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:

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

The general formula is

$$l = \beta_1 + \delta_1 W_1 + \cdots + \delta_k W_k$$

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.

VII. Adjusted Survival Curves Using the Cox PH model

1/30/2023

If there is no model to fit the survival data, use KM curves.

The Cox model allows for adjusted survival curves

$$h(t,X) = e^{\sum_{i=1}^p \beta_i X_i} \Rightarrow S(t,X) = [S_0(t)]^{e^{\sum_{i=1}^p \beta_i X_i}}$$

Estimate by $\hat{S}(t,X) = [\hat{S}_0(t)]^{e^{\sum_{i=1}^p \hat{\beta}_i X_i}}$

found by computer

• General formulas for adjusted survival curves comparing two groups.

$$\hat{S}(t, X_1) = [\hat{S}_0(t)]^{\exp[\hat{\beta}_1(1) + \sum_{i \neq 1} \hat{\beta}_i \bar{X}_i]}$$

$$\hat{S}(t, X_2) = [\hat{S}_0(t)]^{\exp[\hat{\beta}_1(0) + \sum_{i \neq 1} \hat{\beta}_i \bar{X}_i]}$$

Note that we use the mean to adjust for the other covariates.

You can average w/ all variables to get an overall survival curve.

VIII. The Meaning of the PH Assumption

The PH assumption requires that the HR is constant over time

$$\hat{h}(t, X^*) = c\,\hat{h}(t, X) \quad \forall t$$

$$\widehat{HR} = \frac{\hat{h}(t, X^*)}{\hat{h}(t, X)} = \frac{\cancel{\hat{h}_0(t)} \exp[\sum_{i=1}^{p} \hat{\beta}_i X_i^*]}{\cancel{\hat{h}_0(t)} \exp[\hat{\vec{\beta}} \cdot X]} = \exp[\hat{\vec{\beta}} \cdot (X^* - X)]$$

Let $\hat{\theta} = \exp[\hat{\vec{\beta}} \cdot (X^* - X)]$. $\hat{\theta}$ is constant

$c = \hat{\theta}$

You need PH to be satisfied for Cox to be valid.

Ch 5 & 6 discuss non-proportional hazards.

The most sophisticated approach seems to be to use the extended Cox model.

IX. The Cox Likelihood

Likelihood:
- Typically based on outcome distribution
- Outcome distn not specified for Cox model (non-parametric)
- Instead we have a partial likelihood.

Scenario:
- Gary, Larry, Barry have lottery tickets
- Winning tickets chosen at times $t_1, t_2 \ldots$
- Each person eventually chosen
- Can only be chosen once.

Q: What is the probability that the order chosen is Barry, Gary, Larry.

$$\mathbb{P} = \frac{1}{3} \times \frac{1}{2} \times \frac{1}{1} = 1/6$$

Modification:
Barry — 4 tickets
Gary — 1 ticket
Larry — 2 ticket

Q: $\mathbb{P}$ order is Barry, Gary, Larry

$$(4/7)(1/3)(2/2) = 4/21$$

- Subjects number of tickets affects probability.
- Analogously, Cox model likelihood affected by patterns of covariates.

Individual hazards analogous to number of tickets

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

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

$L = L_1 \times L_2 \times L_3$

$L_1 = \left[\frac{h_0(t)e^{\beta}}{h_0(t)e^{\beta} + h_0(t)e^0 + h_0(t)e^0 + h_0(t)e^{\beta}} \right]$

$L_2 = \left[\frac{h_0(t)e^0}{h_0(t)e^0 + h_0(t)e^0 + h_0(t)e^{\beta_1}} \right]$

Harry censored

$L_3 = \left[\frac{h_0(t)e^{\beta_1}}{h_0(t)e^{\beta_1}} \right]$

☆ X. Using Age as the Time Scale

Time-on-study vs. age at follow up.

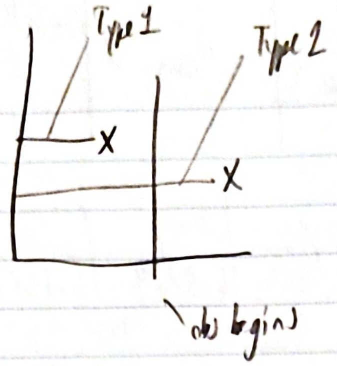
Possible choices for time 0:
- Study entry, birth, treatment time, etc.

ex) Subject enters study at age 45. Their t_0 is 45 years and their survival time is left-truncated at $t_0 = 45$

Left-truncation
- Subject not observed before t_0
- If subject has event before t_0, excluded from study
- If subject has event after t_0, then included in study and assumed not at risk until t_0.

Left truncation

Type 1	Type 2
• event before t_0	• event after t_0
• not included	• included
• effect of E underestimated	



Left censored means event did happen in past but we don't know when.

Time-on-study vs. age-at-follow-up.

If you use age-as-time-scale, then the risk set can increase as subjects are added.

Typically use time-on-study for clinical trials (this applies to my case too!)

For left-truncated data, age as time scale better

I HAVE EVERYTHING I NEED NOW TO START