III. Why the Cox PH Model is Popular

GOOD

- Cox PH model is "robust" in the sense that it closely approximates the correct parametric model.
- Parametric models are preferred only if we're sure we have the correct model.
- Hence Cox is a safe choice.

- Always non-negative.

- To calculate hazard ratios, the baseline gets cancelled out, so you only need to estimate the coefficients.

- $h(t, X)$ and $S(t, X)$ can be estimated without specifying $h_0(t)$

From j-to-i

1/22/23 IV. ML Estimation of the Cox PH Model

$h(t, X) = h_0(t) \exp\left[\sum_{i=1}^{p} \beta_i X_i\right]$ is the Cox model. Maximum likelihood estimation is used to obtain estimates $\hat{\beta}_i$.

The likelihood function L is a mathematical expression which describes the joint probability of obtaining the data actually observed on the subjects in the study as a function of the unknown parameters in the model.

The Cox model likelihood function is a partial likelihood function in that it only considers probabilities for subjects who fail, and not those who are censored.

$$L = L_1 \times L_2 \times \cdots \times L_k = \prod_{j=1}^{k} L_j$$

where L_f = likelihood of failing at time f given survival up to this time.

- Form L from the model
- Maximize $\ln L$ by solving $\frac{\partial \ln L}{\partial \beta_i} = 0$, $i = 1, \ldots, p$

Can solve with gradient descent

Once the ML estimates are obtained, we are usually interested in carrying out statistical inferences about hazard ratios defined in terms of these estimates

V. Computing the hazard ratio

A hazard ratio is defined as the hazard for one individual divided by the hazard for a different individual.

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

If you plug in the Cox model, then you get

$$\boxed{\hat{HR} = \exp\left\{\sum_{i=1}^{p} \hat{\beta}_i (X_i^* - X_i)\right\}} \leftarrow \text{General formula.}$$

VI. Interval estimation: Interaction.

When there are no interaction terms, $\hat{\beta}_1$ is the coefficient and the 95% confidence interval is

$$\exp\left\{\hat{\beta}_1 \pm 1.96\sqrt{\hat{Var}(\hat{\beta}_1)}\right\}$$

It is difficult to calculate the standard error when there are interaction terms.

Model 3

$h(t,X) = h_0(t)\exp\{\beta_1 Rx + \beta_2 \log WBC + \beta_3 (R_x \times \log WBC)\}$

For model 3, the estimated $\hat{}$ hazard ratio reduces to
$\exp\{\underbrace{\beta_1 + \beta_3 \log WBC}_{\hat{l}}\}$

A 95% CI for $\exp[l]$ is gotten by exponentiating ~~the formula~~ for the CI for l.

$\exp\{\hat{l} \pm 1.96\sqrt{Var\,\hat{l}}\}$.

The general formula is

$l = \beta_1 + \delta_1 W_1 + \dots + \delta_k W_k$
where $X = (0,1)$ exposure variable, β_1 = coeff of X_1,
δ_j = coef of $X_1 \times W_j$, $j \in [k]$.

- Because the coefficients in the linear sum are estimated from the same dataset, the coefficients are correlated. so calculation of the variance must consider the (co)variances of the estimated coefficients.
- Computer can handle it.