Contents

- Generally model
  $h(t) \to h(t;\underline{x}) = h_0(t)\exp\{\underline{\beta}'\underline{x}\}$
  where $h_0(t)$ is the 'underlying' hazard rate
  - β_j reflects the effect of x_j on survival
    - if $\beta_j>0$: x_j ↗ ⇒ hazard ↗ ⇒ survival prospect ↘
    - if $\beta_j<0$: x_j ↗ ⇒ hazard ↘ ⇒ survival prospect ↗
    - if $\beta_j=0$: x_j ⇒ no effect on survival.

- **Exponential Model**
  - $h(t;\underline{x}) = \lambda\exp\{\underline{\beta}'\underline{x}\}$
    - $f(t;\underline{x}) = \lambda\exp\{\underline{\beta}'\underline{x}\}\exp\{-\lambda t\exp\{\underline{\beta}'\underline{x}\}\}$
    - $S(t;\underline{x}) = \exp\{-\lambda t\exp\{\underline{\beta}'\underline{x}\}\}$
- **Weibull Model**
  - $h(t;\underline{x}) = \lambda\gamma t^{\gamma-1}\exp\{\underline{\beta}'\underline{x}\}$.
- Estimation by [numerical] ML

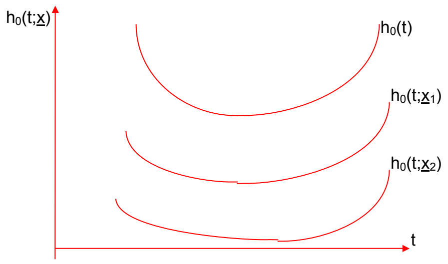
- **Proportional Hazards Model**
  - semiparametric model [Cox, 1972]
  - $h(t,\underline{x}) = h_0(t)\exp\{\underline{\beta}'\underline{x}\}$
  - $h_0(t)$ — baseline hazard
    - i.e. hazard of a patient with $\underline{x}=\underline{0}$
  - Useful for investigating prognostic factors when actual survival distribution not of immediate interest.

- $h(t;\underline{x}) = h_0(t)\exp\{\underline{\beta}'\underline{x}\}$
  - Dependence on factors and covariates is precisely modelled
  - Distribution of failure not specified
    - $h_0(t)$ is not specified
  - Useful in medical situations
    - important to know which prognostic variables have an effect and to what extent

- Can give answers such as
  - Treatment halves hazard rate
  - Smoking trebles hazard rate
- But not answers such as
  - Red wine increases survival by 10 years
  - Coffee shortens life by 6 months

- **special cases**
  - Exponential
    - $h(t;\underline{x})=\lambda \exp\{\beta'\underline{x}\}$
    - $h_0(t)=\lambda$
  - Weibull
    - $h(t;\underline{x}) = \lambda\gamma t^{\gamma-1} \exp\{\beta'\underline{x}\}$
    - $h_0(t)=\lambda\gamma t^{\gamma-1}$

- patients with covariates $\underline{x}_1$ and $\underline{x}_2$
  - $h(t;\underline{x}_1)/ h(t;\underline{x}_2)= h_0(t) \exp\{\beta'\underline{x}_1\}/ h_0(t) \exp\{\beta'\underline{x}_2\}$
  - $= \exp\{\beta'(\underline{x}_1-\underline{x}_2)\}$
  - ***Independent of t***
  - hazard functions proportional over time
  - linear component does not vary with time



- Patients have the same ‘shape’ of hazard function
  - but shifted multiplicatively according to $\underline{x}$.
- ***Hazard functions can never cross***
  - Preliminary check on KM estimates
  - Proportional hazard models are not appropriate if KM estimates clearly cross

- **Parameter Estimation**
  - Observations
    - Survival times $t_1,t_2,\ldots,t_n$
    - Censorings $\delta_1,\delta_2,\ldots\delta_n$
    - Covariates $\underline{x}_1,\underline{x}_2,\ldots\underline{x}_n$
  - $\text{likelihood} = \prod_{i=1}^{n} [h(t_i;\underline{x}_i)]^{\delta_i} S(t_i;\underline{x}_i)$
  - $S(t;\underline{x}_i) = [S_0(t;\underline{x}_i)]^{\exp\{\beta'\underline{x}_i\}}$

- Cannot maximize this without $h_0(t)$
- Instead use ***partial likelihood***
  - Consider time points at which deaths occur and P[individual i dies at time $t_{(i)}$]
  - Construct likelihood from this by a combinatorial argument

- P[it is individual i who dies at time $t_{(i)}$]

$$= \frac{h_0(t_{(i)})e^{\underline{\beta}'\underline{x}_{(i)}}}{\sum_{j\in R(t_{(i)})} h_0(t_{(i)})e^{\underline{\beta}'\underline{x}_{(j)}}}$$

$$= \frac{e^{\underline{\beta}'\underline{x}_{(i)}}}{\sum_{j\in R(t_{(i)})} e^{\underline{\beta}'\underline{x}_{(j)}}}$$

- **by proportional hazards assumption**

- And then obtain partial likelihood

$$L(\beta) = \prod_{i=1}^{n}\left\{\frac{e^{\underline{\beta}'\underline{x}_{(i)}}}{\sum_{j\in R(t_{(i)})} e^{\underline{\beta}'\underline{x}_{(j)}}}\right\}^{\delta_{(i)}}$$

- maximize this w.r.t. $\underline{\beta}$

- **Notes**
  - If no censored observations this is a ***conditional*** likelihood
    - conditional on the observed $t_{(1)},t_{(2)},\ldots,t_{(n)}$
  - With censored observations this is known as a ***partial*** likelihood
  - Cox (1975) showed that the usual likelihood methods apply in this case
    - Also known as ***Cox Regression***

- Asymptotic normality
- Asymptotic variance on second partial derivative of loglikelihood
  - i.e. treat partial likelihood just as if it was a full likelihood
- Various adjustments for ties etc.

- **Example**

Variable	Coefficient	Standard Error	χ^2 statistic (using lrt)	coeff/s.e
treatment (0=A, 1=B)	−1.42	0.64	4.89	**−2.22**
age (years)	−0.004	0.034	0.01	−0.12
sex (1=M,0=F)	0.31	0.72	0.18	0.43
volume of heart (mml)	0.0076	0.0036	4.44	**2.11**
Duration of symptoms (months)	−0.004	0.063	0.00	−0.06
digitalisation	−0.59	0.73	0.66	−0.81

- Can use χ^2 statistic or coeff/s.e.
  - χ^2 statistic if factor with > 2 levels
  - coeff/s.e. if factor with only 2 levels or continuous covariate
    - then χ^2=(coeff/s.e.)^2 for one coefficient
      - [continuous covariate of binary factor]
    - χ^2=Σ(coeff/s.e.)^2 over the k−1 dummy variables
      - k-level factor

- Treatment:
  - = |**–2.22**| >1.96
  - Good evidence of effect of treatment
  - Coeff < 0 so treatment = 1 decreases hazard, i.e. treatment B is 'better'
- Heart volume:
  - coeff/s.e. = +**2.11** >1.96
  - Increased heart volume decreases relapse time

- No evidence that other factors affect relapse time
- **NB Not shewn that other factors have no effect**
- Useful to calculate CIs:
  - 95% CI for β_3 (M/F) is 0.31 ± 2 × 0.72 =(–1.13, 1.75)
    - i.e could be large difference between M & F

Interpretation of β

- Consider model $h(t;\underline{x}) = h_0(t)\exp\{\underline{\beta}'\underline{x}\}$ $= h_0(t)\exp\{\beta_1x_1+\beta_2x_2+\ldots+\beta_kx_k\}$
  - where X_1 is a factor indicating treatment (say)
  - x_1=1 for treatment, x_1=0 for placebo
  - hazard for those on treatment is
    $h_0(t)\exp\{\boldsymbol{\beta_1}+\beta_2x_2+\ldots+\beta_kx_k\}$ &
    $h_0(t)\exp\{\quad\beta_2x_2+\ldots+\beta_kx_k\}$ on placebo
  - So ***hazard ratio*** for treatment is **$\exp\{\beta_1\}$**
    - So of interest to estimate $\exp\{\beta_1\}$ [&/or β_1]
      - With confidence intervals etc.

- **Computer implementation:**
  - Available in **R,** S-PLUS and SPSS (*not Minitab*)
  - All packages produce a table of parameter estimates and standard errors for each factor
  - **R**
    - **Construct `Surv(.)` object first then use `coxph()`**
  - **S-PLUS:**
    - **Statistics>Survival>Cox Proportional Hazards…**
  - **SPSS:**
    - **Analyze>Survival>Cox Regression**

- **Methtrex data:**
  - Data on liver survival:
    - Time variable is **FOLLOWUP**
    - Censoring in **STATUS**
    - Various covariates and prognostic factors
    - **TREATMNT** (0=placebo, 1=methtrex)
    - **MAYO** is a key covariate


```
> library(survival)
Loading required package: splines
> attach(methtrex)
> methtrex[1:4,]
  TREATMNT STATUS FOLLOWUP     MAYO LUDWIG BILIRUBN PROTHROM ALBUMIN AGE AMA
1        0      0       28 4.796029      2       15    12.86     4.2  63   1
2        0      1       32 5.883894      2       74    12.00     4.3  61   1
3        0      1       34 4.868391      4       33    11.76     4.7  60   1
4        0      0       37 4.531851      2       16    12.20     3.6  48   1

> meth.sv<-Surv(FOLLOWUP,STATUS)
> meth.ph<-coxph(meth.sv~TREATMNT+MAYO)
> summary(meth.ph)
```