Survival Analysis

Kaplan-Meier

Table Of Content

Introduction

- Survival analysis
  - analyzing longitudinal data on the occurrence of events
    - death, injury, onset of illness, recovery from illness (binary variables) or transition above or below the clinical threshold of a meaningful continuous variable (e.g. CD4 counts).
  - randomized clinical trial or cohort study design.

Introduction (2)

- Objectives
  - Estimate time-to-event for a group of individuals (KM)
  - To compare time-to-event between two or more groups (LR)
  - To assess the relationship of co-variables to time-to-event (Cox)

Regression vs. Survival Analysis

Technique	Mathematical model	Yields
Linear Regression	$Y=B_1X + B_o$ (linear)	Linear changes
Logistic Regression	$Ln(P/1-P)=B_1X+B_o$ (sigmoidal prob.)	Odds ratios
Survival Analyses	$h(t) = h_o(t)exp(B_1X+B_o)$	Hazard rates

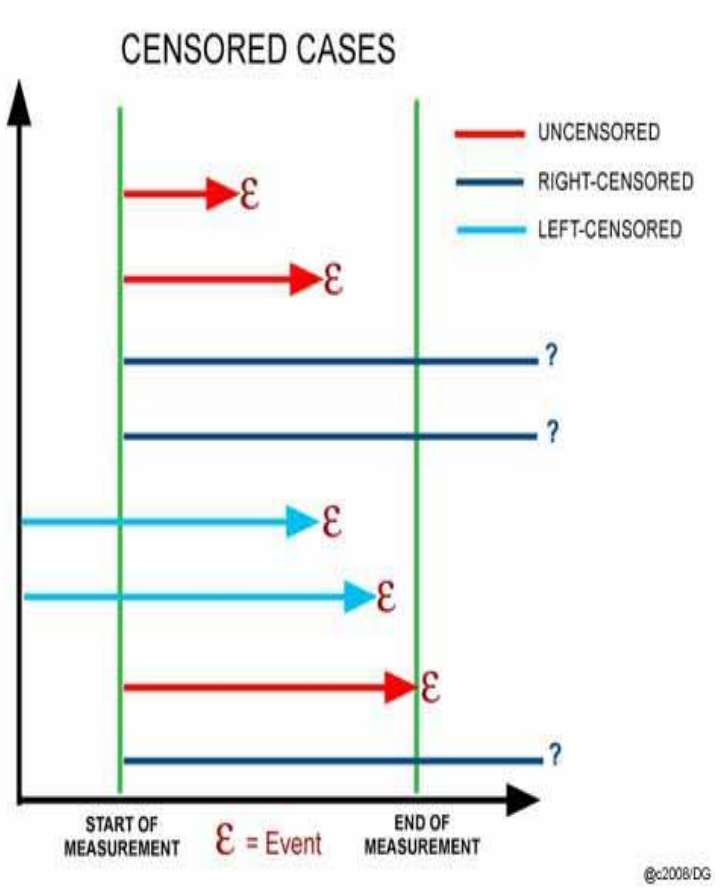
Gage, 2004

Basic Concepts

- Why do we use survival analysis
  - Why not compare mean time-to-event between your groups using a t-test or linear regression?
    - Ignores censoring
  - Why not compare proportion of events in your groups using risk/odds ratios or logistic regression?
    - Ignores time
  - Time-to-event: The time from entry into a study until a subject has a particular outcome

Basic Concepts (2)

- Censored
  - the critical event has not yet occurred
  - lost to follow-up
  - other interventions offered
  - event occurred but unrelated cause



http://faculty.chass.ncsu.edu

Basic Concepts (3)

- For categorical predictor variables eg. drug vs. placebo, drug doses 0, 20, 50, and 100 mg/day.
- Assumptions
  - Independence within the sample
  - Independence of censoring
  - Uniformity within time interval
  - Less censored values
  - Enough sample size
  - No pattern seen in plots of data

(PROPHET StatGuide)

Example

- We have 25 patient with lung cancer and we want to know the survival time after treatment X.
- Eg censored case in this study : One patient recently enrolled 1 month ago, one patient did not return to our study after 5 months, one patient died due to other cause after 13 months, one patient still alive after finish study at 44 months. Only 10 subjects died from lung cancer in our study.
- In dataset 0= censored, 1= died

Example

- Dataset

SN	Survival Time	Status
1	1	0
2	5	0
3	6	1
4	6	1
5	9	0
6	10	1
7	10	0
8	10	1
9	12	1
10	12	1
11	12	1
12	12	1

SN	Survival Time	Status
13	12	1
14	13	0
15	15	0
16	16	0
17	20	0
18	24	1
19	24	0
20	27	0
21	32	1
22	34	1
23	36	1
24	36	1
25	44	0

SPSS Output

Case Processing Summary

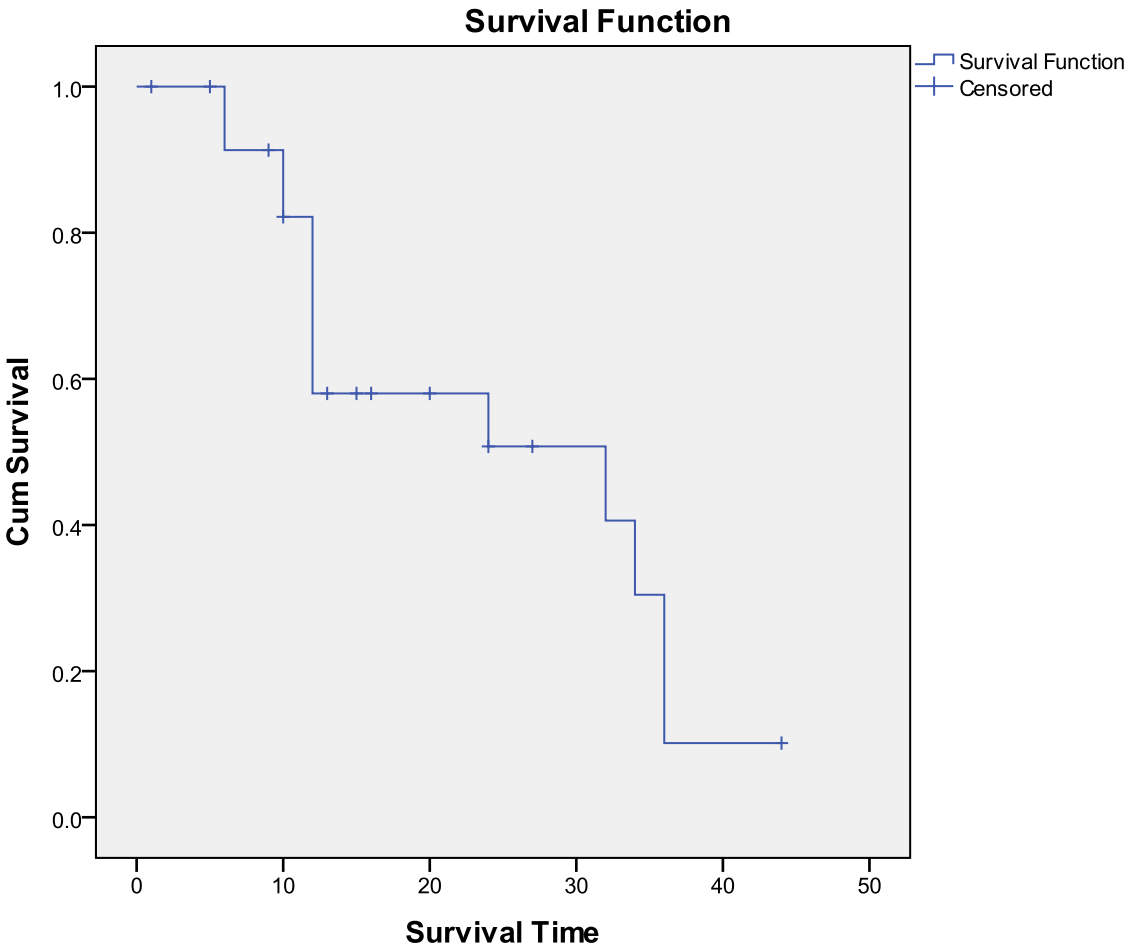
Total N	N of Events	Censored	
		N	Percent
25	14	11	44.0%

Means and Medians for Survival Time

Mean[a]				Median			
		95% Confidence Interval				95% Confidence Interval	
Estimate	Std. Error	Lower Bound	Upper Bound	Estimate	Std. Error	Lower Bound	Upper Bound
24.550	2.920	18.826	30.274	32.000	14.909	2.779	61.221

a. Estimation is limited to the largest survival time if it is censored.

SPSS Output


Survival Function
Survival Function
Censored
Cum Survival
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
50
Survival Time

Interpretation

- 25 subjects with lung cancer enrolled in the study and was treated with treatment X. The follow up ranged from 1 to 44 months. The censored rate was 44%. The median survival for the subjects who had this treatment was 32 months (CI = 2.78 – 61.22)

Kaplan Meier Technique			
	Survival Time	Standard Error	95% CI
Mean	24.55	2.92	18.83,30.27
Median	32.00	14.91	2.78, 61.22

Next Example

- Is there any difference between treatment X and Y in term of shorter time of death?
- We have data for treatment Y

Case Processing Summary

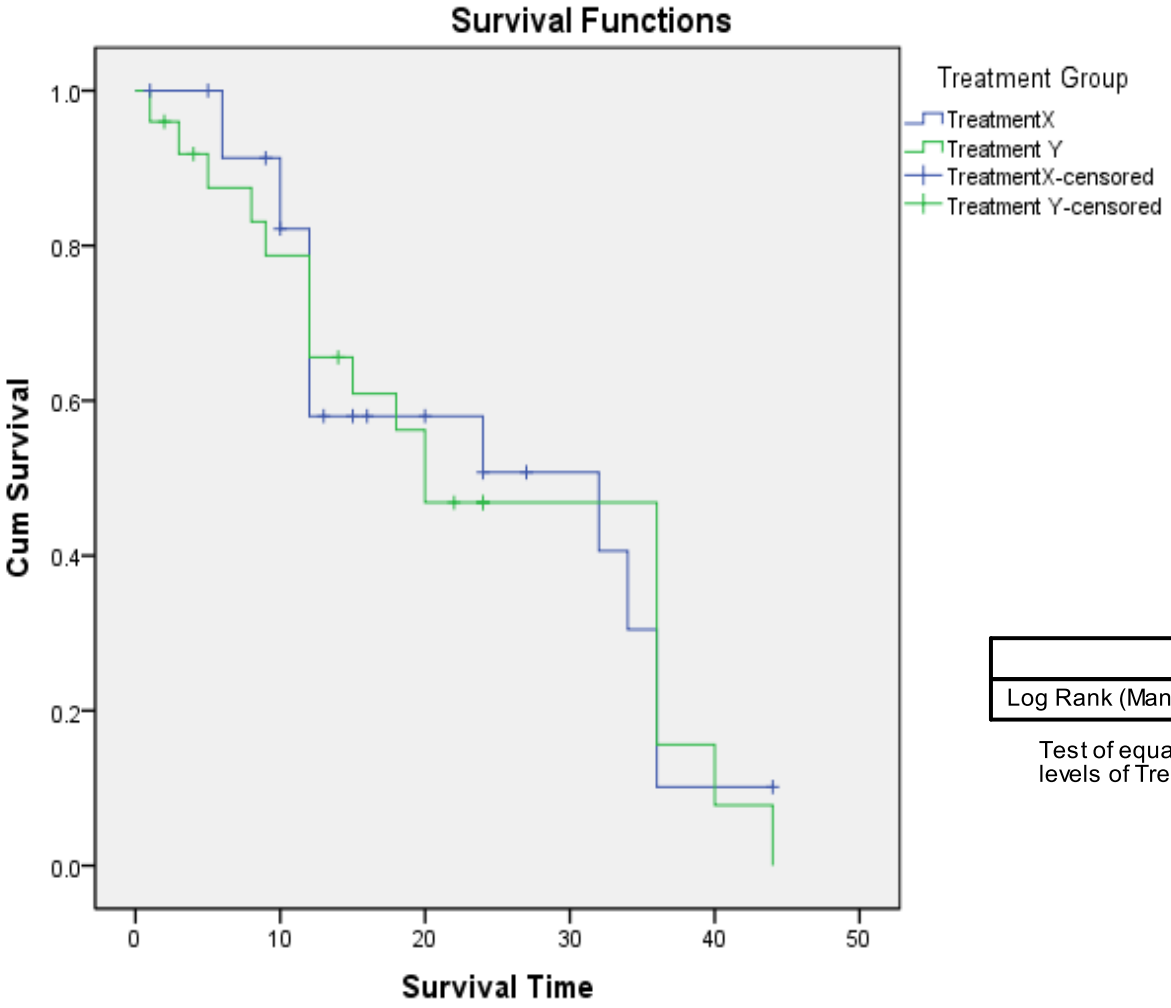
Treatment Group	Total N	N of Events	Censored	
			N	Percent
TreatmentX	25	14	11	44.0%
Treatment Y	25	18	7	28.0%
Overall	50	32	18	36.0%

Means and Medians for Survival Time

Treatment Group	Mean[a]				Median			
			95% Confidence Interval				95% Confidence Interval	
	Estimate	Std. Error	Lower Bound	Upper Bound	Estimate	Std. Error	Lower Bound	Upper Bound
TreatmentX	24.550	2.920	18.826	30.274	32.000	14.909	2.779	61.221
Treatment Y	23.924	3.065	17.918	29.931	20.000	4.702	10.783	29.217
Overall	24.234	2.124	20.071	28.396	24.000	8.545	7.253	40.747

a. Estimation is limited to the largest survival time if it is censored.

SPSS Output



Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	.032	1	.858

Test of equality of survival distributions for the different levels of Treatment Group.

Interpretation

- 50 subjects with lung cancer enrolled in this study. 25 of them had treatment X and another half had treatment Y. The censored rate was 36%. The follow up range from 1 to 44 months.

Kaplan Meier Technique			
Treatment X	Survival Time	Standard Error	95% CI
Mean	24.55	2.92	18.83,30.27
Median	32.00	14.91	2.78, 61.22

Interpretation

Kaplan Meier Technique			
Treatment Y	Survival Time	Standard Error	95% CI
Mean	23.92	3.07	17.92,29.93
Median	20.00	4.70	10.78, 29.22

- Log rank test was used to compare the two groups. There is no significant difference on having shorter time to event(death) between treatment X and treatment Y groups (χ = 0.032, df = 1, p > 0.05)

*One common misconception
of survival analysis is that some researchers interpret the result as one group being more likely to have deaths (this should be given by logistic regression!).

References

- PROPHET StatGuide: Do your data violate Kaplan-Meier assumptions? Retrieved November 20th, http://www.basic.northwestern.edu/statguidefiles/kaplan_ass_viol.html
- Chan, Y. H. (2004). Biostatistics 203. Survival analysis. Singapore Med J, Vol 45(6), 249