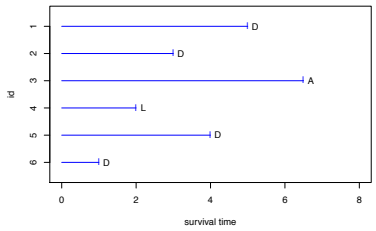
id	Y	δ
1	5	1
2	3	1
3	6.5	0
4	2	0
5	4	1
6	1	1

"Censored" observations give some information about their survival time.

ESTIMATION

- Can we use the partial information in the censored observations?
- Two off-the-top-of-the-head answers:
  - Full sample: Yes. Count them as observations that did not experience the event ever and estimate S(t) as if there were not censored observations.
  - Reduced sample: No. Omit them from the sample and estimate S(t) from the reduced data as if they were the full data.

CENSORED DATA



Problem: How to estimate:

	$\Pr[T > 3.5]$	$\Pr[T > 6]$
Full Sample:	$\frac{4}{6} = .67$	$\frac{2}{6} = .33$
Reduced Sample:	$\frac{2}{4} = .5$	$\frac{0}{4} = 0$

CENSORED DATA

Based on the data and estimates on the previous page,

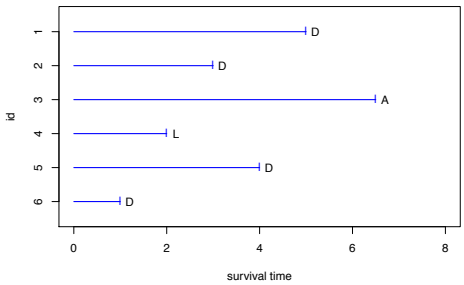
Q: Are the Full Sample estimates biased? Why or why not?

A:

Q: Are the Reduced Sample estimates biased? Why or why not?

A:

CENSORED DATA



Problem: How to estimate:

	$\Pr[T > 3.5]$	$\Pr[T > 6]$	
Full Sample:	$\frac{4}{6} = .67$	$\frac{2}{6} = .33$	← too high
Reduced Sample:	$\frac{2}{4} = .5$	$\frac{0}{4} = 0$	← too low

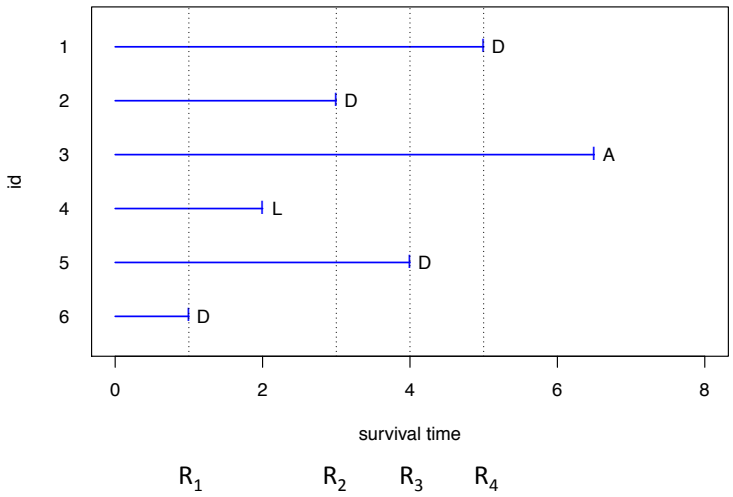
Need a good way to use the partial information in the censored observations.

IMPORTANT ASSUMPTION: Subjects who are censored at time t are representative of all subjects at risk of dying at time t.

OUTLINE

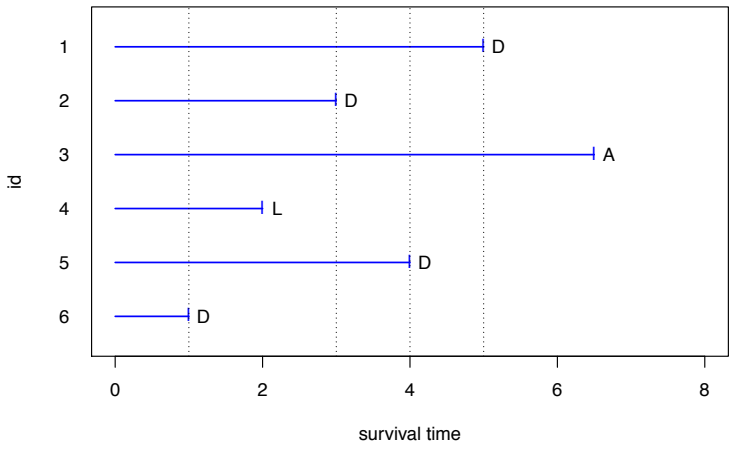
- Session 2:
  - Censored data
  - Risk sets
  - Censoring assumptions
  - Kaplan-Meier Estimator
  - Median estimator
  - Standard errors and CIs

RISK SETS



R_1 R_2 R_3 R_4

RISK SETS



R_1 R_2 R_3 R_4
{1,2,3,4,5,6} {1,2,3,5} {1,3,5} {1,3}

CENSORED DATA ASSUMPTION

- Important assumption: subjects who are censored at time t are at the same risk of dying at t as those at risk but not censored at time t.
  - When would you expect this to be true (or false) for subjects lost to follow-up?
  - When would you expect this to be true (or false) still alive at the time of the analysis?

CENSORED DATA ASSUMPTION

- Important assumption: subjects who are censored at time t are at the same risk of dying at t as those at risk but not censored at time t.
- This means the risk set at time t is an unbiased sample of the population still alive at time t.
- Can use information from the unbiased risk sets to estimate S(t) using the method of Kaplan and Meier (Product-Limit Estimator).

OUTLINE

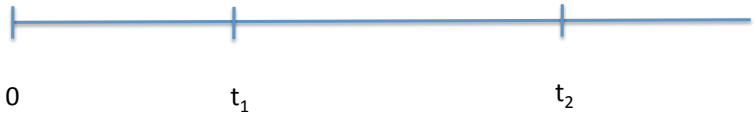
- Session 2:
  - Censored data
  - Risk sets
  - Censoring assumptions
  - Kaplan-Meier Estimator
  - Median estimator
  - Standard errors and CIs

USING RISK SETS INFO TO ESTIMATE S(t)

- Repeatedly use the fact that for $t_2 > t_1$,

$$\Pr[T > t_2] = \Pr[T > t_2 \text{ and } T > t_1] = \Pr[T > t_2 \,|\, T > t_1]\Pr[T > t_1]$$

- An observation censored between t_1 and t_2 can contribute to the estimation of $\Pr[T > t_2]$ by its unbiased contribution to estimation of $\Pr[T > t_1]$.



PRODUCT-LIMIT (KAPLAN-MEIER) ESTIMATE

Notation: Let $t_{(1)}, t_{(2)} \ldots, t_{(J)}$ be the ordered failure times in the sample in ascending order.

$t_{(1)}$ = smallest Y_i for which $\delta_i = 1$ $(t_{(1)} = 1)$

$t_{(2)}$ = 2^{nd} smallest Y_i for which $\delta_i = 1$ $(t_{(2)} = 3)$

$\vdots$

$t_{(J)}$ = largest Y_i for which $\delta_i = 1$ $(t_{(4)} = 5)$

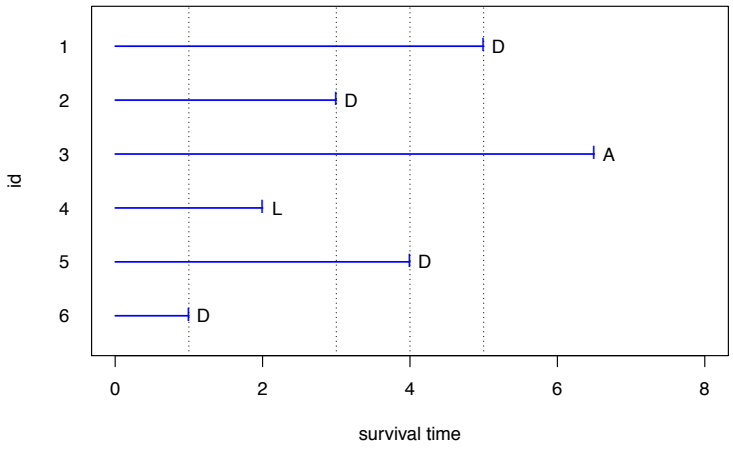
Q: Does J = the number of observed deaths in the sample?

A:

Q:When does J = n?

A:

$t_{(j)}$



MORE NOTATION

For each $t_{(j)}$:

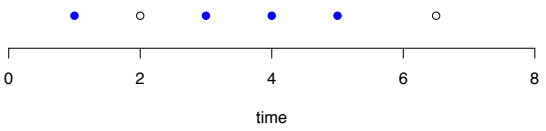
$D_{(j)}$ = number that die at time $t_{(j)}$

$S_{(j)}$ = number known to have survived beyond $t_{(j)}$
(by convention: includes those known to have been censored at $t_{(j)}$)

$N_{(j)}$ = number "at risk" of being observed to die at time $t_{(j)}$
(ie: number still alive and under observation just before $t_{(j)}$)

$S_{(j)} = N_{(j)} - D_{(j)}$

FOR EXAMPLE DATA



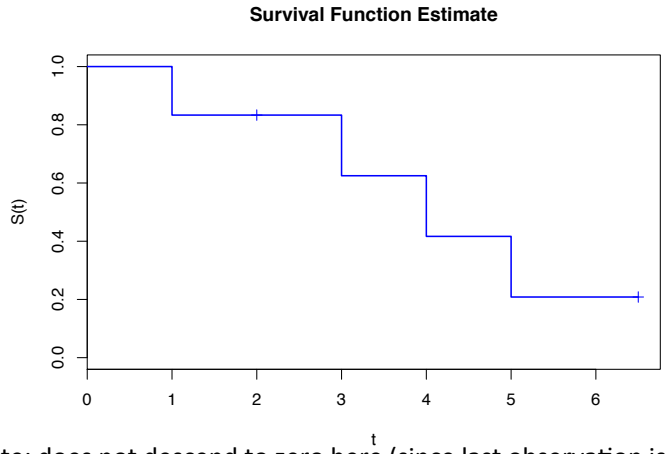
$t_{(j)}$	$N_{(j)}$	$D_{(j)}$	$S_{(j)}$
1	6	1	5
3	4	1	3
4	3	1	2
5	2	1	1

Product-limit (Kaplan-Meier) Estimator:

$$\hat{S}(t) = \Pi_{j:t_{(j)} \le t}(1 - \frac{D_{(j)}}{N_{(j)}}) = \Pi_{j:t_{(j)} \le t}(\frac{S_{(j)}}{N_{(j)}})$$

for t in	$\hat{S}(t)$	
$[0, 1)$	1	(empty product)
$[1, 3)$	$1 \times \frac{5}{6} = .833$	
$[3, 4)$	$1 \times \frac{5}{6} \times \frac{3}{4} = .625$	
$[4, 5)$	$1 \times \frac{5}{6} \times \frac{3}{4} \times \frac{2}{3} = .417$	
$[5, \infty)$	$1 \times \frac{5}{6} \times \frac{3}{4} \times \frac{2}{3} \times \frac{1}{2} = .208$	

K-M ESTIMATOR



Note: does not descend to zero here (since last observation is censored).

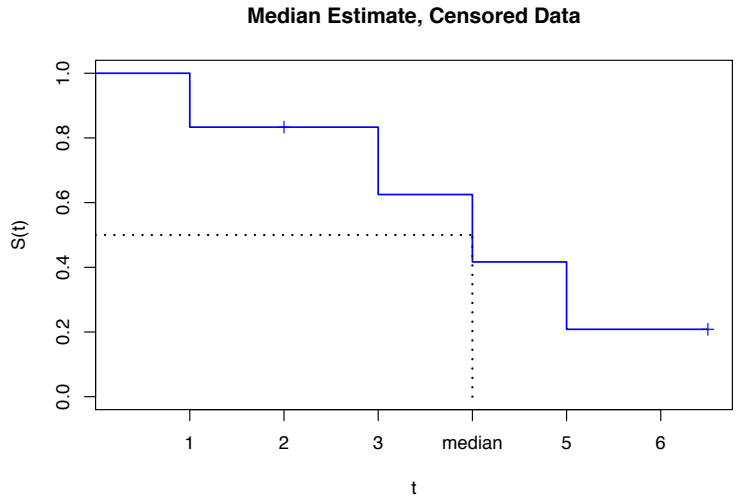
Q: Since the estimate jumps only at observed death times, how does information from the censored observations contribute to it?

A:

OUTLINE

- Session 2:
  - Censored data
  - Risk sets
  - Censoring assumptions
  - Kaplan-Meier Estimator
  - Median estimator
  - Standard errors and CIs

MEDIAN SURVIVAL CENSORED DATA



OUTLINE

- Session 2:
  - Censored data
  - Risk sets
  - Censoring assumptions
  - Kaplan-Meier Estimator
  - Median estimator
  - Standard errors and CIs

KM STANDARD ERRORS

Greenwood's Formula:

- $\widehat{Var}(\hat{S}(t)) = \hat{S}^2(t) \sum_{j:t_{(j)} \le t} \frac{D_{(j)}}{N_{(j)} S_{(j)}}$
- $\text{se}(\hat{S}(t)) = \sqrt{\widehat{Var}(\hat{S}(t))}$
- Pointwise CI: $(\hat{S}(t) - z_{\frac{\alpha}{2}} \text{se}(\hat{S}(t)), \quad \hat{S}(t) + z_{\frac{\alpha}{2}} \text{se}(\hat{S}(t)))$
  - Can include values < 0 or > 1.

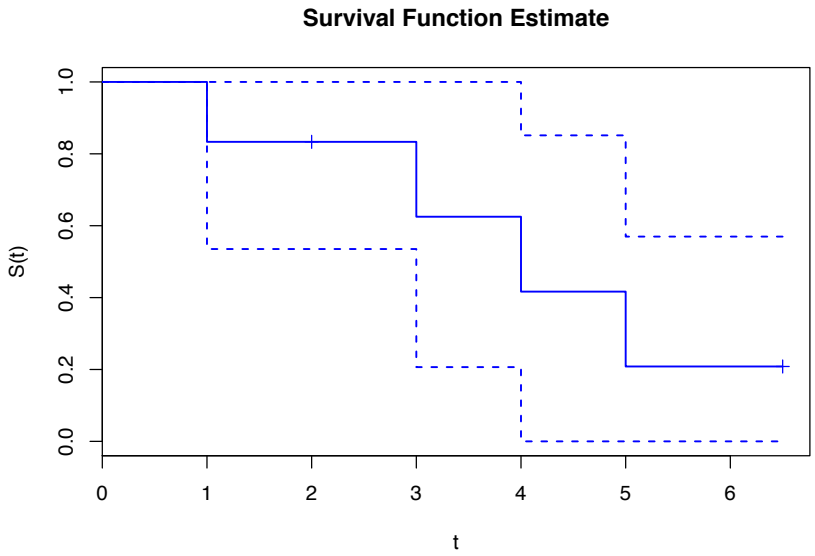
LOG –LOG KM STANDARD ERRORS

Use complementary log log transformation to keep CI within (0,1):

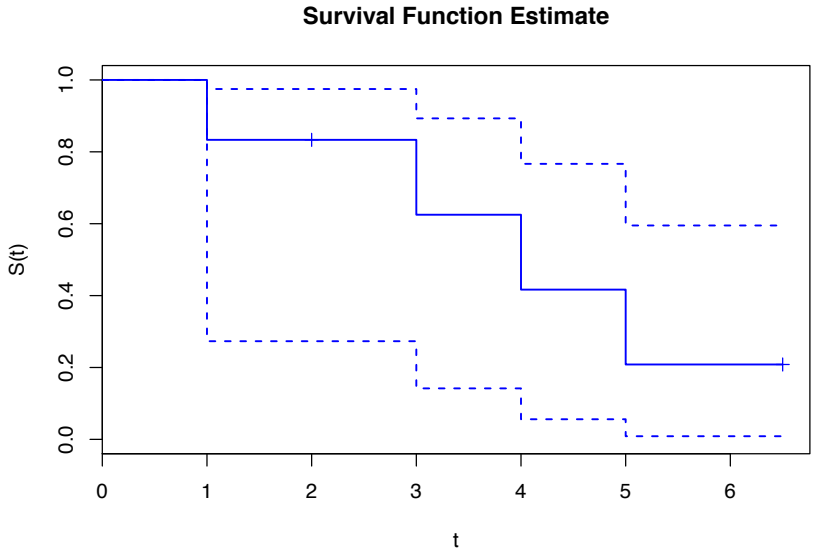
- $\widehat{Var}(\log(-\log(\hat{S}(t)))) = \frac{\sum_{j:t_{(j)} \le t} \frac{D_{(j)}}{N_{(j)} S_{(j)}}}{[\log(\hat{S}(t))]^2}$
- $\text{se} = \sqrt{\widehat{Var}(\log(-\log(\hat{S}(t))))}$
- CI for $\log(-\log(S(t)))$:

  $(\log(-\log(\hat{S}(t))) - z_{\frac{\alpha}{2}} \text{se}, \quad \log(-\log(\hat{S}(t))) + z_{\frac{\alpha}{2}} \text{se})$
- CI for $\hat{S}(t)$: $([\hat{S}(t)]^{e^{z\alpha/2}\text{se}}, [\hat{S}(t)]^{e^{-z\alpha/2}\text{se}})$
  - CI remains within (0,1).

GREENWOOD'S FORMULA



COMPLEMENTARY LOG-LOG



MEDIAN CONFIDENCE INTERVAL

Confidence interval for the median is obtained by inverting the sign test of H_0 : median = M (Brookmeyer and Crowley, 1982).

- With complete data $T_1, T_2, \ldots, T_n$, the sign test of H_0 : median $= M$ is performed by seeing if the observed proportion, $\hat{P}[Y > M]$ is too big (Binomial Distribution or Normal Approximation).
- With censored data $(Y_1, \delta_1), (Y_2, \delta_2), \ldots, (Y_n, \delta_n)$ giving incomplete data about $T_1, T_2, \ldots, T_n$, we cannot always tell whether $T_i > M$:

When $Y_i \leq M, \delta_i = 1$	observed death before M	we know $T_i \leq M$
When $Y_i > M$	death or censored after M	we know $T_i > M$
When $Y_i \leq M, \delta_i = 0$	censored before M	we don't know if $T_i \leq M$ or $T_i > M$

MEDIAN CONFIDENCE INTERVAL

Solution: Following Efron (self-consistency of KM), we estimate $\Pr[T > M]$ when $Y_i \leq M, \delta_i = 0$ using $\frac{\hat{S}(M)}{\hat{S}(Y_i)}$.

- For complete data, we let $U_i = \begin{cases} 1 & T_i > M \\ 0 & T_i \leq M \end{cases}$

  and our test is based on $\sum_{i=1}^{n} U_i$.

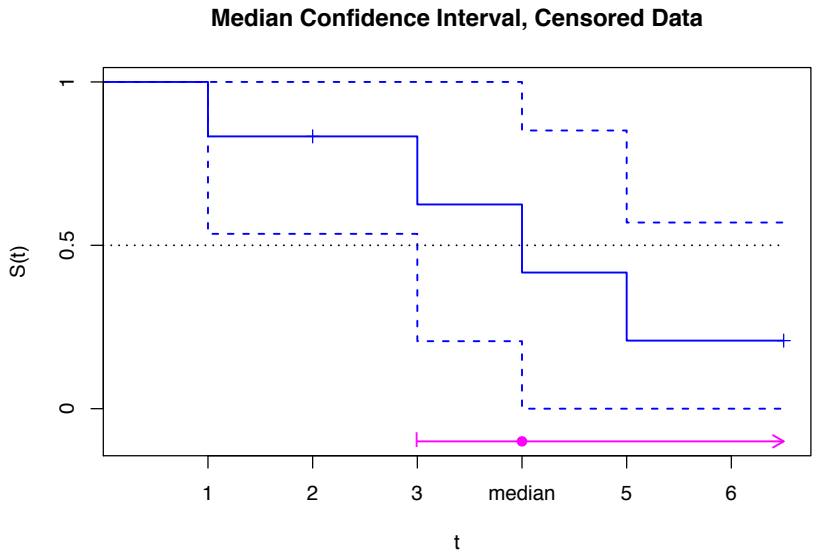
- For censored data, we let $U_i = \begin{cases} 1 & Y_i > M \\ \frac{\hat{S}(M)}{\hat{S}(Y_i)} & Y_i \leq M; \delta_i = 0 \\ 0 & Y_i \leq M; \delta_i = 1 \end{cases}$

  and our test is based on $\sum_{i=1}^{n} U_i$.

MEDIAN CONFIDENCE INTERVAL

- It turns out, this is the same as basing our test of H_0 : median $= M$ on a test of $H_0 : S(M) = \frac{1}{2}$.
- So a 95% CI for the median contains all potential M for which the test of $H_0 : S(M) = \frac{1}{2}$ cannot reject at $\alpha = .05$ (2 sided).
- Since $\hat{S}(M)$ only changes value at observed event times, the test need only be checked at $M = t_{(1)}, t_{(2)}, \ldots, t_{(J)}$.
- Originally proposed for Greenwood's formula CIs for $\hat{S}(M)$, but any good CIs are OK.
- Implemented in many software packages.

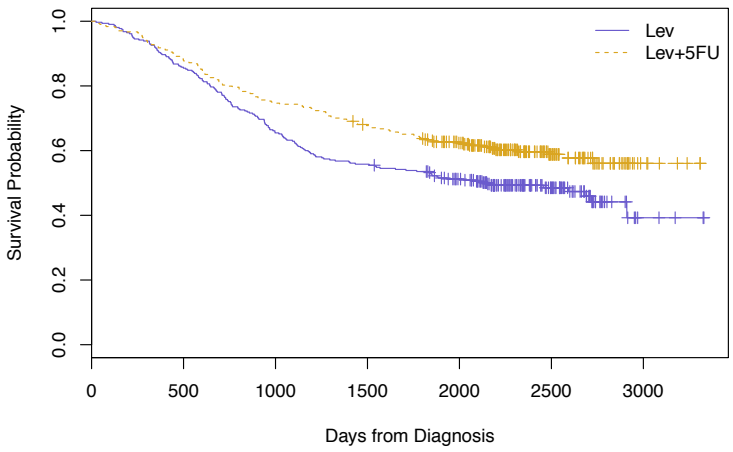
MEDIAN CONFIDENCE INTERVAL



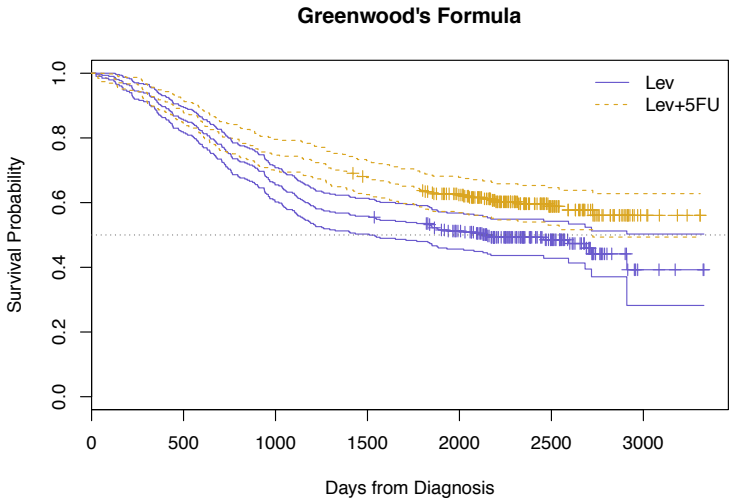
COLON CANCER EXAMPLE

- Clinical trial at Mayo Clinic (Moertel et al. (1990) NEJM)
- Stage B_2 and C colon cancer patients; adjuvant therapy
- Three arms
  - Observation only
  - Levamisole
  - 5-FU + Levamisole
- Stage C patients only
- Two treatment arms only

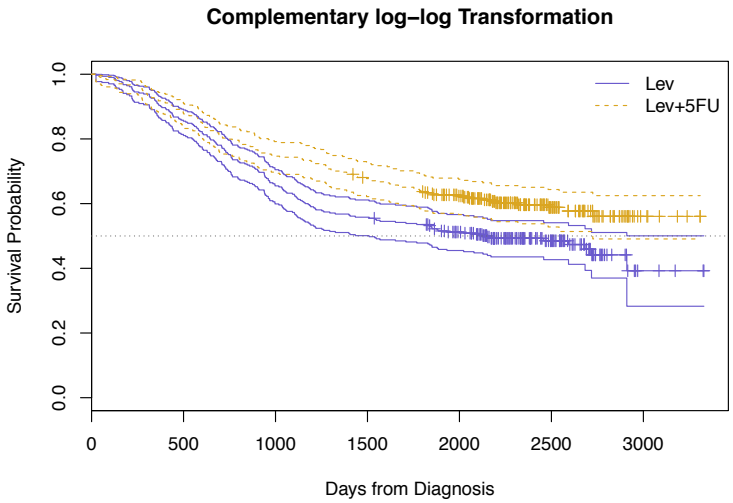
COLON CANCER EXAMPLE



COLON CANCER EXAMPLE



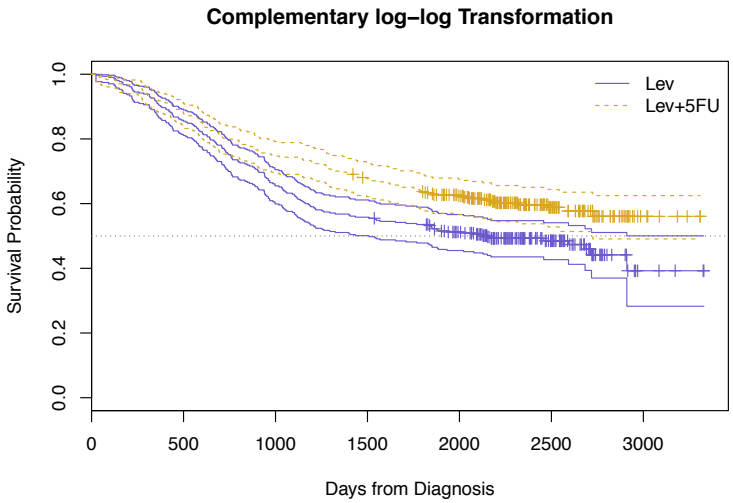
COLON CANCER EXAMPLE



PRESENTATION

	N	Events	Median (days)	95% CI
Levamisole Only	310	161	2152	(1509, ∞)
5FU + Levamisole	304	123	--	(2725, ∞)

COLON CANCER EXAMPLE



ESTIMATION

- Estimate S(t) using KM curve (nonparametric).
  - Pointwise standard errors and Cis
  - Almost always presented
  - Not appropriate when the event of interest happens only to some (more on this this tomorrow)
- Median: based on KM curve: often presented (too often?)

TO WATCH OUT FOR

- Mean survival time hard to estimate without parametric assumptions
  - Censoring means incomplete information about largest times
  - Mean over restricted time interval may be useful in some settings (some on this tomorrow)
- Median estimate more complicated than median of times
- Even with CIs, evaluating differences between curves visually is subjective
- Interpretation of survival function estimates depends on validity of censoring assumptions

In R

Load packages.

```
library(survival)
library(rms)
```

Get data.

```
data(colon) # in survival package
head(colon)
```

```
##   id study      rx sex age obstruct perfor adhere nodes status differ
## 1  1     1 Lev+5FU   1  43        0      0      0     5      1      2
## 2  1     1 Lev+5FU   1  43        0      0      0     5      1      2
## 3  2     1 Lev+5FU   1  63        0      0      0     1      0      2
## 4  2     1 Lev+5FU   1  63        0      0      0     1      0      2
## 5  3     1     Obs   0  71        0      0      1     7      1      2
## 6  3     1     Obs   0  71        0      0      1     7      1      2
##   extent surg node4 time etype
## 1      3    0     1 1521     2
## 2      3    0     1  968     1
## 3      3    0     0 3087     2
## 4      3    0     0 3087     1
## 5      2    0     1  963     2
## 6      2    0     1  542     1
```

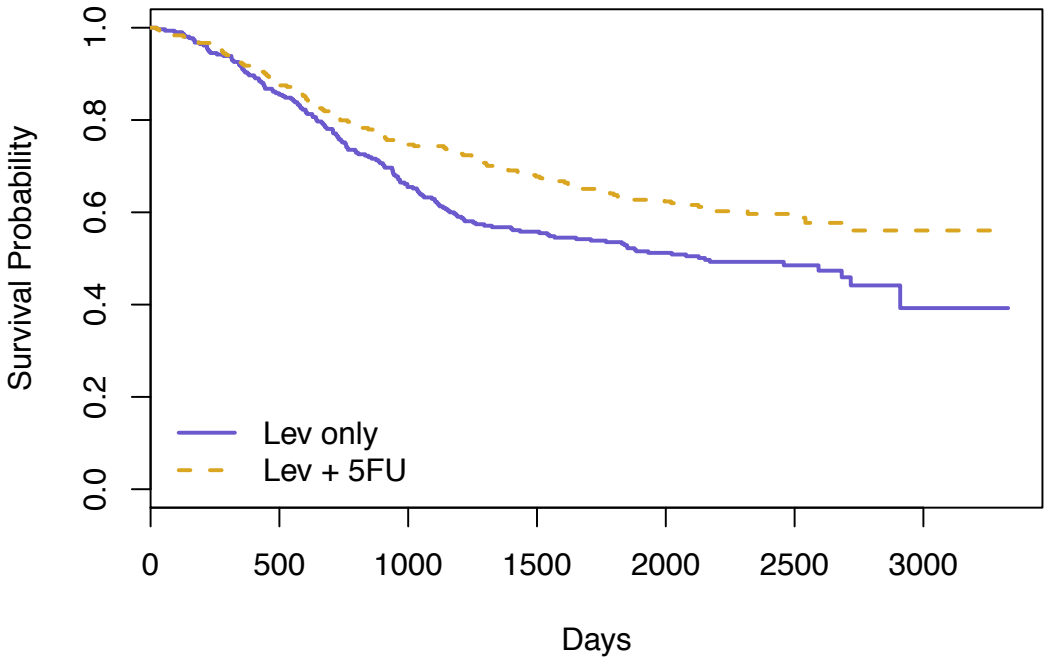
Process data and compute survival curves.

```
df <- colon[colon$etype == 2,] # Use death times.
df <- df[df$rx != "Obs",] # Omit observation only arm.
Y <- with(df, Surv(time, status))
Shats <-survfit(Y ~ rx, data = df)
```

Plot survival curves.

```
colors <- c("slateblue", "goldenrod")
plot(Shats, lty = c(1,2),
     col = colors, lwd = 2,
     xlab = "Days", ylab = "Survival Probability")
legend("bottomleft", lty = c(1,2),
       col = colors, lwd = 2,
       legend = c("Lev only", "Lev + 5FU"), bty = "n")
```

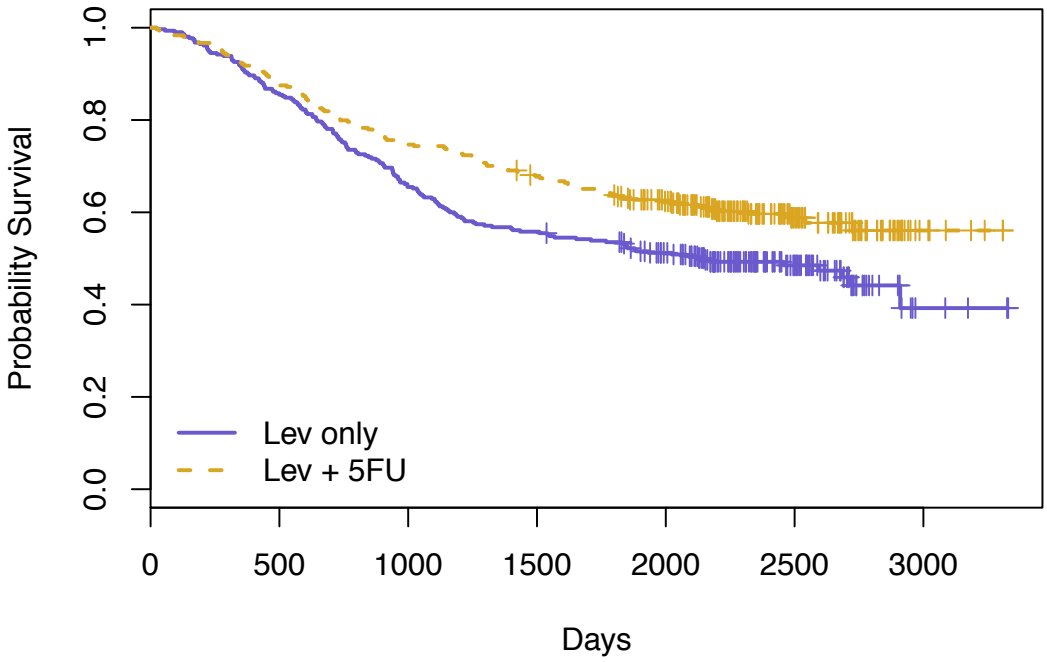
Plot survival curves.



With censoring tick marks

```
plot(Shats, lty = c(1,2),
     col = colors, lwd = 2,
     xlab = "Days", ylab = "Probability Survival",
     mark.time = TRUE)
legend("bottomleft", lty = c(1,2),
       col = colors, lwd = 2,
       legend = c("Lev only", "Lev + 5FU"), bty = "n")
```

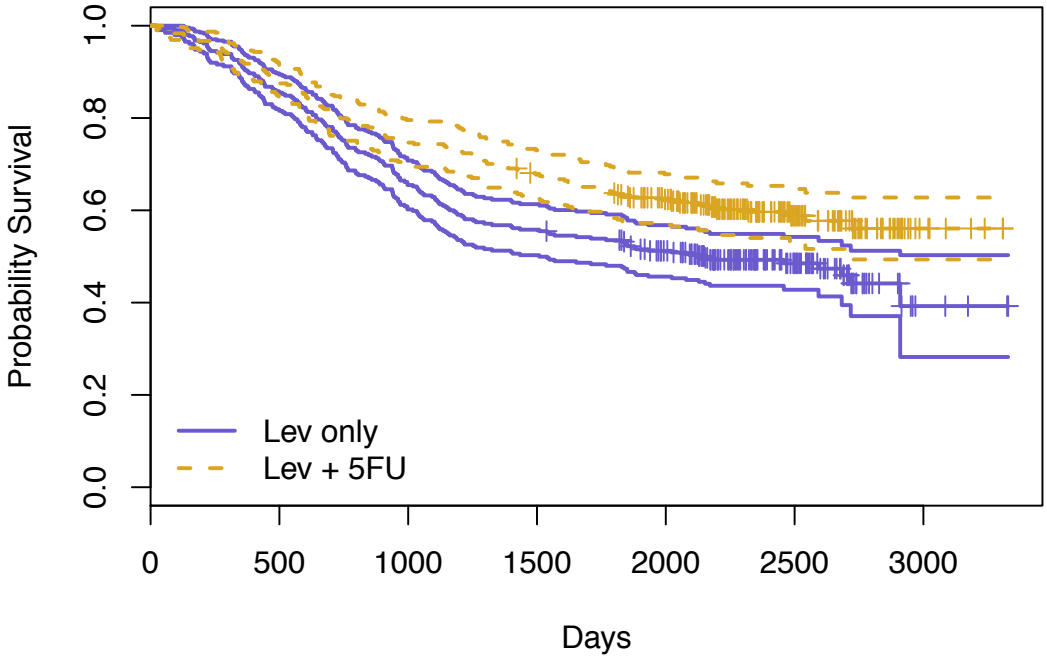
With censoring tick marks



With CIs: Greenwood's formula

```
ShatsG <- survfit(Y ~ rx, data = df, conf.type = "plain")
plot(ShatsG, lty = c(1,2),
     col = colors, lwd = 2,
     xlab = "Days", ylab = "Probability Survival",
     mark.time = TRUE, conf.int = TRUE)
legend("bottomleft", lty = c(1,2),
       col = colors, lwd = 2,
       legend = c("Lev only", "Lev + 5FU"), bty = "n")
```

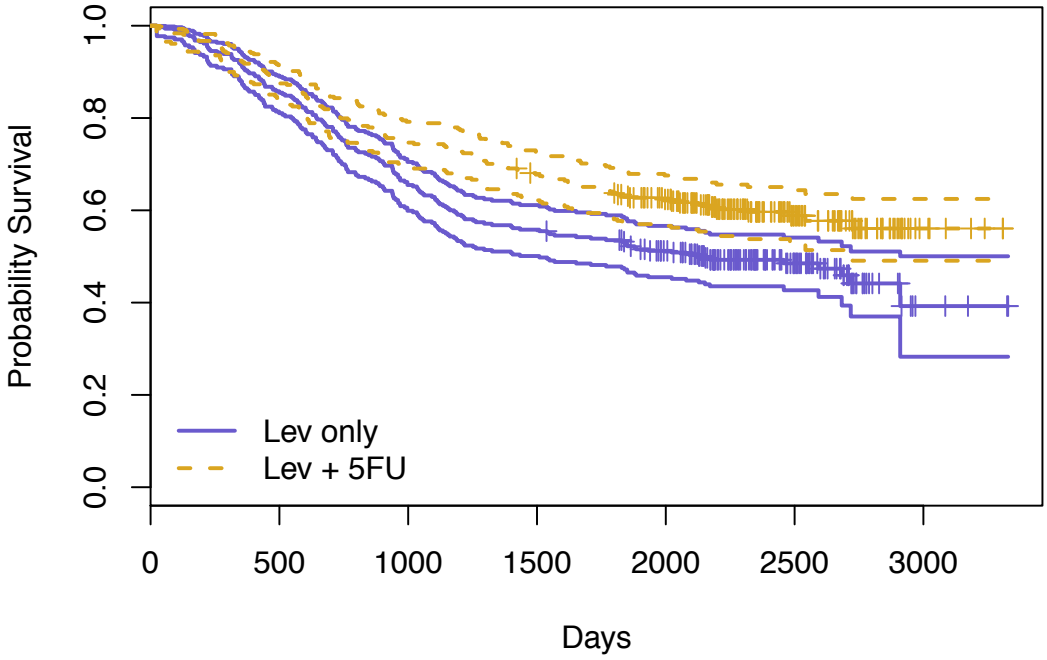
With CIs: Greenwood's formula



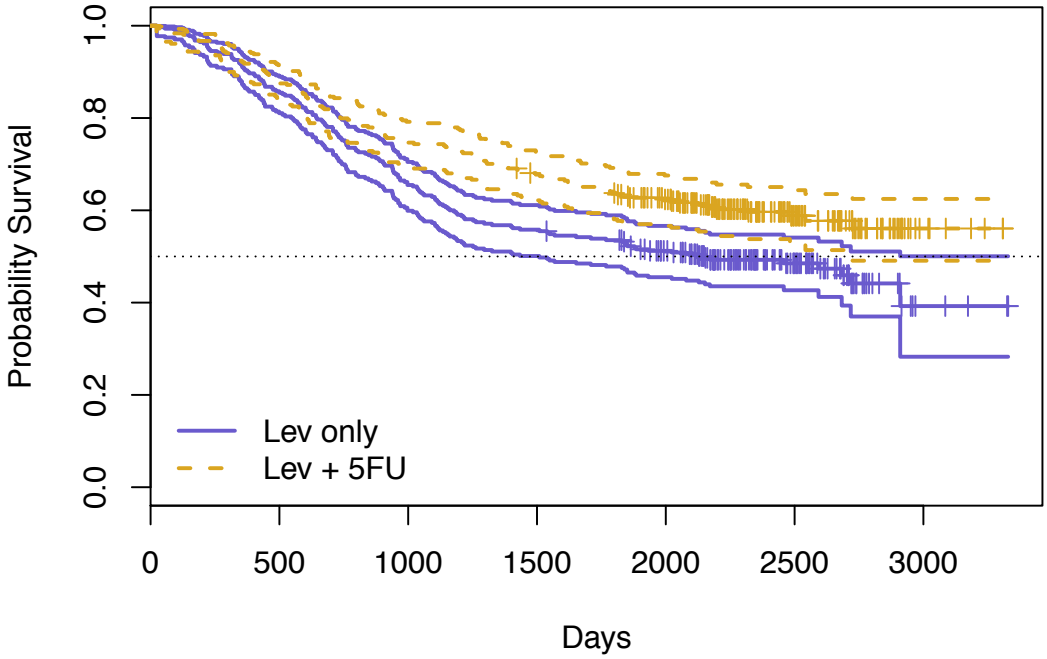
With CIs: Complementary log-log formula

```
ShatsL <-survfit(Y ~ rx, data = df, conf.type = "log-log")
plot(ShatsL, lty = c(1,2),
     col = colors, lwd = 2,
     xlab = "Days", ylab = "Probability Survival",
     mark.time = TRUE, conf.int = TRUE)
legend("bottomleft", lty = c(1,2),
       col = colors, lwd = 2,
       legend = c("Lev only", "Lev + 5FU"), bty = "n")
```

With CIs: Complementary log-log formula



Median CIs: Complementary log-log formula



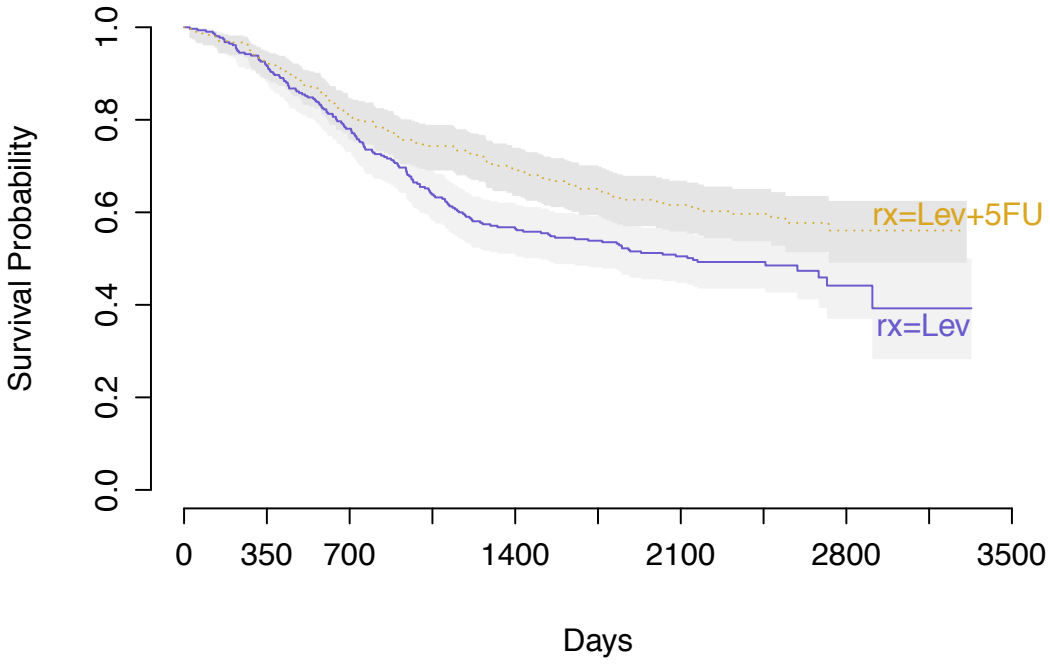
Median CI summary

```
ShatsL
```

```
## Call: survfit(formula = Y ~ rx, data = df, conf.type = "log-log")
##
##               n events median 0.95LCL 0.95UCL
## rx=Lev      310    161   2152    1509      NA
## rx=Lev+5FU 304    123     NA    2725      NA
```

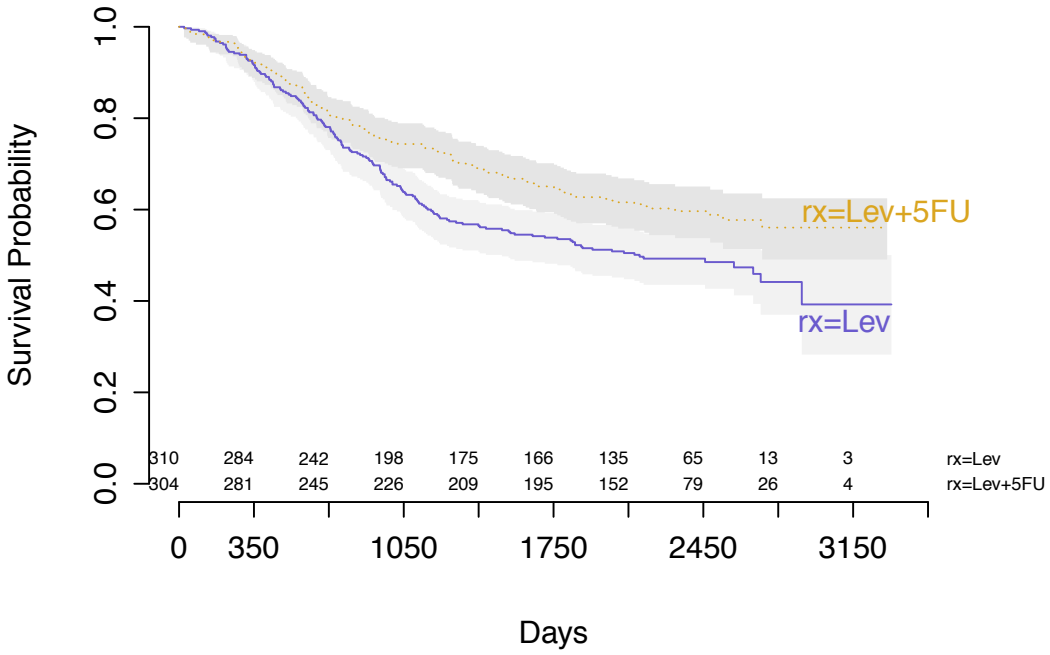
With rms

```
Shat2 <- npsurv(Y ~ rx, data = df, conf.type = "log-log")
survplot(Shat2, xlab = "Days", col = colors)
```



With numbers at risk

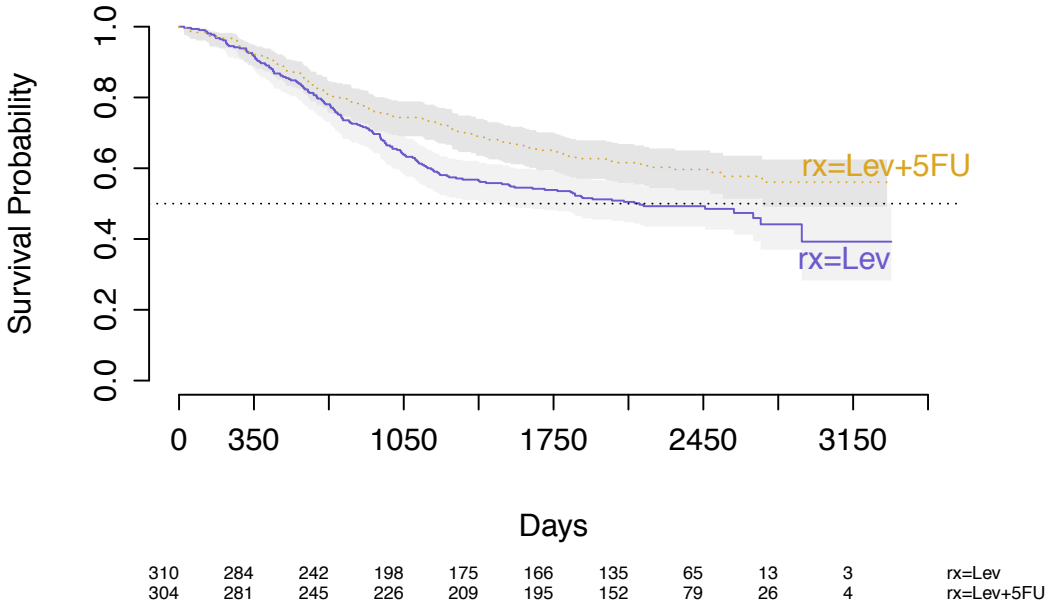
```
survplot(Shat2, xlab = "Days", col = colors, n.risk = TRUE)
```



With numbers at risk below

```
par(mar = c(8, 4,4,2) + .1)
survplot(Shat2, xlab = "Days", col = colors,
         n.risk = TRUE, y.n.risk = -.6)
abline(h = .5, lty = 3)
```

With numbers at risk below



Your turn

1. Using the colon cancer data, plot treatment group survival curves comparing Observation only arm to Levamisole only and Levamisole + 5 FU arms.
2. Compute median survival times and CIs for each group.

SESSION 3:
TWO AND K-SAMPLE METHODS

Module 4: Introduction to Survival Analysis
Summer Institute in Statistics for Clinical Research
University of Washington
June, 2016

Barbara McKnight, Ph.D.
Professor
Department of Biostatistics
University of Washington

OVERVIEW

- Session 1
  - Introductory examples
  - The survival function
  - Survival Distributions
  - Mean and Median survival time

- Session 2
  - Censored data
  - Risk sets
  - Censoring Assumptions
  - Kaplan-Meier Estimator and CI
  - Median and CI

- Session 3
  - Two-group comparisons: logrank test
  - Trend and heterogeneity tests for more than two groups

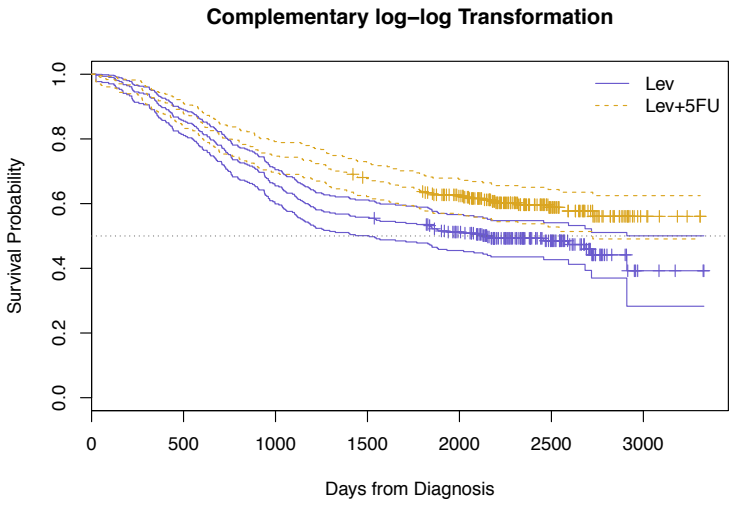
- Session 4
  - Introduction to Cox regression

TESTING

- Group comparisons
  - Two groups
  - k- group heterogeneity
  - k- group trend

- Assume, H_0 : no differences between groups

COLON CANCER EXAMPLE



THE P-VALUE QUESTION

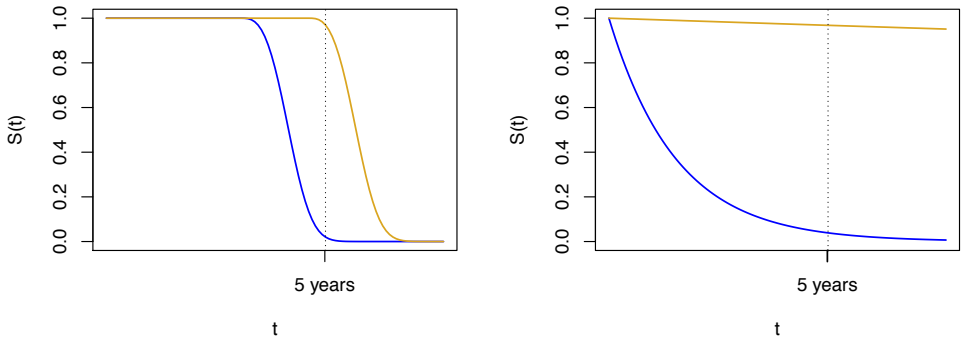
- Statistical significance?

COMPARING SURVIVAL DISTRIBUTIONS

- Two-sample data: comparing $S_1(t)$ and $S_2(t)$
  - (Y_{1i}, δ_{1i}), $i=1,\ldots,n_1$, $T \sim S_1(t)$
  - (Y_{2i}, δ_{2i}), $i=1,\ldots,n_2$, $T \sim S_2(t)$

- Could look at $S_2(t) - S_1(t)$ at a single time *t*, but this might be misleading unless all you care about is survival at that time.

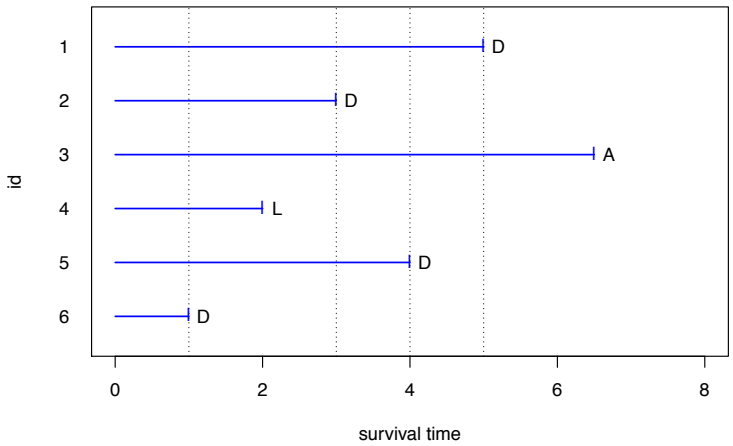
COMPARISON AT 5 YEARS



COMPARING SURVIVAL DISTRIBUTIONS

- There are many ways to measure $S_2(t) - S_1(t)$, the distance between two functions of time
- Here: focus on most commonly used test: the logrank test, which compares consistent ratios of hazard functions
- Module 12 will consider other tests

RISK SETS



R_1 {1,2,3,4,5,6} R_2 {1,2,3,5} R_3 {1,3,5} R_4 {1,3}

LOGRANK TEST

- The test is based on a 2x2 table of group by current status at each observed failure time (ie for each risk set)
- $T_{(j)}$, $j=1,\ldots m$, as shown in the Table below.

Event/Group	1	2	Total
Die	$d_{1(j)}$	$d_{2(j)}$	$D_{(j)}$
Survive	$n_{1(j)}-d_{1(j)}= s_{1(j)}$	$n_{2(j)}-d_{2(j)} = s_{2(j)}$	$N_{(j)}-D_{(j)} = S_{(j)}$
At Risk	$n_{1(j)}$	$n_{2(j)}$	$N_{(j)}$

TWO-GROUP COMPARISONS

- The contribution to the test statistic at each event time is obtained by calculating the expected number of deaths in one group, assuming that the risk of death at that time is the same in each of the two groups.
- This yields the usual *"row total times column total divided by grand total"* estimator. For example, for group 1, the expected number is $$\hat{E}_{1(j)} = \frac{n_{1(j)} D_{(j)}}{N_{(j)}}$$
- Most software packages base their estimator of the variance on the hypergeometric distribution, defined as follows:

$$\hat{V}_{(j)} = \frac{n_{1(j)} n_{2(j)} D_{(j)} \left(N_{(j)} - D_{(j)} \right)}{N_{(j)}^2 \left(N_{(j)} - 1 \right)}$$

LOGRANK TWO-GROUP COMPARISONS

- Each test may be expressed in the form of a ratio of sums over the observed survival times as follows

$$Q = \frac{[\sum_{j=1}^{J} (d_{1(j)} - \hat{E}_{1(j)})]^2}{\hat{V}_{(j)}} = \frac{[\sum_{j=1}^{J} \left(\frac{n_{1(j)} n_{2(j)}}{n_{1(j)} + n_{2(j)}} \right) \left(\frac{d_{1(j)}}{n_{1(j)}} - \frac{d_{2(j)}}{n_{2(j)}} \right)]^2}{\hat{V}_{(j)}}$$

- Where t_j, *j = 1,…, J,* are the unique ordered event times
- Under the null hypothesis of no difference in survival distribution, the *p*-value for *Q* may be obtained using the chi-square distribution with one degree-of-freedom, when the expected number of events is large.

$$p = \Pr\left(\chi^2_1 \geq Q \right)$$

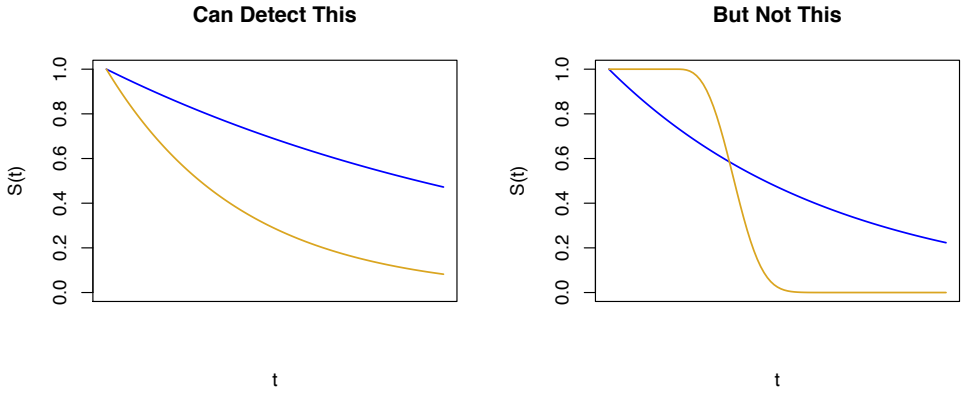
COLON CANCER EXAMPLE

- Comparing Lev and Lev+5FU:

Group	N	Obs	Exp
Lev	310	161	136.9
Lev+5FU	304	123	147.1
Total	614	284	284.0

- Log-rank test: $\chi^2_1 = 8.2$, p-value = 0.0042

LOGRANK TEST



Other tests (generalized Wilcoxon and others) can give more weight to early or late differences.

LOGRANK TEST

- Detects consistent differences between survival curves over time.
- Best power when:
  - H_0: $S_1(t) = S_2(t)$ for all t vs H_A: $S_1(t) = [S_2(t)]^c$, or
  - H_0: $\lambda_1(t) = \lambda_2(t)$ for all t vs H_A: $\lambda_1(t) = c\,\lambda_2(t)$
- Good power whenever survival curve difference is in consistent direction

STRATIFIED LOGRANK TEST

- In a large-enough clinical trial, confounding bias due to imbalance between treatment arms is unlikely.
- However, better power can be obtained by adjusting for strongly prognostic variables.
- One way to adjust: stratified logrank test
- Can also use Cox regression (Modules 8 and 12)

STRATIFIED LOGRANK TEST

- Assume R strata ($r = 1,\ldots,R$)
- Recall (non-stratified) log-rank test statistic

$$Q = \frac{[\sum_{j=1}^{J}(d_{1(j)} - \hat{E}_{1(j)})]^2}{\hat{V}_{(j)}}$$

- Stratified log-rank test

$$Q = \frac{\left[\sum_{j_1=1}^{J_1}\left(d_{1,1(j)} - \hat{E}_{1,1(j)}\right) + \ldots + \sum_{j_r=1}^{J_r}\left(d_{1r(j)} - \hat{E}_{1r(j)}\right) + \ldots + \sum_{j_R=1}^{J_R}\left(d_{1R(j)} - \hat{E}_{1R(j)}\right)\right]^2}{\sum_{j_1=1}^{J_1}\hat{V}_{1(j)} + \ldots + \sum_{j_r=1}^{J_r}\hat{V}_{r(j)} + \ldots + \sum_{j_R=1}^{J_R}\hat{V}_{R(j)}}$$

STRATIFIED LOG-RANK TEST

- H_0: $\lambda_{1r}(t) = \lambda_{2r}(t)$ for all t and r = 1,…,R
- H_A: $\lambda_{1r}(t) = c\lambda_{2r}(t)$, $c \neq 1$, for all t and $r = 1,\ldots,R$

- Under H_0 test statistic ~ χ^2_1 when the number of events is large
- The $d_{1r(j)}, \hat{E}_{1r(j)}$ and $\hat{V}_{r(j)}$ are based solely on subjects from the r^{th} stratum
- Will be powerful when direction of group difference is consistent across strata and over time.

EXAMPLE - WHAS

- Example: The Worcester Heart Attack Study (WHAS)
- Goal: study factors and time trends
  associated with long term survival
  following acute myocardial infarction (MI)
  among residents of the Worcester, Massachusetts
  Standard Metropolitan Statistical Area (SMSA)
- Study began in 1975
- Data collection approximately every other year
- Most recent cohort: subjects who experienced an MI in 2001
- The main study: over 11,000 subjects
- Here:a small sample from the main study with n = 100

EXAMPLE - WHAS

- t_0: time of hospital admission following an acute myocardial infarction (MI)
- Event: Death from any cause following hospitalization for an MI
- Time: Time from hospital admission to
  - Death
  - End of study
  - Last contact
- Interest in effect of gender adjusted for age

TESTING GENDER BY AGE GROUP

```
survdiff(formula = Yw[age_trend == 46] ~ gender[age_trend ==
    46], data = whas100)

n=25, 1 observation deleted due to missingness.

                               N Observed Expected (O-E)^2/E (O-E)^2/V
gender[age_trend == 46]=Male   20        5     6.53     0.357      1.95
gender[age_trend == 46]=Female  5        3     1.47     1.584      1.95

 Chisq= 1.9  on 1 degrees of freedom, p= 0.163
```


TESTING GENDER BY AGE GROUP

```
> survdiff(Yw[age_trend == 65] ~ gender[age_trend == 65], data = whas100)
Call:
survdiff(formula = Yw[age_trend == 65] ~ gender[age_trend ==
    65], data = whas100)

n=23, 1 observation deleted due to missingness.

                               N Observed Expected (O-E)^2/E (O-E)^2/V
gender[age_trend == 65]=Male   17        4      5.6     0.458      2.41
gender[age_trend == 65]=Female  6        3      1.4     1.833      2.41

 Chisq= 2.4  on 1 degrees of freedom, p= 0.121
```

TESTING GENDER BY AGE GROUP

```
> survdiff(Yw[age_trend == 75] ~ gender[age_trend == 75], data = whas100)
Call:
survdiff(formula = Yw[age_trend == 75] ~ gender[age_trend ==
    75], data = whas100)

n=22, 1 observation deleted due to missingness.

                               N Observed Expected (O-E)^2/E (O-E)^2/V
gender[age_trend == 75]=Male   15       10     9.07    0.0947     0.273
gender[age_trend == 75]=Female  7        4     4.93    0.1743     0.273

 Chisq= 0.3  on 1 degrees of freedom, p= 0.602
```


TESTING GENDER BY AGE GROUP

```
> survdiff(Yw[age_trend == 86] ~ gender[age_trend == 86], data = whas100)
Call:
survdiff(formula = Yw[age_trend == 86] ~ gender[age_trend ==
    86], data = whas100)

n=30, 1 observation deleted due to missingness.

                               N Observed Expected (O-E)^2/E (O-E)^2/V
gender[age_trend == 86]=Male   13        9     8.83   0.00318   0.00574
gender[age_trend == 86]=Female 17       13    13.17   0.00213   0.00574

 Chisq= 0  on 1 degrees of freedom, p= 0.94
```

STRATIFIED TEST

```
> survdiff(Yw ~ gender + strata(age_trend), data = whas100)
Call:
survdiff(formula = Yw ~ gender + strata(age_trend), data = whas100)

n=100, 1 observation deleted due to missingness.

               N Observed Expected (O-E)^2/E (O-E)^2/V
gender=Male   65       28       30     0.138     0.402
gender=Female 35       23       21     0.197     0.402

 Chisq= 0.4  on 1 degrees of freedom, p= 0.526
```

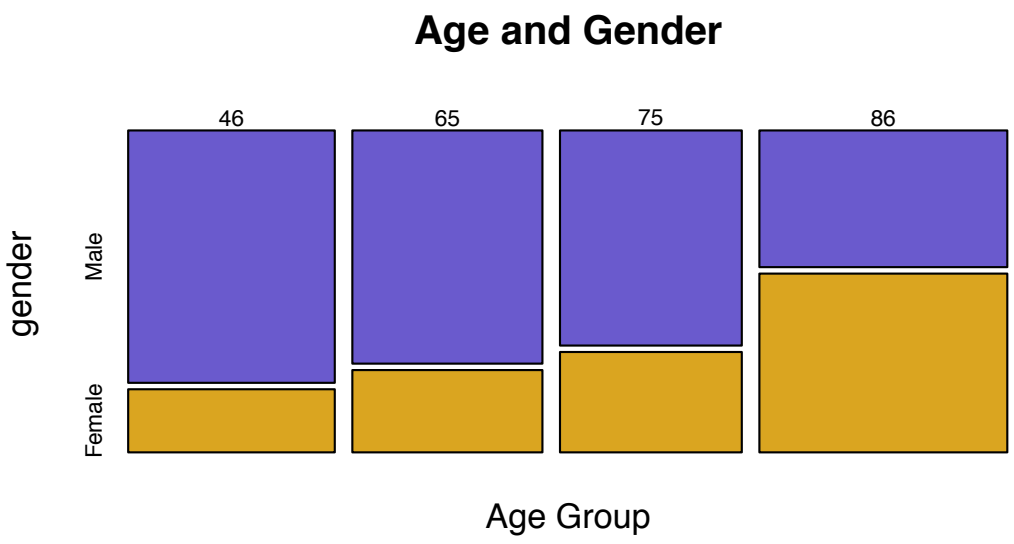

UN-STRATIFIED TEST

```
> survdiff(Yw ~ gender, data = whas100)
Call:
survdiff(formula = Yw ~ gender, data = whas100)

n=100, 1 observation deleted due to missingness.

               N Observed Expected (O-E)^2/E (O-E)^2/V
gender=Male   65       28     34.7      1.29      4.06
gender=Female 35       23     16.3      2.74      4.06

 Chisq= 4.1  on 1 degrees of freedom, p= 0.044
```

WHY?

```
> with(whas100,table(age_trend, gender))
         gender
age_trend Male Female
       46   20      5
       65   17      6
       75   15      7
       86   13     17
```


WHY?



HETEROGENEITY

- When there are more than two groups, can test for difference somewhere between groups:
- Null hypothesis: $\lambda_1(t) \equiv \lambda_2(t) \equiv \ldots \equiv \lambda_k(t)$
- Alternative hypothesis: $\not\equiv$ somewhere

COLON DATA: THREE TREATMENT GROUPS

	Observed Events	Expected Events
Obs	161	146.1
Lev	123	157.5
Lev+5FU	168	148.4
	452	452

- $\chi^2_2 = 11.7$ (df = one fewer than number of groups)
- P-value: 0.003

TREND

- When there are more than two "ordered" groups, it is sometimes of interest to test the null hypothesis of no difference against a "trend" alternative
- $\lambda_1(t) \le \lambda_2(t) \le \ldots \le \lambda_k(t)$ with < somewhere, or
- $\lambda_1(t) \ge \lambda_2(t) \ge \ldots \ge \lambda_k(t)$ with > somewhere
- Placebo and two or more doses of a therapeutic agent
- Pre-hypothesized

TREND

- The test statistic for trend uses "scores": $s_1, s_2, \ldots, s_k$

$$\frac{\left(\sum_{i=1}^{k} s_j \sum_{j=1}^{J_k} \left(d_{ij} - E_{ij}\right)\right)^2}{s'Vs}$$

- Null hypothesis: $\lambda_1(t) \equiv \lambda_2(t) \equiv \ldots \equiv \lambda_k(t)$
- Specific alternative hypothesis:

$$c^{s_1}\lambda_1(t) \equiv c^{s_2}\lambda_2(t) \equiv \ldots \equiv c^{s_k}\lambda_k(t), c \ne 1$$

- Good power when average difference between observed and expected events grows or diminishes with increasing s_i

TREND

Tumor differentiation and all-cause mortality:

Group	N	Observed	Expected
Well Differentiated	93	42	47.5
Moderately Differentiated	663	311	334.9
Poorly Differentiated	150	88	58.6

Tarone trend test: $\chi_1^2 = 11.57$, $P = 6.6 \times 10^{-4}$

SUMMARY

- Can use logrank test to detect consistent differences (over time) in the hazard of dying (the event occurring) using censored survival data
  - Can stratify on prognostic variables
- Can test for differences between more than two groups
- When alternative is ordered by prior hypothesis, can test for trend rather than heterogeneity

TO WATCH OUT FOR:

- Only ranks are used for “standard” tests
- Observations with time = 0
- Crossing hazard functions
- P-value not valid if you decide between trend and heterogeneity test after looking at the data
  - Data told you what your hypothesis was

In R

Load packages.

```
library(survival)
library(rms)
library(survMisc)
library(foreign)
```

Get data.

```
data(colon) # in survival package
head(colon)
```

```
##   id study      rx sex age obstruct perfor adhere nodes status differ
## 1  1     1 Lev+5FU   1  43        0      0      0     5      1      2
## 2  1     1 Lev+5FU   1  43        0      0      0     5      1      2
## 3  2     1 Lev+5FU   1  63        0      0      0     1      0      2
## 4  2     1 Lev+5FU   1  63        0      0      0     1      0      2
## 5  3     1     Obs   0  71        0      0      1     7      1      2
## 6  3     1     Obs   0  71        0      0      1     7      1      2
##   extent surg node4 time etype
## 1      3    0     1 1521     2
## 2      3    0     1  968     1
## 3      3    0     0 3087     2
## 4      3    0     0 3087     1
## 5      2    0     1  963     2
## 6      2    0     1  542     1
```

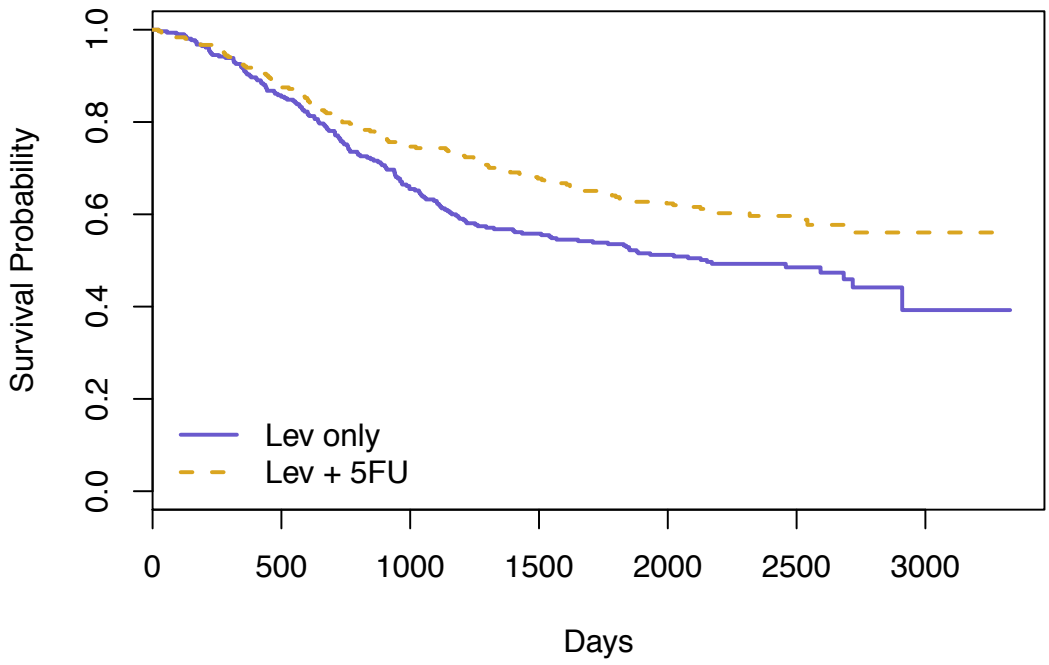
Process data and compute survival curves.

```
df <- colon[colon$etype == 2,] # Use death times.
df <- df[df$rx != "Obs",] # Omit observation only arm.
Y <- with(df, Surv(time, status))
Shats <-survfit(Y ~ rx, data = df)
```

Plot survival curves.

```
colors <- c("slateblue", "goldenrod")
plot(Shats, lty = c(1,2),
     col = colors, lwd = 2,
     xlab = "Days", ylab = "Survival Probability")
legend("bottomleft", lty = c(1,2),
       col = colors, lwd = 2,
       legend = c("Lev only", "Lev + 5FU"), bty = "n")
```

Plot survival curves.



Logrank test

```
survdiff(Y ~ rx, data = df)
```

```
## Call:
## survdiff(formula = Y ~ rx, data = df)
##
##               N Observed Expected (O-E)^2/E (O-E)^2/V
## rx=Lev      310      161      137      4.24      8.21
## rx=Lev+5FU 304      123      147      3.95      8.21
##
##  Chisq= 8.2  on 1 degrees of freedom, p= 0.00417
```

Stratified logrank test

Get data.

```
whas100 <- read.dta("/Users/barb1/Documents/Biostat/Class/SIB/SISCR2016/Module4
Yw <- with(whas100, Surv(surv, fstat == "Dead"))
```

Stratum-specific test

```
survdiff(Yw[age_trend == 46] ~ gender[age_trend == 46], data = whas100)
```

```
## Call:
## survdiff(formula = Yw[age_trend == 46] ~ gender[age_trend ==
##     46], data = whas100)
##
## n=25, 1 observation deleted due to missingness.
##
##                                 N Observed Expected (O-E)^2/E (O-E)^2/V
## gender[age_trend == 46]=Male   20        5     6.53     0.357      1.95
## gender[age_trend == 46]=Female  5        3     1.47     1.584      1.95
##
##  Chisq= 1.9  on 1 degrees of freedom, p= 0.163
```

Stratum-specific test

```
survdiff(Yw[age_trend == 65] ~ gender[age_trend == 65], data = whas100)
```

```
## Call:
## survdiff(formula = Yw[age_trend == 65] ~ gender[age_trend ==
##     65], data = whas100)
##
## n=23, 1 observation deleted due to missingness.
##
##                                 N Observed Expected (O-E)^2/E (O-E)^2/V
## gender[age_trend == 65]=Male   17        4      5.6     0.458      2.41
## gender[age_trend == 65]=Female  6        3      1.4     1.833      2.41
##
##  Chisq= 2.4  on 1 degrees of freedom, p= 0.121
```

Stratum-specific test

```
survdiff(Yw[age_trend == 75] ~ gender[age_trend == 75], data = whas100)
```

```
## Call:
## survdiff(formula = Yw[age_trend == 75] ~ gender[age_trend ==
##     75], data = whas100)
##
## n=22, 1 observation deleted due to missingness.
##
##                                 N Observed Expected (O-E)^2/E (O-E)^2/V
## gender[age_trend == 75]=Male   15       10     9.07    0.0947     0.273
## gender[age_trend == 75]=Female  7        4     4.93    0.1743     0.273
##
##  Chisq= 0.3  on 1 degrees of freedom, p= 0.602
```

Stratum-specific test

```
survdiff(Yw[age_trend == 86] ~ gender[age_trend == 86], data = whas100)
```

```
## Call:
## survdiff(formula = Yw[age_trend == 86] ~ gender[age_trend ==
##     86], data = whas100)
##
## n=30, 1 observation deleted due to missingness.
##
##                                 N Observed Expected (O-E)^2/E (O-E)^2/V
## gender[age_trend == 86]=Male   13        9     8.83   0.00318   0.00574
## gender[age_trend == 86]=Female 17       13    13.17   0.00213   0.00574
##
##  Chisq= 0  on 1 degrees of freedom, p= 0.94
```

Stratified logrank test

```
survdiff(Yw ~ gender + strata(age_trend), data = whas100)
```

```
## Call:
## survdiff(formula = Yw ~ gender + strata(age_trend), data = whas100)
##
## n=100, 1 observation deleted due to missingness.
##
##                 N Observed Expected (O-E)^2/E (O-E)^2/V
## gender=Male    65       28       30     0.138     0.402
## gender=Female 35       23       21     0.197     0.402
##
##  Chisq= 0.4  on 1 degrees of freedom, p= 0.526
```

Un-stratified logrank test

```
survdiff(Yw ~ gender, data = whas100)
```

```
## Call:
## survdiff(formula = Yw ~ gender, data = whas100)
##
## n=100, 1 observation deleted due to missingness.
##
##                 N Observed Expected (O-E)^2/E (O-E)^2/V
## gender=Male    65       28     34.7      1.29      4.06
## gender=Female 35       23     16.3      2.74      4.06
##
##  Chisq= 4.1  on 1 degrees of freedom, p= 0.044
```

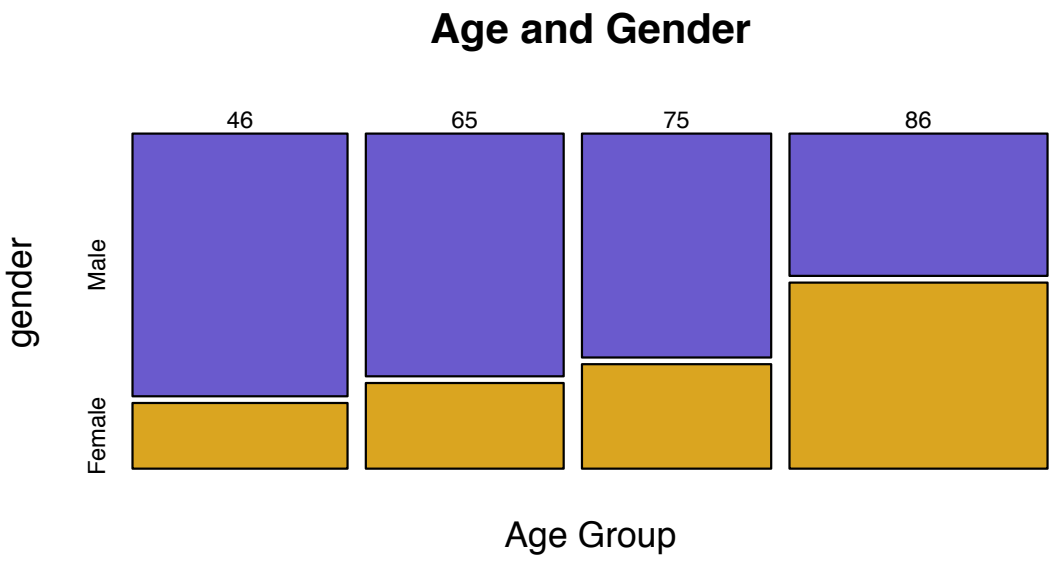
Why?

```
with(whas100,table(age_trend, gender))
```

```
##           gender
## age_trend Male Female
##        46   20      5
##        65   17      6
##        75   15      7
##        86   13     17
```

Why?

```
mosaicplot(age_trend ~ gender, data = whas100, dir = "v",
           col = c("slateblue", "goldenrod"), main = "Age and Gender",
           xlab = "Age Group")
```



Three group test data

```
df2 <- colon[colon$etype == 2,] # Use death times.
Y2 <- with(df2, Surv(time, status))
Shats3 <-survfit(Y2 ~ rx, data = df2, conf.type = "log-log")
```

Three group test

```
survdiff(Y2 ~ rx, data = df2)
```

```
## Call:
## survdiff(formula = Y2 ~ rx, data = df2)
##
##              N Observed Expected (O-E)^2/E (O-E)^2/V
## rx=Obs     315      168      148      2.58      3.85
## rx=Lev     310      161      146      1.52      2.25
## rx=Lev+5FU 304      123      157      7.55     11.62
##
##  Chisq= 11.7  on 2 degrees of freedom, p= 0.0029
```

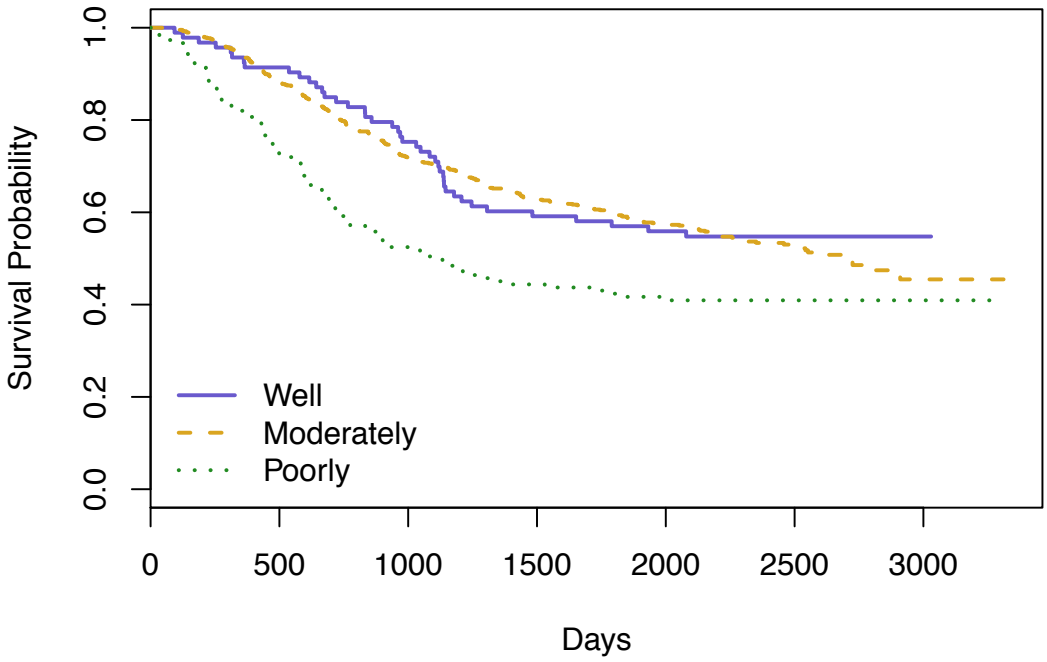
Trend test function (Dan Gillen)

```
## survtrend() : Function to compute the Tarone trend test formula : Surv()
## response and covariate for trend test data : dataset containing response
## and predictor print.table : if TRUE, prints observed and expected failures
## under H_0 in each group
survtrend <- function(formula, data, print.table = TRUE) {
    lrfit <- survdiff(formula, data = data)
    df <- length(lrfit$n) - 1
    score <- coxph(formula, data = data)$score
    if (print.table) {
        oetable <- cbind(lrfit$n, lrfit$obs, lrfit$exp)
        colnames(oetable) <- c("N", "Observed", "Expected")
        print(oetable)
    }
    cat("\nLogrank Test : Chi(", df, ") = ", lrfit$chisq, ", p-value = ", 1 -
        pchisq(lrfit$chisq, df), sep = "")
    cat("\nTarone Test Trend : Chi(1) = ", score, ", p-value = ", 1 - pchisq(sc
        1), sep = "")
}
```

Trend test

```
Shats3t <-survfit(Y2 ~ differ, data = df2, conf.type = "log-log")
colors <- c("slateblue", "goldenrod", "forestgreen")
plot(Shats3t, lty = c(1:3),
     col = colors, lwd = 2,
     xlab = "Days", ylab = "Survival Probability")
legend("bottomleft", lty = c(1:3),
       col = colors, lwd = 2,
       legend = c("Well", "Moderately", "Poorly"), bty = "n")
```

Trend test



Trend test

```
survtrend(Y2 ~ differ, data = df2)
```

```
##            N Observed Expected
## differ=1  93       42  47.5287
## differ=2 663      311 334.9173
## differ=3 150       88  58.5540
##
## Logrank Test : Chi(2) = 17.18909, p-value = 0.0001851124
## Tarone Test Trend : Chi(1) = 11.57379, p-value = 0.0006688778
```

Your turn

Using the colon cancer data in the survival package with overall survival (after loading survival package type ?colon for documentation):

1. Perform the logrank test of whether the hazard ratio for all-cause mortality associated with having more than 4 lymph nodes positive for cancer at diagnosis is one.
2. Perform the logrank test of whether the hazard ratio for all-cause mortality associated with having more than 4 lymph nodes positive for cancer at diagnosis is one after stratification adjustment for treatment arm.
3. Perform the logrank test of whether the all-cause-mortality hazard depends on extent of disease at diagnosis (heterogeneity test).
4. Perform the logrank test of whether the all-cause-mortality hazard is higher or lower for greater extent of disease at diagnosis (trend test).

Write a "results" sentence or two for each of these analyses.

SESSION 4:
INTRODUCTION TO COX REGRESSION

Module 4: Introduction to Survival Analysis
Summer Institute in Statistics for Clinical Research
University of Washington
July, 2016

Barbara McKnight, Ph.D.
Professor
Department of Biostatistics
University of Washington

OVERVIEW

- Session 1
  - Introductory examples
  - The survival function
  - Survival Distributions
  - Mean and Median survival time

- Session 2
  - Censored data
  - Risk sets
  - Censoring Assumptions
  - Kaplan-Meier Estimator and CI
  - Median and CI

- Session 3
  - Two-group comparisons: logrank test
  - Trend and heterogeneity tests for more than two groups

- Session 4
  - Introduction to Cox regression

OUTLINE

- Motivation:
  - Confounding and stratified randomization designs
- Cox Regression model
  - Coefficient interpretation
  - Estimation and testing
  - Relationship to 2- and K-sample tests
- Examples throughout

OUTLINE

- **Motivation:**
  - **Confounding and stratified randomization designs**
- Cox Regression model
  - Coefficient interpretation
  - Estimation and testing
  - Relationship to 2- and K-sample tests
- Examples throughout

CONFOUNDING

- Observational data: sometimes observed associations between an explanatory variable and outcome can be due to their joint association with another variable.
  - Age related to both sex and risk of death.
  - Other examples?

PRECISION IN RCTS

- Because of randomization, confounding/imbalance usually not an issue except in small trials.
- As in linear regression, regression models for censored survival data allow group comparisons among subjects with similar values of adjustment or "precision" variables (more later).
- Fairer and possibly more powerful comparison as long as adjustment variables are not the result of treatment.

STRATIFIED RANDOMIZATION

- For strong predictors: concern about possible randomization imbalance
  - Clinic or center
  - Stage of disease
  - Sex
  - Age
- Adjust for stratification variables in analysis
  - More powerful if predictors are strong
  - Same conditioning as the sampling

OUTLINE

- Motivation:
  - Confounding and stratified randomization designs
- **Cox Regression model**
  - **Coefficient interpretation**
  - **Estimation and testing**
  - **Relationship to 2- and K-sample tests**
- Examples throughout

COX REGRESSION MODEL

- Usually written in terms of the hazard function
- As a function of independent variables $x_1, x_2, \ldots x_k$,

$$\lambda(t) = \lambda_0(t)e^{\beta_1 x_1 + \cdots + \beta_k x_k}$$

↑ relative risk / hazard ratio

$$\log \lambda(t) = \log \lambda_0(t) + \beta_1 x_1 + \cdots + \beta_k x_k$$

↑ intercept

RELATIVE RISK / HAZARD RATIO

$$\lambda(t|x_1, \ldots, x_k) = \lambda_0(t)e^{\beta_1 x_1 + \cdots + \beta_k x_k}$$

$$\frac{\lambda(t|x_1,\ldots,x_k)}{\lambda(t|0,\ldots,0)} = e^{\beta_1 x_1 + \cdots + \beta_k x_k}$$

REGRESSION MODELS

LS Linear Regression: $Y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + \epsilon$

Linear:	$Y \sim N(\mu, \sigma^2)$	$\mu = EY = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k$
Cox:	$T \sim S(t)$	$\lambda(t) = \lambda_0(t) e^{\beta_1 x_1 + \cdots + \beta_k x_k}$
	↑	↑
	Distribution of outcome variable	Dependence of distribution on $x_1, \ldots x_k$

PROPORTIONAL HAZARDS MODEL

$$\lambda(t) = \lambda_0(t) e^{\beta_1 x_1 + \cdots + \beta_k x_k}$$

Interpretation of e^{β_1} in general:

"Relative risk (or hazard ratio) associated with a one unit higher value of x_1, holding $x_2, \ldots, x_k$ constant".

$\lambda(t)$ for $x_1 + 1$:	$\lambda_0(t) e^{\beta_1 (x_1 + 1) + \cdots + \beta_k x_k}$
$\lambda(t)$ for x_1:	$\lambda_0(t) e^{\beta_1 x_1 + \cdots + \beta_k x_k}$
ratio:	$e^{\beta_1 (x_1 + 1 - x_1)} = e^{\beta_1}$

EXAMPLE

Single binary x:

$$x = \begin{cases} 1 & \text{Test treatment} \\ 0 & \text{Standard treatment} \end{cases}$$

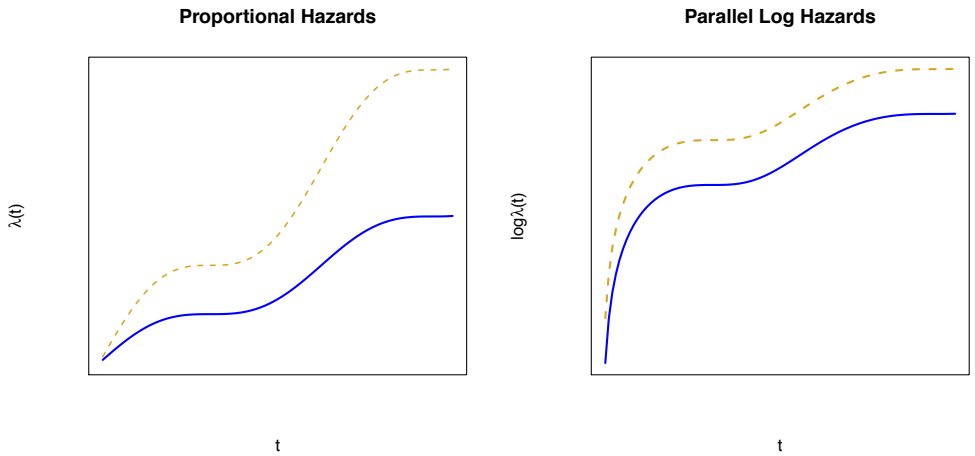
$$\lambda(t) = \lambda_0(t)e^{\beta x}$$

Interpretation of e^{β}:

"Relative risk (or hazard ratio) comparing test treatment to standard".

$$\begin{aligned} \lambda(t) \text{ for } x = 1: \quad \lambda_0(t)e^{\beta \cdot 1} &= \lambda_0(t)e^{\beta} \\ \lambda(t) \text{ for } x = 0: \quad \lambda_0(t)e^{\beta \cdot 0} &= \lambda_0(t) \\ \text{ratio:} \quad e^{\beta(1-0)} &= e^{\beta} \end{aligned}$$

EXAMPLE



RELATIONSHIP TO SURVIVAL FUNCTION

Single binary x:

$$x = \begin{cases} 1 & \text{Test treatment} \\ 0 & \text{Standard treatment} \end{cases}$$

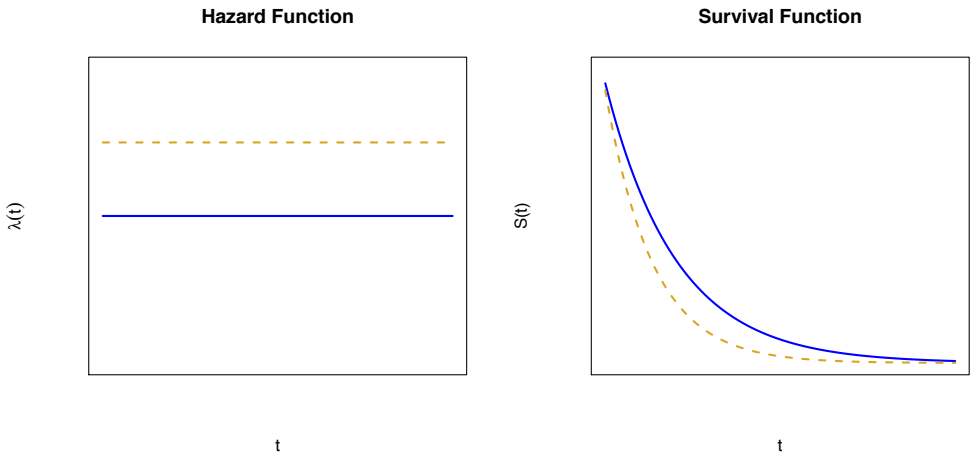
$$\lambda(t) = \lambda_0(t)e^{\beta x} \implies S(t) = [S_0(t)]^{e^{\beta x}}$$

In terms of $S_0(t)$:

$S(t)$ for $x = 1$: $\quad [S_0(t)]^{e^{\beta \cdot 1}} = [S_0(t)]^{e^{\beta}}$

$S(t)$ for $x = 0$: $\quad [S_0(t)]^{e^{\beta \cdot 0}} = [S_0(t)]^1 = S_0(t)$

PICTURE



ESTIMATES AND CONFIDENCE INTERVALS

- We estimate β by maximizing the "partial likelihood function"
- Requires iteration on computer
- $\hat{\beta}$ is a MPLE (Maximum Partial Likelihood Estimator)
- We do not need to estimate $\lambda_0(t)$ to do this

- Most packages will estimate se($\hat{\beta}$) using the information matrix from this PL.
- 95% CI for β: $(\hat{\beta} - 1.96\text{se}(\hat{\beta}),\ \hat{\beta} + 1.96\text{se}(\hat{\beta}))$
- 95% CI for RR = e^{β} : $(e^{\hat{\beta}-1.96\text{se}(\hat{\beta})}, e^{\hat{\beta}+1.96\text{se}(\hat{\beta})})$

PARTIAL LIKELIHOOD

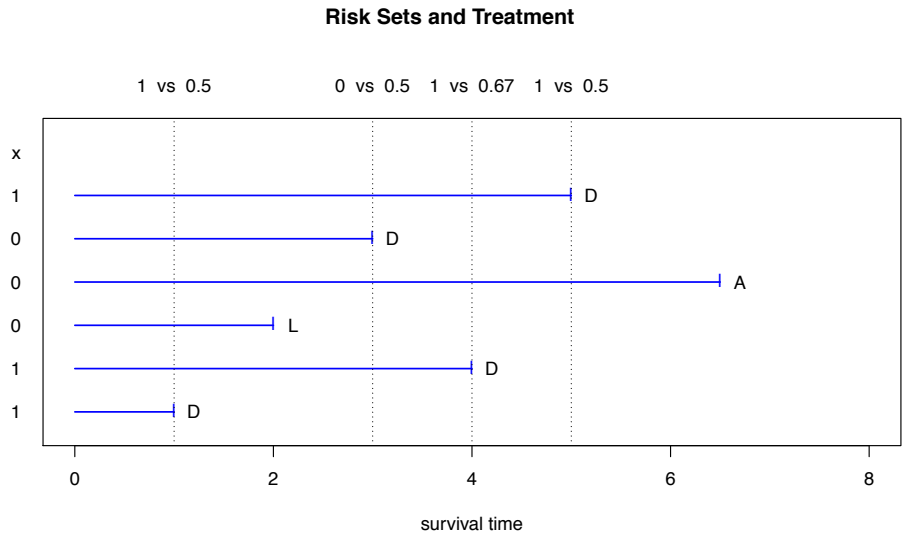
Data for the i^{th} subject: $(t_i, \delta_i, x_{1i}, \ldots x_{ki})$

For subject with the j^{th} ordered failure time : $(t_{(j)}, 1, x_{1(j)}, \ldots, x_{k(j)})$

$$\text{PL}(\beta_1, \ldots, \beta_k) = \prod_{j=1}^{J} \frac{e^{\beta_1 x_{1(j)} + \cdots + \beta_k x_{k(j)}}}{\sum_{i:t_i \geq t_{(j)}} e^{\beta_1 x_{1i} + \cdots + \beta_k x_{ki}}}$$

- $(\hat{\beta}_1, \ldots, \hat{\beta}_k)$ are the values of $(\beta_1, \ldots, \beta_k)$ that maximize $\text{PL}(\beta_1, \ldots, \beta_k)$. (MPLEs)
- Compares x values for the subject who failed at time $t_{(j)}$ to those of all subjects at risk at time $t_{(j)}$.
- Does not depend on the values of the t_i, only on their order.
- Does not depend on $\lambda_0(t)$.

RISK SET PICTURE



FULL LIKELIHOOD

$$
\begin{aligned}
L(\beta, \lambda_0(t)) &= \prod_{\text{Failures}} \Pr[T = t_i] \prod_{\text{Censorings}} \Pr[T > t_i] \\
&= \prod_{\text{Failures}} \lambda(t_i|x_i) S(t_i|x_i) \prod_{\text{Censorings}} S(t_i|x_i) \\
&= \prod_{i=1}^{n} [\lambda(t_i|x_i)]^{\delta_i} S(t_i|x_i) \\
&= \prod_{i=1}^{n} [\lambda_0(t_i) e^{\beta x_i}]^{\delta_i} e^{-\int_0^{t_i} \lambda_0(s) e^{\beta x} ds}
\end{aligned}
$$

PARTIAL LIKELIHOOD

Let H_t represent the entire history of failure, censoring and x in the sample before time t.

Then the likelihood can be rewritten as follows:

$$
\begin{aligned}
L(\beta, \lambda_0(t)) &= \prod_{j=1}^{J} \Pr[i^{th} \text{ subject fails at } t_{(j)} | H_{t_{(j)}}, \text{ some subject fails at } t_{(j)}] \cdot \\
&\qquad \Pr[H_{t_{(j)}}, \text{ some subject fails at } t_{(j)}] \\
&= \prod_{j=1}^{J} \frac{\lambda(t_{(j)}|x_{(j)})}{\sum_{i:t_i \geq t_{(j)}} \lambda(t_{(j)}|x_i)} \cdot \prod_{j=1}^{J} \Pr[H_{t_{(j)}}, \text{ some subject fails at } t_{(j)}] \\
&= \prod_{j=1}^{J} \frac{\lambda_0(t_{(j)}) e^{\beta x_{(j)}}}{\sum_{i:t_i \geq t_{(j)}} \lambda_0(t_{(j)}) e^{\beta x_i}} \cdot \prod_{j=1}^{J} \Pr[H_{t_{(j)}}, \text{ some subject fails at } t_{(j)}] \\
&= \underbrace{\prod_{j=1}^{J} \frac{e^{\beta x_{(j)}}}{\sum_{i:t_i \geq t_{(j)}} e^{\beta x_i}}}_{\text{Partial Likelihood}} \cdot \underbrace{\prod_{j=1}^{J} \Pr[H_{t_{(j)}}, \text{ some subject fails at } t_{(j)}]}_{\text{Depends on } \lambda_0(\cdot) \text{ and } \beta} \\
&= \qquad \text{Depends only on } \beta
\end{aligned}
$$

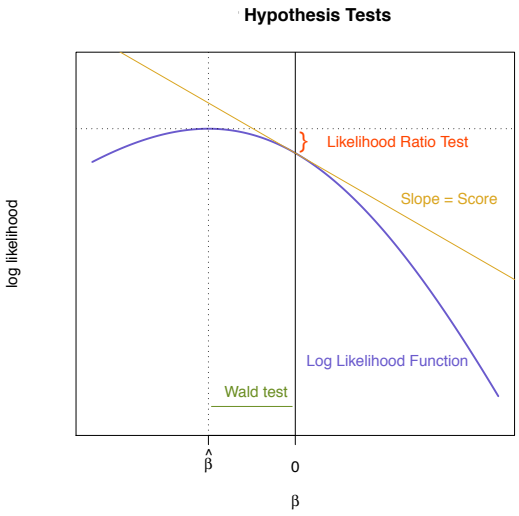
HYPOTHESIS TESTS

Three tests of $H_0 : \beta = 0$ are possible:

1. Wald test: $\dfrac{\hat{\beta}}{\text{se}(\hat{\beta})}$
2. (Partial) Likelihood ratio test
3. Score test: ($\approx$ logrank test)

Likelihood ratio test is best, but requires fitting full ($\beta = \hat{\beta}$) and reduced ($\beta = 0$) models.

LIKELIHOODS AND TESTS



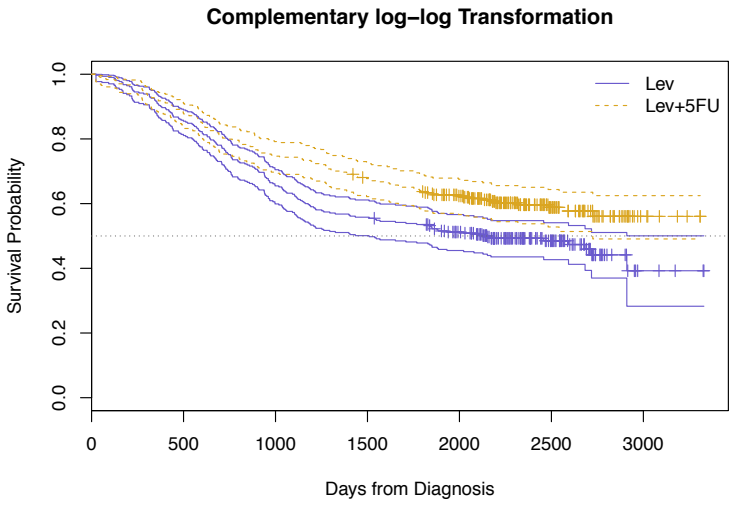
COLON CANCER EXAMPLE

- Levamisole and Fluorouracil for adjuvant therapy of resected colon carcinoma
  - Moertel et al. *New England Journal of Medicine*. 1990;322(6):352–358.
  - Moertel et al. *Annals of internal medicine*. 1995;122(5): 321–326.
- 1296 patients
- Stage B_2 or C
- 3 unblinded treatment groups
  - Observation only
  - Levamisole (oral, 1yr)
  - Levamisole (oral, 1yr) + 5 fluorouracil (intravenous 1yr)

COLON CANCER EXAMPLE

- Clinical trial at Mayo Clinic
- Stage B_2 and C colon cancer patients; adjuvant therapy
- Three arms
  - Observation only
  - Levamisole
  - 5-FU + Levamisole at Mayo Clinic
- Stage C patients only
- Two treatment arms only

COLON CANCER EXAMPLE



COLON CANCER EXAMPLE

Variable	n	Deaths	Hazard ratio	CI	P-value
Levamisole Only	310	161	1.0 (reference)	--	--
Levamisole + 5FU	304	123	0.71	(0.56, 0.90)	.004

Q: Which group has better survival?

A:

TEST COMPARISON

Test	Statistic	P-value
Wald's	8.13	.004
Score	8.21	.004
Likelihood Ratio	8.21	.004

Two-sided tests

ANOTHER EXAMPLE

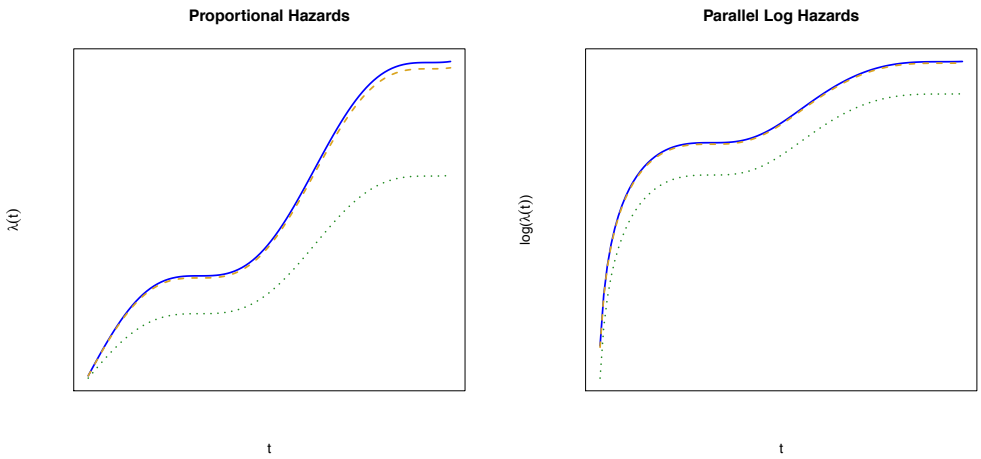
Three groups: use indicators for two

$$x_1 = \begin{cases} 1 & \text{Levamisole Only} \\ 0 & \text{otherwise} \end{cases} \qquad x_2 = \begin{cases} 1 & \text{Levamisole + 5FU} \\ 0 & \text{otherwise} \end{cases}$$

Model: $\lambda(t) = \lambda_0(t)e^{\beta_1 x_1 + \beta_2 x_2}$

RRs:	Levamisole Only	vs.	Observation	e^{β_1}
	Levamisole + 5FU	vs.	Observation	e^{β_2}
	Levamisole + 5FU	vs.	Levamisole Only	$e^{\beta_2 - \beta_1}$

HEURISTIC HAZARDS



COLON CANCER

Variable	n	Deaths	Hazard Ratio	95% CI	P-value
Observation Only	315	168	1.0 (reference)	--	--
Levamisole Only	310	161	0.97	(0.78, 1.21)	0.81
Levamisole + 5FU	204	123	0.69	(0.55, 0.87)	0.002

Q: Which group has best survival?

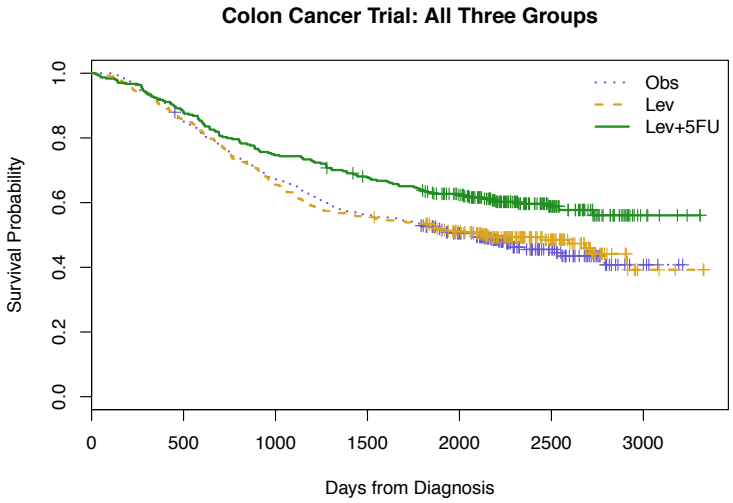
A:

TEST COMPARIOSN

Test	Statistic	P-value
Wald's	11.56	.003
Score	11.68	.003
Likelihood Ratio	12.15	.002

Same hypothesis as 3-group heterogeneity test. Score test is same in large samples.

COLON CANCER TRIAL DATA



TREND

- When there are several groups, it is sometimes of interest to test whether risk increases from one group to the next:
  - Several dose groups
  - Other ordered variable
  - Example: tumor differentiation
- For $x = \begin{cases} 1 & \text{well differentiated} \\ 2 & \text{moderately differentiated} \\ 3 & \text{poorly differentiated} \end{cases}$

$$\text{Model: } \lambda(t) = \lambda_0(t)e^{\beta x}$$

- Score test is the same as the trend test
- Could use other values for x (actual dose levels)

TREND

$$\text{For } x = \begin{cases} 1 & \text{well differentiated} \\ 2 & \text{moderately differentiated} \\ 3 & \text{poorly differentiated} \end{cases}$$

$$\text{Model: } \lambda(t) = \lambda_0(t)e^{\beta x}$$

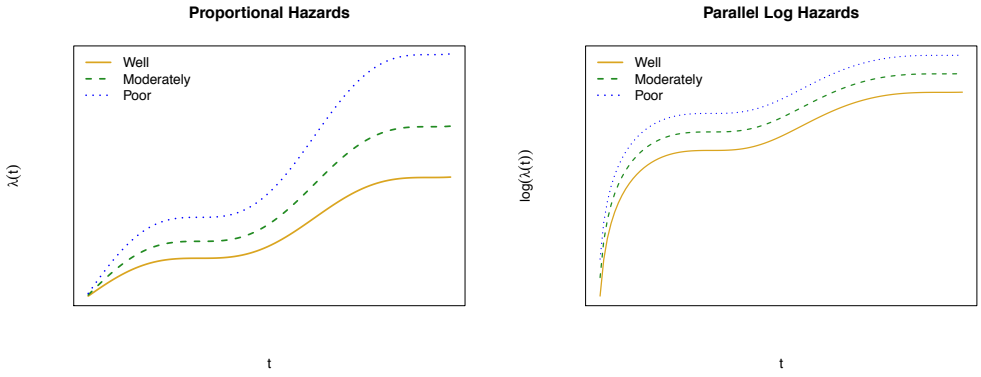
Interpretation of e^{β}: HR associated with the comparison of one worse differentiation group to one better:

- poorly differentiated to moderately differentiated, or
- moderately differentiated to well differentiated

Q: What is HR comparing poorly differentiated to well differentiated?

A:

TREND



TREND WITH DIFFERENTIATION

One presentation based entirely on trend ("grouped linear") model:

	Hazard Ratio	95% CI
One category worse differentiation (well, moderately, poor)	1.4	(1.1, 1.8)
P = .003 (trend)		

I prefer presenting hazard ratios and CI's based on dummy variable model, and providing P-value for trend.

TREND WITH DIFFERENTIATION

My preferred presentation based on dummy variable mode with trend P-value:

	n	Deaths	Hazard Ratio	95% CI
Well differentiated	66	26	1.0 (reference)	--
Moderately differentiated	434	196	1.2	(0.80, 1.8)
Poorly differentiated	98	54	1.8	(1.2, 3.0)
P = .003 (trend)				

I usually would not present this for an *a priori* trend hypothesis, but for comparison here, the heterogeneity P-value (2 df) is 0.009.

TO WATCH OUT FOR:

- Coefficients in Cox regression are positively associated with risk, not survival.
  - Positive β means large values of x are associated with shorter survival.
- Without certain types of time-dependent covariates, Cox regression does not depend on the actual times, just their order.
  - Can add a constant to all times to remove zeros (some packages remove observations with time = 0) without changing inference
- For LRT, nested models must be compared based on same subjects.
  - If some values of variables in larger model are missing, these subjects must be removed from fit of smaller model.

In R

Load packages.

```
library(survival)
library(rms)
library(survMisc)
library(foreign)
```

Get data.

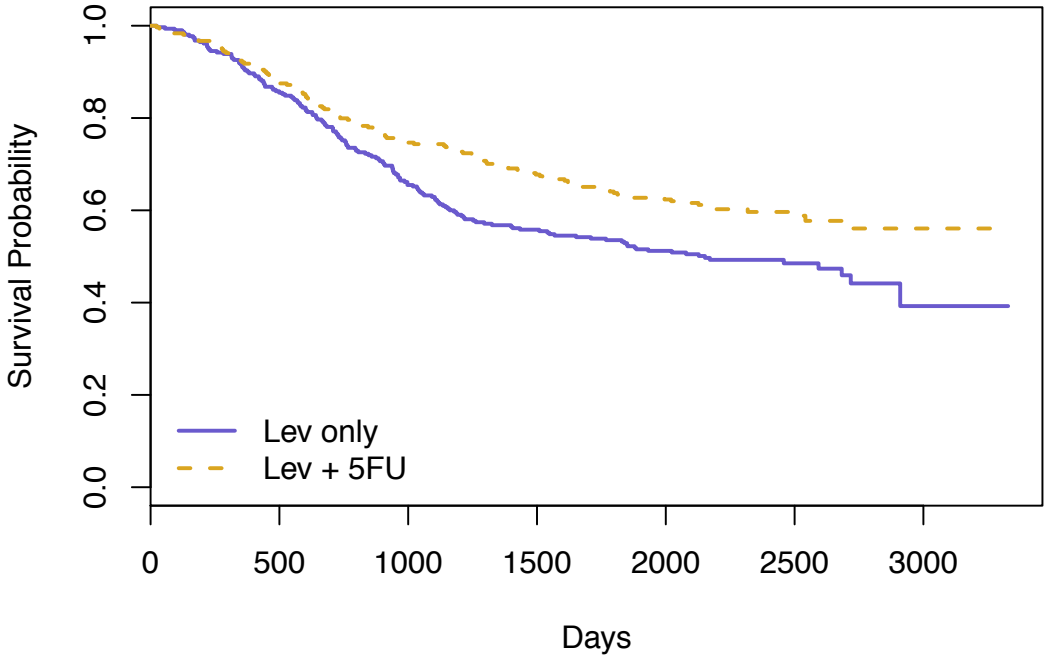
```
data(colon) # in survival package
head(colon)
```

```
##   id study      rx sex age obstruct perfor adhere nodes status differ
## 1  1     1 Lev+5FU   1  43        0      0      0     5      1      2
## 2  1     1 Lev+5FU   1  43        0      0      0     5      1      2
## 3  2     1 Lev+5FU   1  63        0      0      0     1      0      2
## 4  2     1 Lev+5FU   1  63        0      0      0     1      0      2
## 5  3     1     Obs   0  71        0      0      1     7      1      2
## 6  3     1     Obs   0  71        0      0      1     7      1      2
##   extent surg node4 time etype
## 1      3    0     1 1521     2
## 2      3    0     1  968     1
## 3      3    0     0 3087     2
## 4      3    0     0 3087     1
## 5      2    0     1  963     2
## 6      2    0     1  542     1
```

Process data and compute survival curves.

```
df <- colon[colon$etype == 2,] # Use death times.
df <- df[df$rx != "Obs",] # Omit observation only arm.
Y <- with(df, Surv(time, status))
Shats <-survfit(Y ~ rx, data = df)
```

Plot survival curves.



Fit Cox model

```
model1 <- coxph(Y ~ rx, data = df)
```

```
## Warning in coxph(Y ~ rx, data = df): X matrix deemed to be singular;
## variable 2
```

Results

```
summary(model1)
```

```
## Call:
## coxph(formula = Y ~ rx, data = df)
##
##   n= 614, number of events= 284
##
##             coef exp(coef) se(coef)     z Pr(>|z|)
## rxLev     0.3417    1.4073   0.1199 2.851  0.00436 **
## rxLev+5FU     NA        NA   0.0000    NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##           exp(coef) exp(-coef) lower .95 upper .95
## rxLev         1.407     0.7106     1.113      1.78
## rxLev+5FU        NA         NA        NA        NA
##
## Concordance= 0.541  (se = 0.015 )
## Rsquare= 0.013   (max possible= 0.996 )
## Likelihood ratio test= 8.21  on 1 df,   p=0.00416
## Wald test            = 8.13  on 1 df,   p=0.00436
## Score (logrank) test = 8.21  on 1 df,   p=0.004174
```

Data with All Three Groups

```
df2 <- colon[colon$etype == 2,] # Use death times.
Y2 <- with(df2, Surv(time, status))
```

Dummy variable model

```
model2 <- coxph(Y2 ~ rx, data = df2)
```

Summary

```
summary(model2)
```

```
## Call:
## coxph(formula = Y2 ~ rx, data = df2)
##
##   n= 929, number of events= 452
##
##               coef exp(coef) se(coef)      z Pr(>|z|)
## rxLev     -0.02664   0.97371  0.11030 -0.241  0.80917
## rxLev+5FU -0.37171   0.68955  0.11875 -3.130  0.00175 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##           exp(coef) exp(-coef) lower .95 upper .95
## rxLev        0.9737      1.027    0.7844    1.2087
## rxLev+5FU    0.6896      1.450    0.5464    0.8703
##
## Concordance= 0.536  (se = 0.013 )
## Rsquare= 0.013   (max possible= 0.998 )
## Likelihood ratio test= 12.15  on 2 df,   p=0.002302
## Wald test            = 11.56  on 2 df,   p=0.003092
## Score (logrank) test = 11.68  on 2 df,   p=0.002906
```

Trend model

```
model3 <- coxph(Y2 ~ differ, data = df2)
```

Summary

```
summary(model3)
```

```
## Call:
## coxph(formula = Y2 ~ differ, data = df2)
##
##   n= 906, number of events= 441
##    (23 observations deleted due to missingness)
##
##           coef exp(coef) se(coef)     z Pr(>|z|)
## differ 0.32788   1.38803  0.09618 3.409 0.000651 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##        exp(coef) exp(-coef) lower .95 upper .95
## differ     1.388     0.7204      1.15     1.676
##
## Concordance= 0.544  (se = 0.011 )
## Rsquare= 0.013   (max possible= 0.998 )
## Likelihood ratio test= 11.51  on 1 df,   p=0.0006916
## Wald test            = 11.62  on 1 df,   p=0.0006515
## Score (logrank) test = 11.57  on 1 df,   p=0.0006689
```

Dummy Variables for Differentiation

```
model4 <- coxph(Y2 ~ factor(differ), data = df2)
```

Summary

```
summary(model4)
```

```
## Call:
## coxph(formula = Y2 ~ factor(differ), data = df2)
##
##   n= 906, number of events= 441
##    (23 observations deleted due to missingness)
##
##                     coef exp(coef) se(coef)     z Pr(>|z|)
## factor(differ)2 0.04963   1.05088  0.16441 0.302  0.76275
## factor(differ)3 0.53196   1.70226  0.18764 2.835  0.00458 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                 exp(coef) exp(-coef) lower .95 upper .95
## factor(differ)2     1.051     0.9516    0.7614     1.450
## factor(differ)3     1.702     0.5875    1.1784     2.459
##
## Concordance= 0.544  (se = 0.011 )
## Rsquare= 0.017   (max possible= 0.998 )
## Likelihood ratio test= 15.25  on 2 df,   p=0.0004872
## Wald test            = 16.85  on 2 df,   p=0.0002195
## Score (logrank) test = 17.19  on 2 df,   p=0.0001855
```

Your turn

Using all-cause mortality as the outcome for the colon data in the survival package in R:

1. Fit a Cox model with a binary treatment indicator relating whether more than 4 lymph nodes were positive for disease at diagnosis is related to the hazard of all-cause mortality.
2. Fit a Cox model with dummy-variable indicators for whether extent of disease at diagnosis is related to the hazard of all-cause mortality.
3. Fit a Cox model with "grouped-linear" measure for how extent of disease at diagnosis is related to the hazard of all-cause mortality.

Write a "results" sentence or two for each of these analyses.