2024 Summer Institute In Statistics for Clinical & Epidemiological Research

Module 3:

Design, Conduct, and Analysis of Randomized Clinical Trials with Time to Event Primary Endpoints

Lecture 11:

Semi-Parametric Inference with Time to Event Data

Scott S. Emerson, M.D., Ph.D.
Professor Emeritus of Biostatistics
University of Washington

1

Semiparametric Models

- Exact form of within group distributions are unknown, but related to each other by some finite dimensional parameter vector
- Full inference only for comparing distributions
- One group's distn can be found from another group's and a finite dimensional parameter
- (Most often: Distributions equal under H_0)

(My definition of semiparametric models is a little stronger than some statisticians', but agrees with commonly used semiparametric survival models)

2

General Analysis Models

A Useful Analogy

3

Urn Model

- Balls in an urn of various colors and patterns
- Balls might represent people in a study
  - At any given time, the balls that are in the urn are therefore the risk set
- Colors and patterns represent risk factors

4

Death Process

- Periodically, I come in and choose a ball from the urn and take it
- When a ball is chosen it fails
- My predilection for choosing certain colors or patterns identifies true risk factors
- Characteristics of the balls that I do not notice have no effect on survival probabilities

5

Evidence for Risk Factors

- A certain color/pattern must be my favorite if
  - (Time based observations)
    - I come in more often when that color/pattern is in the urn
      - You need not consider what else is in the urn
  - (Risk set based observations)
    - I choose that color/pattern with a frequency disproportionate to its frequency in the urn
      - If I am blind to a characteristic, my choices should look like random sampling
      - You need not consider the times that I come in

6

(Semi)parametric Models

- Two general (semi)parametric probability models used in survival analysis
- Accelerated failure time models
  - Consider time of failure
- Proportional hazards models
  - Consider relations among hazards
  - (Additive hazards models also used, but less frequently)

7

Accelerated Failure Time Models

- Two groups that differ in some risk factor have survivor functions related by a parameter measuring acceleration or deceleration of time

$$S(t;\theta) = S_0(\theta t)$$

  - E.g.,
    - A smoker ages twice as fast as a nonsmoker
    - Each human year is seven dog years

8

Proportional Hazards Models

- Two groups that differ in some risk factor have survivor functions related by a parameter measuring increased hazard

$$\lambda_1(t) = \theta\lambda_0(t)$$
$$S_1(t) = [S_0(t)]^\theta$$

  - E.g.,
    - At any given time, a smoker is ten times more likely to develop lung cancer as a nonsmoker

9

Scientific Studies

- As a scientist you may
  - Observe
    - When I come into the room and take a ball,
    - The colors/patterns on all the balls in the urn, and
    - The color/patterns on the ball that I take
  - Experiment
    - Change the compostion in the urn and see
      - Whether I come in the room more or less often, and
      - The lengths to which I might go to find balls with certain colors or patterns by restricting my choices

10

Altering the Risk Set

- Censoring and time-varying covariates are analogous to changes in the composition of the urn
- Censoring = removing balls from the urn
- Time-varying covariates = repainting the balls or adding different balls

11

Caveats: Informative Censoring

- Altering the risk set can be problematic
- Recall that in order for survival estimates to be consistent, the risk set in the sample must look like a random sample from the population
- You should not selectively remove or change balls that were (for their risk factors) particularly more likely or less likely to be chosen
- If you notice that I search the urn from top to bottom,
  - Don't just change the balls sitting at the top of the urn
  - Make sure you stir the urn after each change

12

Caveats: Time-varying Covariates

- Time-varying covariates are far more easily implemented in the hazard based models
  - Risk set approach makes this easy
- However, scientifically we run the risk of overfitting our data using variables we are less interested in
  - A priest delivering last rites is highly predictive of death and that may obscure that it was a gunshot wound that led to the death

13

Semiparametric Models: Notation

For group k : $F_k(t) = \Psi\left(t, \vec{\Phi}_k\right)$

where :

$\Psi(\cdot,\cdot)$ has unknown form (in t)

$\vec{\Phi}_0 = \vec{0}$ for identifiability of $\Psi(\cdot,\cdot)$

$\vec{\Phi}_k$ is finite dimensional and unknown
(estimable by comparing two or more groups)

14

Semiparametric Survival Models

Accel failure : $F_k(t) = F_0(t\theta_k)$

Prop hzd : $S_k(t) = [S_0(t)]^{\theta_k}$

where in a regression problem

$$g(\theta_k) = \vec{X}_k^T \vec{\beta}$$

for some link function $g(\cdot)$

15

Semiparametric Inference

- Semiparametric inference generally proceeds through estimating equations
- Estimates found by iterative search
- Asymptotic distributions from special theory

16

PH Partial Likelihood

- Proportional hazards regression based on hazard of observed failure relative to sum of hazards in the risk set
- Often referred to as "rank based method" because no information is used about observation time except its order in the sample

$$S_i(t) = [S_0(t)]^{\theta_i}; \lambda_i(t) = \theta_i \lambda_0(t) \quad \text{where } \log(\theta_i) = \vec{X}_i^T \vec{\beta}$$

$$\text{Partial likelihood}: L(\vec{\beta}) = \prod_{i=1}^{n} \left[\frac{\lambda_i(T_i)}{\sum_{j:T_j \geq T_i} \lambda_j(T_i)} \right]^{D_i} = \prod_{i=1}^{n} \left[\frac{\theta_i}{\sum_{j:T_j \geq T_i} \theta_j} \right]^{D_i}$$

17

Partial Likelihood

- Covariate vector for the i-th subject: $\vec{X}_i$

$$\lambda_i(t) = \lambda_0(t) \exp\{\vec{X}_i \vec{\beta}\}$$

$$L(\vec{\beta}) \propto \prod_{i=1}^{n} \left\{ \frac{\exp\{\vec{X}_i \vec{\beta}\}}{\sum_{j:T_j \geq T_i} \exp\{\vec{X}_j \vec{\beta}\}} \right\}^{D_i}$$

$$\log L(\vec{\beta}) = \sum_{i=1}^{n} D_i \left\{ \vec{X}_i \vec{\beta} - \log \sum_{j:T_j \geq T_i} \exp\{\vec{X}_j \vec{\beta}\} \right\}$$

18

18

Partial Likelihood Based Score

- Appears as
  - The covariate value **observed** for the individual that had an event
  - Minus value **expected** among risk set as weighted by relative hazard

$$U_k(\beta) = \frac{\partial}{\partial \beta_k} \log L(\vec{\beta}) = \sum_{i=1}^{n} D_i \left\{ X_{ik} - \frac{\sum_{j:T_j \geq T_i} X_{jk} \exp\{\vec{X}_j \vec{\beta}\}}{\sum_{j:T_j \geq T_i} \exp\{\vec{X}_j \vec{\beta}\}} \right\}$$

19

19

Partial Likelihood Based Score: Two Samples

- For a two sample problem, $X_i = 0, 1$
  - For group x, let d_{ix} be events and n_{ix} be number at risk at time t_i

$$U_k(\beta) = \sum_{i=1}^{n} \left\{ d_{i1} - \frac{n_{i1} e^{\beta}}{n_{i0} + n_{i1} e^{\beta}} (d_{i0} + d_{i1}) \right\}$$

$$U_k(\beta) = \sum_{i=1}^{n} \left\{ \frac{n_{i0} n_{i1}}{n_{i0} + n_{i1} e^{\beta}} \left(\hat{\lambda}_{i1} - e^{\beta} \hat{\lambda}_{i0} \right) \right\}$$

- Under the null hypothesis $e^{\beta} = 1$, and with equal censoring distributions, number at risk will tend to reflect the randomization ratio
  - Relative weighting of observed differences in hazard over time by size of risk group

20

20

Partial Likelihood Based Information

$$I_{k\ell}(\beta) = \frac{\partial^2}{\partial\beta_k \partial\beta_\ell} \log L(\vec{\beta}) = \frac{\partial}{\partial\beta_k} U_\ell(\beta)$$

$$= \sum_{i=1}^{n} D_i \left\{ \frac{\sum_{j:T_j \geq T_i} X_{jk} X_{j\ell} \exp\{\vec{X}_j \vec{\beta}\}}{\sum_{j:T_j \geq T_i} \exp\{\vec{X}_j \vec{\beta}\}} - \frac{\sum_{j:T_j \geq T_i} X_{jk} \exp\{\vec{X}_j \vec{\beta}\} \sum_{j:T_j \geq T_i} X_{j\ell} \exp\{\vec{X}_j \vec{\beta}\}}{\left[\sum_{j:T_j \geq T_i} \exp\{\vec{X}_j \vec{\beta}\}\right]^2} \right\}$$

21

21

Semiparametric Summary Measures

- Estimation of summary measures is generally limited to the parameter fundamental to the semiparametric model
  - Proportional hazards
    - Only make inference about hazard ratio
  - Accelerated failure time
    - Only make inference about ratio of quantiles
- Methods do exist for estimating the "baseline" survival curve using the estimated parameters from the semiparametric model
  - Such are primarily used descriptively
  - Some have used such estimates for prediction models

22

General Analysis Models

Time-varying Covariates

- The proportional hazards model is widely used, even when we cannot be sure the hazard function is proportional over all time
- Because it relies so heavily on estimation through the hazards, it does allow us to consider time-varying covariates

23

Fixed Covariates

- In a typical study, we compare the distribution of some outcome across groups defined at the start of the study
- Example: Risk of hang gliding
  - Identify two groups
    - Hang gliders
    - Cowards
  - Follow survival experience over time

24

Problem

- What if a coward obtains courage?
- Misclassification will attenuate the true effect of hang gliding on survival
  - Biased estimates
  - Less precision

25

A Wrong Approach

- We cannot divide the sample into groups according to lifetime habits
- Suppose we consider
  - Ever hang glided (hung glide?) vs Constant coward
- We might detect spurious associations due to “survivorship”
  - If we started study at birth, we might find hang gliding is beneficial
  - Most people don’t start hang gliding until teenaged
  - We would detect the fact that hang gliders survived at least that long

•26

26

A Correct Approach

- Let each subject contribute observation time to the appropriate group according to covariate at the relevant time
  - And this is the best way to consider changes in treatment regimens
- Proportional hazards model
  - Easily done, if noninformative censoring results
- Accelerated failure time model
  - Difficult due to need to integrate hazards over disjoint intervals

27

Issues

- Issues related to the use of time-varying covariates are analogous to those when deciding to adjust for any variable
- Can regard measurements made at different times as different covariates
- Need to consider
  - Causal pathway of interest
  - Confounding (bias)
  - Precision
- Time aspect does increase the dimensionality

28

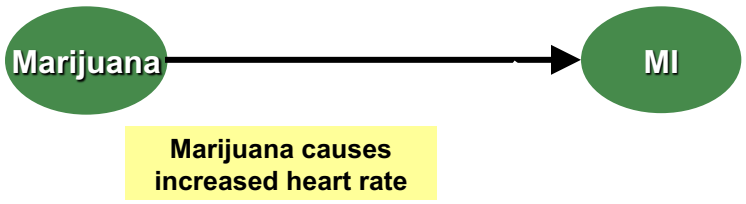
Issues: Informative Censoring

- Possibility that impending event causes informative censoring (confounding?)
- Types of variables
  - Extrinsic: Unaffected by individual decisions
    - As a rule, time-varying extrinsic variables will not cause informative censoring
    - E.g., Air pollution on a given day in an asthma study
      - (providing it does not affect relocation)
  - Intrinsic: Potentially affected by impending event
    - E.g., Marijuana use

29

Causation versus Association

- Example: Scientific interest in causal pathways between marijuana use and heart attacks (MI)
  - Pictorial representation of hypothetical causal effect of marijuana on MI that might be of scientific interest

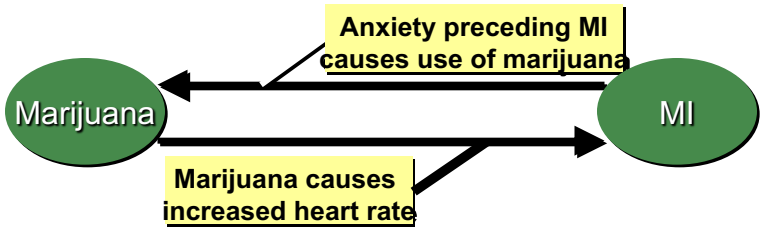


Marijuana causes increased heart rate

30

Causation versus Association

- In an observational study, we cannot thus be sure which causative mechanism an association might represent
  - Either of these mechanisms will result in an association between marijuana use and MI



31

Issues: Obscuring Effect of Interest

- With time-varying covariates, we have increased opportunity to measure short term effects
- This is good if that is our interest
  - Immediate effects of blood pressure on hemorrhagic stroke
- This is bad if we wanted to assess long acting risk factors
  - Chronic effect of asbestos on lung cancer
    - A former asbestos worker is still at high risk
- Capability for modeling time-varying covariates also increases chances for modeling a variable in the causal pathway of interest

32

Issues: Summary Measure

- As illustrated previously, the interpretation of some of the statistics commonly used in survival analysis is heavily dependent upon the censoring distribution
- It is very difficult to explore how the changing size of risk sets might be altering the interpretation of the time-averaged hazard ratio in a proportional hazards model
- Nonetheless, the Cox PH model has seen wide application, and we have gotten used to it
  - We can most easily justify its use based on a Weibull approximation
  - But that might be less valid when there is a chance of nonmonotonic hazard functions

33

Issues: Final Comments

- Time-varying covariates are definitely of scientific interest
  - Frequently of interest in the setting of non-adherence or changes in treatment regimen
  - Time varying covariates is the best way to address such questions, as opposed to censoring subjects or subsetting as if changes occurred at baseline
- However, they should not be used casually
- Usually, my first choice is to try to address scientific questions with fixed covariates
  - I will put up with some misclassification, to avoid making mistakes that are due to incorrect, untestable assumptions

34

Semiparametric Models: Issues

- Advantages
  - Can handle sparse data
  - More robust than any single parametric model
- Disadvantages
  - Not as easily interpreted when semiparametric model does not hold
    - But the Cox PH estimated hazard ratio can be interpreted as directly standardized hazard rates
    - The weights used, however, can be a little obscure
  - Little reason to suggest a given risk factor would affect distribution in only one way

35

Inflammatory Assertion

- (Semi)parametric models are not typically in keeping with the state of knowledge as an experiment is being conducted
- The assumptions are more detailed than the hypothesis being tested, e.g.,
  - Question: How does the intervention affect the first moment of the probability distribution?
  - Assumption: We know how the intervention affects the 2nd, 3rd, …, ∞ central moments of the probability distribution.

36

The Problem

- Incorrect parametric assumptions can lead to incorrect statistical inference
- Precision of estimators can be over- or understated
  - Hypothesis tests do not attain the nominal size
- Hypothesis tests can be inconsistent
  - Even an infinite sample size may not detect the alternative
- Interpretation of estimators can be wrong

37

(Semi)parametric Example

- Survival cure model (Ibrahim, 1999, 2000)
- Probability model
  - Proportion π_i is cured (survival probability 1 at ∞) in the i-th treatment group
  - Noncured group has survival distribution modeled parametrically (e.g., Weibull) or semiparametrically (e.g., proportional hazards)
  - Treatment effect is measured by $\theta = \pi_1 - \pi_0$
- The problem as I see it: Incorrect assumptions about the nuisance parameter can bias the estimation of the treatment effect

38

Foundational Issues: Null

- Which null hypothesis should we test?
  - Strong Null: The intervention has no effect whatsoever

$$H_0 : F(t) = G(t), \forall t$$

  - Weak Null: The intervention has no effect on some summary measure of the distribution

$$H_0 : \theta = \theta_0$$

39

Foundational Issues: Alternative

- What should the distribution of the data under the alternative represent?
- Counterfactual
  - An imagined form for *F(t), G(t)* if something else were true
- Empirical
  - The most likely distribution of the data if the alternative hypothesis about θ were true

40

My Views

- The null hypothesis of greatest interest is rarely that a treatment has no effect
  - Bone marrow transplantation
  - Women's Health Initiative
  - National Lung Screening Trial
- The empirical alternative is most in keeping with inference about a summary measure

41

An Aside

- The above views have important ramifications regarding the computation of standard errors for statistics under the null
- Permutation tests (or any test which presumes $F=G$ under the null) will generally be inconsistent

42

Glass Half Empty: Problem with (Semi)parametrics

- Many mechanisms would seem to make it likely that the problems in which a fully parametric model or even a semiparametric model is correct constitute a set of measure zero
- Treatments are often directed to outliers
- Treatments are often only effective in subsets
- Factors affect rates; outcomes measure cumulative effects

43

Glass Half Full: Value of (Semi)parametrics

- The most commonly used regression models have estimating equations that lend themselves to well understood properties
- Of course, some semiparametric models similarly lead to interpretable estimating equations
- In particular, over the past 50 years, we have gained a wealth of experience with the Cox proportional hazards model in particular
  - We can learn where it give us good insight
  - We can learn what to watch out for

44

A Non-Solution: Model Checking

- Model checking is apparently used by many to allow them to believe that their models are correct.
- From a recent referee's report:
  - "I know of no sensible statistician (frequentist or Bayesian) who does not do model checking."
- Apparently the referee believes the following unproven proposition:
  - If we cannot tell the model is wrong, then statistical inference under the model will be correct

45

A Non-Solution: Model Checking

- Counter example: Exponential vs Lognormal medians
- Pretest with Kolmogorov-Smirnov test (n=40)
  - Power to detect wrong model
    - 20% (exp); 12% (lnorm)
  - Coverage of 95% CI under wrong model
    - 85% (exp); 88% (lnorm)

46

A Non-Solution: Model Checking

- Model checking particularly makes little sense in a regulatory setting
- Commonly used null hypotheses presume the model fits in the absence of a treatment effect
  - Frequentists would be testing for a treatment effect as they do model checking
- Bayesians should model any uncertainty in the distribution
  - Interestingly, if one does this, the estimate indicating parametric family will in general vary with the estimate of treatment effect

47