COX REGRESSION MODEL

Dr. Aric LaBarr
Institute for Advanced Analytics
MSA Class of 2014

PROPORTIONAL HAZARDS

Accelerated Failure Time Model

Initial

- Investigate Survival Curves
- Investigate Hazard Function

Distributions

- Find "best" distribution
  - Graphical methods
  - Statistical Tests

Model Building

- Select significant variables
- Finalize model

Accelerated Failure Time Model

Initial
- Investigate Survival Curves
- Investigate Hazard Function

Distributions
- Find "best" distribution
  - Graphical methods
  - Statistical Tests

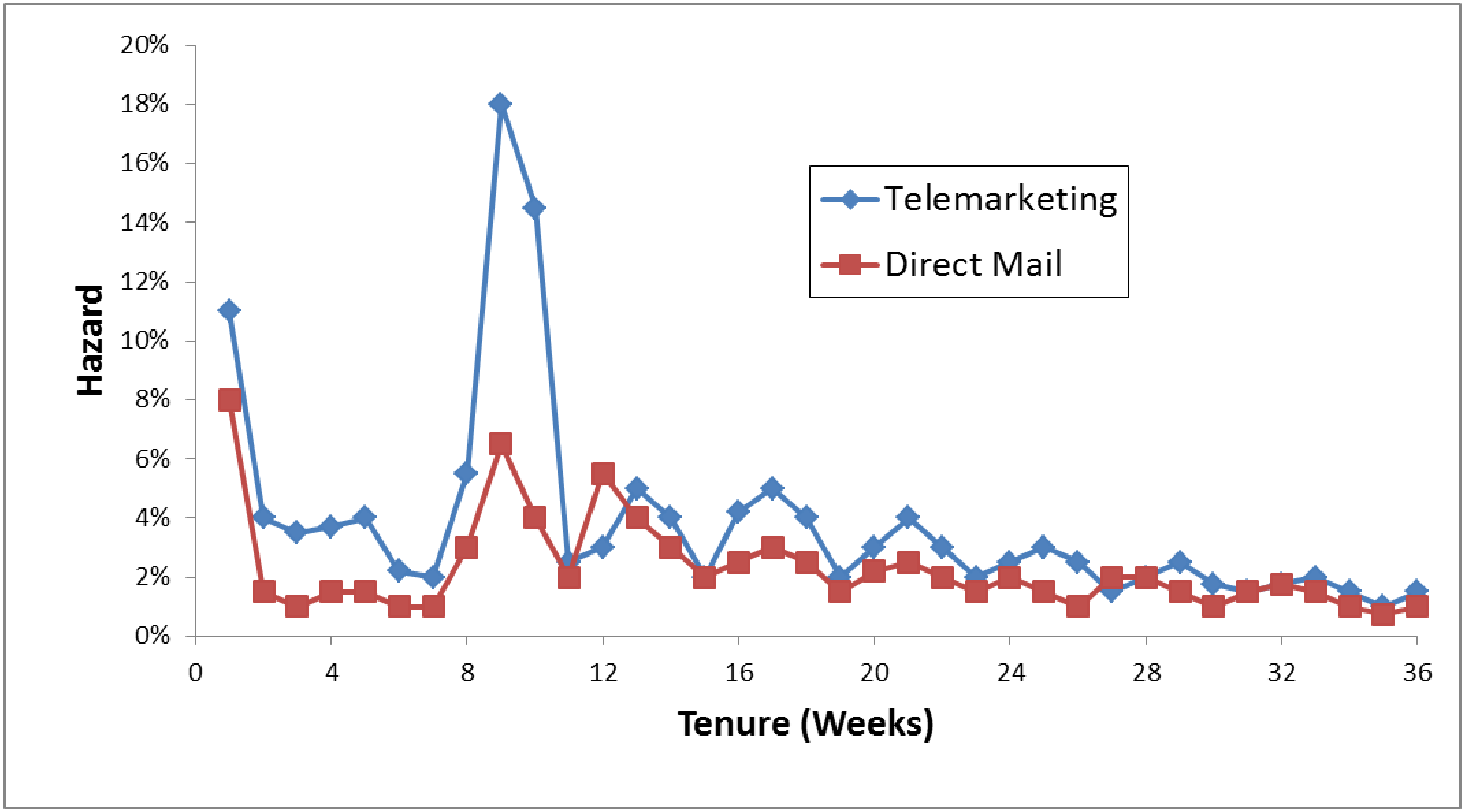
Proportional Hazards changes this!!!

Model Building
- Select significant variables
- Finalize model

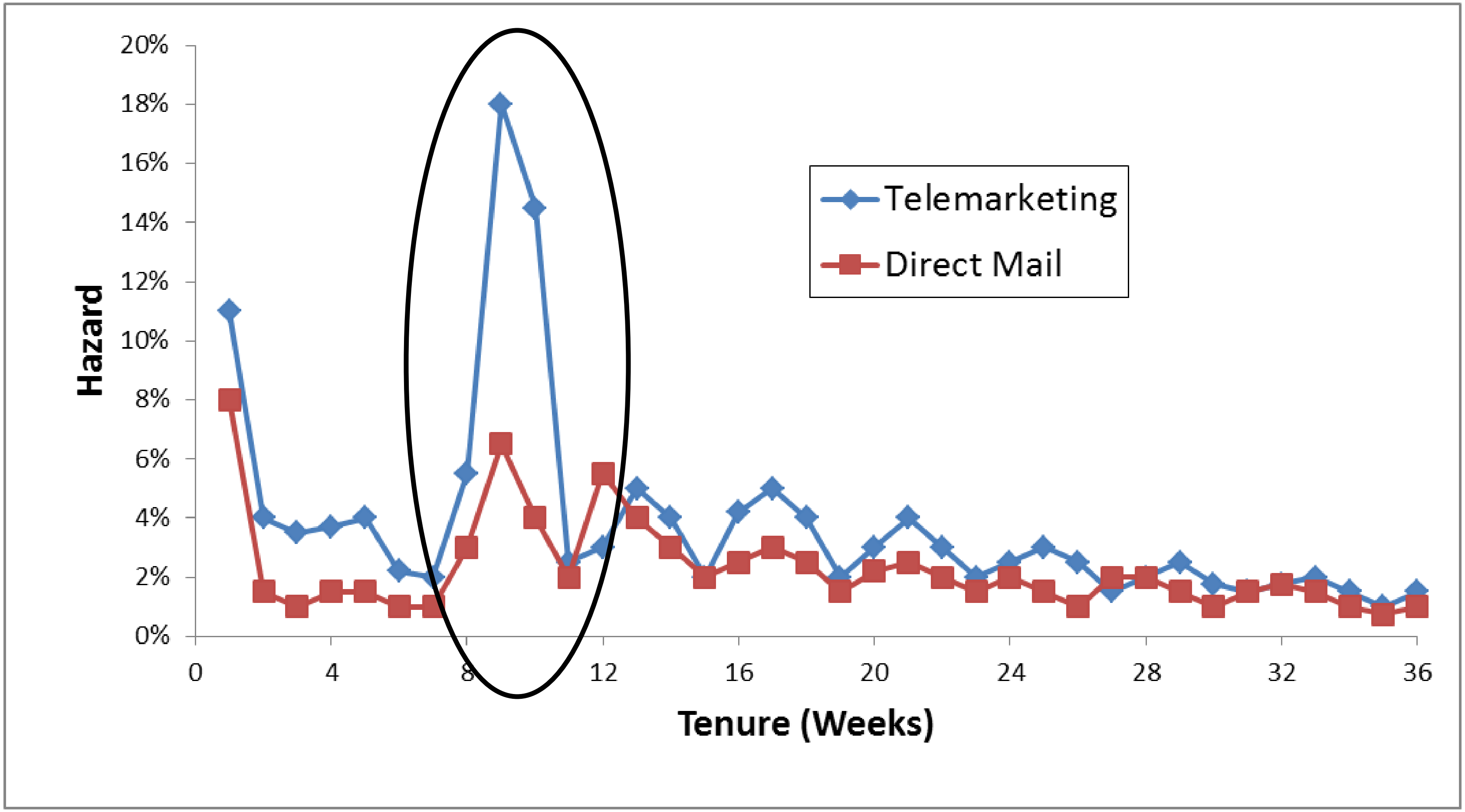
Cox Regression Model – Example

- "On average, a customer who signed up via direct mail lasts 4 years longer compared to a customer who signed up via telemarketing."
- Results do not say how long someone will last, only relative length of tenure between two groups.
- Cox assumed that factors measured at an initial time point have a uniform proportional effect on hazards between individuals (or groups).

Cox Regression Model – Example



Cox Regression Model – Example



Cox Regression Model – Example

- Cox assumed that factors measured at an initial time point have a **uniform** proportional effect on hazards between individuals (or groups).
- Cox developed a model that calculated the "average" effect of a variable across all tenures on the hazard.

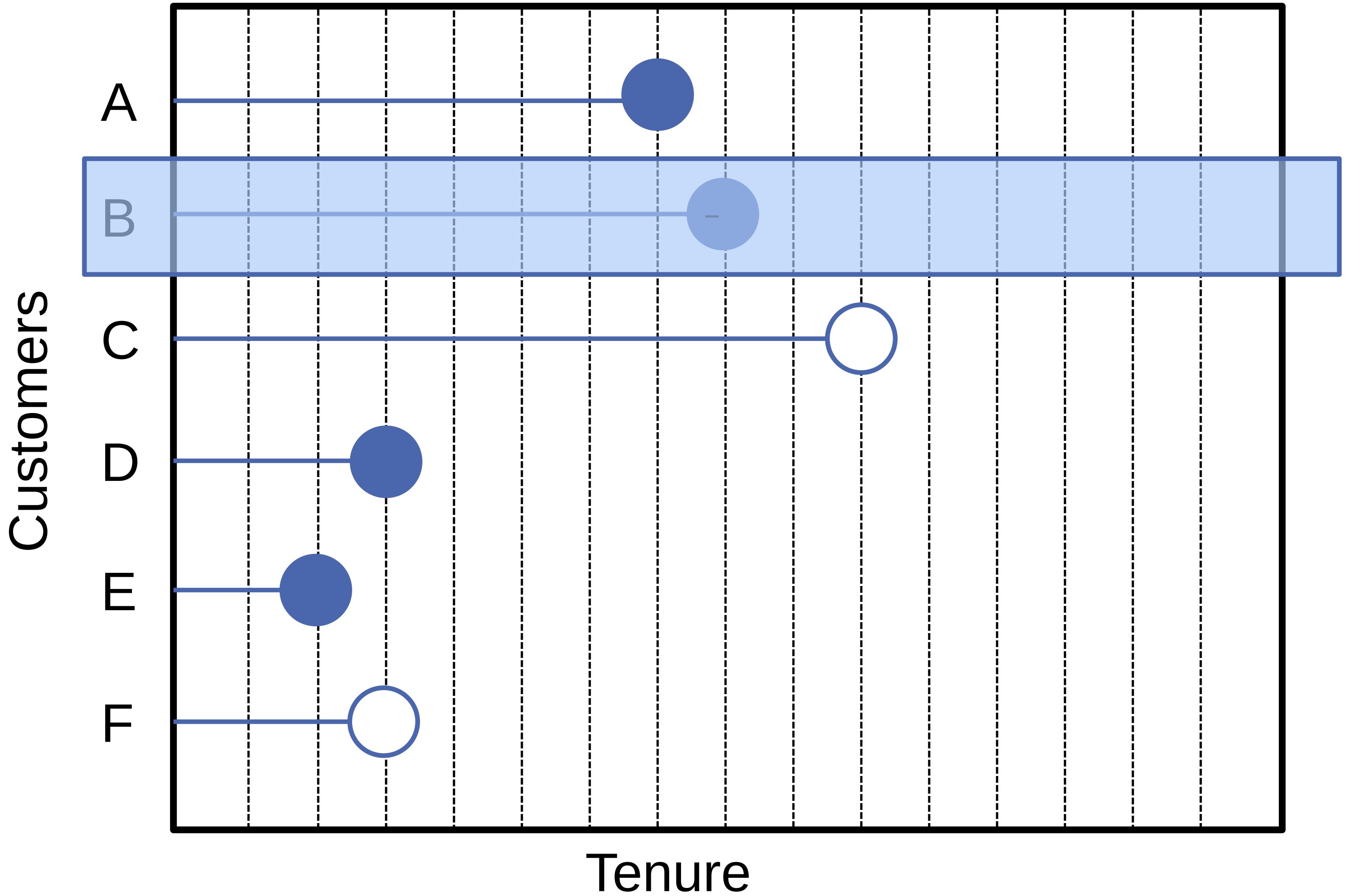
- What effect do the initial conditions have on hazards?

Cox Regression Model – Example

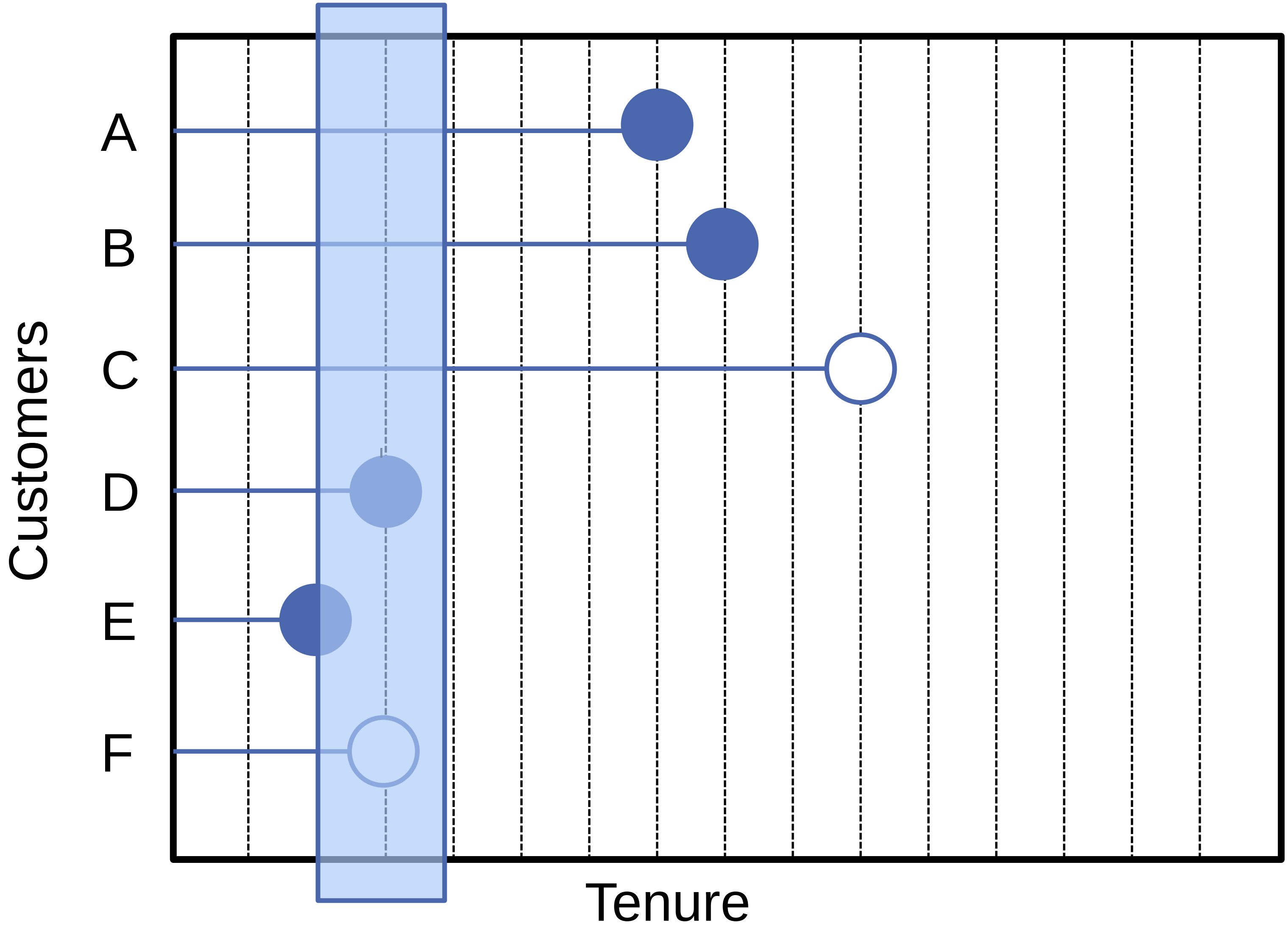
- Cox assumed that factors measured at an initial time point have a **uniform** proportional effect on hazards between individuals (or groups).
- Cox developed a model that calculated the “average” effect of a variable across all tenures on the hazard.
- What effect do the initial conditions have on hazards?

NOT FOCUSED ON PREDICTION!!!

Accelerated Failure Time Model



Cox Regression Model



Proportional Hazards Model

- Focus on the most basic version of the Cox regression model – proportional hazards with no time-varying covariates.

$$h(t) = h_0(t)e^{\beta_1 x_{i,1} + \cdots + \beta_k x_{i,k}}$$

Proportional Hazards Model

- Focus on the most basic version of the Cox regression model – proportional hazards with no time-varying covariates.

$$h(t) = h_0(t)e^{\beta_1 x_{i,1} + \cdots + \beta_k x_{i,k}}$$

Baseline hazard function

Proportional Hazards Model

- Focus on the most basic version of the Cox regression model – proportional hazards with no time-varying covariates.

$$h(t) = h_0(t)e^{\beta_1 x_{i,1}+\cdots+\beta_k x_{i,k}}$$

Covariates used to predict hazards

Proportional Hazards Model

- Focus on the most basic version of the Cox regression model – proportional hazards with no time-varying covariates.

$$h(t) = h_0(t)e^{\beta_1 x_{i,1} + \cdots + \beta_k x_{i,k}}$$

$$\log h(t) = \alpha(t) + \beta_1 x_{i,1} + \cdots + \beta_k x_{i,k}$$

Proportional Hazards Model

- Focus on the most basic version of the Cox regression model – proportional hazards with no time-varying covariates.

$$h(t) = h_0(t) e^{\beta_1 x_{i,1} + \cdots + \beta_k x_{i,k}}$$

$$\log h(t) = \alpha(t) + \beta_1 x_{i,1} + \cdots + \beta_k x_{i,k}$$

Proportional Hazards Model

- Remember that the Weibull (and Exponential) model was a rare case where we already defined the proportional hazards model.
- Here is the proportional hazards model:

$$\log h(t) = \alpha(t) + \beta_1 x_{i,1} + \cdots + \beta_k x_{i,k}$$

- If $\alpha(t) = \alpha$, then we have the Exponential model:

$$\log h(t) = \alpha + \beta_1 x_{i,1} + \cdots + \beta_k x_{i,k}$$

- If $\alpha(t) = \alpha \log t$, then we have the Weibull model:

$$\log h(t) = \alpha \log t + \beta_1 x_{i,1} + \cdots + \beta_k x_{i,k}$$

Proportional Hazards Model

- Remember that the Weibull (and Exponential) model was a rare case where we already defined the proportional hazards model.
- Here is the proportional hazards model:

$$\log h(t) = \alpha(t) + \beta_1 x_{i,1} + \cdots + \beta_k x_{i,k}$$

- If $\alpha(t) = \alpha$, then we have the Exponential model:

$$\log h(t) = \alpha + \beta_1 x_{i,1} + \cdots + \beta_k x_{i,k}$$

- If $\alpha(t) = \alpha \log t$, then we have the Weibull model:

$$\log h(t) = \alpha \log t + \beta_1 x_{i,1} + \cdots + \beta_k x_{i,k}$$

Proportional Hazards Model

- Why is the proportional hazard model so popular?
- The hazard for any one individual (group) is a fixed proportion of the hazard for any other individual (or group).

$$h_i(t) = h_0(t)e^{\beta_1 x_{i,1} + \cdots + \beta_k x_{i,k}}$$

$$h_j(t) = h_0(t)e^{\beta_1 x_{j,1} + \cdots + \beta_k x_{j,k}}$$

Proportional Hazards Model

- Why is the proportional hazard model so popular?
- The hazard for any one individual (group) is a fixed proportion of the hazard for any other individual (or group).

$$h_i(t) = h_0(t)e^{\beta_1 x_{i,1} + \cdots + \beta_k x_{i,k}}$$

$$h_j(t) = h_0(t)e^{\beta_1 x_{j,1} + \cdots + \beta_k x_{j,k}}$$

$$\frac{h_i(t)}{h_j(t)} = e^{\beta_1 (x_{i,1} - x_{j,1}) + \cdots + \beta_k (x_{i,k} - x_{j,k})}$$

Proportional Hazards Model

```
proc phreg data=Survival.Recid;
     model week*arrest(0)=fin age race wexp mar paro prio;
run;
```

Hazard Ratio

- If a parameter estimate is positive, increases in that variable increase the expected hazard.
- If a parameter estimate is negative, increases in that variable decrease expected hazard.
- $100 \times \left(e^{\beta} - 1\right)$ is the % increase in the expected hazard for each one-unit increase in the variable.
- e^{β} is the hazard ratio – the ratio of the hazards for each one-unit increase in the variable.

?

PARTIAL MAXIMUM LIKELIHOOD ESTIMATION

Partial Likelihood Estimation

- This is the more important piece of the work done by Sir David Cox in his original article.
- In the proportional hazards model, the likelihood can be split into two pieces:
  - 1st piece: depends on $h_0(t)$ and the parameters
  - 2nd piece: **only** depends on the parameters
- Basically, Cox disregarded the first piece and maximized the second piece.

Partial Likelihood Estimation

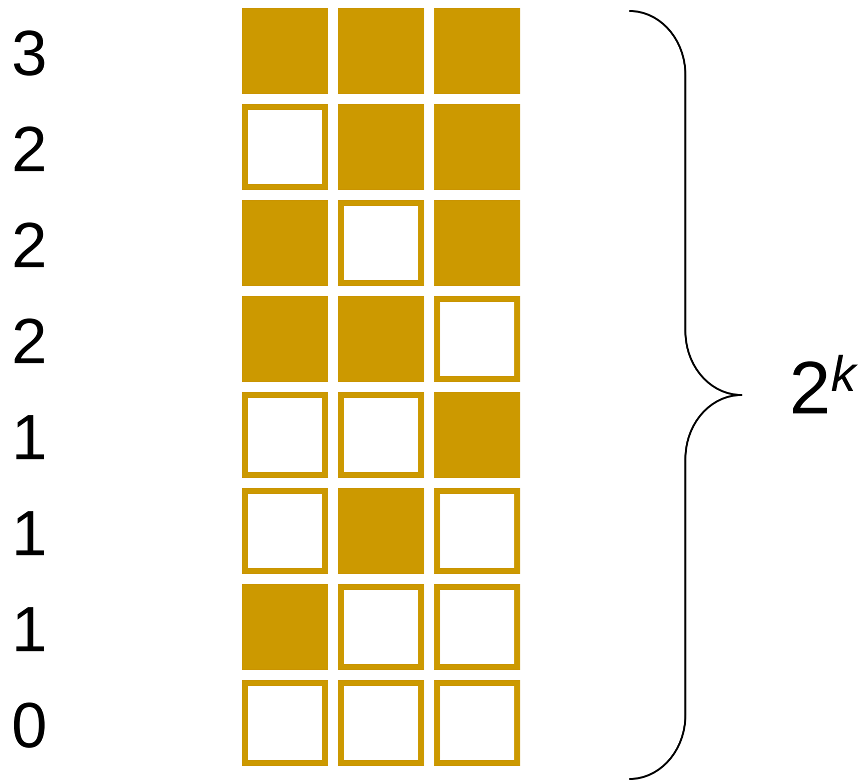
- Some information about the parameters is lost due to the partial likelihood estimation – inefficient estimates.
- Inefficiency is rather small.
- Estimates still have some desired properties:
  - Unbiased
  - Estimates can be tested in the same way as before.

AUTOMATIC SELECTION TECHNIQUES

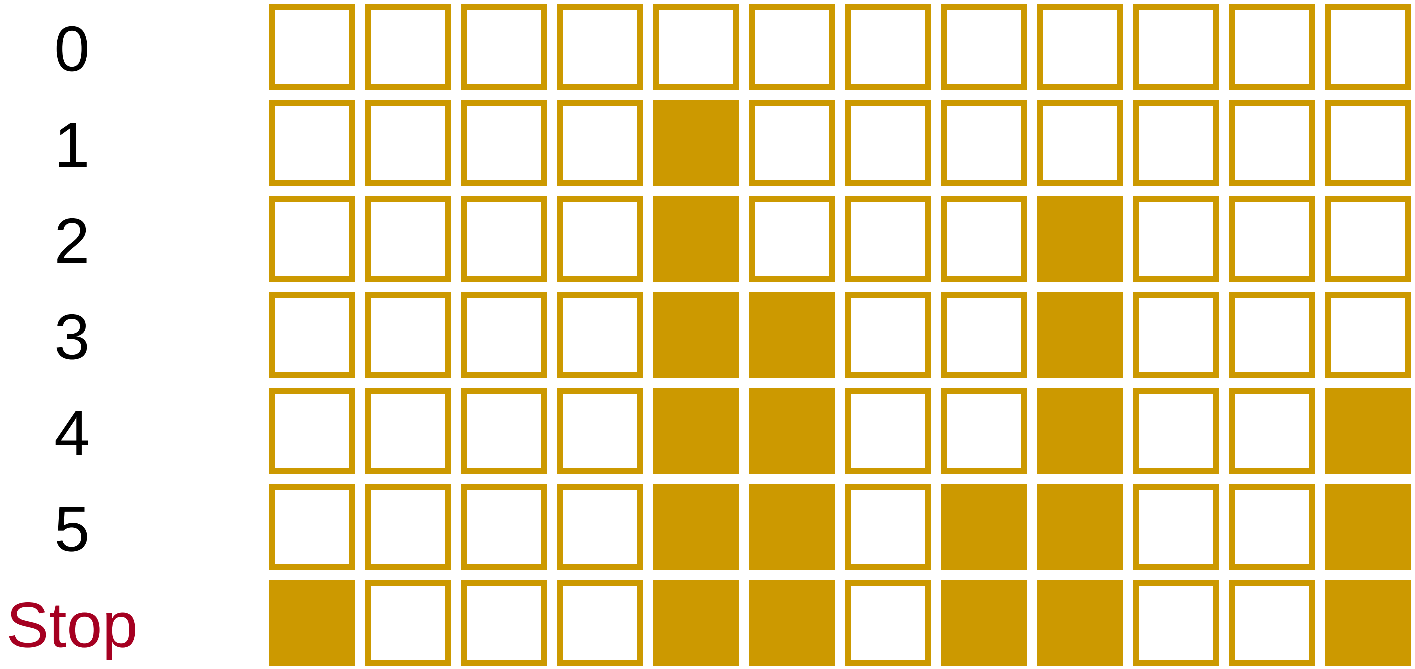
Automatic Selection Techniques

- One of the benefits of PROC PHREG is the automatic selection techniques that it employs.
- Has similar selection techniques as PROC LOGISTIC:
  - Best
  - Forward
  - Backward
  - Stepwise

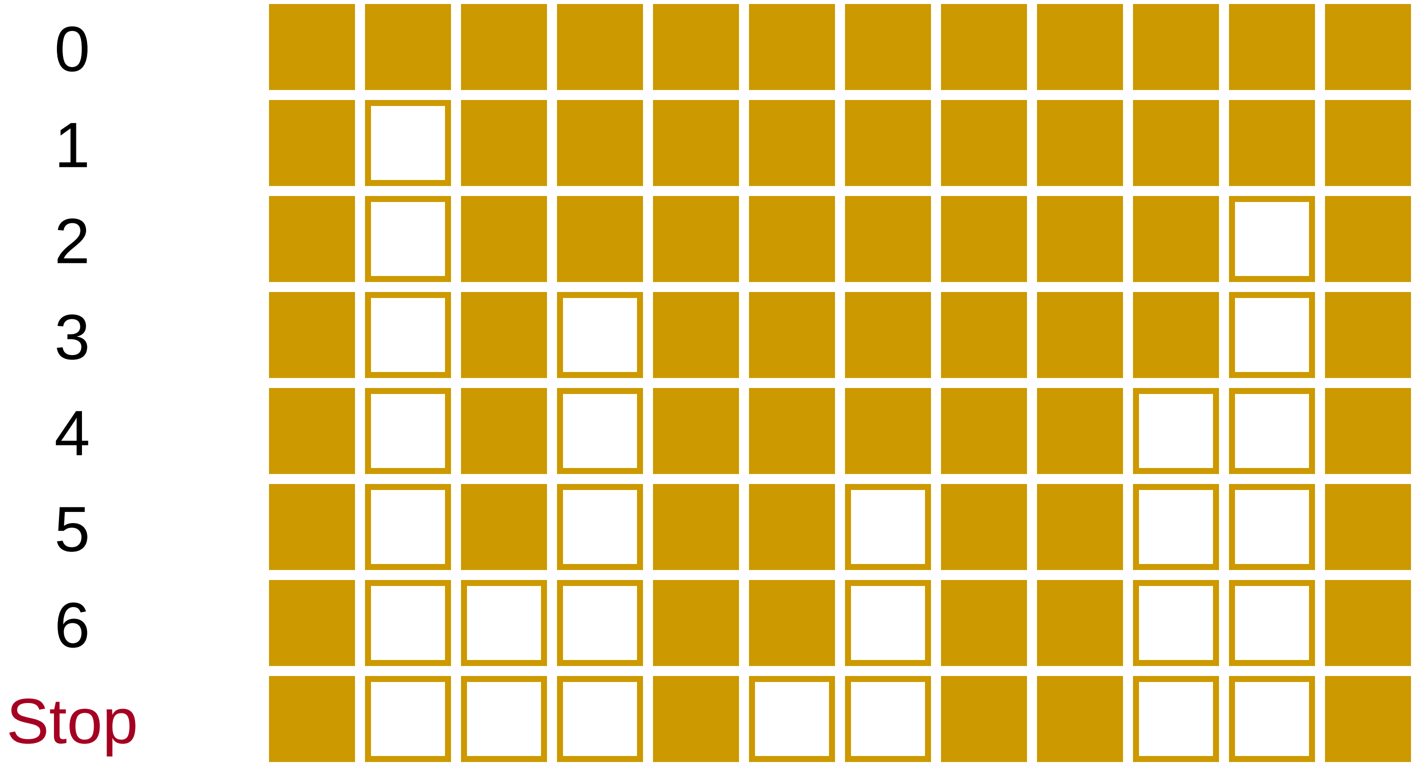
Best Subsets



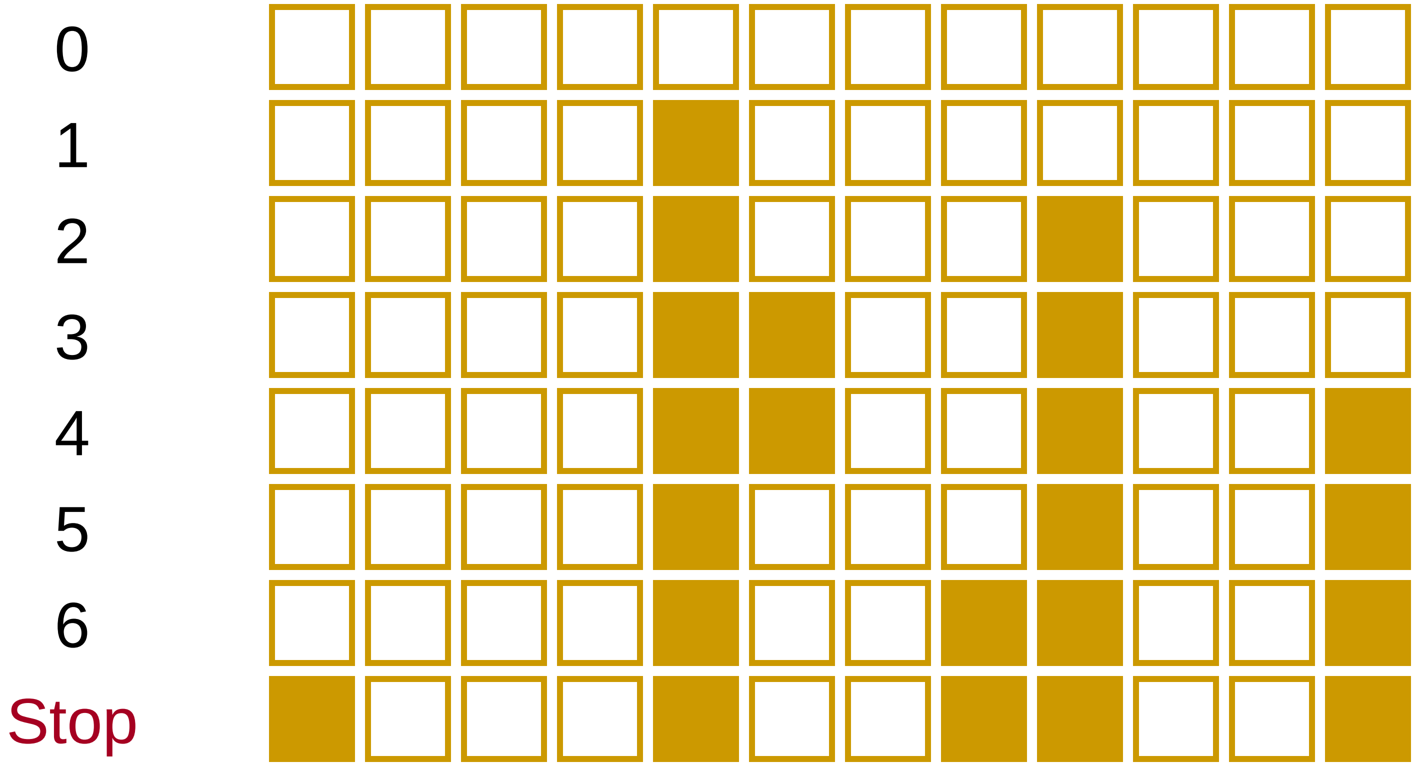
Forward Selection



Backward Elimination



Stepwise Selection



Automatic Selection Techniques

```
proc phreg data=Survival.Recid;
     model week*arrest(0)=fin age race wexp mar paro prio
                          / selection=score;
run;

proc phreg data=Survival.Recid;
     model week*arrest(0)=fin age race wexp mar paro prio
                          / selection=forward;
run;

proc phreg data=Survival.Recid;
     model week*arrest(0)=fin age race wexp mar paro prio
                          / selection=backward;
run;

proc phreg data=Survival.Recid;
     model week*arrest(0)=fin age race wexp mar paro prio
                          / selection=stepwise;
run;
```

TIME-DEPENDENT COVARIATES

Time-Dependent Covariates

- Time-dependent covariates are predictor variables that could change their value across time.
- The Cox regression model can account for these changing values of input variables.
- The following equation has one fixed variable and one time-dependent variable:

$$\log h(t) = \alpha(t) + \beta_1 x_{i,1} + \beta_2 x_{i,2}(t)$$

Time-Dependent Covariates

- Time-dependent covariates are predictor variables that could change their value across time.
- The Cox regression model can account for these changing values of input variables.
- The following equation has one fixed variable and one time-dependent variable:

$$\log h(t) = \alpha(t) + \beta_1 x_{i,1} + \beta_2 x_{i,2}(t)$$

Value stays fixed for all time points!

Time-Dependent Covariates

- Time-dependent covariates are predictor variables that could change their value across time.
- The Cox regression model can account for these changing values of input variables.
- The following equation has one fixed variable and one time-dependent variable:

$$\log h(t) = \alpha(t) + \beta_1 x_{i,1} + \beta_2 x_{i,2}(t)$$

Value changes across time!

Time-Dependent Covariates

- There are some potential problems with time-dependent variables:
  - Variables measured at different regular intervals than response variable.
  - Variables measured at irregular time intervals.
  - Variables that are undefined for certain intervals of time.

Time-Dependent Covariates

- Prisoner Recidivism Data:
  - EMP1 ~ EMP52 variables
    - Measure the full-time employment status during that week.
  - Variables measured at same regular interval as response variable week of recapture.

Time-Dependent Covariates

```
proc phreg data=Survival.Recid;
     model week*arrest(0)=fin age race wexp mar paro
                          prio employed;
     array emp(*) emp1-emp52;
     employed = emp[week];
run;
```

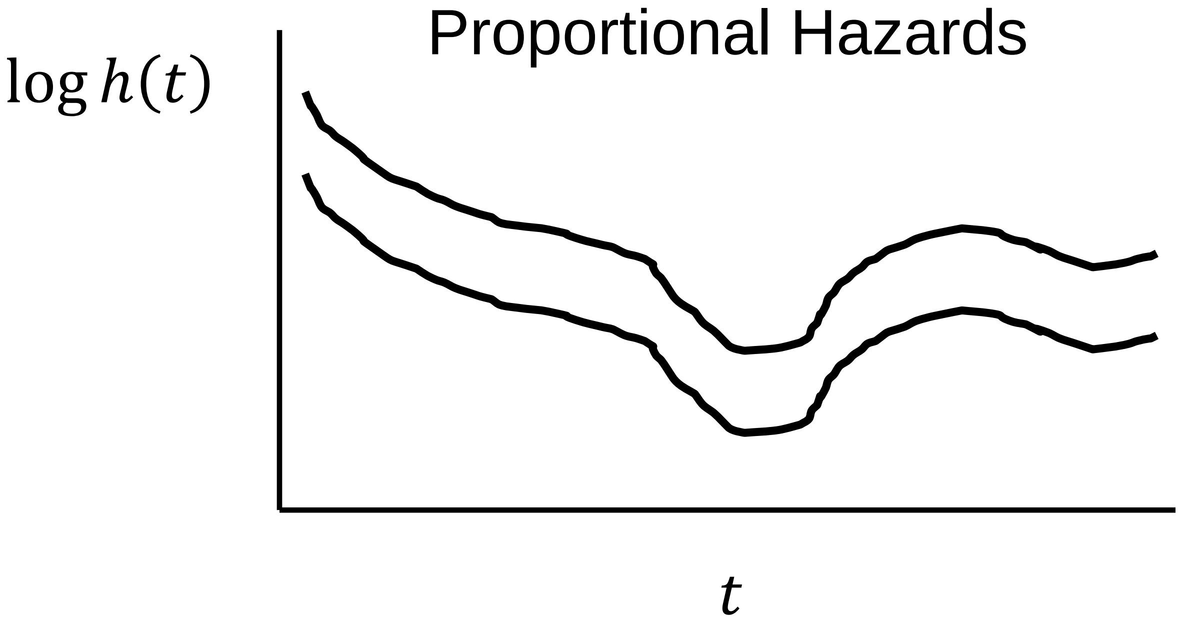
Time-Dependent Covariates

- The book goes through examples of dealing with time-dependent variables that are measured at irregular time points and that are undefined for certain intervals.
- Basic intuition is used for these calculations.

NON-PROPORTIONAL HAZARD MODELS

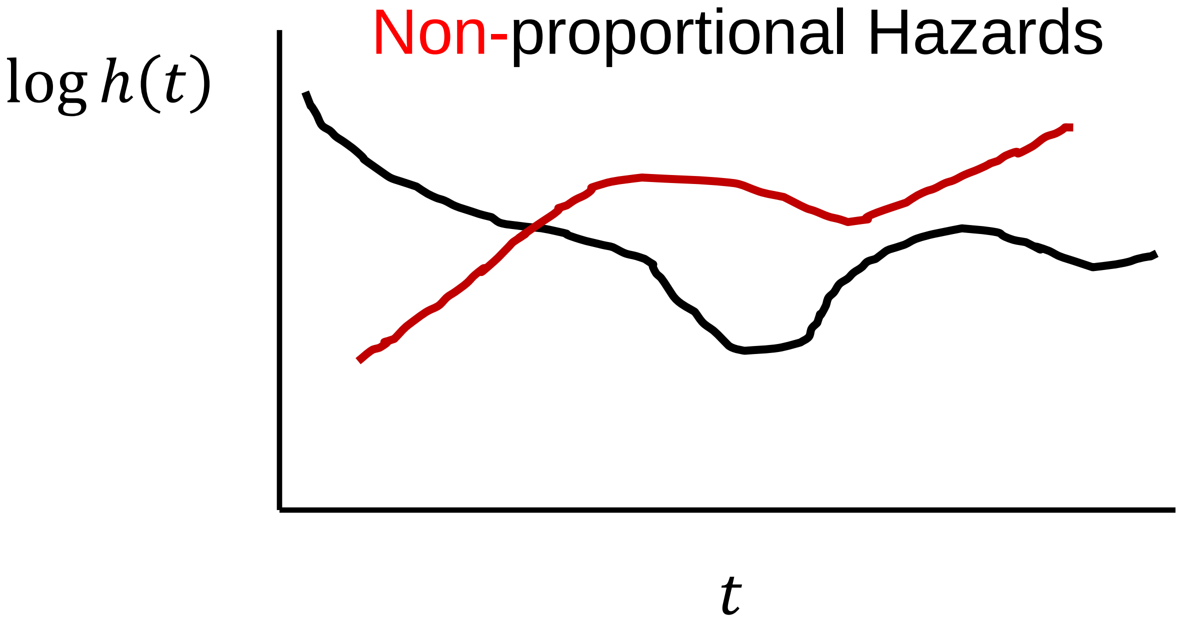
Non-proportional Hazard Models

- The Cox regression model can also be extended into the non-proportional hazards model.
- This occurs when the hazard functions between two individuals are not parallel.



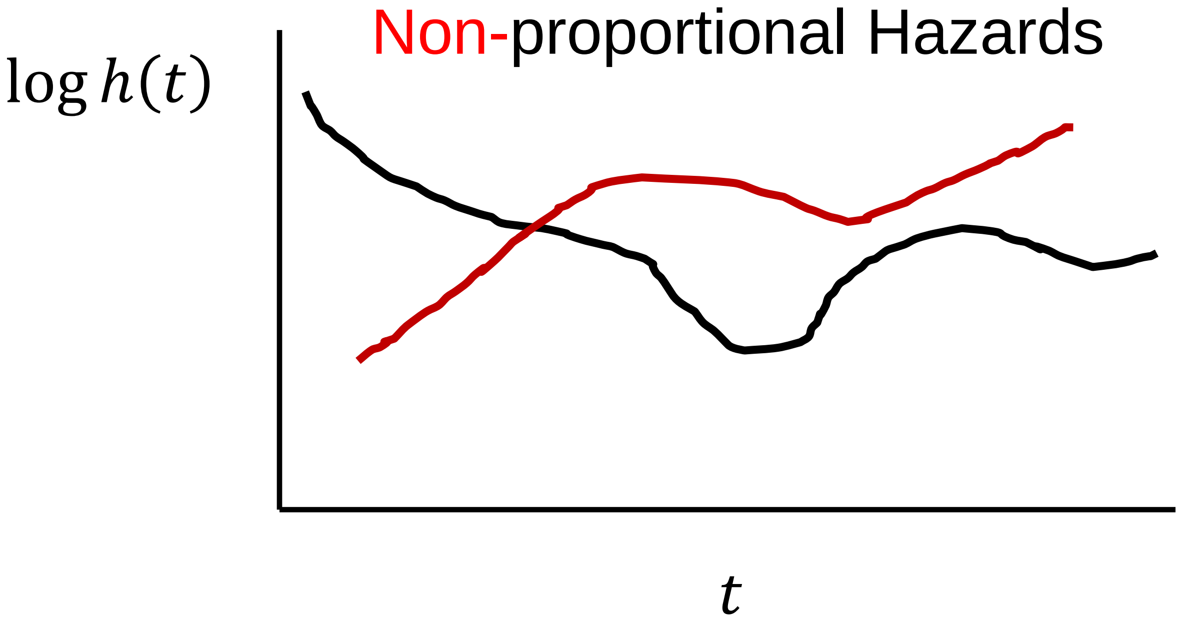
Non-proportional Hazard Models

- The Cox regression model can also be extended into the non-proportional hazards model.
- This occurs when the hazard functions between two individuals are not parallel.



Non-proportional Hazard Models

- Non-proportional hazards are the same as interactions of one or more variables with time.



Non-proportional Hazard Models

- Non-proportional hazards are the same as interactions of one or more variables with time.
- An example of this would be the following equation:

$$\log h(t) = \alpha(t) + \beta_1 x + \beta_2 x t$$

or

$$\log h(t) = \alpha(t) + (\beta_1 + \beta_2 t) x$$

Non-proportional Hazard Models

```
proc phreg data=Survival.Recid;
     model week*arrest(0)=fin age race wexp mar paro
                          prio fintime;
     fintime = fin*week;
run;
```

RESIDUAL ANALYSIS

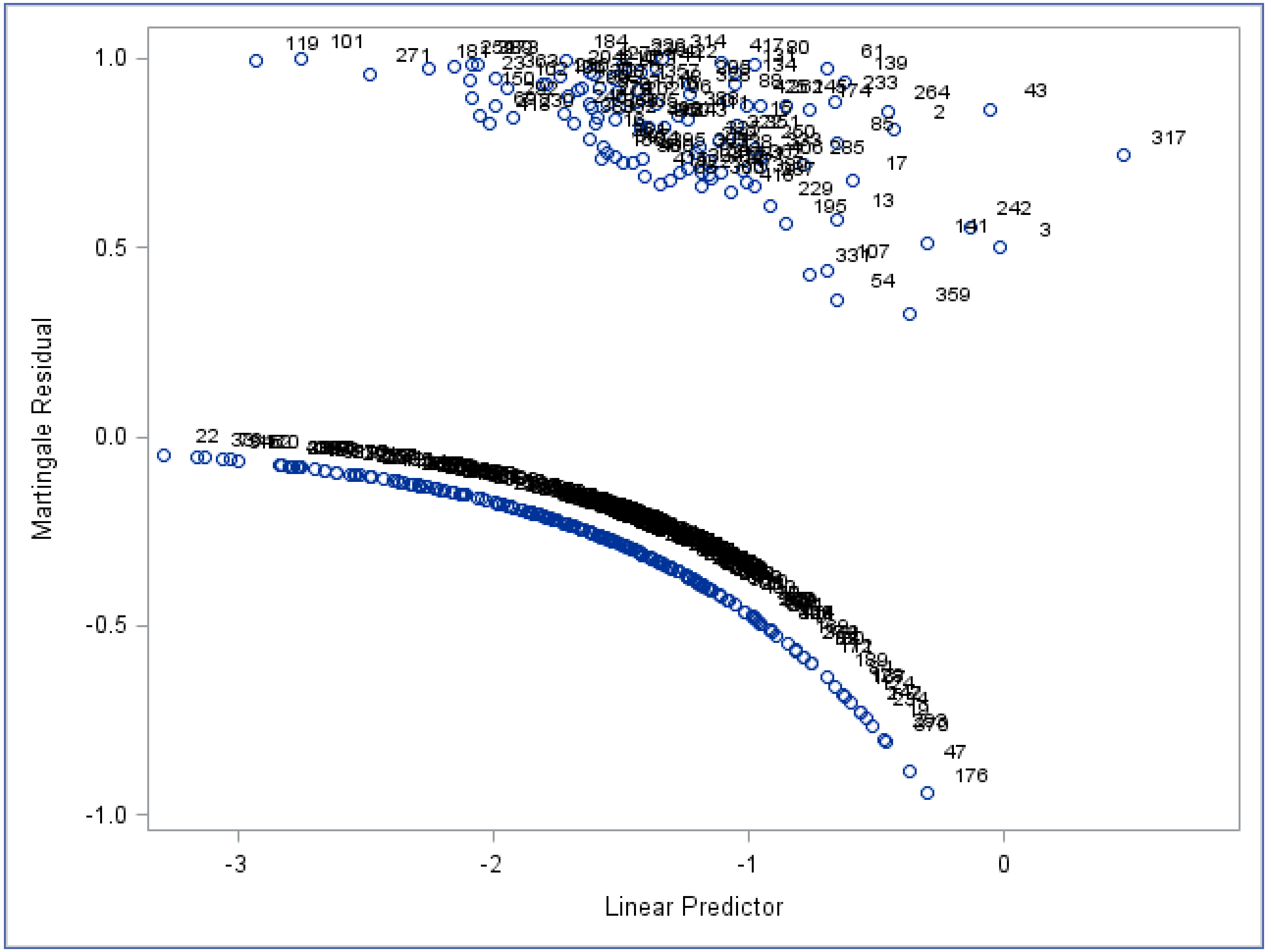
Residuals and Influence Statistics

- Similar to all other types of analysis we have done, we need to check for outliers and influential observations.
- We will discuss three types of residuals:
  - Martingale Residuals
  - Deviance Residuals
  - Schoenfeld Residuals

Martingale Residuals

- Martingale residuals are the difference between the observed number of events and the expected number of events at a specific point in time.
- These are **not** symmetrical around zero!

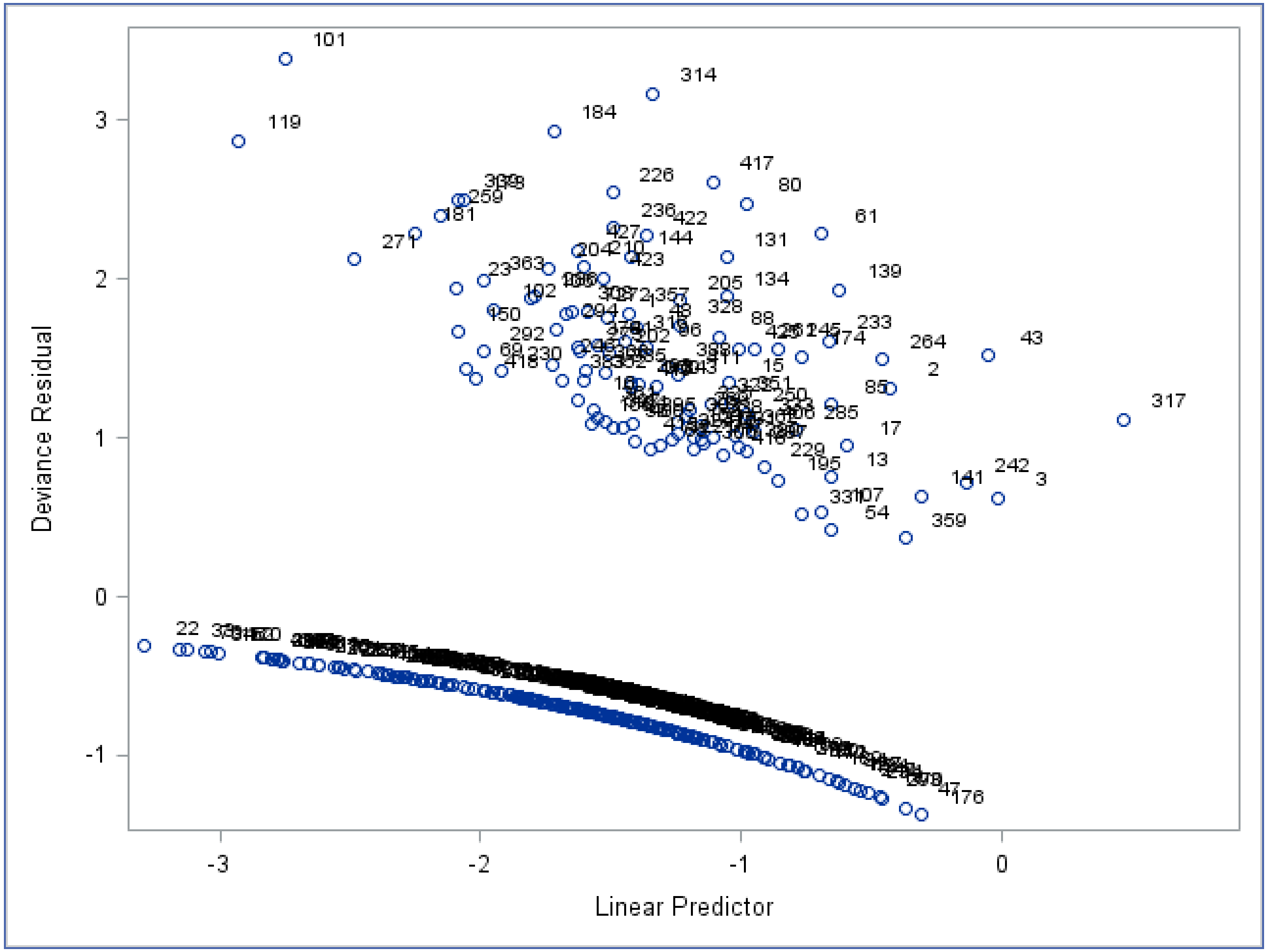
Martingale Residuals



Deviance Residuals

- Deviance residuals transform the martingale residuals into a more symmetric distribution.
- Values above 3 should be investigated.

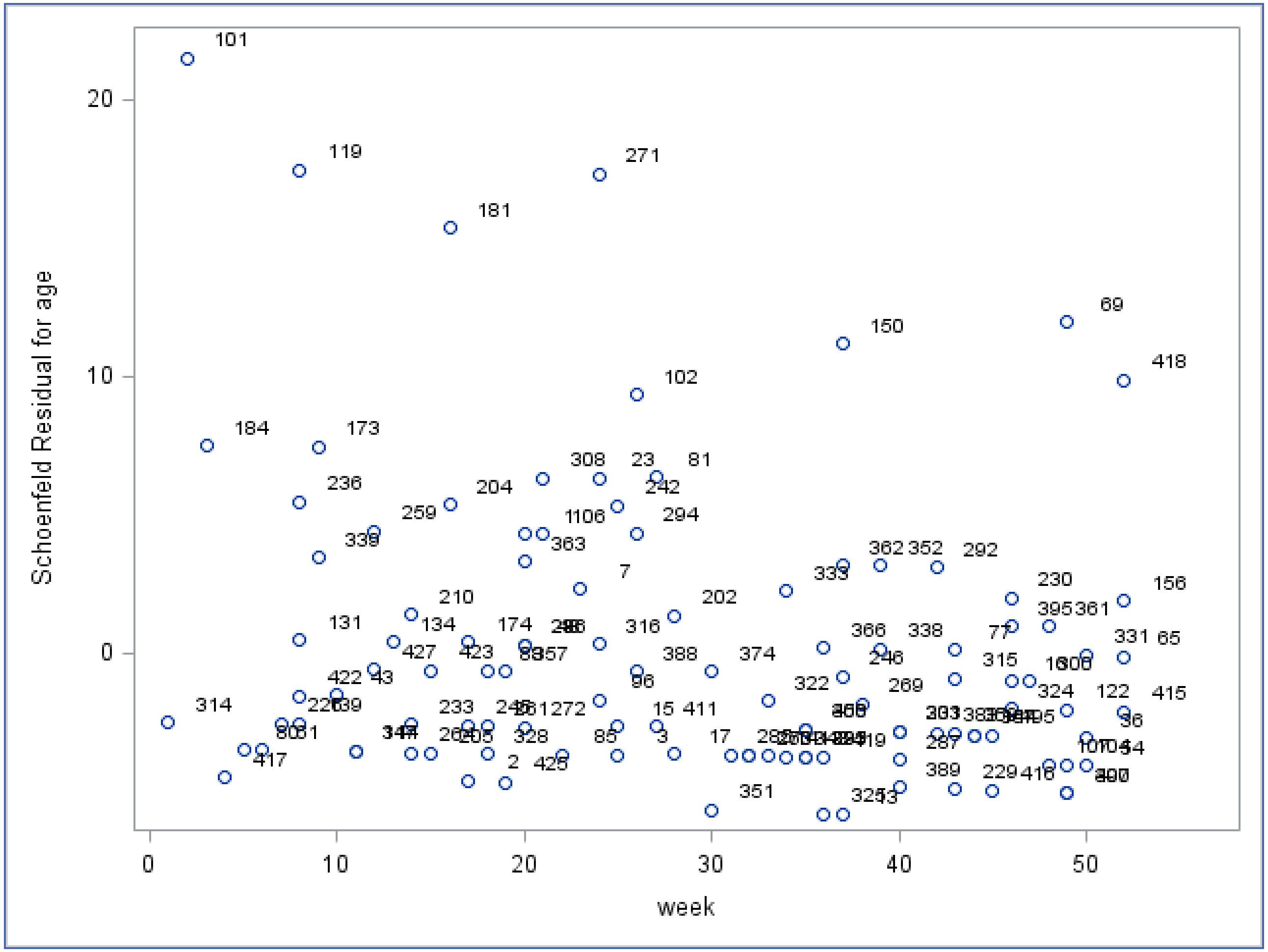
Deviance Residuals



Schoenfeld Residuals

- Schoenfeld residuals are calculated for each variable for each individual.
- They are the difference between the actual value of the variable and the expected value for someone who had the event occur at that time.

Schoenfeld Residuals



Influence Statistics

- In the Cox regression model there are two common techniques to measuring the influence of individual observations on the model:
  - Likelihood displacement
  - DFBETA

Residuals & Influential Observations

```
proc phreg data=Survival.Recid;
     model week*arrest(0)=fin age prio;
     output out=outres xbeta=xb resmart=mart
                       resdev=dev
                       ressch=schfin schage schprio ld=ld
                       dfbeta=dfbfin dfbage dfbprio;
run;
```