SAS® **GLOBAL FORUM** 2018

USERS PROGRAM

Using Cox Proportional Hazard Model To Predict Failure: Practical Applications in Multiple Scenarios

Sudeep Kunhikrishnan

April 8 – 11 | Denver, CO

#SASGF

Using Cox Proportional Hazard Model To Predict Failure: Practical Applications in Multiple Scenarios

Sudeep Kunhikrishnan

ABSTRACT

- This presentation focuses on business applications of survival analysis – using Cox Proportional Hazard Modeling in multiple scenarios
- Survival analysis
    - defined as a set of methods for analyzing data where the outcome variable is the time until the occurrence of an event of interest
    - helps in estimating time to event for a group of individuals, between two or more groups, and assess the relationship of co-variables to time-to-event.
    - Advantages over t-test or regression and logistic regression
- The model results can be used in these scenarios as we model the effect of predictors and covariates on the hazard rate but leaves the baseline hazard rate unspecified. There is no assumption or knowledge of absolute risk requirement and the model can be used on the total population of study. This also helps in ease of comparison between different groups.
- Sample results and illustration discussed

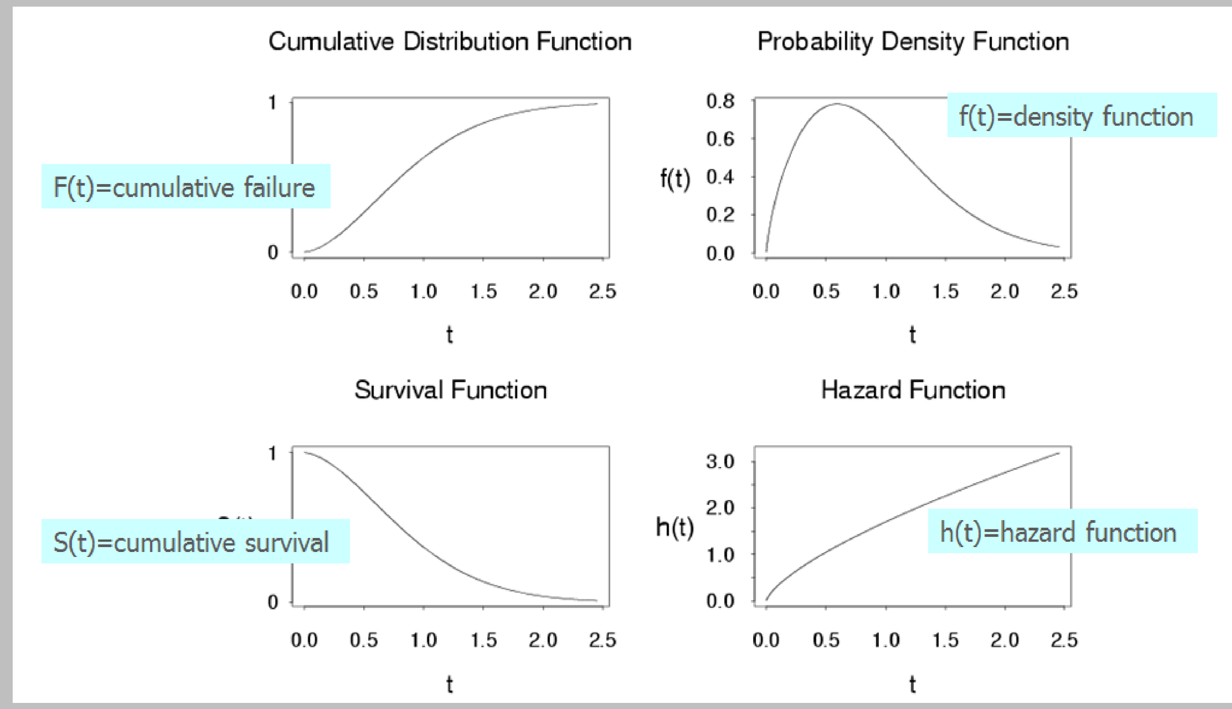
INTRODUCTION TO SURVIVAL ANALYSIS

- Survival analysis is used to model factors that influence the time to an event
- Cox Proportional Hazard Regression Model allowed us to identify factors that increased likelihood of the event happening as compared to baseline parameters we select
    - **Estimate time-to-event for a group of individuals**
    - **To compare time-to-event between two or more groups**
    - **To assess the relationship of co-variables to time-to-event**
    - Parametric, Semi-parametric and Non-parametric Models
    - *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

Note: expected time-to-event = 1/incidence rate

COX PROPORTIONAL HAZARD MODEL

- Semi – parametric (also called proportional hazards regression)
- Models the effect of predictors and covariates on the hazard rate but leaves the baseline hazard rate unspecified.
- Do NOT assume knowledge of absolute risk - no base hazard rate
- Can use total population & estimates relative risk - ease of comparison
- Key Terms: Time to Event and Censoring



SAS PROCS

- **LIFETEST** - Produces life tables and Kaplan-Meier survival curves. Is primarily for univariate analysis of the timing of events.
- **LIFEREG** – Estimates regression models with censored, continuous-time data under several alternative distributional assumptions. Does not allow for time-dependent covariates.
- **PHREG** – Uses Cox's partial likelihood method to estimate regression models with censored data. Handles both continuous-time and discrete-time data and allows for time-dependent covariables

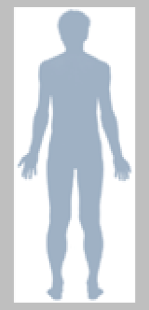
SAMPLE CODE

```
proc phreg data=overall outest=population;
model month*churn(0)=debtornum tenure csat level no_of_times_in_collections
no_of_complaints avg_consumption avg_bill no_calls no_chats no_login/ties=efron
;
baseline out=a survival=s logsurv=ls loglogs=lls;
run;
```

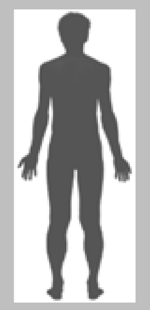
Key Terms: Time to Event (in this case Month) and Censoring applied on Churn variable

EPOSTERBOARDS TEMPLATE

ILLUSTRATION OF HAZARD RATIO



Patient Attribute	Model Reference Category	Hazard Ratio	Patient Overall Hazard Ratio
Age	Age 65+	1.0	1.0 * 1.0 * 1.0 * 1.0 = 1
Gender	Female	1.0	
Diabetic Status	Diabetic	1.0	
Therapy Type	Reference Category	1.0	

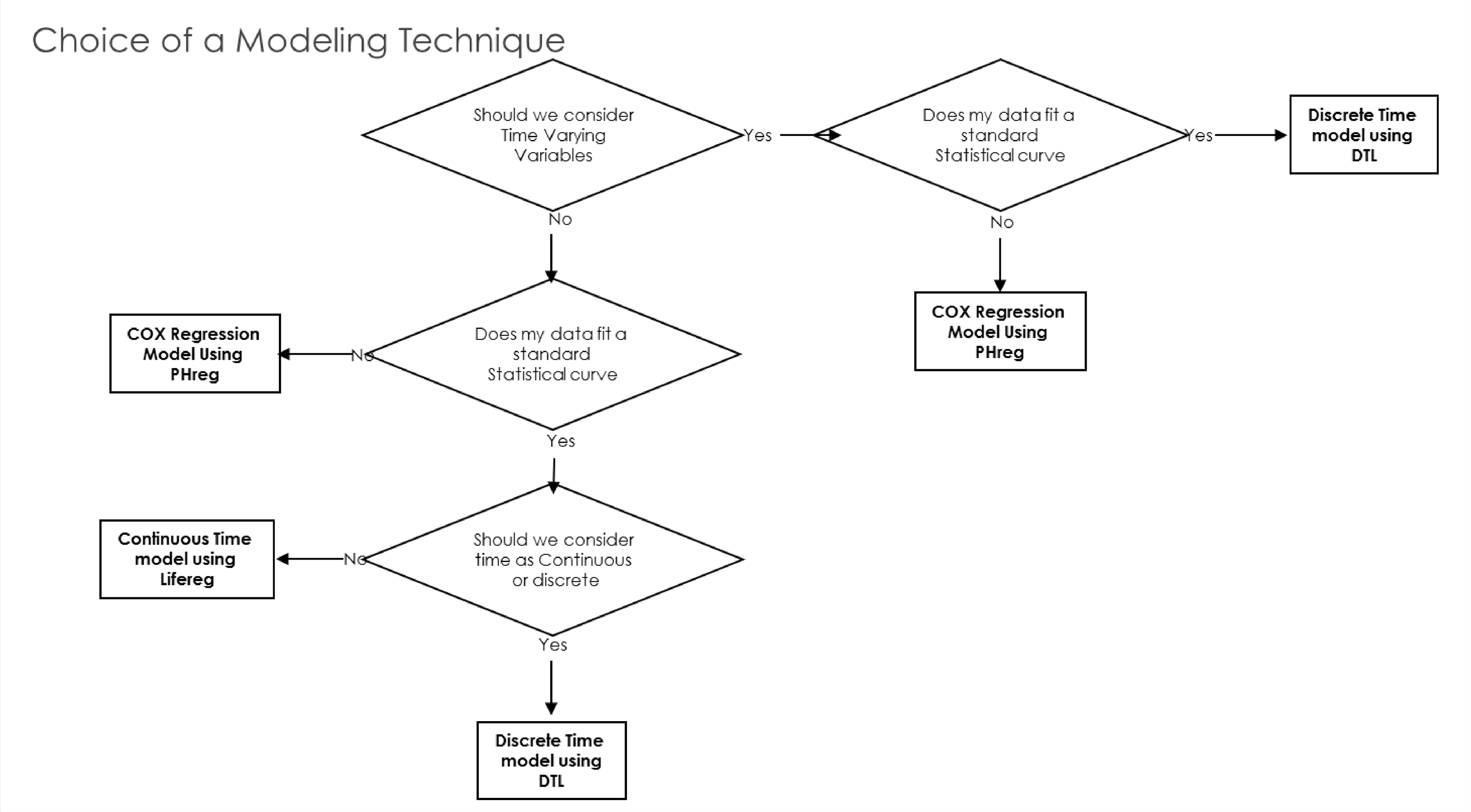


Patient Attribute	Details	Hazard Ratio	Patient Overall Hazard Ratio
Age	32	0.75	**0.75 * 1.0 * 1.0 * 0.82 = 0.615⁺**
Gender	Female	1.0	
Diabetic Status	Diabetic	1.0	
Therapy Type	Exposure to new study	0.82	

With respect to reference category, this patient has 38.5% more chance to survive from time t_n to t_{n+1}, of which an additional 13.5% is augmented to age factor by the exposure to client drug within 180 days of treatment

⁺*(based on entire population analysis using survival model)*

SUMMARY AND PRACTICAL APPLICATIONS



- Practical application scenarios
  - Identify what factors would cause a customer to churn
  - Evaluate the impact of a drug on treatment longevity
  - Identify project outcome – will a change be breakthrough or incremental innovation
  - Identify patient survival chances – based on historical data
  - Identify chances of failure for equipment

REFERENCES

https://www.sas.com/content/dam/SAS/en_ca/User%20Group%20Presentations/Toronto-Data-Mining-Forum/Chen-SurvivalModel-March2012.pdf

SAS® GLOBAL FORUM 2018

April 8 – 11 | Denver, CO
Colorado Convention Center

#SASGF