COX PROPORTIONAL HAZARDS MODEL AND ITS CHARACTERISTICS

OUTLINE

- Introduction to Linear Regression
- The Formula for the Cox PH Model
- Example
- Why the Cox PH Model is Popular
- Computing the Hazard Ratio
- ML Estimation of the Cox PH Model
- Adjusted Survival Curves
- The Meaning of the PH Assumption
- Summary

INTRODUCTION TO LINEAR REGRESSION

Linear Regression

- Describes a relation between some explanatory (predictor) variables and a variable of special interest, called the response variable

- Example: response var.: apartment rent
  predictor vars.: size, location, furnishing,…

- Goals with regression
  - Understanding the relation between explanatory and response vars.
  - Prediction of the value of the response var. for new explanatory vars.

Simple Linear Regression

- Only 1 predictor

- Model:

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \ , \qquad i = 1, \ldots, n$$

- Where: Y is the response

  x is the predictor

  β_0 and β_1 regression coefficients

  ε is an error term, s.t. $E[\varepsilon_i] = 0$ for all i

  $Var(\varepsilon_i) = \sigma^2$ for all i

  $\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = 0$ if $i \neq j$

Simple Linear Regression

- Parameter estimation
  - Estimates of β_0 and β_1 denoted by $\hat{\beta}_0$ and $\hat{\beta}_1$
  - Determined by least square approach
  - Interpretation: x increases by 1 unit $\Rightarrow Y$ increases by $\hat{\beta}_1$

- Inference on parameters
  - Is there a statistically significant relation btw. x and Y ?
  - Is $\hat{\beta}_1$ signifficantly different from 0 ?

- Prediction

$$\hat{Y}^* = \hat{\beta}_0 + \hat{\beta}_1 x^*$$

Multiple Linear Regression

- *p* predictors, *p* > 1

- Model:

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip} + \varepsilon_i\ , \qquad i = 1, \ldots, n$$

Where: Y is the response

$x_1, \ldots, x_p$ are predictors

$\beta_0, \ldots, \beta_p$ are regression coefficients

ε is an error term, with the same assumptions as before

Multiple Linear Regression

- Parameters estimation
  - Done in the same way as before, but interpretation of coefficients slightly different: $\hat{\beta}_j$ is the increase in *Y* if the predictor x_j increases by 1 unit **and all other predictors are held constant**

- Inference & prediction
  - Analogous to the case of simple linear regression

Can a Multiple Regression be Substituted by Many Simple Regressions?

- Consider the following models:

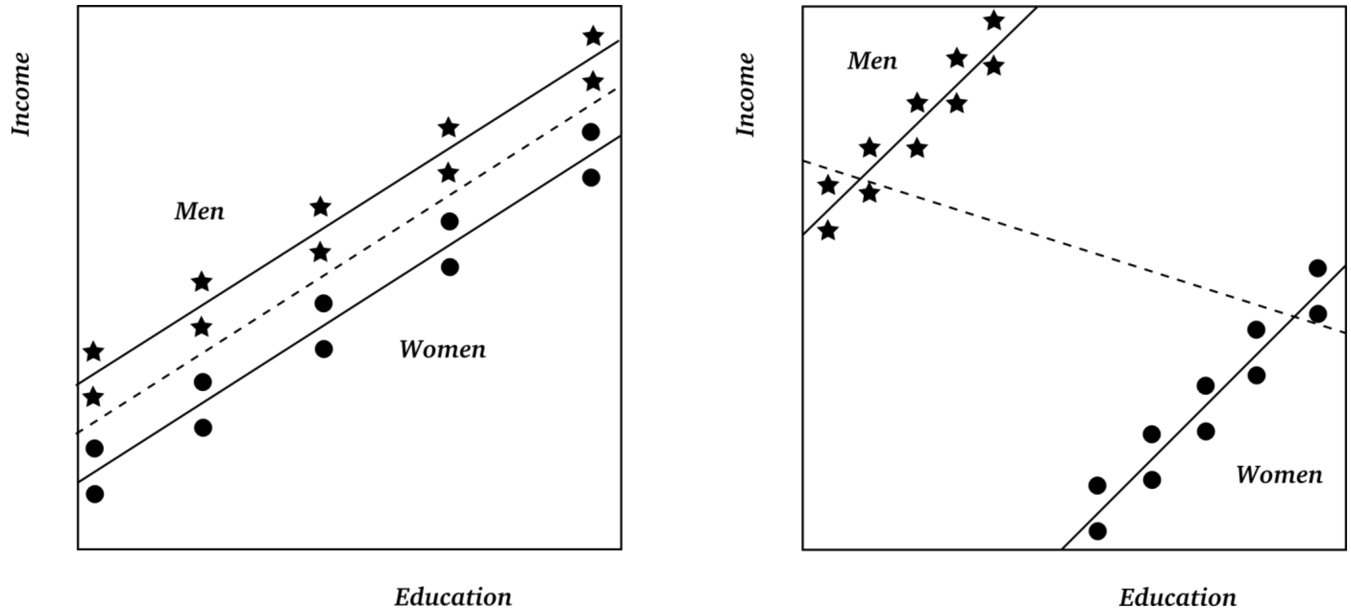
$$\begin{aligned}
&(1) \quad Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i \\
&(2) \quad Y_i = \beta_0' + \beta_1' x_{i1} + \varepsilon_i' \\
&(3) \quad Y_i = \beta_0'' + \beta_2'' x_{i2} + \varepsilon_i''
\end{aligned}$$

- Assume x_1 and x_2 are correlated and $\hat{\beta}_1$ and $\hat{\beta}_2$ are significantly different from zero. Then if we use model (2), a part of the effect of x_2 will be mistakenly attributed to x_1. Hence $\hat{\beta}_1 \neq \hat{\beta}_1'$ in general. Similarly, $\hat{\beta}_2 \neq \hat{\beta}_2''$.

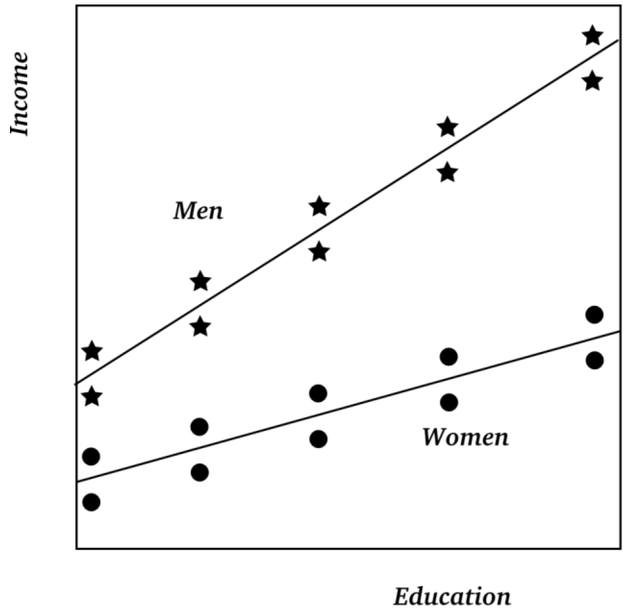
Confounding and Interaction

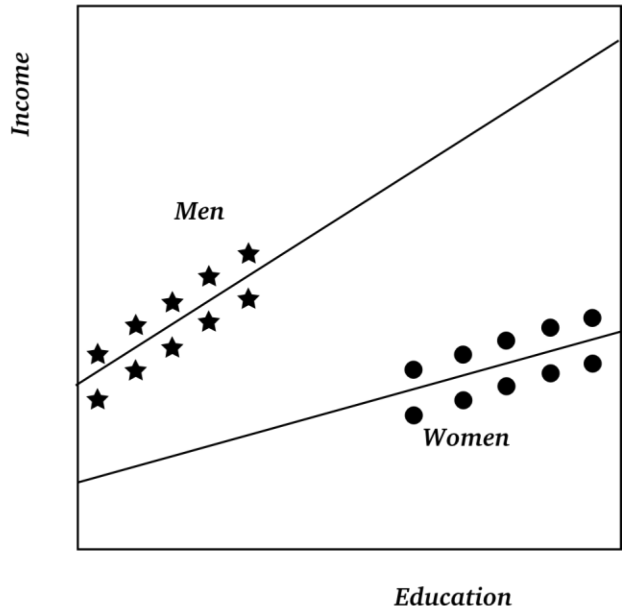
- Confounding:
  - Extraneous variable correlated with both dependent and independent variable
  - May lead to wrong conclusions about causal relationship of independent and dependent variable
- Interaction:
  - Independent variables combine to affect a dependent variable
  - Not to be confused with correlation

Example of Confounding



Example of Interaction





Precision gain

- Assume a multiple regression with two regressors X_1 and X_2
- X_1 is the variable of interest and X_2 is not statistically significant
- Yet confidence interval for X_1 is narrower when X_2 is present
- We prefer to keep X_2 in the model to have a better estimate for X_1

THE FORMULA FOR THE COX PH MODEL

The Formula for the Cox PH Model

- The formula for the Cox PH model is

$$h(t, \mathbf{X}) = h_0(t) \exp\left(\sum_{i=1}^{p} \beta_i X_i \right)$$

where

$$\mathbf{X} = \left(X_1, X_2, \ldots, X_p \right)$$

are the explanatory/predictor variables.

Explanation of the Formula

$$h(t, \mathbf{X}) = h_0(t)\exp\left(\sum_{i=1}^{p} \beta_i X_i\right)$$

- Product of two quantities:
  - $h_0(t)$ is called the baseline hazard
  - Exponential of the sum of β_i and X_i
- X 's zero (no X 's): reduces to baseline hazard
- Baseline hazard is an unspecified function
  - Semi-parametric model
  - Reason for Cox model being popular

Parametric Models

- Functional form is completely specified
- Example: Weibull

$$h(t, \mathbf{X}) = \lambda p t^{p-1}$$

where

$$\lambda = \exp\left(\sum_{i=1}^{p} \beta_i X_i\right)$$

and

$$h_0(t) = p t^{p-1}$$

- Parameters: p, β_i (more in chapter 7)

Important Properties of the Cox PH Formula

$$h(t,\mathbf{X}) = h_0(t)\exp\left(\sum_{i=1}^{p}\beta_i X_i\right)$$

- The baseline hazard $h_0(t)$ does not depend on $\mathbf{X}$ but only on t .
- The exponential involves the X 's but not t .
- The X are time-independent
- Proportional Hazard assumption follows

Time Independent Variables

- Not changing over time
  - Example: sex
- Values are set at time $t = 0$
- Variables unlikely to change are often considered time independent
  - Example: smoking status
- Also other variables are sometimes treated as time independent
  - Examples: age, weight

Extension to Time Dependent X

- Doesn’t satisfy PH assumption
- Need extended Cox model (chapter 6)

NUMERICAL EXAMPLE

Example: Data

- T = weeks until going out of remission
- X_1 = group status
- X_2 = log WBC (confounder, effect modifier)

  Interaction?
- $X_3 = X_1 \times X_2$ = group status × log WBC

Same dataset for each model
$n = 42$ subjects
T = time (weeks) until out of remission

Model 1: *Rx* only

Model 2: *Rx* and log WBC

Model 3: *Rx*, log WBC, and
Rx × log WBC

EXAMPLE

Leukemia Remission Data

Group 1($n = 21$)		Group 2($n = 21$)	
t(weeks)	log WBC	t(weeks)	log WBC
6	2.31	1	2.80
6	4.06	1	5.00
6	3.28	2	4.91
7	4.43	2	4.48
10	2.96	3	4.01
13	2.88	4	4.36
16	3.60	4	2.42
22	2.32	5	3.49
23	2.57	5	3.97
6+	3.20	8	3.52
9+	2.80	8	3.05
10+	2.70	8	2.32
11+	2.60	8	3.26
17+	2.16	11	3.49
19+	2.05	11	2.12
20+	2.01	12	1.50
25+	1.78	12	3.06
32+	2.20	15	2.30
32+	2.53	17	2.95
34+	1.47	22	2.73
35+	1.45	23	1.97

+ denotes censored observation

Example: R Output Model 1

```
Call:
coxph(formula = Surv(time, event) ~ Rx, data = Data, method = "breslow")
```

```
n= 42

       coef     exp(coef)      se(coef)       z        Pr(>|z|)
Rx     1.5092   4.5231         0.4096         3.685    0.000229 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

       exp(coef)       exp(-coef)      lower .95       upper .95
Rx     4.523           0.2211          2.027           10.09

Rsquare= 0.304   (max possible= 0.989 )
Likelihood ratio test     = 15.21  on 1 df,   p=9.615e-05
Wald test                 = 13.58  on 1 df,   p=0.0002288
Score (logrank) test      = 15.93  on 1 df,   p=6.571e-05

> model1$loglik[2]
[1] -86.37962
```

Example: R Output Model 3

```
Call:
coxph(formula = Surv(time, event) ~ Rx * logWBC, data = Data, method =
"breslow")
```

```
 n= 42

               coef     exp(coef) se(coef)  z       Pr(>|z|)
Rx             2.3549   10.5375   1.6810    1.401   0.161
logWBC         1.8028   6.0665    0.4467    4.036   5.45e-05 ***
Rx:logWBC     -0.3422   0.7102    0.5197   -0.658   0.510
---

          exp(coef) exp(-coef) lower .95 upper .95
Rx        10.5375   0.0949     0.3907    284.201
logWBC    6.0665    0.1648     2.5275    14.561
Rx:logWBC 0.7102    1.4080     0.2564    1.967

Rsquare= 0.648   (max possible= 0.989 )
Likelihood ratio test= 43.8  on 3 df,   p=1.633e-09
Wald test            = 30.6  on 3 df,   p=1.030e-06
Score (logrank) test = 45.9  on 3 df,   p=5.95e-10

> model3$loglik[2]
[1] -72.06572
```

$P = 0.510$: $\frac{-0.342}{-0.520} = -0.66 = Z$ Wald statistic

LR statistic: uses Log likelihood $= -72.066$

$-2 \ln L$ (log likelihood statistic) $= -2 \times (-72.066)$
$= 144.132$

LR (interaction in model 3)
$= -2 \ln L_{\text{model 2}} - (-2 \ln L_{\text{model 3}})$
$= (-2 \times -72.280) - (-2 \times -72.066)$
$= 144.550 - 144.132 = 0.428$

(LR is χ^2 with 1 d.f. under H_0:
no interaction.)
$0.40 < P < 0.50$, **not significant**
Wald test $P = 0.510$

Example: R Output Model 2

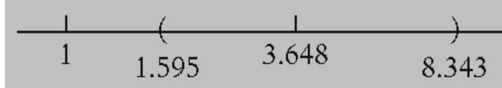
```
Call:
coxph(formula = Surv(time, event) ~ Rx + logWBC, data = Data, method =
"breslow")
```

```
n= 42

        coef     exp(coef)se(coef)z       Pr(>|z|)
Rx      1.2941   3.6476   0.4221   3.066    0.00217 **
logWBC  1.6043   4.9746   0.3293   4.872    1.11e-06 ***
---

        exp(coef) exp(-coef) lower .95 upper .95
Rx       3.648      0.2742     1.595     8.343
logWBC   4.975      0.2010     2.609     9.486

Rsquare= 0.644   (max possible= 0.989 )
Likelihood ratio test= 43.41  on 2 df,   p=3.744e-10
Wald test            = 31.78  on 2 df,   p=1.254e-07
Score (logrank) test = 42.94  on 2 df,   p=4.743e-10

> model2$loglik[2]
[1] -72.27926
```

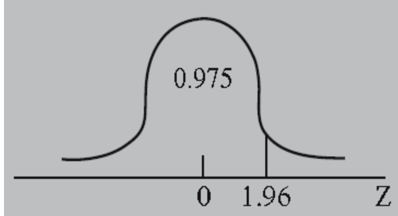
Point estimate: $\widehat{HR} = 3.648 = e^{1.294}$

Coefficient of treatment variable

95% confidence interval for the *HR*: (1.595, 8.343)



95% CI for β_1: 1.294 ± (1.96) (0.422)



95% CI for $HR = e^{\beta_1}$:

$$\exp[\hat{\beta}_1 \pm 1.96 s_{\hat{\beta}_1}] = e^{1.294 \pm 1.96(0.422)}$$

Example: Continued

Reasons to include logWBC in the model

Confounding: crude versus adjusted HR are meaningfully different → must control for logWBC

Precision of confidence intervals: even if no confounding we might prefer to keep logWBC if CI is smaller

Model 1:

	Coef.	Std. Err.	p > |z|	Haz. Ratio
Rx	1.509	0.410	0.000	$\boxed{4.523}$

No. of subjects = 42 Log likelihood = –86.380

Model 2:

	Coef.	Std. Err.	p > |z|	Haz. Ratio
Rx	1.294	0.422	0.002	$\boxed{3.648}$
log WBC	1.604	0.329	0.000	4.975

No. of subjects = 42 Log likelihood = –72.280

	[95% Conf. Interval]	
Rx model 1	2.027	10.094
	width = **8.067**	
	width = **6.748**	
Rx model 2	1.595	8.343
log WBC	2.609	9.486

WHY IS THE COX PH MODEL POPULAR?

Reasons for the Popularity of the Model

- Robustness
  - Cox model is a "safe" choice of a model in many situations

- Because of the model form:

$$h(t,\mathbf{X}) = \underbrace{h_0(t)}_{\geq 0} \times \underbrace{\exp\left(\sum_{i=1}^{p} \beta_i X_i\right)}_{\geq 0}$$

the estimated hazards are always non-negative.

- Even though $h_0(t)$ is unspecified we can estimate β_i's and thus compute the hazard ratio.

Reasons for the Popularity of the Model

- $h(t, \mathbf{X})$ and $S(t, \mathbf{X})$ can be estimated for a Cox model using a minimum of assumptions.

- In survival analysis the Cox model is preferred to a logistic model, since the latter one ignores survival times and censoring information.

COMPUTING THE HAZARD RATIO

Definition of the Hazard Ratio

- The Hazard Ratio is defined as

$$HR = \frac{\hat{h}(t, \mathbf{X}^*)}{\hat{h}(t, \mathbf{X})}$$

where

$$\mathbf{X}^* = (X_1^*, X_2^*, \ldots, X_p^*)$$

and

$$\mathbf{X} = (X_1, X_2, \ldots, X_p)$$

Interpretation of the Hazard Ratio

- Hazard for one individual divided by the hazard for a different individual
- For sake of interpretation we usually want $HR \geq 1$ i.e.

$$\hat{h}\left(t, \mathbf{X}^*\right) \geq \hat{h}\left(t, \mathbf{X}\right)$$

- We thus typically take
  - $\mathbf{X}^*$: group with larger hazard (e.g. placebo group)
  - $\mathbf{X}$: group with smaller hazard (e.g. treatment group)

Simplification of the Hazard Ratio

- Baseline hazard cancels out

$$HR = \frac{\hat{h}(t, \mathbf{X}^*)}{\hat{h}(t, \mathbf{X})} = \frac{\hat{h}_0(t)\exp\left(\sum_{i=1}^{p}\hat{\beta}_i X_i^*\right)}{\hat{h}_0(t)\exp\left(\sum_{i=1}^{p}\hat{\beta}_i X_i\right)} = \exp\left(\sum_{i=1}^{p}\hat{\beta}_i\left(X_i^* - X_i\right)\right)$$

Example: Remission Data, Model 1

- Only one variable of interest: exposure status
  - Placebo group: $X_1^* = 1$
  - Treatment group: $X_1 = 0$
- Hazard Ratio simplifies to

$$HR = \exp\left(\hat{\beta}_1\left(X_1^* - X_1\right)\right) = e^{\hat{\beta}_1}$$

- Since $\hat{\beta}_1 = 1.509$

  we have $HR = 4.523$

Example: Remission Data, Model 2

- Two variables of interest: exposure status and logWBC
  - Placebo group: $X_1^* = 1$
  - Treatment group: $X_1 = 0$
  - logWBC is held constant
- No product terms

$$HR = \exp\left(\hat{\beta}_1\left(X_1^* - X_1\right) + \hat{\beta}_2\left(X_2^* - X_2\right)\right)$$

$$= \exp\left(\hat{\beta}_1(1-0) + \hat{\beta}_2(\log WBC - \log WBC)\right)$$

$$= e^{\hat{\beta}_1}$$

Example: Remission Data, Model 2

- Since $\hat{\beta}_1 = 1.294$

  we have $HR = 3.648$

- Hazard Ratio is independent of logWBC
- Hazard Ratio different from model 1 because estimates change

Example: Remission Data, Model 3

- Three variables of interest
- Product terms

$$HR = \exp\left(\sum_{i=1}^{3} \hat{\beta}_i \left(X_i^* - X_i\right)\right)$$

$$= \exp\left(\hat{\beta}_1 - \hat{\beta}_3 \left(1 \times \log WBC - 0 \times \log WBC\right)\right)$$

$$= \exp\left(\hat{\beta}_1 - \hat{\beta}_3 \log WBC\right)$$

- Hazard Ratio depends on logWBC

ML ESTIMATION OF THE COX PH MODEL

Full Likelihood and Baseline Hazard Estimation

- The full Likelihood can be written in the following form

$$L_n(F) = \prod_{j=1}^{n} f(T_j)^{\delta_j} (1 - F(T_j))^{1-\delta_j}$$

$$= \prod_{j=1}^{n} h(T_j \mid X_j)^{\delta_j} S(T_j \mid X_j) =$$

$$= \prod_{j=1}^{n} h_0(T_j)^{\delta_j} \exp\left(\beta' X_j\right)^{\delta_j} \exp\left(- H_0(T_j) \exp\left(\beta' X_j\right)\right)$$

- This allows to derive estimator for the baseline hazard

$$\hat{h}_{0i} = \frac{1}{\sum_{j \in R(t_i)} \exp(\beta' X_j)}$$

$$\hat{H}_0(t) = \sum_{t_i \le t} \frac{1}{\sum_{j \in R(t_i)} \exp(\beta' X_j)}$$

The Cox Likelihood

- Assume *k* different failure times $t_{(1)} < t_{(2)} < \ldots < t_{(k)}$ s.t. there is exactly one failure at each $t_{(i)}, i = 1, \ldots, k.$

- Let $[i]$ denote the subject with an event at time $t_{(i)}$ and $R(t)$ the risk set at time *t*

- The Cox likelihood is given by

$$L(\boldsymbol{\beta}) = \prod_{j=1}^{k} \frac{\exp\left(\sum_{i=1}^{p} \beta_i X_{[j]i}\right)}{\sum_{l \in R(t_{(j)})} \exp\left(\sum_{i=1}^{p} \beta_i X_{li}\right)}$$

The Cox Likelihood

- *L* is also called “partial” likelihood
  - *Considers probabilities for subject who fail*
  - *Does not consider probabilities for censored subjects explicitly*
    - *Censored subjects are taken into account in the risk set*

- Estimates of β_i's denoted by $\hat{\beta}_i$'s
  - $\hat{\beta}_i$ solves $\dfrac{\partial \log L}{\partial \beta_i} = 0, \quad i = 1, \ldots, p$
  - i.e. $\hat{\beta}_i$'s maximize the Cox likelihood

Properties of the Estimates

- $\hat{\boldsymbol{\beta}} \xrightarrow{p} \boldsymbol{\beta}$ as $k \to \infty$

- $Var\left(\hat{\boldsymbol{\beta}}\right) = I^{-1}$ where I is the Fisher-information matrix given by

$$I_{i,j} = E\left(\left(\frac{\partial}{\partial \beta_i} \log L(\boldsymbol{\beta})\right) \times \left(\frac{\partial}{\partial \beta_j} \log L(\boldsymbol{\beta})\right)\right)$$

- $\hat{\boldsymbol{\beta}}$ is asymptotically normal

The Cox Likelihood: Example

- Gary, Larry, Barry have lottery tickets
- Winning tickets chosen at times t_1, t_2, . . .
- Each person ultimately chosen
- Can be chosen only once
- What is the probability that the order chosen is as follows: Barry, Gary, Larry?

Answer:

$$\text{Probability} = \underset{\text{Barry}}{\frac{1}{3}} \times \underset{\text{Gary}}{\frac{1}{2}} \times \underset{\text{Larry}}{\frac{1}{1}} = \frac{1}{6}$$

The Cox Likelihood: Example

- New scenario: Barry has 4 tickets, Gary has 1 ticket, Larry has 2 tickets
- What is the probability that the order chosen is again: Barry, Gary, Larry?

Answer:

$$\text{Probability} = \frac{4}{7} \times \frac{1}{3} \times \frac{2}{2} = \frac{4}{21}$$

- Subject‘s number of tickets affects probability
- For Cox model subject’s pattern of covariates affects likelihood of ordered events

The Cox Likelihood: Example

- Data

ID	TIME	STATUS	SMOKE
Barry	2	1	1
Gary	3	1	0
Harry	5	0	0
Larry	8	1	1

- Cox PH Model $h(t) = h_0(t)e^{\beta_1 SMOKE}$

ID	Hazard
Barry	$h_0(t)e^{\beta_1}$
Gary	$h_0(t)e^{0}$
Harry	$h_0(t)e^{0}$
Larry	$h_0(t)e^{\beta_1}$

The Cox Likelihood: Example

- The likelihood

$$L = \left[\frac{h_0(t)e^{\beta_1}}{h_0(t)e^{\beta_1} + h_0(t)e^{0} + h_0(t)e^{0} + h_0(t)e^{\beta_1}}\right]$$

$$\times \left[\frac{h_0(t)e^{0}}{h_0(t)e^{0} + h_0(t)e^{0} + h_0(t)e^{\beta_1}}\right] \times \left[\frac{h_0(t)e^{\beta_1}}{h_0(t)e^{\beta_1}}\right]$$

- The baseline hazard cancels out and does not play any role in estimation
- Likelihood determined only by the order of events

ADJUSTED SURVIVAL CURVES USING THE COX PH MODEL

Estimation of Survival Curves

- No Model: Kaplan-Meier method (chapter 2)
- Cox model: adjusted survival curves
  - Adjust for explanatory variables used as predictors
  - Like KM curves plotted as step functions

Converting Hazard Functions to Survival Functions

- Hazard Function:

$$h(t,\mathbf{X}) = h_0(t)\exp\left(\sum_{i=1}^{p}\beta_i X_i\right)$$

- Survival Function:

$$S(t,\mathbf{X}) = \left[S_0(t)\right]^{\exp\left(\sum_{i=1}^{p}\beta_i X_i\right)}$$

Estimated Survival Function

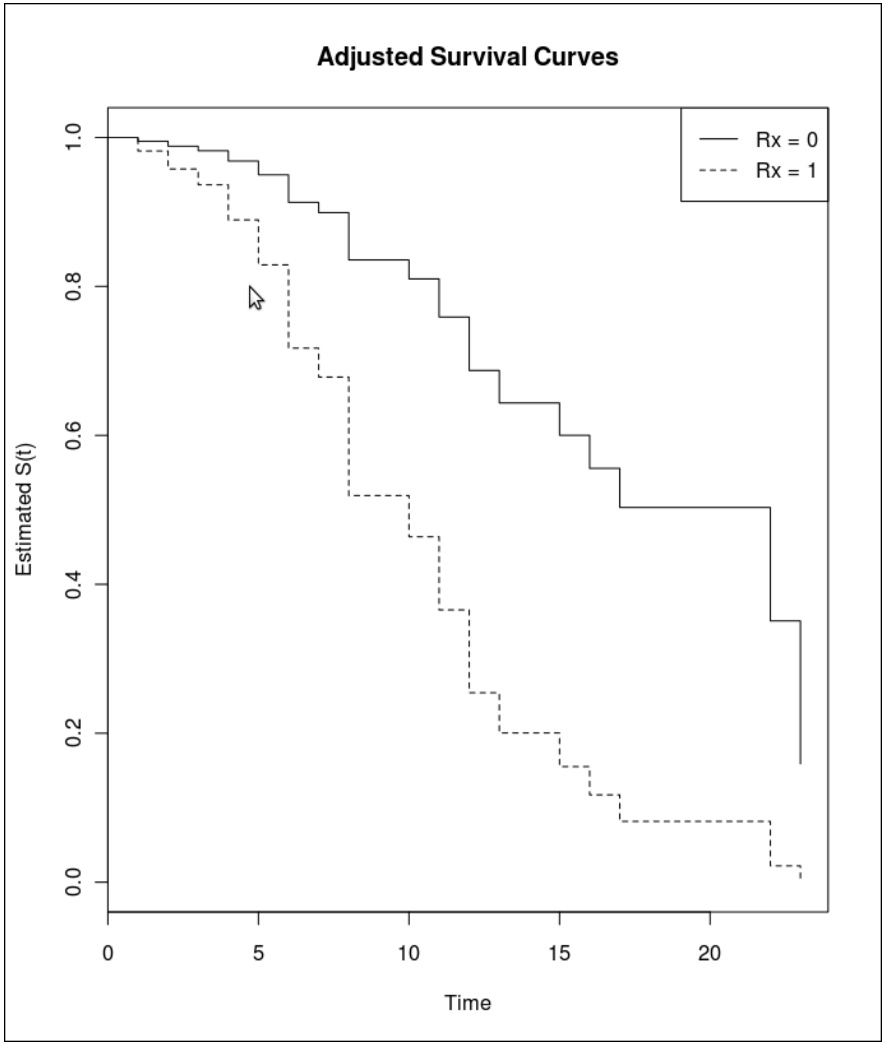
- Estimated survival function:

$$\hat{S}(t,\mathbf{X}) = \left[\hat{S}_0(t)\right]^{\exp\left(\sum_{i=1}^{p}\hat{\beta}_i X_i\right)}$$

- Estimated quantities: $\hat{S}_0(t), \hat{\beta}_i$

Example: Remission Data

```
### R Code
## Read the data
[…]
## Fit model 2 and plot
model2 <- coxph(Surv(time,event)
   ~ Rx + logWBC,
   method="breslow", data=Data)

plot(survfit(model2,
   newdata=data.frame(Rx=c(0,1),
   logWBC=rep(mean(logWBC),2))),
   lty=c(1,2), xlab="Time",
   ylab="Estimated S(t)",
   main="Adjusted Survival
   Curves")
```



MEANING OF THE PROPORTIONAL HAZARDS ASSUMPTION

Meaning of the PH Assumption

- Remember that the PH assumption requires that the HR is constant over time

$$HR = \exp\left(\sum_{i=1}^{p} \hat{\beta}_i \left(X_i^* - X_i \right) \right)$$

Meaning of the PH Assumption: Example

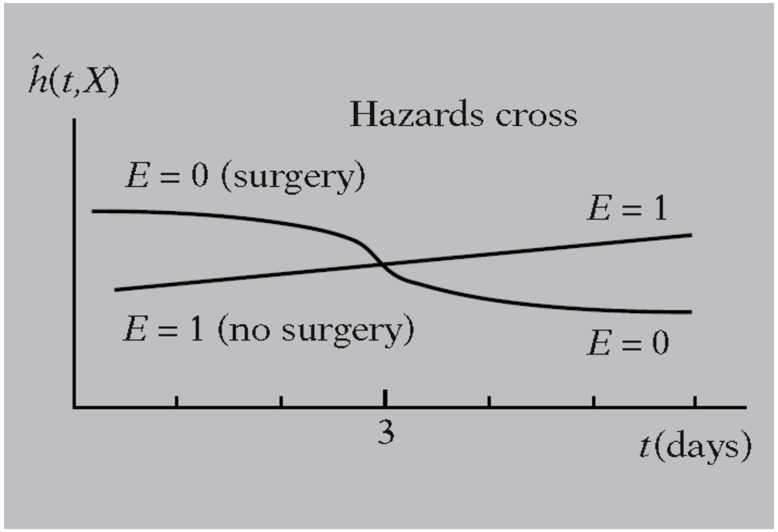
- A study in which cancer patients are randomized to either surgery or radiation therapy without surgery
- (0,1) exposure variable denoting surgery status, with 0 if a patient receives surgery and 1 if not
- exposure variable is the only variable of interest
- Is this model appropriate?
- No. Why?

Meaning of the PH Assumption: Example

- High risk for complications from surgery or perhaps even death early in the recovery process

- We expect to see hazard functions for each group that cross

- It is therefore inappropriate to use a CoxPH model for this situation



General rule: If the hazards cross, then a Cox PH model is not appropriate

Meaning of the PH assumption: What if Cox PH Model is Inappropriate?

- Start analysis using data after HR curves cross
- Fit PH model data before HR crossing and after crossing; get HR estimates (before crossing) and HR estimates (after crossing)
- Stratify by exposure (use KM curves)
- Use extended Cox model
- More on this in chapters 5 and 6

SUMMARY

Summary

- Introduction to Linear Regression
  - Simple vs. multiple
  - Confounding and interaction
  - Precision gain
- The Formula for the Cox PH Model
  - Formula: $h(t,\mathbf{X}) = h_0(t)\exp\left(\sum_{i=1}^{p}\beta_i X_i\right)$
  - Semi-parametric model
  - Leukemia example

Summary

- Why is the Model Popular?
  - Robustness
  - Gives non-negative hazards
  - Can calculate hazard ratio
  - Can estimate $h(t, \mathbf{X})$ and $S(t, \mathbf{X})$

- Computing the Hazard Ratio
  - Formula: $$\frac{h(t, \mathbf{X}^*)}{h(t, \mathbf{X})} = \exp\left[\sum_{i=1}^{p} \hat{\beta}_i \left(X_i^* - X_i\right)\right]$$

Summary

- ML Estimation for Cox PH Model
  - Full likelihood
  - Partial likelihood
  - Example
- Adjusted Survival Curves Using the Cox PH Model
  - Survival curve formula obtained from hazard function
  - $S(t, \mathbf{X}) = [S_0(t)]^{\exp(\sum \beta_i X_i)}$
  - To get adjusted curve usually use mean values for the covariates

Summary

- The Meaning of the PH Assumptions
  - Hazard ratio independent of time
  - Baseline hazard not involved in the HR formula
  - An example when PH assumption does not hold (crossing hazards)

References

- [1] D. G. Kleinbaum & M. Klein, Survival Analysis - A Self-Learning Text. Springer, Second Edition, 2005.
- [2] D. R. Cox, Regression Models and Life-Tables. Imperial College, London, 1972.
- [3] J. Fox, Applied Regression Analysis, Linear Models, and Related Methods. Sage Publications, 1997.
- [4] M. Dettling, Lecture Notes on Applied Statistical Regression. ETH Zurich, 2010. Available at http://stat.ethz.ch/education/semesters/as2010/asr/ASR-HS10-Scriptum.pdf
- [5] J. P. Klein & M. L. Moeschberger, Survival Analysis: Techniques for Censored and Truncated Data. Springer, Second Edition, 2003.