Proportional Hazards Models

Central Idea

What if instead of a **mean model** for Y, we directly model **the hazard function**?

Intuitive Idea

What if, inspired from GLMs, we consider the **log hazard** to be a linear model?

$$\log h_i(t) = x_i'\beta = \beta_0 + x_i'\beta.$$

This gives that

$$h_i(t) = \exp(\beta_0)\exp(x_i'\beta).$$

Relaxing the Restriction

Note that this looks like a **baseline hazard** which exists for $x_i = 0$, that is **scaled multiplicatively** based on the coefficients.

Relaxing the Restriction

Note that this looks like a **baseline hazard** which exists for $x_i = 0$, that is **scaled multiplicatively** based on the coefficients.

The **baseline hazard** is $\exp(\beta_0)$ which is constant, and thus implies an **exponential distribution**.

What if we replaced that?

Formally

A **proportional hazards** model is a model for the hazard function directly which takes the form

$$h_i(t; x_i) = h_0(t) \exp(x_i' \beta).$$

Here $h_0(t)$ is **any** baseline hazard ratio, corresponding to the hazard for $x_i = 0$.

Note, there can be **no intercept** contained in the linear predictor.

Interpretation and Model Fitting

Hazard Ratios

Consider two individuals, i and i', with $x_i = x_{i'}$ except for $x_{ij} = x_{i'j} + 1$.

The **hazard ratio** gives

$$\frac{h_i(t|x_{ij} = x + 1)}{h_{i'}(t|x_{i'j} = x)} = \exp(\beta_j).$$

This provides a mechanism to **interpret** the coefficients.

Likelihood Expression

Recall that the likelihood can be expressed as a function of the hazard.

If a fully **parametric model** is given for $h_i(t)$, then the loglikelihood can be maximized as

$$\ell = \sum_{i=1}^{n} \left[\delta_i \log h_i(t_i) - \int_0^{t_i} h_i(s) ds \right].$$

Weakly Parametric Assumptions

If a parametric baseline hazard is **too restrictive**, we can specify a **piecewise constant** baseline.

Take $0 = a_0 < a_1 < a_2 < \cdots < a_{K-1} < a_K = \infty$, then for $k = 1, \ldots, K$,

$$h_0(t) = h_{0k} \quad a_{k-1} \leq t < a_k.$$

This implies an **exponential distribution** on each interval.

Weibull Regression

PH or AFT Model

In general **proportional hazards** and **accelerated failure time** models are incompatible with one another.

However, if $h_0(t)$ is taken to be **Weibull**, then the model can be interpreted as either a PH or an AFT model.

PH or AFT Model

In general **proportional hazards** and **accelerated failure time** models are incompatible with one another.

However, if $h_0(t)$ is taken to be **Weibull**, then the model can be interpreted as either a PH or an AFT model.

If β are the coefficients for the PH model, and β^* for the AFT model, then

$$\beta^* = -\frac{\beta}{\kappa},$$

based on a Weibull(ρ, κ) distribution.

Summary

- **Proportional hazards models** are models that directly specify a form for the hazard function.

Summary

- **Proportional hazards models** are models that directly specify a form for the hazard function.
- They are based on a **parametric baseline hazard** function, which is scaled multiplicatively by the covariates.

Summary

- **Proportional hazards models** are models that directly specify a form for the hazard function.
- They are based on a **parametric baseline hazard** function, which is scaled multiplicatively by the covariates.
- **Weakly parametric** baseline hazards are possible for added flexibility.

Summary

- **Proportional hazards models** are models that directly specify a form for the hazard function.
- They are based on a **parametric baseline hazard** function, which is scaled multiplicatively by the covariates.
- **Weakly parametric** baseline hazards are possible for added flexibility.
- A PH model can be fit using **maximum likelihood** estimation.

Summary

- **Proportional hazards models** are models that directly specify a form for the hazard function.
- They are based on a **parametric baseline hazard** function, which is scaled multiplicatively by the covariates.
- **Weakly parametric** baseline hazards are possible for added flexibility.
- A PH model can be fit using **maximum likelihood** estimation.
- PH and AFT models are generally incompatible, except for **Weibull regression** which can be seen as both.