APH 431: Essential Math for Pharmacy

Class14

Module Leader: Xiaoyu Tang, PhD

Agenda

- What are survival data (time-to-event data)?
  - Censoring
  - Survival function
  - Hazard function
  - How to summarize/describe survival data
- Kaplan-Meier estimator, log-rank test(Non-parametric)
- Proportional hazards (Cox) regression, Hazard Ratio
- Test for proportional hazard assumption

What are survival data (time-to-event data)?

- Survival data focus on the time until an event of interest happens.
  - Example: time to death (overall survival); time to coronary heart disease
- Interest is not only how long it will take until the subject has an event, but also whether a subject will have an event

1) Subject with an event $\begin{cases} \text{Event} = 1 \\ \text{Time: time to event} \end{cases}$
2) Subject without an event $\begin{cases} \text{Event} = 0 \\ \text{Time: time followed in the study} \end{cases}$

Survival Data Example

	institution	time	status	age	sex
1	3	306	1	74	1
2	3	455	1	68	1
3	3	1010	0	56	1
4	5	210	1	57	1
5	1	883	1	60	1
6	12	1022	0	74	1

data(cancer, package="survival")
Survival in patients with advanced lung cancer from the North Central Cancer Treatment Group.
Event of interest: Death
Time: observed survival time in days
Status: 1 dead; 0 no event (censored)

Censoring

- Definition: The value of a measurement is only partially known

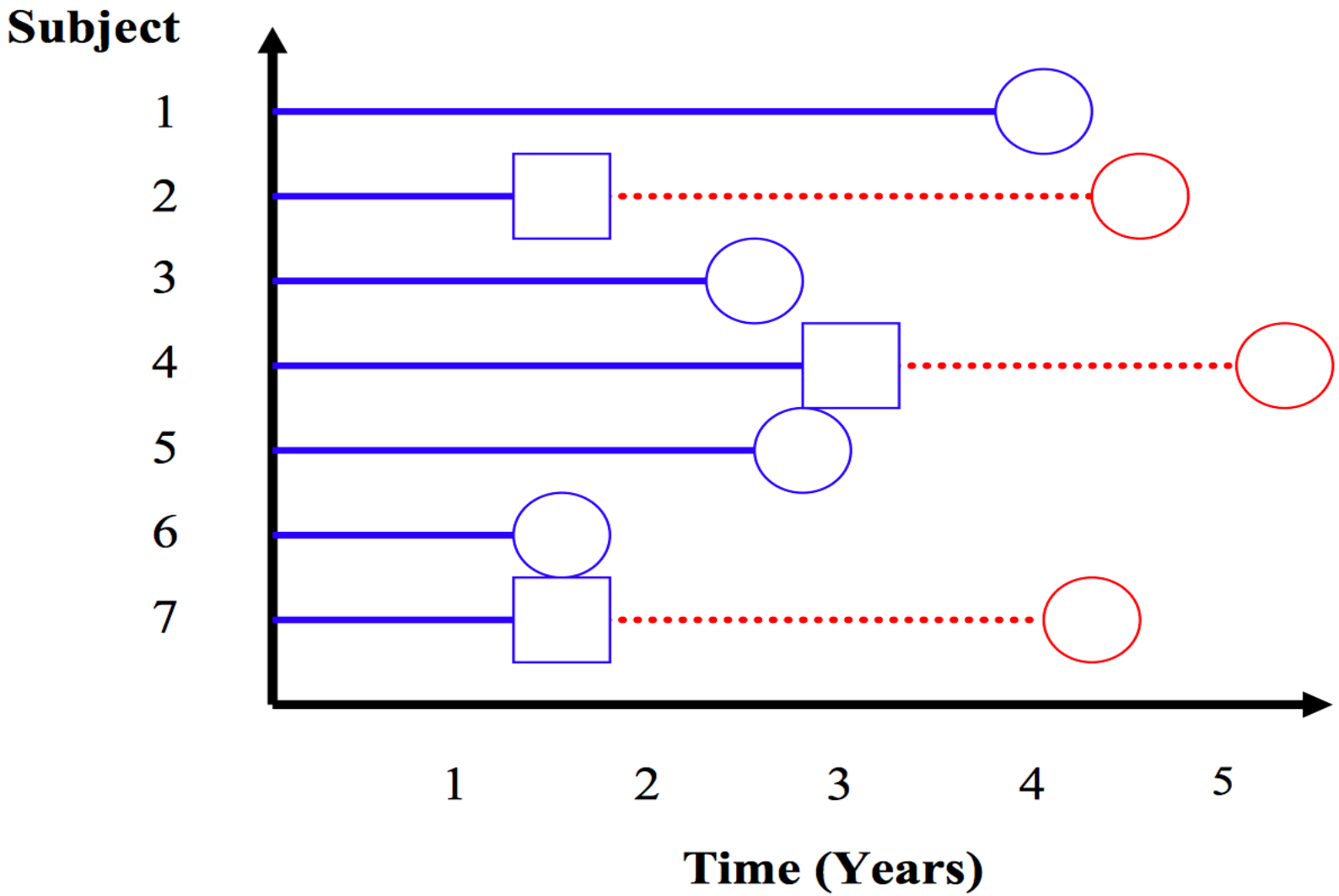
- **Type of censoring:**
  - **Right-censored: we have a lower limit on the time to event for censored subjects (the event of interest does not happen to that individual during the study)**
  - Left-censored: Event occurred before the study
  - Interval-censored: Event occurred between two time points in the study
- Not just “missing data”, informative, subject’s event time may be greater than their follow-up time

- Assumption: Censoring is independent to the likelihood of developing the event of interest (non-informative censoring)

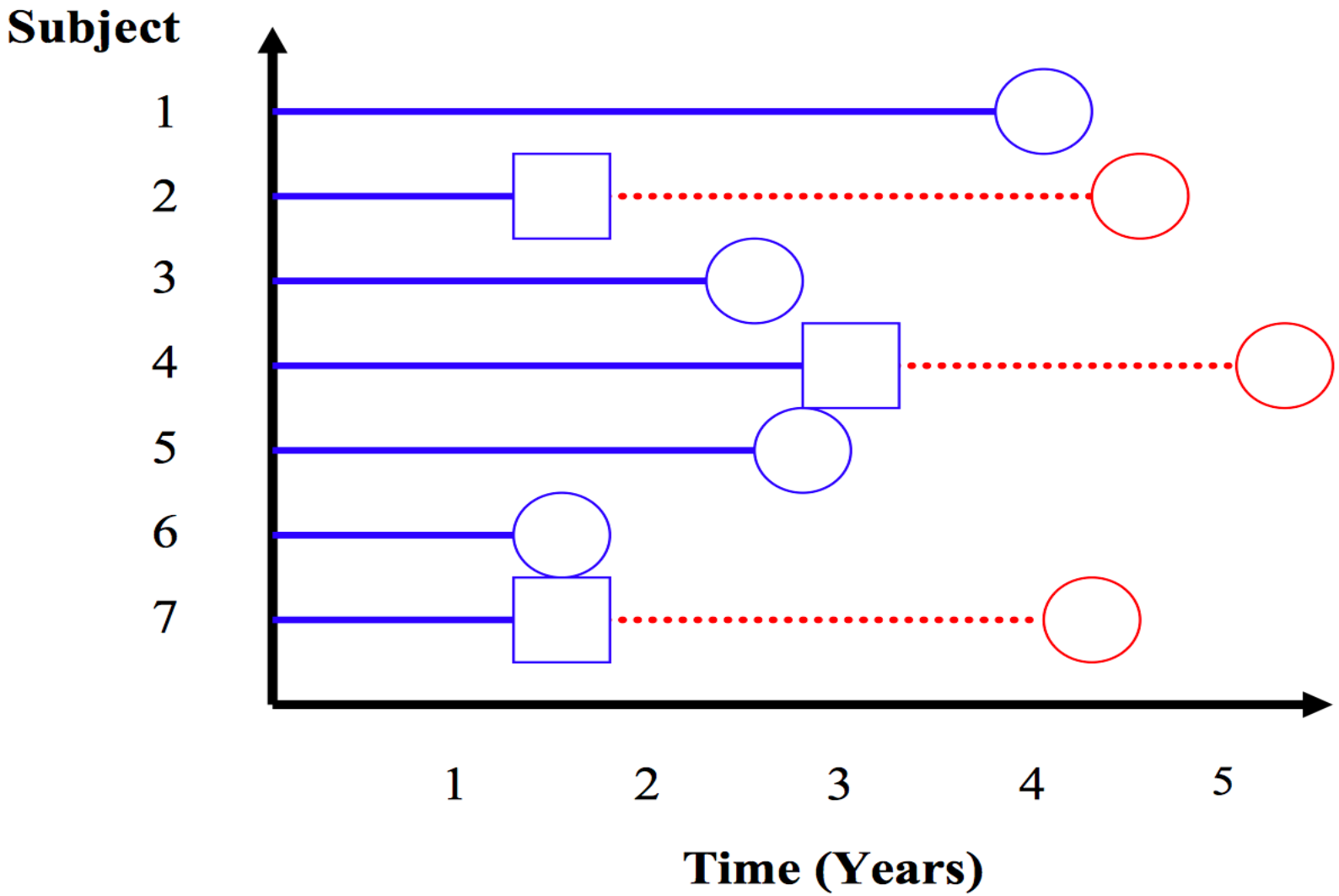
Follow-up Time and Survival Time

- Follow-up time (observed): length of time from a subject's entry into a study until the event of interest or censoring
  - Subjects may have different follow-up times

- Survival time: length of time from a subject's entry into a study until the event of interest



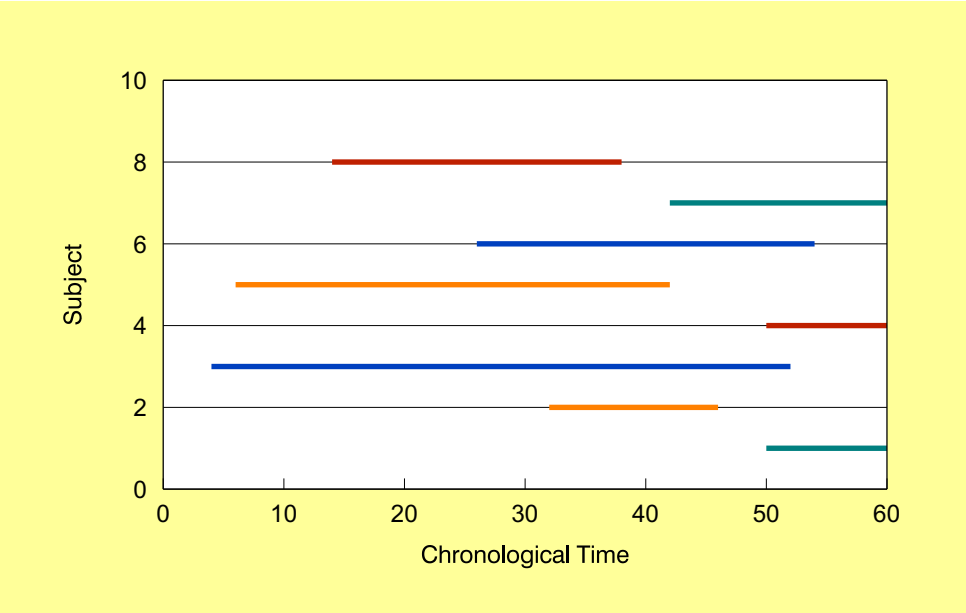
Circle: Event; Square: Censored

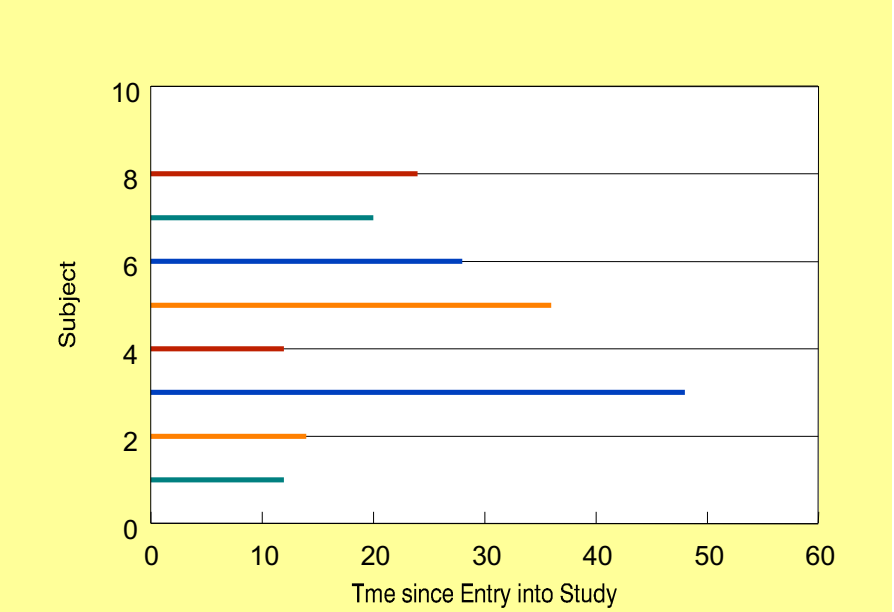


Circle: Event; Square: Censored
Blue Lines: Follow-up time (observed);
Red dotted lines: Time from censoring until the event (unobserved)
Survival times: Blue times for 1,3,4,6; Blue Lines+ Red dotted lines for 2,4,7

Chronological Time and Time since entry

- Example: Clinical trial, late stage cancer patients, patients enrolled as they are diagnosed





Describing Survival Data

- Cumulative Distribution Function: The probability that the time to event is less than or equal to a fixed time t. $F(t) = \Pr(T \leq t)$

- Survival Function $S(t)$: The probability that the time to event is greater than a fixed time t. $S(t) = \Pr(T > t) = 1 - F(t)$

- Hazard Function: instantaneous rate of occurrence of the event , given the individual survival to t.

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}$$

- Cumulative Hazard: $H(t) = \int_0^t h(u)du\,, t > 0$

Mathematical Relationships

- $h(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt}\ln\big(S(t)\big), f(t) = \frac{dF(t)}{dt} = \lim_{\Delta t \to 0} \frac{P(t \leq T < t+\Delta t)}{\Delta t}$

- $H(t) = -\ln(S(t))$

$$H(t) = \int_0^t h(u)du = \int_0^t \frac{f(u)}{S(u)}du = \int_0^t \frac{d\{1 - S(u)\}}{S(u)} = -\log\{S(t)\}$$

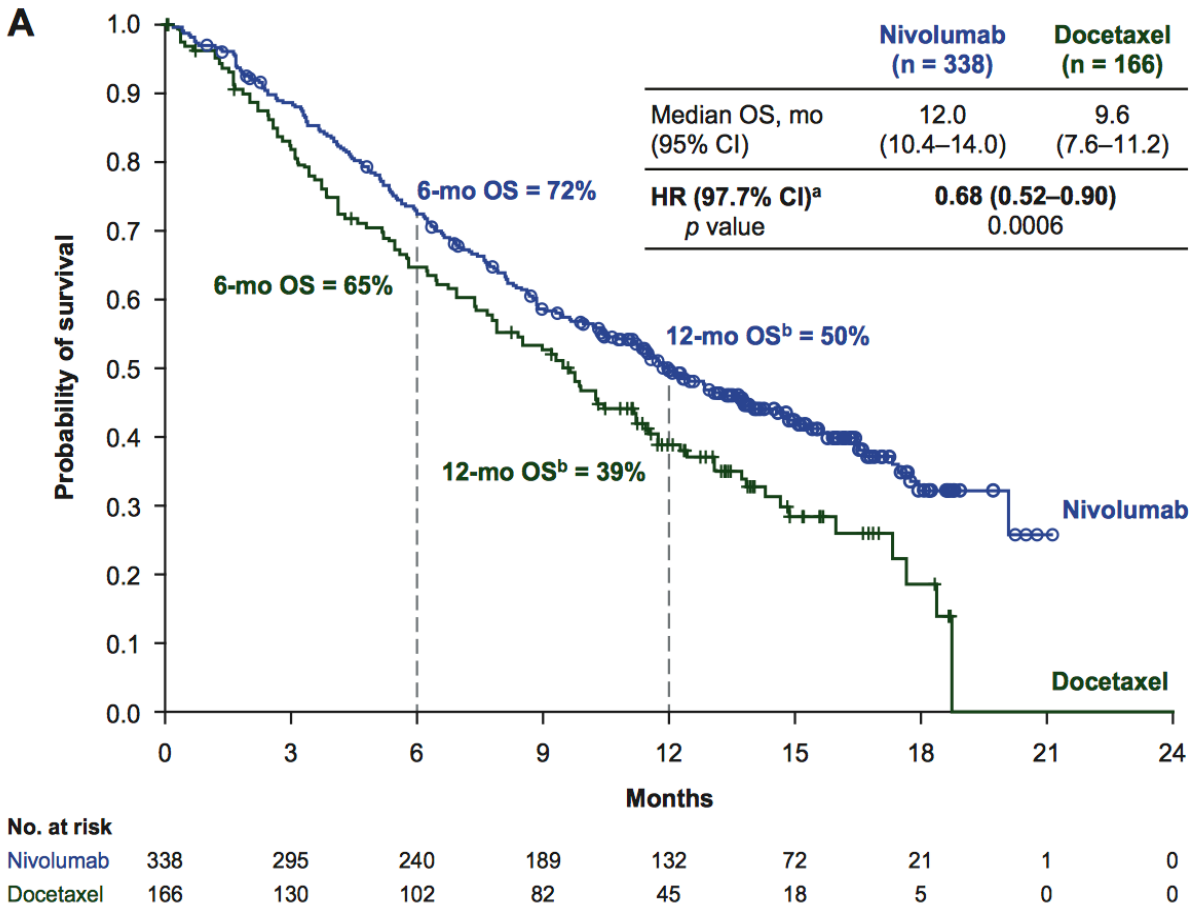
- $\exp\big(-H(t)\big) = S(t)$

Summarize Survival Data

- Calculating mean survival time problematic
  - Mean may not be able to be calculated because of censoring
  - Survival time often not normal, mean less useful
- Usually summarize with median survival time
  - Median survival time may not have been reached in the study
- Can give survival probability at specified times
  - survival probability at 2 years (2 year survival)
  - survival probability at 5 years (5 year survival)

Example in published Phase 3 Clinical Trial[1]

Primary endpoint: OS, defined as the time from randomization to the date of death



	Nivolumab (n = 338)	Docetaxel (n = 166)
Median OS, mo (95% CI)	12.0 (10.4–14.0)	9.6 (7.6–11.2)
HR (97.7% CI)[a]	**0.68 (0.52–0.90)**	
p value	0.0006	

[1]CheckMate 078, Wu et al *Journal of Thoracic Oncology* 2019

Kaplan-Meier Estimator (Product Limit estimator)

- Draws the survival curve, accounting for censoring
- Non-parametric procedure – makes no assumption about the distribution, shape of the survival curve

$$\hat{S}(t) = \prod_{i:t_i \leq t} (1 - \frac{d_i}{n_i})$$

- t_i is the time when at least one event happened, d_i is the number of events occurred at t_i and n_i is the number of individuals who are at risk of the event up to t_i

- At risk set: exclude subjects who have
  - Already had the event
  - Been censored before the time of interest

Life Table Using KM Approach

Time, Years	Number at Risk(ni)	Number of Events(di)	Number Censored	Survival Probability
1	20	1		(1-1/20)=0.950
2	19		1	0.950* (1-0/19)=0.950
3		1		
5		1		
6			1	
9			1	
10			1	
11			1	
12			1	
13			1	
14		1		
17		1	1	
18			1	
…	…	…	…	…

Life Table Using KM Approach

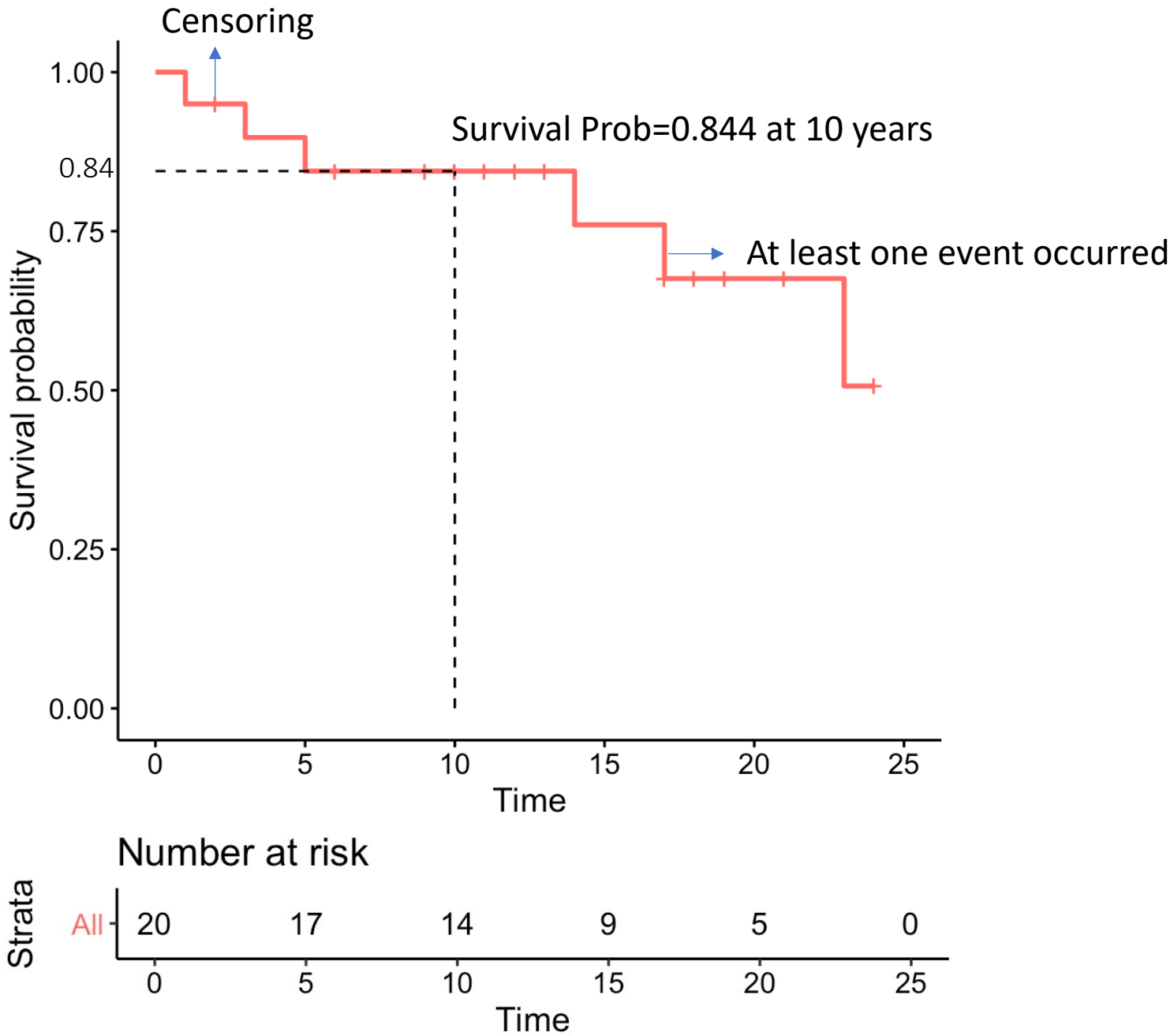
Time, Years	Number at Risk(ni)	Number of Events(di)	Number Censored	Survival Probability
1	20	1		(1-1/20)=0.950
2	19		1	0.950* (1-0/19)=0.950
3	18	1		0.950* ((1-1/18)=0.897
5	17	1		0.897* (1-1/17)=0.844
6	16		1	0.844* (1-0/16)=0.844
9	15		1	0.844
10	14		1	0.844
11	13		1	0.844
12	12		1	0.844
13	11		1	0.844
14	10	1		0.844* (1-1/10)=0.760
17	9	1	1	0.760* ((1-1/9)=0.676
18	7		1	0.676* (1-0/7)=0.676
…	…	…	…	…

KM Table in R

- library(survival)
- Smodel<-
  survfit(Surv(Time,Status)~1,data=mydata)
- summary(Smodel, censored=TRUE)

time	n.risk	n.event	survival
1	20	1	0.950
2	19	0	0.950
3	18	1	0.897
5	17	1	0.844
6	16	0	0.844
9	15	0	0.844
10	14	0	0.844
11	13	0	0.844
12	12	0	0.844
13	11	0	0.844
14	10	1	0.760
17	9	1	0.676
18	7	0	0.676
19	6	0	0.676
21	5	0	0.676
23	4	1	0.507
24	3	0	0.507

KM plot


Censoring
Survival Prob=0.844 at 10 years
At least one event occurred
1.00
0.84
0.75
0.50
0.25
0.00
Survival probability
0
5
10
15
20
25
Time
Number at risk
Strata
All
20
17
14
9
5
0
0
5
10
15
20
25
Time

KM plot

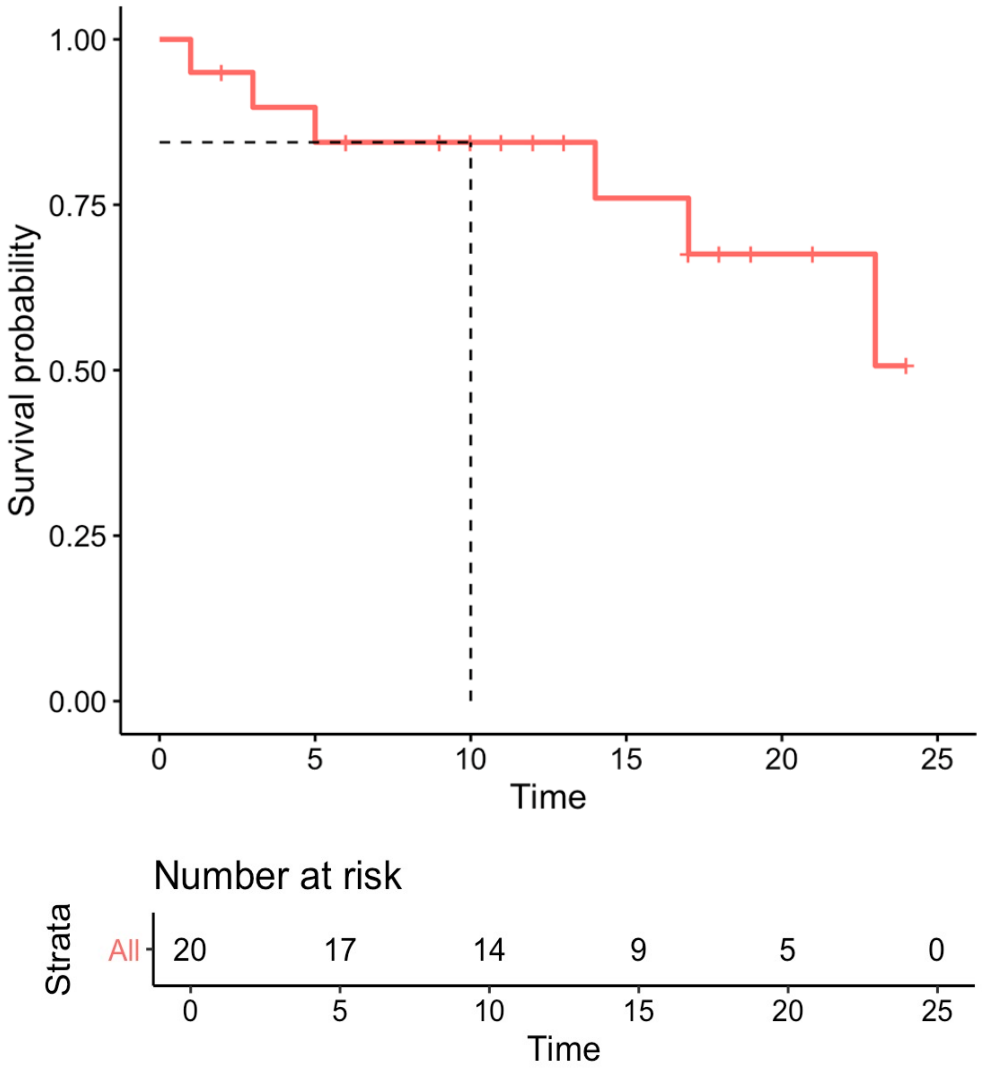
R code

- plot(Smodel)

- library(survminer)
  library(ggplot2)
  ggsurvplot(Smodel,data=mydata, conf.int = FALSE, risk.table = TRUE)

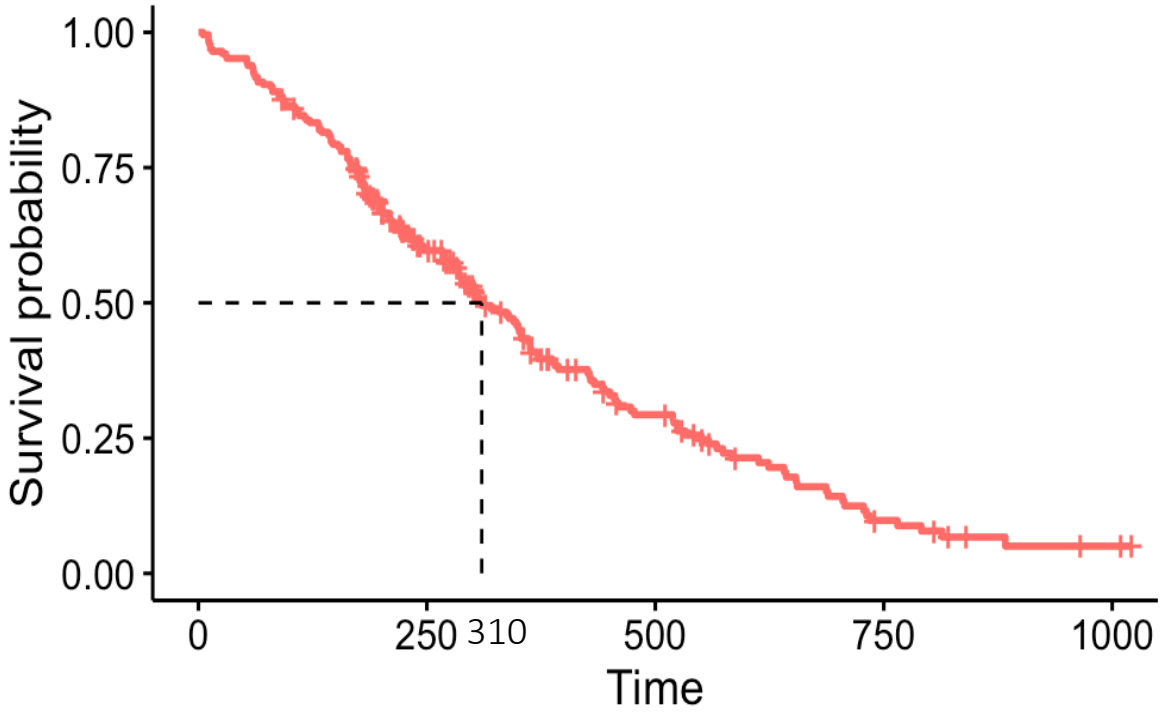
- summary(Smodel,times=c(10))

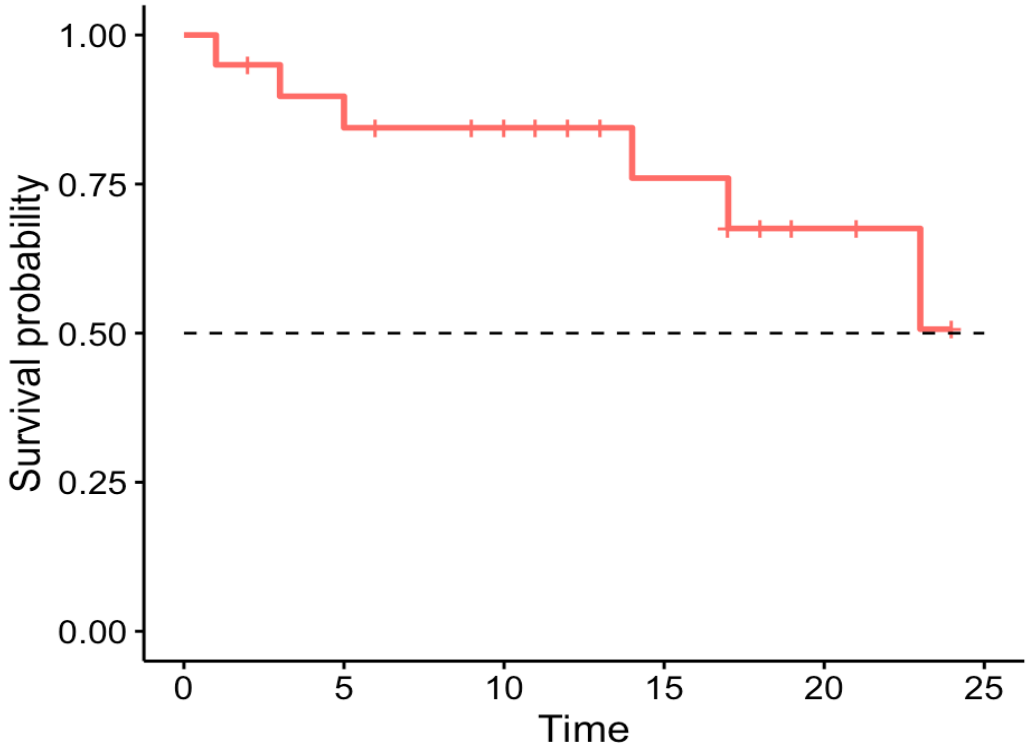
```
time n.risk n.event survival
  10     14       3    0.844
```



Median Survival Time

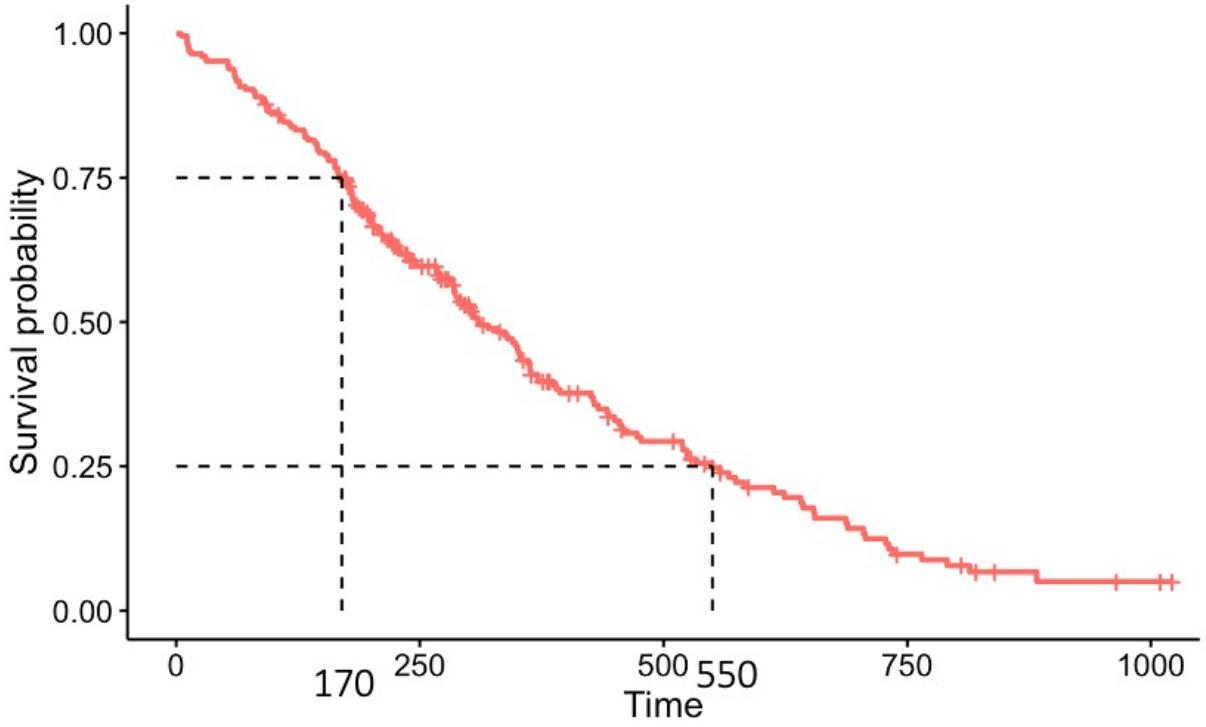
- The smallest time for which the survival probability is equal to 50%, but may not be estimable.
- The time it takes for 50% of patients to have the event
- survfit(Surv(Time,Status)~1,data=mydata)
- ggsurvplot(Smodel,data=mydata,conf.int = FALSE, risk.table = TRUE, surv.median.line = c("hv"))





Percentiles of Survival Time

- 100α percentile: smallest time for which survival probability is equal to 1- α.
- The time it takes for α % of patients to have the event

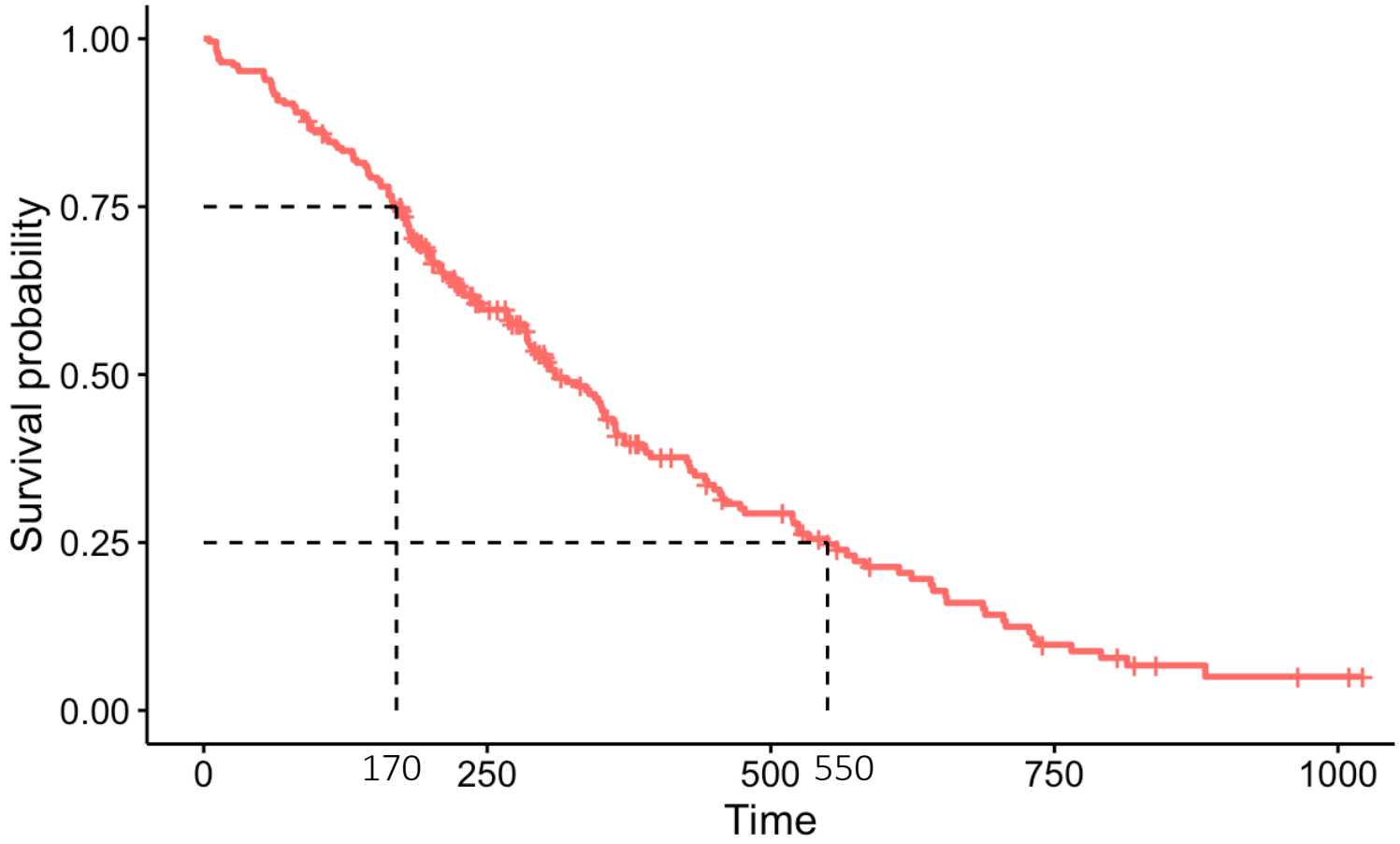


Percentiles of Survival Time

- 100α percentile: smallest time for which survival probability is equal to **1- α**.

quantile(Smodel,probs=c(0.25,0.5,0.75))

25%	50%	75%
170	310	550



Inference for the KM estimator

Greenwood's formula

- Variability in the KM estimate of survival probability

$$\widehat{Var}\left(\hat{S}(t)\right) = \hat{S}(t)^2 \sum_{i:t_i \le t} \frac{d_i}{n_i(n_i - d_i)}$$

- Standard error of KM estimate of survival probability:

$$\widehat{SE}\left(\hat{S}(t)\right) = \hat{S}(t) \sqrt{\sum_{i:t_i \le t} \frac{d_i}{n_i(n_i - d_i)}}$$

Standard Errors of Survival Estimates

Time, Years	Number at Risk	Number of Events	Survival Probability	$\frac{d_i}{n_i(n_i-d_i)}$	$\sum_{i:t_i \le t} \frac{d_i}{n_i(n_i-d_i)}$	$\hat{S}(t)\sqrt{\sum_{i:t_i \le t} \frac{d_i}{n_i(n_i-d_i)}}$
1	20	1	0.950	1/(20*(20-1))=0.0026	0.0026	0.950*$\sqrt{0.0026}$=0.049
2	19		0.950	0.000	0.0026	0.049
3	18	1	0.897	1/(18*(18-1))=0.0032	0.0026+0.0032= 0.0058	0.897*$\sqrt{0.0058}$=0.068
…	…	…	…	…	…	…

SE of KM estimates

- library(survival)
- Smodel<-survfit(Surv(Time,Status)~1,data=mydata)
- summary(Smodel, censored=TRUE)

time	n.risk	n.event	survival	std.err
1	20	1	0.950	0.0487
2	19	0	0.950	0.0487
3	18	1	0.897	0.0689
5	17	1	0.844	0.0826
6	16	0	0.844	0.0826
9	15	0	0.844	0.0826
10	14	0	0.844	0.0826
11	13	0	0.844	0.0826
12	12	0	0.844	0.0826
13	11	0	0.844	0.0826
14	10	1	0.760	0.1093
17	9	1	0.676	0.1256
18	7	0	0.676	0.1256
19	6	0	0.676	0.1256
21	5	0	0.676	0.1256
23	4	1	0.507	0.1740
24	3	0	0.507	0.1740

95% Confidence Interval for $\hat{S}(t)$

- $\hat{S}(t)$ is approximately normal
- 95% CI: [$\hat{S}(t) - 1.96 * \widehat{SE}\left(\hat{S}(t)\right)$, $\hat{S}(t) + 1.96 * \widehat{SE}\left(\hat{S}(t)\right)$]
- survfit(Surv(Time,Status)~1,data=mydata, conf.type="plain")

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
1	20	1	0.950	0.0487	0.854	1.000
2	19	0	0.950	0.0487	0.854	1.000
3	18	1	0.897	0.0689	0.762	1.000
5	17	1	0.844	0.0826	0.682	1.000

- **BUT**
- problematic when $\hat{S}(t)$ is close to 0 or 1
- Could yield values >1 or <0

Complementary log(-log) transformation

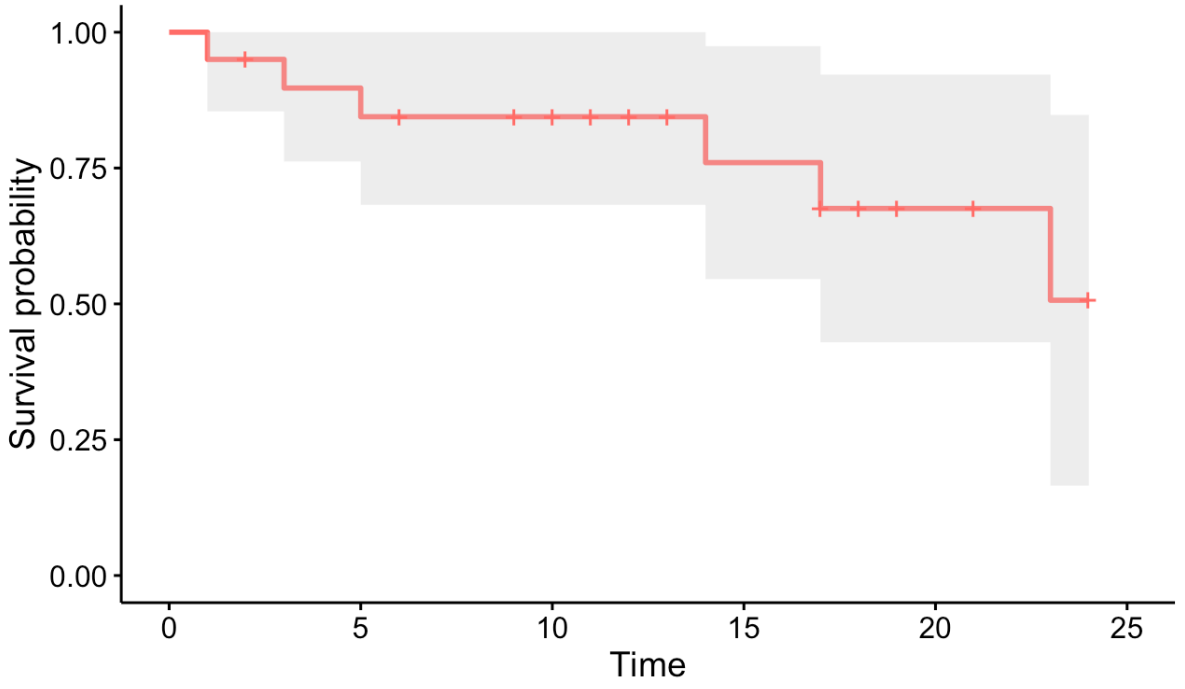
- $\log(-\log(S(t))$: transform $\hat{S}(t)$ on to (-∞, ∞)
- $\eta = \widehat{Var}\left(\log(-\log\left(\hat{S}(t)\right)\right) = \frac{1}{\left[\log\left(\hat{S}(t)\right)\right]^2}\sum_{i:t_i \le t}\frac{d_i}{n_i(n_i-d_i)}$
- 95% CI: $[\hat{S}(t)^{\exp(\,1.96*\sqrt{\eta})}, \hat{S}(t)^{\exp(-1.96*\sqrt{\eta})}]$
- survfit(Surv(Time,Status)~1,data=mydata,conf.type="log-log")

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
1	20	1	0.950	0.0487	0.695	0.993
2	19	0	0.950	0.0487	0.695	0.993
3	18	1	0.897	0.0689	0.648	0.973
5	17	1	0.844	0.0826	0.591	0.947

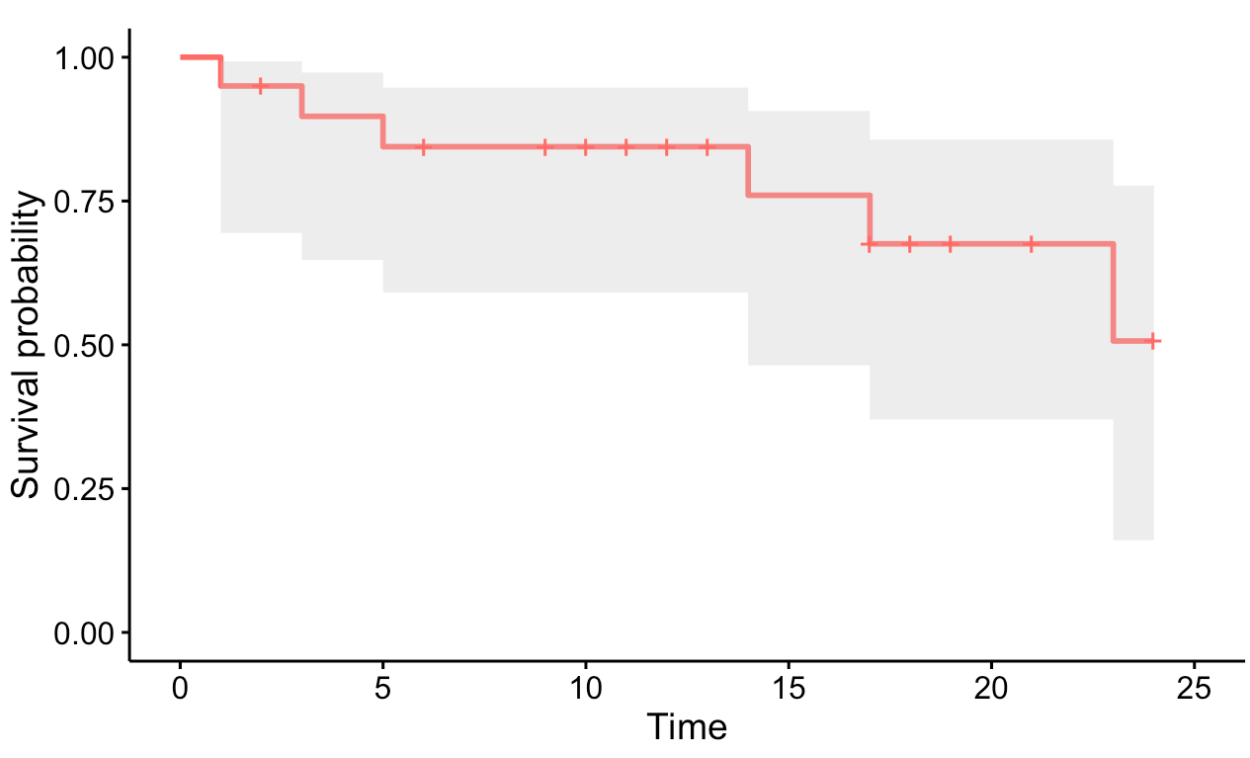
KM plot with 95% CI band

- ggsurvplot(Smodel,data=,mydata,risk.table = TRUE, conf.int=TRUE)

CI with original form:



CI with log-log transformation



Summary of 95% CI for KM estimates

- Original form:
  - Replace negative lower bounds by 0 and upper bounds greater than 1 by 1
  - Not very satisfactory
- log(-log) transformation
  - Complicated, but yields proper bounds
  - Usually used in actual clinical trials

Comparing survival data between two groups

Log-rank test

- Compares survival curves across groups
- Null hypothesis. is survival curves are identical across groups
- Can be used to compare any number of groups
  - Chi-square statistic, k-1 df, where k is the number of groups being compared

Example

```
> survdiff(Surv(time, status) ~ sex,data=lung2)
Call:
survdiff(formula = Surv(time, status) ~ sex, data = lung2)

        N Observed Expected (O-E)^2/E (O-E)^2/V
sex=0 138      112     91.6      4.55      10.3
sex=1  90       53     73.4      5.68      10.3

 Chisq= 10.3  on 1 degrees of freedom, p= 0.001
```

- Example:
  - sex –1 female, 0 male
  - status – event variable (death or not), must be coded 1/0 for event/non-event
  - time – survival time (in days), time to death (for those who had death) or time followed for those censored

- The Chi-Squared test statistic is 10.3 with 1 degree of freedom and the corresponding p-value is 0.001. Since this p-value is less than 0.05, we reject the null hypothesis. In other words, we have sufficient evidence to say that there is a statistically significant difference in survival between the two groups.

Cox Proportional Hazards Regression Model

$$h(t \mid X_1, X_2 \ldots, X_k) = h_0(t) e^{\beta_1 X_1 + \beta_2 X_2 + \ldots + (\beta_k + \beta_{k+1} \cdot t) X_k}$$

- Time is in the baseline hazard: h_0 (t)
- Covariates are in the exponentiated multiplier of the baseline hazard
- The baseline hazard plays the role of the intercept

Example

- Dummy variables for treatment group:
  X = 0 for control and X=1 for experimental
  Let h(t I X = 0) denote hazard at time t for control group
  Let h(t I X = 1) denote hazard at time t for experimental group

Cox model: $h(t|X) = h_0(t) * e^{\beta_1 X}$

where:

$h_0(t) = h(t|X = 0)$ is the baseline hazard function

Assumptions of the model

- The Cox proportional hazards model is called a **semi-parametric model**, because there are no assumptions about the shape of the baseline hazard function. There are however, other assumptions

1) Independent observations randomly drawn from the population

2) Proportional hazards

Proportional hazard model

$$\frac{h(t|X=1)}{h(t|X=0)} = \frac{h_0(t)e^{\beta*1}}{h_0(t)e^{\beta*0}} = \frac{h_0(t)e^{\beta}}{h_0(t) * 1}$$

Hazard ratio

$$\frac{h(t|X=1)}{h(t|X=0)} = e^{\beta}$$

Log hazard ratio

$$\log[\frac{h(t|X=1)}{h(t|X=0)}] = \beta$$

- Hazard ratio is the same over time
- The hazard for individuals with X=1 is a fixed proportion of the hazard of individuals with X=0 (hence “proportional hazards”model)

Proportional hazard model

- More generally, suppose X is any type of variable, and

$$h(t, X) = h_0(t)e^{\sum_{i=1}^{p} \beta_i X_i}$$

- The hazard ratio can be described as

$$HR = \frac{h(t,X^*)}{h(t,X)} = \frac{h_0(t)e^{\sum_{i=1}^{p} \beta_i X_i^*}}{h_0(t)e^{\sum_{i=1}^{p} \beta_i X_i}}$$

$$HR = e^{\sum_{i=1}^{p} \beta_i (X_i^* - X_i)}$$

Hypothesis testing

Hazard ratio scale

$$\frac{h(t|X=1)}{h(t|X=0)} = e^{\beta}$$

Log hazard ratio scale

$$\log[\frac{h(t|X=1)}{h(t|X=0)}] = \beta$$

- Hypothesis testing for treatment difference:

H_0: HR = 1

vs

H_1: HR ≠ 1

H_0: β = 0

vs

H_1: β ≠ 0

- Interpretation of hazard ratios and log-hazard ratios in Cox models:
  - **HR<1 (or β<0)**: lower hazard of the event compared to reference group
  - **HR>1 (or β>0):** larger hazard of the event compared to reference group

Partial Likelihood

- Likelihood function for Cox models can be factored into:
  - One part that depends on $h_0(t)$ and β, which contains relatively little information about β
  - The other part that depends on β alone (partial likelihood)

- Estimate values of β that maximize partial likelihood

Example

- Dependent variable specified as time_variable, event_variable
- Event variable must be coded 1/0 for event/no event
- The hazard of death for females is 0.588 times the hazard for males (95% CI 0.424, 0.816). The p value is 0.00149<0.05, which indicates that the difference in hazard between males and females is statistically significant.

```
> res.cox <- coxph(Surv(time, status) ~ sex, data = lung2)
> summary(res.cox)
Call:
coxph(formula = Surv(time, status) ~ sex, data = lung2)

  n= 228, number of events= 165

       coef exp(coef) se(coef)      z Pr(>|z|)
sex -0.5310    0.5880   0.1672 -3.176  0.00149 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

    exp(coef) exp(-coef) lower .95 upper .95
sex     0.588      1.701    0.4237     0.816
```

Check proportional hazard assumption

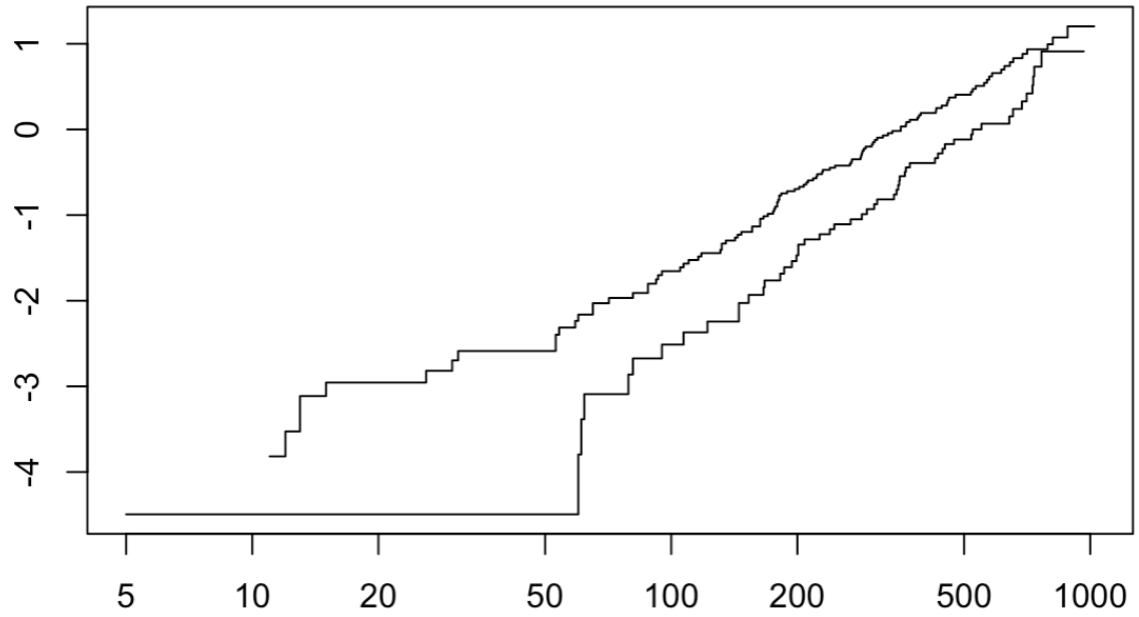
- In K-M analysis, plot the complementary log-log vs ln t and look for approximate parallel lines. The complementary log-log is ln[-ln S(t)]

- In PH regression, plot Schoenfeld residuals by time. Under PH, residuals are uncorrelated; mean is 0; plot should give random scatter around 0 over time
  - Test Schoenfeld residuals for pattern over time using zph test

Checking the PH assumption

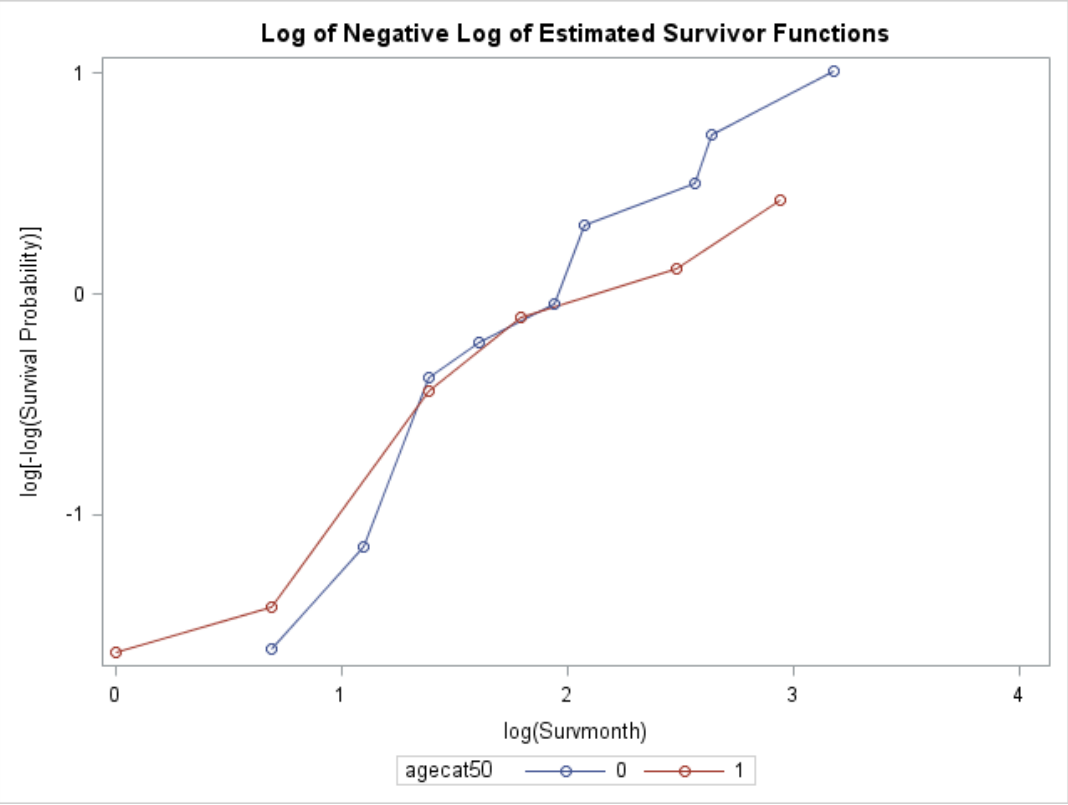
- Plot log(-log(S(t))) for K-M curves
  - Parallel lines indicate proportional hazards

```
> fit<-survfit(Surv(time, status) ~ sex, data = lung2)
> plot(fit,fun="cloglog")
```



Checking the PH assumption

- Plot log(-log(S(t))) for K-M curves
  - Parallel lines indicate proportional hazards



Checking the PH assumption

- Null hypothesis: no correlation between Schoenfeld residuals and time
- Schoenfeld Residuals
  - A Schoenfeld residuals are the *differences* between that individual's covariate values at the event time and the corresponding risk-weighted average of covariate values among all those then at risk.
  - Grambsch and Therneau (1994) showed that the scaled Schoenfeld residual measures the deviation of a time-dependent log hazard ratio $\beta(t)$ from time-constant log hazard ratio β.
  - If model correct (PH holds) residuals should be randomly scattered about 0, no pattern over time
  - A plot that shows a non-random pattern against time is evidence of violation of the PH assumption

Example

Stanford Heart Transplant Data

- Survival of patients on the waiting list for the Stanford heart transplant program.

start, stop, event:	Entry and exit time and status for this interval of time
age:	age-48 years
year:	year of acceptance (in years after 1 Nov 1967)
surgery:	prior bypass surgery 1=yes
transplant:	received transplant 1=yes
id:	patient id

R: Testing for Non-proportional Hazards

```
> jmodel <- coxph(Surv(start,stop,event)~age+year+
+                   transplant+surgery, data=heart)
> summary(jmodel)
Call:
coxph(formula = Surv(start, stop, event) ~ age + year + transplant +
    surgery, data = heart)

  n= 172, number of events= 75

               coef exp(coef) se(coef)      z Pr(>|z|)
age          0.02717   1.02754  0.01371  1.981   0.0476 *
year        -0.14635   0.86386  0.07047 -2.077   0.0378 *
transplant1 -0.01025   0.98980  0.31375 -0.033   0.9739
surgery     -0.63721   0.52877  0.36723 -1.735   0.0827 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

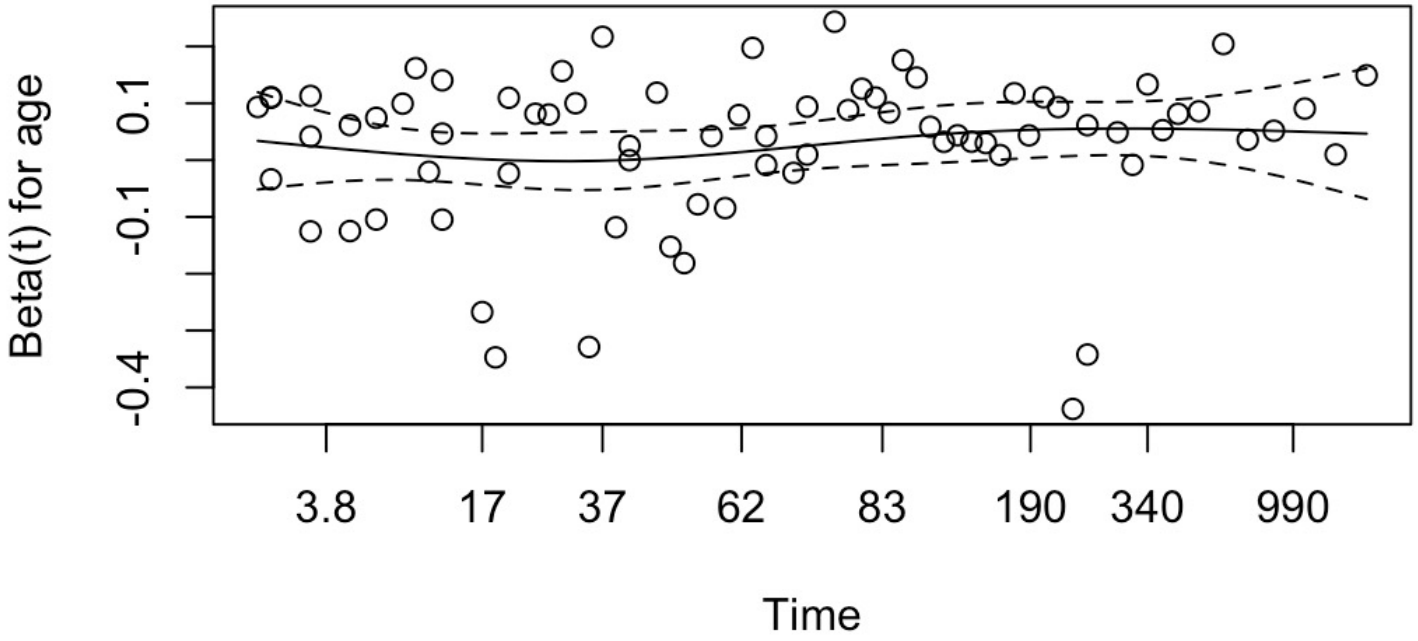
```
> Phtest <- cox.zph(jmodel,global=TRUE)
> Phtest
             chisq df    p
age         0.8669  1 0.35
year        1.7164  1 0.19
transplant 0.2029  1 0.65
surgery     0.0974  1 0.75
GLOBAL      3.4731  4 0.48
```

- The test is not statistically significant for each of the covariates. These variables do not violate the PH assumption

R: Testing for Non-proportional Hazards

- Schoenfeld residuals by time

```
> plot(Phtest[1])
```



- Flat line at 0 indicates PH assumption holds
- No signal of violation for PH assumption.

How to handle non-proportional hazards

- Background covariate, not primary interest
  - Stratify by covariate

- Covariate of interest
  - Include time*covariate interaction
  - Separate analyses based on time

Example

- Patients diagnosed with a particular disease are randomized either to a new treatment or standard of care, and followed until their disease progresses or they are censored.

- IntervGroup – intervention, coded 1 for those receiving the new treatment, 0 for those receiving standard of care
- SexF – coded 1 for females, 0 for males
- Survmonth – time to progression of disease or censoring, in months
- Event – coded 1 for those with disease progression, 0 for those with no progression of disease (censored)
- age50 variable that is coded 0 for those with age<50, 1 for those with age >= 50.

Example

```
> ph.out <- coxph(Surv(Survmonth,Event) ~ IntervGroup + SexF + age50,data=datset)
> summary(ph.out)
Call:
coxph(formula = Surv(Survmonth, Event) ~ IntervGroup + SexF +
    age50, data = datset)

  n= 50, number of events= 36

                coef exp(coef) se(coef)      z Pr(>|z|)
IntervGroup -2.03472   0.13072  0.47378 -4.295 1.75e-05 ***
SexF         0.09765   1.10258  0.37331  0.262    0.794
age50        0.16624   1.18086  0.37169  0.447    0.655
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
> PH.test <- cox.zph(ph.out,global=TRUE)
> PH.test
               chisq df     p
IntervGroup 0.009182  1 0.924
SexF        0.000605  1 0.980
age50       3.331450  1 0.068
GLOBAL      4.198150  3 0.241
```

The hazard of disease progression for those in intervention group is 0.131 times the hazard of disease progression for those in control group, adjusting for other variables. The hazard difference is statistically significant as p value<0.001. The test for PH assumption shows the p value for age50=0.068, which is around the borderline significance, suggesting (but not showing significant evidence because it >0.05) that the PH assumption may be violated here.

Stratified PH Analysis : **Variables Not of Interest**

- This allows the baseline hazard to differ for each stratum, while assuming the other non-strata variables follow the PH assumption
- No HR for the stratified variable

```
> ph.out2 <- coxph(Surv(Survmonth,Event)  ~ IntervGroup + strata(age50) + SexF,data=datset)
> summary(ph.out2)
Call:
coxph(formula = Surv(Survmonth, Event) ~ IntervGroup + strata(age50) +
    SexF, data = datset)

  n= 50, number of events= 36

               coef exp(coef) se(coef)      z Pr(>|z|)
IntervGroup -2.0127    0.1336   0.4766 -4.223 2.41e-05 ***
SexF         0.1308    1.1397   0.3784  0.346     0.73
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

The hazard of disease progression for those in intervention group is 0.134 times the hazard of disease progression for those in control group, adjusting for sex and stratified by age. the hazard difference is statistically significant as p value<0.001.

Interaction with time

- Include time-dependent variable and deal with complicated conclusion of HR varying over time
- Interpretation of association is more complex
- tt( ) is a function, input a variable and the function calculates that variable times ln(time)

```
> ph.out3 <- coxph(Surv(Survmonth,Event) ~ IntervGroup + age50 + SexF + tt(age50),
+                  tt=function(x,t,...) x*log(t),data=datset)
> summary(ph.out3)
Call:
coxph(formula = Surv(Survmonth, Event) ~ IntervGroup + age50 +
    SexF + tt(age50), data = datset, tt = function(x, t, ...) x *
    log(t))

  n= 50, number of events= 36

               coef exp(coef) se(coef)      z Pr(>|z|)
IntervGroup -2.0474    0.1291   0.4758 -4.303 1.68e-05 ***
age50        1.2659    3.5462   0.7121  1.778   0.0755 .
SexF         0.1188    1.1261   0.3706  0.320   0.7486
tt(age50)   -0.8138    0.4432   0.4405 -1.847   0.0647 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Interaction with time

```
> ph.out3 <- coxph(Surv(Survmonth,Event) ~ IntervGroup + age50 + SexF + tt(age50),
+                  tt=function(x,t,...) x*log(t),data=datset)
> summary(ph.out3)
Call:
coxph(formula = Surv(Survmonth, Event) ~ IntervGroup + age50 +
    SexF + tt(age50), data = datset, tt = function(x, t, ...) x *
    log(t))

  n= 50, number of events= 36

               coef exp(coef) se(coef)      z Pr(>|z|)
IntervGroup -2.0474    0.1291   0.4758 -4.303 1.68e-05 ***
age50        1.2659    3.5462   0.7121  1.778   0.0755 .
SexF         0.1188    1.1261   0.3706  0.320   0.7486
tt(age50)   -0.8138    0.4432   0.4405 -1.847   0.0647 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

The p value of the interaction term in 0.0754>0.05 is of borderline significance, suggesting (but not showing significant evidence because it >0.05) that the PH assumption may be violated here.

The HR for age is decreasing with time because the coefficient of the interaction term is negative.