Medical Imaging
UNIVERSITY OF TORONTO

Surviving Time to Event Analysis: Why Am I Censored and who is Kaplan-Meier?

presented by:

Pascal N. Tyrrell
Tristal (Yanlin) Li
Ernest (Khashayar) Namdar

I have no idea where to start,
but I have heard survival analysis is cool!

Keywords: Risk, Prognosis

The aim is to predict the risk for future

$$Risk\ Score = c_1 \times f(predictor_1) + c_2 \times f(predictor_2) + \ldots + c_N \times f(predictor_N)$$

$$f = \begin{cases} x \\ \ln(x) \\ \log(x) \end{cases}$$

$Note: We\ can\ also\ add\ interaction\ terms: c_{N+1} \times f(predictor_1 \times predictor_2) + \ldots$

How should the risk models be evaluated?

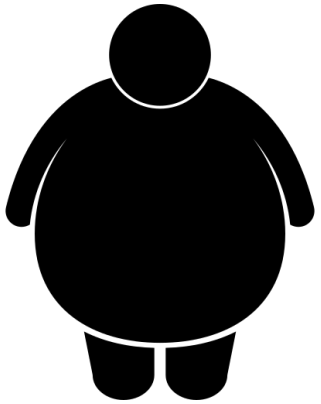
Two key questions:

Is the patient dead? binary

How much is the risk? continues

Patient 1

Reliable Model: Risk 1 < Risk 2



Patient 2

C-Index

Consider **pairs** of patients

Permissible Pair

Patient I has a worse outcome & higher risk	>>>	The pair is concordant	+1
Patient I has a worse outcome & lower risk	>>>	The pair is not concordant	0
Patient I has a worse outcome & equal risk	>>>	Tie in Risk	+0.5
Patient I has the same outcome	>>>	Tie in Outcome	

$$C - Index = \frac{N_{concordant} + 0.5 \times N_{risk\ ties}}{N_{permissible}}$$

Survival Models

$$S(t) = \prod_{0}^{t} 1 - \frac{N_{died\ at\ i}}{N_{known\ to\ survivived\ to\ i}}$$

Probability that patients survive until t

$$S(t) = \Pr(T > t)$$

Patient	Time of Death
1	30
2	60
3	90
4	120
5	120
6	180+
7	210
8	210+

Censored Data

$S(0) = \frac{8}{8} = 1$

$S(30) = \frac{7}{8} = 0.88$

$S(60) = \frac{6}{8} = 0.75$

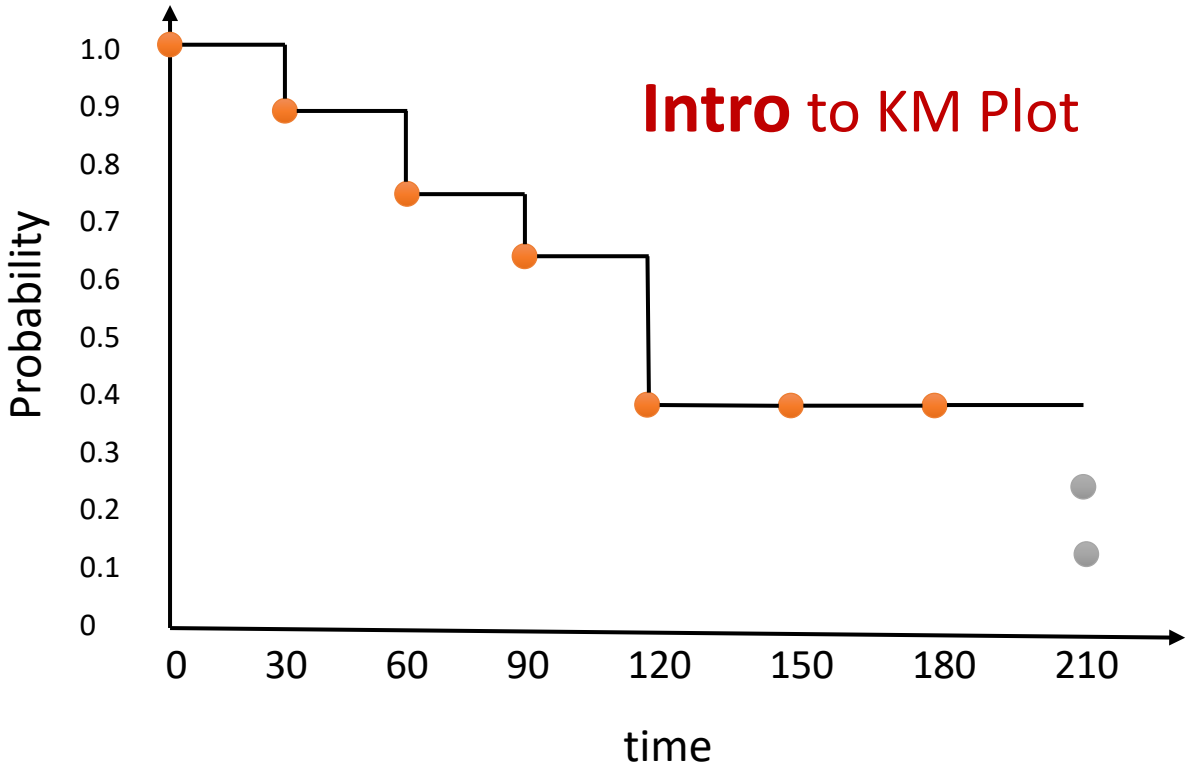
$S(90) = \frac{5}{8} = 0.63$

$S(120) = \frac{3}{8} = 0.38$

$S(150) = \frac{3}{8} = 0.38$

$S(180) = \frac{3}{8} = 0.38$

$S(210) = ?$



Plan for the Day

Python

R

https://drive.google.com/drive/folders/1Wuk5c7Apu1PZ1_OVeUkRoZaIZX8qQj_V?usp=share_link

Dataset

"These are data from one of the first successful trials of adjuvant chemotherapy for colon cancer."

id	study	rx	sex	age	obstruct	perfor	adhere	nodes	status	differ	extent	surg	node4	time	etype
1	1	Lev+5FU	1	43	0	0	0	5.0	1	2.0	3	0	1	1521	2
1	1	Lev+5FU	1	43	0	0	0	5.0	1	2.0	3	0	1	968	1
2	1	Lev+5FU	1	63	0	0	0	1.0	0	2.0	3	0	0	3087	2
2	1	Lev+5FU	1	63	0	0	0	1.0	0	2.0	3	0	0	3087	1
3	1	Obs	0	71	0	0	1	7.0	1	2.0	2	0	1	963	2

Source: R Package 'survival'

We will load the csv file to start from absolute scratch

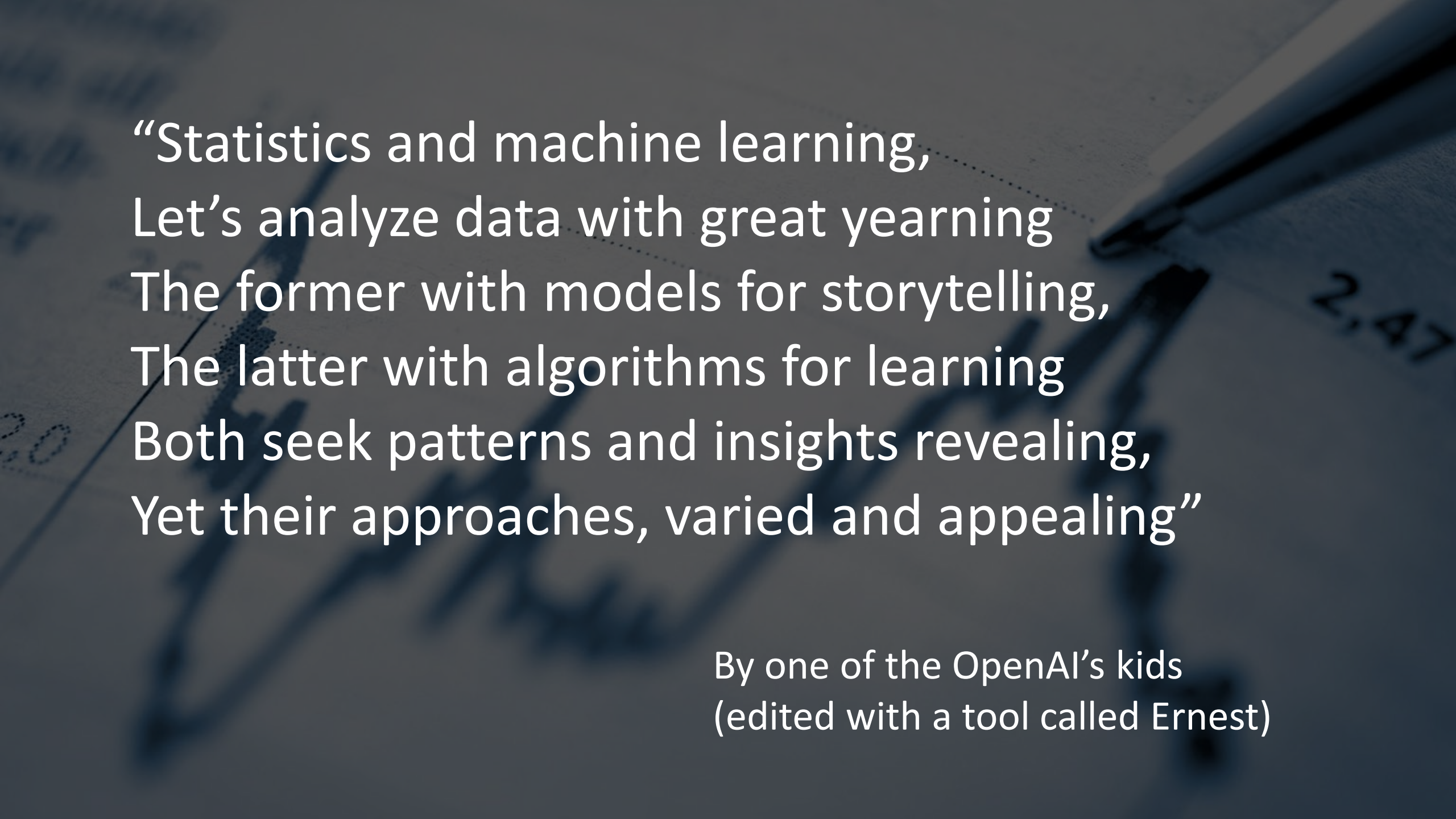
https://github.com/knamdar/R_survial_datasets

status: censoring status

etype: event type
1=recurrence,**2=death**

rx: treatment

time: days until event or censoring



“Statistics and machine learning,
Let’s analyze data with great yearning
The former with models for storytelling,
The latter with algorithms for learning
Both seek patterns and insights revealing,
Yet their approaches, varied and appealing”

By one of the OpenAI’s kids
(edited with a tool called Ernest)