Survival Analysis Part II

Luke Keele

February 8, 2022

Overview

- In Part I, we focused on data issues and basic bivariate statistics for survival data.

Overview

- In Part I, we focused on data issues and basic bivariate statistics for survival data.
- Today, we focus on multivariate models for survival data.

Overview

- In Part I, we focused on data issues and basic bivariate statistics for survival data.
- Today, we focus on multivariate models for survival data.
- Multivariate methods, however, require some background in causal inference principles.

Confounding

- Simple survival comparisons by treatment group are generally problematic. Why?

Confounding

- Simple survival comparisons by treatment group are generally problematic. Why?
- Differences in survival may be due to group composition.

Confounding

- Simple survival comparisons by treatment group are generally problematic. Why?
- Differences in survival may be due to group composition.
- E.g. treated group may have higher frailty *before* treatment.

Confounding

- Simple survival comparisons by treatment group are generally problematic. Why?
- Differences in survival may be due to group composition.
- E.g. treated group may have higher frailty *before* treatment.
- Treatment may appear to be ineffectual, but survival would be lower without treatment.

Confounding

- Simple survival comparisons by treatment group are generally problematic. Why?
- Differences in survival may be due to group composition.
- E.g. treated group may have higher frailty *before* treatment.
- Treatment may appear to be ineffectual, but survival would be lower without treatment.
- Treatment decisions are almost always confounded – correlated with patient characteristics.

The Answers

- Two (main) solutions:

The Answers

- Two (main) solutions:
- Randomization.

The Answers

- Two (main) solutions:
- Randomization.
- Multivariate statistical models.

Key Assumption

- We attempt to make treated and control groups similar on observed data, i.e. control for confounders.

Key Assumption

- We attempt to make treated and control groups similar on observed data, i.e. control for confounders.
- Key assumption: there are no unobserved differences between treated and control group.

Key Assumption

- We attempt to make treated and control groups similar on observed data, i.e. control for confounders.
- Key assumption: there are no unobserved differences between treated and control group.
- This assumption is untestable, i.e. you can never rule out the possibility that differences are due to unobserved differences in groups.

Key Assumption

- The key risk:

Key Assumption

- The key risk:
- Confounding by indication is present when patients are selected to receive medical treatments based on prognostic factors that indicate which patient would benefit from the treatment but those factors are not recorded.

Is this plausible?

- When you control for variables, you should be able to answer the question following question:

  How could it be that two units that are identical in all meaningful background characteristics nonetheless receive different treatments?

Is this plausible?

- When you control for variables, you should be able to answer the question following question:

  How could it be that two units that are identical in all meaningful background characteristics nonetheless receive different treatments?

- Are all the reasons for treatment fully recorded in the data?

Is this plausible?

- When you control for variables, you should be able to answer the question following question:

  How could it be that two units that are identical in all meaningful background characteristics nonetheless receive different treatments?

- Are all the reasons for treatment fully recorded in the data?
- Ask yourself if that is plausible?

Data Guidelines

- X represents baseline covariates – age, sex, etc.

Data Guidelines

- X represents baseline covariates – age, sex, etc.
- Control for all X's that predict treatment or outcome.

Data Guidelines

- X represents baseline covariates – age, sex, etc.
- Control for all X's that predict treatment or outcome.
- Clearly delineate when X's are measured.

Data Guidelines

- X represents baseline covariates – age, sex, etc.
- Control for all X's that predict treatment or outcome.
- Clearly delineate when X's are measured.
- Don't adjust for X's measured after the treatment occurred.

Data Guidelines

- X represents baseline covariates – age, sex, etc.
- Control for all X's that predict treatment or outcome.
- Clearly delineate when X's are measured.
- Don't adjust for X's measured after the treatment occurred.
- Only interpret effect for exposure variable.

Statistical Adjustment

- Regression models.

Statistical Adjustment

- Regression models.
- Propensity scores.

The Cox Model

- The work-horse regression model in medical statistics for survival data.

The Cox Model

- The work-horse regression model in medical statistics for survival data.
- RHS: exposure + control variables.

The Cox Model

- The work-horse regression model in medical statistics for survival data.
- RHS: exposure + control variables.
- Exponentiated coefficients are hazard ratios.

The Cox Model: Interpretation

The general formula: $exp[(X_j - X_k)\hat{\beta}]$.

The Cox Model: Interpretation

- $D = 0, 1$
- $\hat{\beta} = .5$

The Cox Model: Interpretation

- $D = 0, 1$
- $\hat{\beta} = .5$
- Hazard ratio is $exp[(1 - 0) \times .5] = 1.65$
- "A one–unit change in D increases the hazard of the event (death) by 65 percent."

The Cox Model: Interpretation

- $X = 18 - 90$ (age)
- $\hat{\beta} = 0.05$
- Increase in age by 20 years.

The Cox Model: Interpretation

- $X = 18 - 90$ (age)
- $\hat{\beta} = 0.05$
- Increase in age by 20 years.
- Hazard ratio is $exp[(20 - 0) \times .05] = 2.71$
- "A 20–unit change in age increases the hazard of the event by 271 percent."

The Cox Model: Interpretation

- $X = 18 - 90$ (age)
- $\hat{\beta} = 0.05$
- Increase in age by 20 years.
- Hazard ratio is $exp[(20 - 0) \times .05] = 2.71$
- "A 20–unit change in age increases the hazard of the event by 271 percent."
- Note: Hazard ratio is $exp[(60 - 40) \times .05] = 2.71$

The Cox Model: Interpretation

- $D = 0, 1$
- $\hat{\beta} = -.15$

The Cox Model: Interpretation

- $D = 0, 1$
- $\hat{\beta} = -.15$
- Hazard ratio is $exp[(1-0) \times -.15] = 0.86$
- "A one–unit change in D reduces the hazard of the event by 14 percent."

The Cox Model: Problems

- Cox model has some key drawbacks.

The Cox Model: Problems

- Cox model has some key drawbacks.
- Cox model assumes hazards are constant over time.
- Should effect be the same from 20 to 40 as 60 to 80?

The Cox Model: Problems

- Cox model has some key drawbacks.
- Cox model assumes hazards are constant over time.
- Should effect be the same from 20 to 40 as 60 to 80?
- Cox model can lead to bias in a number of scenarios.
- See: *Miguel Hernan, 2010. "The Hazard of Hazard Ratios." Epidimiology 21:1, 13–15.*

Propensity Score Weighting

- The main alternative is a propensity score analysis.

Propensity Score Weighting

- The main alternative is a propensity score analysis.
- Matching is one alternative: use log-rank test on matched data.

Propensity Score Weighting

- The main alternative is a propensity score analysis.
- Matching is one alternative: use log-rank test on matched data.
- Inverse propensity score weighting has some advantages with survival data.

Propensity Score Weighting

- The main alternative is a propensity score analysis.
- Matching is one alternative: use log-rank test on matched data.
- Inverse propensity score weighting has some advantages with survival data.
- Primarily, it allows one to retain hazard ratio interpretation.

The Propensity Score

- $A = 0, 1$ (treatment indicator).
- The propensity score: $\pi(X) = Pr(A = 1|X)$.

The Propensity Score

- $A = 0, 1$ (treatment indicator).
- The propensity score: $\pi(X) = Pr(A = 1|X)$.
- This is the conditional probability of being treated given a set of observed covariates.

Balancing Property

- The probability of treatment should be the same for people with the same propensity score.

Balancing Property

- The probability of treatment should be the same for people with the same propensity score.
- Alternatively, units with similar propensity scores should look similar in terms of all their observed characteristics.

Confounding and Sampling

- Confounding can be viewed as a problem of biased sampling.

Confounding and Sampling

- Confounding can be viewed as a problem of biased sampling.
- If units were randomly selected, X would be independent of A.

Confounding and Sampling

- Confounding can be viewed as a problem of biased sampling.
- If units were randomly selected, X would be independent of A.
- Random assignment is a form of random sampling.

Confounding and Sampling

- Confounding can be viewed as a problem of biased sampling.
- If units were randomly selected, X would be independent of A.
- Random assignment is a form of random sampling.
- i.e.

$$P(A=1|X=1)=P(A=1|X=0)$$

$$P(A=0|X=1)=P(A=0|X=0)$$

Confounding and Sampling

- Assume that older people are more likely to be treated.

Confounding and Sampling

- Assume that older people are more likely to be treated.
- Therefore older people are over-represented in the treated group.

Confounding and Sampling

- Assume that older people are more likely to be treated.
- Therefore older people are over-represented in the treated group.
- The propensity score $\pi(X)$ tells us which types of units are over and under-represented.

Inverse Probability Weighting

- Using IPW, we up-weight units that are under-represented ($X = 0$).

Inverse Probability Weighting

- Using IPW, we up-weight units that are under-represented ($X = 0$).
- And we down-weight units that are over-represented ($X = 1$).
- Thus we weight by the inverse of the propensity score.

IPW Weights

- More typically, we estimate the propensity score, $P(A|X) = \hat{\pi}(X)$, using logit or probit.

IPW Weights

- More typically, we estimate the propensity score, $P(A|X) = \hat{\pi}(X)$, using logit or probit.
- For $a_i = 1$, $w_i = 1/\hat{\pi}(L)$.

IPW Weights

- More typically, we estimate the propensity score, $P(A|X) = \hat{\pi}(X)$, using logit or probit.
- For $a_i = 1$, $w_i = 1/\hat{\pi}(L)$.
- For $a_i = 0$, $w_i = 1/(1 - \hat{\pi}(L))$.

Extreme Weights

- Sometimes weights are quite large, this causes IPW estimator to perform poorly.

Extreme Weights

- Sometimes weights are quite large, this causes IPW estimator to perform poorly.
- Use stabilized weights

Stabilized IPW

The IP stabilized weight is:

- For $a_i = 1$, $w_i = Pr(\hat{a})/\hat{\pi}(X)$.

Stabilized IPW

The IP stabilized weight is:

- For $a_i = 1$, $w_i = Pr(\hat{a})/\hat{\pi}(X)$.
- For $a_i = 0$, $w_i = (1 - Pr(\hat{a}))/(1 - \hat{\pi}(X))$.

Marginal Structural Model

The IPW estimate of the ATE via a marginal structural model:

1. Estimate p-score using logistic regression.

Marginal Structural Model

The IPW estimate of the ATE via a marginal structural model:

1. Estimate p-score using logistic regression.
2. Generate IP weights.

Marginal Structural Model

The IPW estimate of the ATE via a marginal structural model:

1. Estimate p-score using logistic regression.
2. Generate IP weights.
3. Estimate regression of Y on A, weighting by IP weights.

Outcome-Treatment Methods

1. IPW models treatment not outcome.

Outcome-Treatment Methods

1. IPW models treatment not outcome.
2. Can be combined with outcome model.

Outcome-Treatment Methods

1. IPW models treatment not outcome.
2. Can be combined with outcome model.
3. Should further reduce bias and increase precision.

Outcome-Treatment Methods

1. IPW models treatment not outcome.
2. Can be combined with outcome model.
3. Should further reduce bias and increase precision.
4. These models are called doubly robust.

Diagnostics

1. Estimate p-score using logistic regression.

Diagnostics

1. Estimate p-score using logistic regression.
2. Generate IP weights.

Diagnostics

1. Estimate p-score using logistic regression.
2. Generate IP weights.
3. Estimate regression of X on A, weighting by IP weights.
4. Should be no difference: treatment should be uncorrelated with covariates.

Conclusion

1. Most journals will expect a Cox model analysis.

Conclusion

1. Most journals will expect a Cox model analysis.
2. IPW improves on Cox regression.

Conclusion

1. Most journals will expect a Cox model analysis.
2. IPW improves on Cox regression.
3. You can also generate risk adjusted survival curves.

Conclusion

1. Most journals will expect a Cox model analysis.
2. IPW improves on Cox regression.
3. You can also generate risk adjusted survival curves.
4. Stata routines are limited and not best practice.