CH1 Intro to Survival Analysis

- Survival Analysis

  T: event / failure happens

- Main Goal:

  $X \rightarrow$ $h(t)$ Hazard Rate

  $X \rightarrow$ $S(t)$ Survivor Func

  $X \rightarrow$ HR Hazard Ratio

  ↑ predictors

- Descriptions (refined)

$$S(t) = P(T > t)$$

$$h(t) = \lim_{\Delta \to 0} \frac{P(t \le T \le t + \Delta t \mid T > t)}{\Delta t}$$

Analogous to "speed", Range $[0, \infty)$

$$HR = \frac{h_1(t)}{h_2(t)}$$

$h_1(t) \rightarrow$ expose to X_i

$h_2(t) \rightarrow$ unexposed to X_i

Func of t

Not T !!

$$F(t) = 1 - S(t)$$ Failure func

$$f(t) = \frac{dF(t)}{dt}$$ Failure Rate

$h(t)$ v.s. $f(t)$:

$h(t)$ is conditional version of $f(t)$

- Relationships

$$h(t) = \frac{f(t)}{S(t)}$$

$$h(t) = -\frac{dS(t)/dt}{S(t)}$$

$$S(t) = \exp\left[-\int_0^t h(u)\,du\right]$$

Know one $\Rightarrow$ Know other two

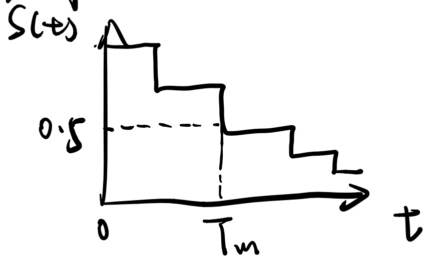
- Descriptions (coarse/overall)

  $$\bar{T} = \frac{\sum_i T_i}{N_i}$$ → # patients

  $$\bar{h} = \frac{\# \text{Failure}}{\sum_i T_i}$$

  Median survival time:

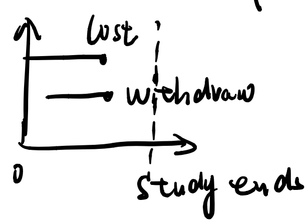
  from survival curve



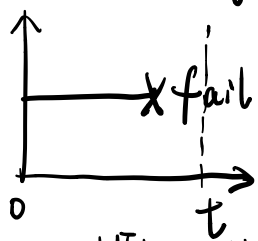
- **Censor**

- Censoring: Don't know exact failure time T

- Right censor
  (most common)

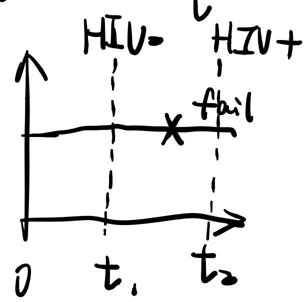


  Left censor



  only observe at 0 and t

  Interval censor



  only observe at t_1, t_2

- Assumptions
  - Random:

    censored   Not censored

$$h_{ce}(t) = h_{Nce}(t)$$

- Independence:

subgroup A $\quad h_{A,ce}(t) = h_{A,Nce}(t)$

subgroup B $\quad h_{B,ce}(t) = h_{B,Nce}(t)$

- Non-informative

$$\text{Failure} \overset{\text{Not}}{\underset{\text{informative}}{\Longleftrightarrow}} \text{Censor}$$

(e.g. A&B from same family.
A fails ⇒ B drops / censors)
Informative.

- How to deal with censored samples
contain them in the risk set until
it censors.

CH2 Kaplan-Meier Survival Curves and the Log-Rank Test.

- KM method:

Estimate $\hat{S}(t) = \dfrac{\#\text{ patients surviving passing } t}{\#\text{ patients}}$

for each t

- General KM Form:

$$\hat{S}(t_{(f)}) = \hat{S}(t_{(f-1)}) \cdot \hat{P}(T > t_{(f)} \mid T \geq t_{(f)})$$

where $\hat{S}(t_{(f-1)}) = \prod_{i=1}^{f-1} \hat{P}(T > t_{(i)} \mid T \geq t_{(i)})$

if we consider censorship, then after someone censored, it must be dropped from the Risk set.

- Log-Rank Test
  - if there are (≥2) groups. to test if their KM curves are "statistically equivalent".
  - Basic Idea:
    - Chi-square test
    - observed v.s. expected

CH5 Stratified Cox Model

- variables : PH Assump ?
  - √ : Include in the power of e
  - X : stratify

    $h_g(t, X) = h_{0g}(t)\, e^{\sum \beta_i X_i}$ (Non-Interaction)

    $g = 1, 2, \cdots, k^*$.

    β_i : same for each stractum

    $h_{0g}(t)$: Diff — — — — —

- # of variables (Z_i) ✗ PH Assump
  - = 1 : # stratum = # of value Z can take
  - > 1 : # strutum = $\prod_i$ (# of value Z_i can take)

- Interaction Between two kinds of variables ?
  - X : Non-Interaction Form :

    $h_g(t, X) = h_{0g}(t)\, e^{\sum \beta_i X_i}$

    β_i is same for Diff g !!
  - √ : Interaction Form :

$$\begin{cases} h_g(t,X) = h_{0g}(t)\, e^{\sum_i \beta_{gi} X_i} \\ \text{Alternative: } h_g(t,X) = h_{0g}(t)\, e^{\sum_i \beta_i X_i + \sum_j \alpha_j X_i Z_j} \end{cases}$$

β_i is diff for Diff g!!

β_i, α_j is same for Diff g!

- How to Determine whether there is Interaction or Not??

$\boxed{\text{Likelihood Ratio Test}}$

- How to Inference parameters β and α ??

① For Each Stratum, compute L_k using Naive partial Likelihood.

② $L = L_1 \cdot L_2 \cdot \cdots \cdot L_k$

CH6 Cox PH Model With Time-dependent Variables

- Some variables $\not\propto$ PH Assump:

① Stratification

② Consider/Make it as a time-dependent variable

- Cox pH Model with Time-dependent:
  - To test whether a variable is time-dependent or not (∝ pH Assump)
  - To Model True time-dependent variables.
- Extend Cox pH Model (Two forms)
  - To Model True time dependent variables

$$h(t, X(t)) = h_0(t) \exp\left[\sum_{i=1}^{p_1} \beta_i \mathbf{X}_i + \sum_{j=1}^{p_2} \delta_j X_j(t)\right]$$

$\mathbf{X}$ are time-invariant variables.

$X_j(t)$ is time-dependent.

- To access pH Assump/Model Time-independent variables that does not comply with pH Assump

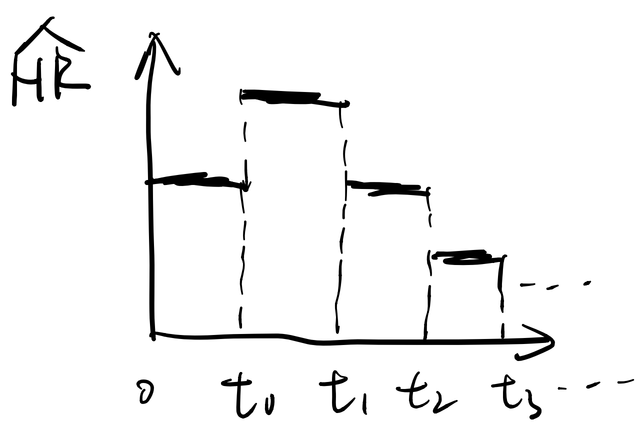
$$h(X,t) = h_0(t)\exp\left[\sum_{i=1}^{p_1}\beta_i X_i + \sum_{i=1}^{p_2}\delta_i X_i g_i(t)\right]$$

$g_i(t)$ can be:

$$\begin{cases} t \\ \log t \\ \begin{cases} 0 & t < t_0 \\ 1 & t \geq t_0 \end{cases} \quad \text{Heaviside Func} \\ \text{and other func w.r.t. } t. \end{cases}$$

- Heaviside Func:

can be used to model the Hazard ratio like



- How to Compute β's using partial ML for Cox Model with Time-dependent variables?

Still use the formula:

$$L_k = \frac{\left(e^{\sum \beta_i X_i + \sum \delta_j X_j(t)}\right)_k}{\sum_{m=1}^{N} \left(e^{\sum \beta X_i + \sum \beta_j X_j(t)}\right)_m}$$

Note that we Need to Know $X_j(t)$ for Each sample in Risk set at time t!!

- Note when compute $\widehat{HR}$, it is also Time-dependent!

CH7 Parametric Survival Models

- parametric v.s. semiparametric
  - parametric: has full description for $h(t)$

- Semiparameteric

  has some unspecified term in the final Equation.

  For Example,

  $$h(t) = \underbrace{h_0(t)}_{\text{unspecified}} e^{\sum \beta_i X_i}$$

- Why we want parameteric?
  - More consistent with theoritical Result $S(t)$
  - Simplicity
  - Completeness - $h(t)$, $S(t)$ has concrete forms.

- Three Assumptions for Survival Models we Met until Now
  - PH Assump: $\frac{h_1(t)}{h_2(t)} = \text{const}$
  - AST (Accelerated Failure Time) Assump: $\frac{T_1}{\ } = \text{const}$

T_2
failure time from 0

- PO (proportional odds) Assump:

$$\boxed{\text{Survival odds} = \frac{S(t)}{1-S(t)}}$$

$$\frac{\text{Survival odds 1}}{\text{Survival odds 2}} = \text{const}$$

- How to get specified and parameterized Equations based on Three Assump?

① specify the Form/Distribution of $h(t)$ or $S(t)$

② Compute $h(t)$ (for PH)
T (for AFT)
$\frac{S(t)}{1-S(t)}$ (for PO)

③ Use $\exp \sum \beta_i X_i$ to Reparameterize parameters in those equations so that we can merge with X_i's

④ After get the Reparameterized Form of $h(t)$ or $S(t)$, we can use Maximum likelihood and

$f(t) = h(t)S(t)$ to estimate all parameters

- Important Relationship

$$F(t) = P(T \le t)$$

$$f(t) = \frac{dF(t)}{dt}$$

$$S(t) = P(T > t) = \int_t^{\infty} f(u)\,du$$

$$h(t) = -\frac{\frac{dS(t)}{dt}}{S(t)} = \frac{f(t)}{S(t)}$$

$$S(t) = \exp\left(-\int_0^t h(u)\,du\right)$$

$$f(t) = h(t)\,S(t)$$

Know one of ($S(t)$ $h(t)$ $f(t)$), we know other two.

- Exponential Model

$$h(t) = \lambda$$

$$S(t) = \exp(-\lambda t)$$

PH Assump: $\lambda = \exp(\sum_i \beta_i X_i)$

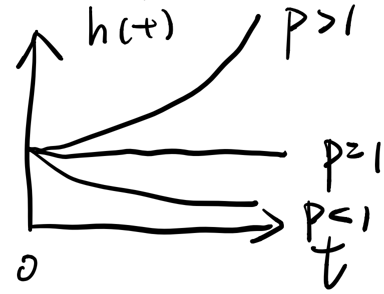
AFT Assump: $\frac{1}{\lambda} = \exp(\sum_i \beta_i X_i)$

- Weibull Model

Do not reparameterize

$h(t) = \lambda p t^{p-1}$ (p: shape parameter)

Note: $\lambda, p > 0$



$S(t) = \exp(-\lambda t^p)$

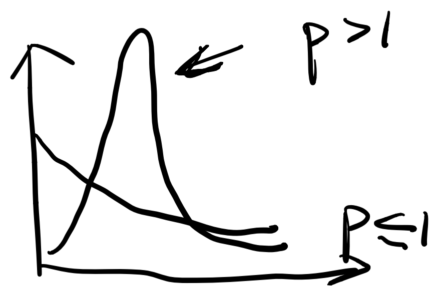
pH Assump: $\lambda = \exp(\sum \beta_i X_i)$

AFT Assump: $\frac{1}{\lambda^{1/p}} = \exp(\sum \beta_i X_i)$

$\ln[-\ln S(t)]$: linear with $\ln(t)$

- Log-Logistic Model:

$h(t) = \frac{\lambda p t^{p-1}}{1+\lambda t^p}$ $(p, \lambda > 0)$



$S(t) = \frac{1}{1+\lambda t^p}$

AFT Assump: $\frac{1}{\lambda^{1/p}} = \exp(\sum \beta_i X_i)$

po Assump: $\lambda = \exp(\sum_i \beta_i X_i)$

$\ln\left[\frac{1-S(t)}{S(t)}\right]$: linear with $\ln(t)$

- How to Infer parameters ??
  - using Parametric Likelihood

    $$L = L_1 \cdot L_2 \cdot \cdots \cdot L_N$$

  - Handles Censored data very well
    - Fail at t_0 $h(t_0)S(t_0)$

      $$L_k = f(t_0)$$

    - Right Censored:

      $$L_k = \int_{t_0}^{\infty} f(u)\,du$$

    - Left Censored:

      $$L_k = \int_0^{t_0} f(u)\,du$$

    - Interval Censored:

      $$L_k = \int_{t_1}^{t_2} f(u)\,du$$

- Assumptions:

  ① No competing Risks

  ② Subjects Independence

  ③ Follow-up time continuous.

- Frailty Models

- Individual v.s. population
  - previous h(t) only considers all predictors, but does not account for individual variation.
  - Frailty accounts for variation of all Individuals (Such as Noise in LR Model)
    - $h(t|\alpha) = \alpha h(t)$
    - $S(t|\alpha) = S(t)^{\alpha}$
    - Note $\alpha \propto g(\alpha)$:

      g ↳ Distribution

      $\mu = E(\alpha) = 1$

      $Var(\alpha) = \theta$ ⇗ Needs to be estimated.

- Population (unconditional)

  $$S_u(t) = \int_0^{\infty} S(t|\alpha)\, g(\alpha)\, d\alpha$$

  $$h_u(t) = \frac{-d[S_u(t)]/dt}{S_u(t)}$$

- Note that individual level pH Assumption might not be generalized to population level

- Inference:

  compute $f_u(t) = h_u(t) S_u(t)$

  ↑ unconditional!!

Then use ML to Estimate all β's and $Var(\alpha) = \theta$. Note that α_j for individual j cannot be estimated due to huge num of parameters! !

- Shared Frailty:

  same α_j for each cluster and population includes α_j $j = 1, \dots K$

  special case: Recurrent Events:

  Each patient has the same α for all Failure time (see individual as group)

CH8: Recurrent Event

- Two types of Recurrent Events
  - Each time has identical event
  - Diff times have diff events for the Same patient (e.g. second disease will happen only after the first disease)

- Methods for Diff types
  - Identical: Counting process
  - Diff: Stratified Cox { Stratified CP, Gap Time, Marginal
    parametric Model with shared frailty.

- Counting Process
  - Data layout:

ID	Status	Start	stop	X_1	X_2 ...

    Diff patients can have diff # of records

    ⇓

  - To apply Cox Model, we also need to lay the data based on failure times.

failure time	Risk set	# failed	# censored in $[t_f, t_{f+1}]$

- Two Assumptions
  1. CP treats each failure time event independently, even if they come from the same patient.
  2. the patient will always be at the Risk set until after his/her last failure or censorship (the patient

cannot be dropped off unless we go through all his/her records)

- Model: Cox Model

  (Not necessarily pH Model. we can also use seratified or extended model for Non pH variables and time-variant variables, Also can include interation as we leart before)

  $$h(t;X) = h_0(t)\, e^{\sum \beta_i X_i}$$

- Inference: partial Likelihood

  $$L = L_1 \cdot L_2 \cdots L_{(\text{last failure time})}$$

  where $$L_f = \frac{\sum_{m \in \text{failed samples}} \exp\left(\sum \beta_{im} X_{im}(t_{(f)})\right)}{\sum_{s \in \text{Riskset}} \exp\left(\sum \beta_{is} X_{is}(t_{(f)})\right)}$$

- Robust Estimation
  - to adjust for correlation between records from the same object
  - Not adjust for $\hat{\beta}_i$, instedd, for $Var(\hat{\beta}_i)$
  - $\hat{R}(\hat{\beta}) = \hat{Var}(\hat{\beta})[\hat{R}_s' \hat{R}_s] \hat{Var}(\hat{\beta})$

- Stratified Cox
  - Assumption: treat time interval # as stratum (e.g. stratum 1: the first time interval data of all patients. stratum 2: the second time interval data ...)
  - Model: Stratified Cox Model where stratum is time interval # which one int
  - Three Methods:
    Same: Model & Inference
    Diff: Risk set for stratum t (see P_{382})

- Stratified CP:
  - Stratum 1 $\xrightarrow{\text{affect}}$ stratum 2 $\xrightarrow{\text{affect}}$ stratum 3 $\xrightarrow{\text{affect}}$ .... $\xrightarrow{\text{affect}}$ stratum N
  - The Influence from stratum t to $t+1$ is:
    only when the patient finish the stratum t (go through the entire time interval), it can be counted into the Risk set of stratum $t+1$

∵ To determine the Risk set at each time for stratum f+1 we need to be cautious whether the patient ends stratum f at this time t.

- ==Gap Time==
  - Focus more on time gap between two consecutive event.

    ∵ start time: always 0
    stop time: the time interval length since the previous event.

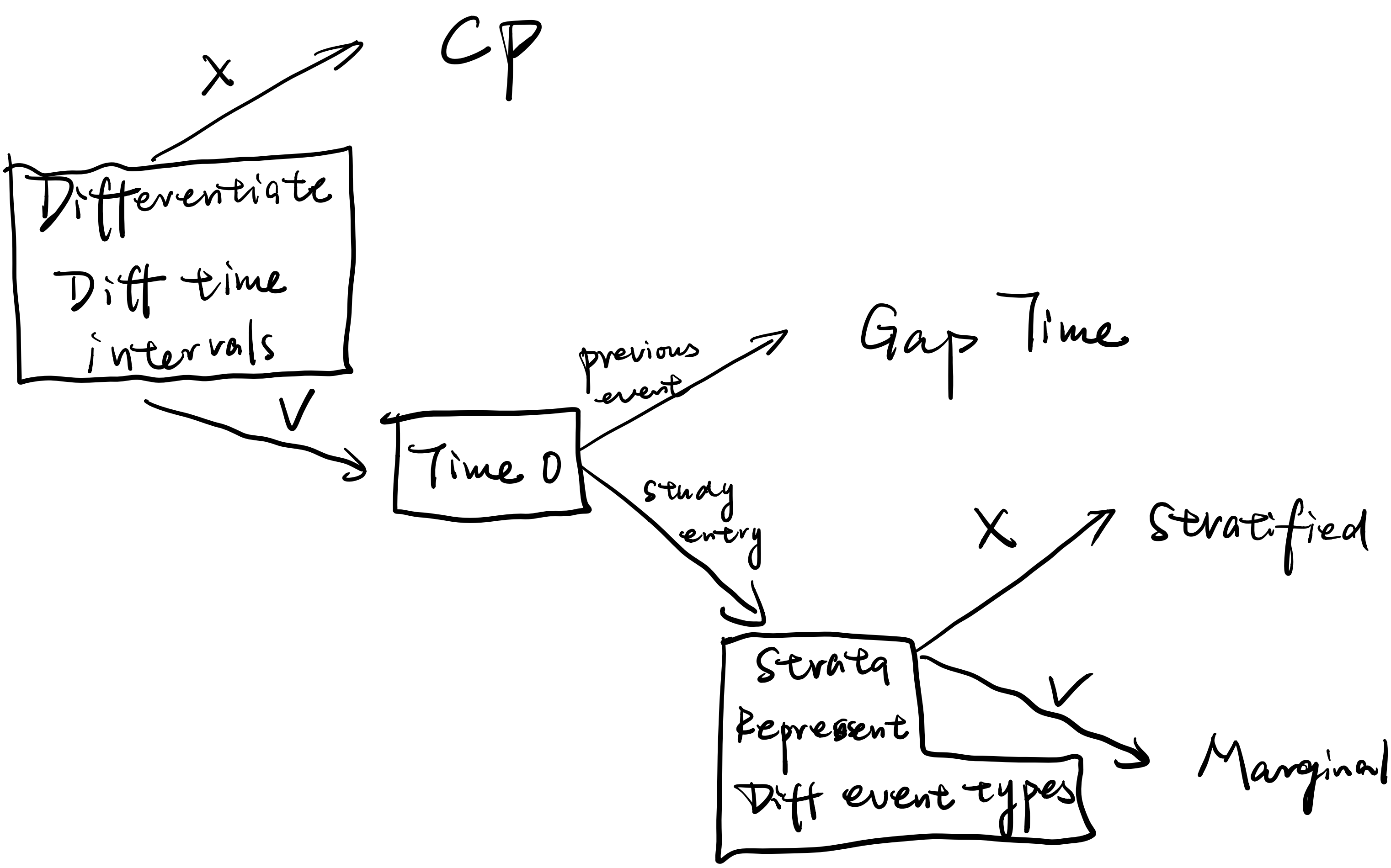
    ∵ for each stratum, it reset the start time and use the same method to determine the Risk set

- ==Marginal==
  - start time for each stratum is always the time

patients enter the study.

- It considers each stratum as a separate process
  ∴ Don't care about the relative order between diff strata.

- How to choose which model to use



- Parametric Model with Shared frailty

$$h_i(t \mid \alpha_i, X_i) = \alpha_i \, h(t \mid X_i)$$

Diff Records from the same patient

share the same α_i

- Survival Curves with Recurrent Events.
  - one plot for one stratum

    $S_1(t) = P(T_1 > t)$

    $S_2(t) = P(T_2 > t)$ survival time up to the 2nd event

    $\vdots$

  - Two types
    - stratified: use survival time from time from (k-1)th to kth event, restricting data to subjects with k-1 events
    - Marginal: use survival time from the study entry to k-th event, ignoring previous events.
  - Then after get event time and Risk set, we can use KM method to plot each Survival Curve.

(CH2 continue)

- Two group case:

  H_0: No diff between two group curves

  Log Rank test statistics:

$$Z = \frac{\sum_{i=1}^{2} (O_i - E_i)}{\sqrt{\sum_{i=1}^{2} Var(O_i - E_i)}} \sim \chi^2 (1\, df)$$

Expected:

proportion in total risk set

\# total failure

group 1: $e_{1f} = \left(\frac{n_{1f}}{n_{1f} + n_{2f}} \right) \times (m_{1f} + m_{2f})$

$$e_{2f} = \left(\frac{n_{2f}}{n_{1f} + n_{2f}} \right) \times (m_{1f} + m_{2f})$$

understand: for expectation. we expect they have the same failure rate. therefore. the expected version is that their failure # should be related to the proportion of each group over total samples

Residual:

$$O_i - E_i = \sum_{f=1} \underline{(m_{if} - e_{if})}$$

diff between observed and expected # failure

Variance:

$$Var(O_i - E_i) = \frac{n_{1f} n_{2f} (m_{1f} + m_{2f})(n_{1f} + n_{2f} - m_{1f} - m_{2f})}{(n_{1f} + n_{2f})^2 (n_{1f} + n_{2f} - 1)}$$

same for $i = 1, 2$

- Several Group Case
  - log rank statistics $\sim \chi^2$ with $\underline{G} - 1$ df
    G ← # groups.
  - Details see P96

- Approximate case

$$\chi^2 \approx \sum_{i=1}^{\#\text{groups}} \frac{(O_i - E_i)^2}{E_i}$$

- Stratified case:
  each group is divided into same strata.

  Diff between previous version:

$$O_i - E_i = \sum_{s} \sum_{f} (m_{ifs} - e_{ifs})$$

stratum failure time

- Alternative to Log Rank Test
  - limitation to Log Rank:
    each time has same weight
  - Alternative:
    Wilcox, Tarone-Ware. Peto.
    Flemington Harrington.

- CI for KM curves

  95% CI: $\hat{S}_{KM}(t) \pm 1.96\sqrt{\hat{Var}[\hat{S}_{KM}(t)]}$

  where $\hat{Var}[\hat{S}_{KM}(t)] = (\hat{S}_{KM}(t)) \sum_{f: t_{(f)} \le t} \left[\frac{m_f}{n_f(n_f - m_f)}\right]$

  # failure

  # in risk set

  Cummulative summation.

- CI for the Median survival time

  find observed failure times satisfy:

  $$(\hat{S}_{KM}(t) - 0.5)^2 < 3.84\, \hat{Var}[\hat{S}_{KM}(t)]$$

CH3 PH Model

- why pH Model
  - vs. KM: Involve Math model which can connect with predictors X_i
  - Robust and Safe: if don't know the true parametric model, then use PH Model can closely approximate the true resulte
  - Semi-parametric: don't fully specify the form of baseline func.

- Cox PH Model

No t

predictor variable

$$h(t,X) = \underbrace{h_0(t)}_{\text{Baseline func: no } X_i\text{'s}} e^{\sum_i \beta_i X_i}$$

$$S(t,X) = [S_0(t)]^{e^{\sum_i \beta_i X_i}}$$

$$HR = \frac{h(t,X^*)}{h(t,X)} = e^{\sum_i \beta_i (X_i^* - X_i)}$$

- Two special cases
    - Adjust for a covariant E (especially for confounders)

$$h(t,X) = h_0(t)\exp[\underbrace{\beta E}_{\text{adjusted}} + \underbrace{\textstyle\sum_i \delta_i W_i}_{\text{other predictors}}]$$

$\because$ Ajusted verson is still a func of all other predictors W_i and t

if we want to plot adjusted curve for each adjusted value, then we can specify other predictors $W_i = \overline{W_{ij}}$ or $W_i = \text{median}(W_{ij})$, where they are over all samples in the dataset.

    - Interaction between X_i's

$$h(t,X) = h_0(t)\exp[\sum_i \beta_i X_i + \sum_{j,k} \delta_j X_j X_k]$$

- pH Assumption

$$\hat{HR} = \exp[\sum_i \beta_i (X^*_i - X_i)]$$

should be constant w.r.t t

- **partial Likelihood**

$$L = \prod_{f=1}^{N} L_f$$ ← failure time

$$L_f = \frac{\sum_{i \in \text{Failure set}} h(f, x_i)}{\sum_{j \in \text{Risk set}} h(f, x_j)}$$

Note that Risk set at f includes

① samples whose $T \geq t$

② censored samples should be dropped from the Risk set

Interpretability of L_f:

at f, the Likelihood that among all samples in the Risk set, how likely $i \in$ Failure set fails.

- Estimate β_i's

$\frac{\partial \ln L}{\partial \beta_i} = 0 \Rightarrow$ should use iterative method to compute e.g. SGD. Newton, etc

(1) Diff initialization

(2) $Var[\beta_i]$ exists.

- Note: sometimes, due to diff ways to treat time, we need to pay attention to include which patients are in the Risk set. (e.g. time-on-study v.s. Age-as-time)

CH4 Test PH Assumption

- graphical
  - log-log plot
  - observed vs. expected
- statistical
  - The Goodness of Fit
  - Time-Dependent covariant.

- log-log plot
  - Intuition: $S = [S_0(t)]^{e^{\sum \beta_i x_i}}$

$$-\ln[-\ln S(t,X_1)] + \ln[-\ln S(t,X_2)]$$

$$= \sum_i \beta_i (X_{1i} - X_{2i})$$

nothing to do with t.

Range: $(-\infty, +\infty)$

$\Rightarrow$ $-\log[-\log \hat{S}]$ should parallel on $\ln[\ln \hat{S}] - t$ plot

- Method 1
  - examin variables one by one
  - For each variable,
    - categorical
    - Continuous — transfer (result in small # of categories)

plot KM curve $\hat{S}(t)$ stratified on the variable we want assess, if $\hat{S}(t)$ are parallel between all values of this variable $\Rightarrow$ pH satisfied

- what is parallel:

  assume pH is OK unless strong evidence of nonparallel.

- Method 2
  - Examine several predictors at same time
  - Stratify on $\prod_i$ (# values X_i can take)

    $\Downarrow$ drawback: too small for some stratum.

  - Alternative:

    1. when assessing for X_i, assume other variables satisfy pH Assump

    2. Fit $h_g(t, X) = h_{0g}(t)\, e^{\sum_{j \neq i} \beta_j X_j}$ for each stratum - X_i

       since we assume X_j ($j \neq i$) satisfy pH

    3. plot log-log $\hat{S}$ derived from $\boxed{h_g(t, \bar{X}) = h_{0g}(t)\, e^{\sum_{j \neq i} \beta_j \bar{X}_j}}$ parallel $\Rightarrow$ pH Assump ✓

       $\bar{X}_j$: over all data

- Observed v.s. expected
  - Method 1 (one-at-a-time)
    - observed: KM curve for each

value the assessed variable can take

- expected:
  - categorical:

    fit $h_0(t)e^{\beta X}$ on overall data

    $\downarrow$ var we want to assess

    and plot $\hat{S}(t, X_i) = [\hat{S}_0(t)]^{e^{\beta X_i}}$

    for all values X can take

  - Continuous:

    Continuous $\Rightarrow$ categorical

    option 1:

    $$h(t, X) = h_0(t)\exp\left[\sum_{i=1}^{k-1} \beta_i X_{ci}\right]$$

    X_{ci}: dummy var X_c can take

    (# X_{ci} can take is k)

    option 2:

    $$h(t, X) = h_0(t)\exp(\beta X)$$

    fit on all data

    then for each category:

    $$S(t, X) = [S_0(t)]^{e^{\beta \bar{X}_c}}$$

    $\downarrow$ mean value within category c

- if observed curves $\Updownarrow$ close expected curves

  $\Longrightarrow$ pH satisfied

- GOF Testing

  use Schoenfeld Residual

  (See P 181 → 183)

- Test for time-dependent

  - For one variable

    $$h(t,X) = h_0(t)\exp[\beta X + \delta(X \cdot g(t))]$$

    $H_0: \delta = 0$

    Wald statistics or likelihood ratio statistics : χ^2 with 1 df under H_0

  - several predictor at the same time

    $$h(t,X) = h_0(t)\exp[\sum_i \beta_i X_i + \delta_i X_i g_i(t)]$$

    $H_0: \delta_1 = \delta_2 = \cdots = \delta_K = 0$

    LR stats $= -2\ln L_{pH} - (-\ln L_{\text{extended pH}})$

$\sim \chi^2_p$ under H_0

if p is significant $\Rightarrow$ H_0 X

$\Rightarrow$ eliminate variable one-by-one to find which var $\not\propto$ pH Assump.

- Assess pH for a given predictor adjusted for other predictors which already satisfy pH Assump

want to assess ↑

$$h(t,X) = h_0(t)\exp\left[\sum_{i=1}^{p-1}\beta_i X_i + \beta^* X^* + \delta^*(X^* \times g(t))\right]$$

$H_0: \delta^* = 0$

wald/LR statistics $\sim \chi^2$ with 1 df

** To test PH Assump, should use at least two methods.

CH9 Competing Risks

2 Method → Separate Model for diff Event Types (assume Indep)
→ Lunn-McNeil (LM) Approach. (X assume Indep)

3 plots → KM (assume Indep)
→ CIC (X assume Indep)
→ CPC

- Competing events:
  only one event can happen among all possibilities (up to now we don't assume Recurrence, consider it as modeling on only the first failure)

- Goal:
  ① Get Failure rate for $X_1, X_2 \cdots X_K$
  ② Compare HR_A vs. HR_B

Method 1: Separate Models for Diff Event Types

Idea: each event type is indep, treat others as censored.

Step 1: use Cox PH Model to estimate separate hazard for each failure type

$$h_c(t) = \lim_{\Delta t \to 0} \frac{P(t \le T_c \le t+\Delta t \mid T_c \ge t)}{\Delta t}$$

$$h_c(t, X) = h_{0c}(t) \exp\left[\sum_i \beta_{ic} X_i\right]$$

Note $h_{0c}(t)$, T_c, β_{ic} are diff for diff event type.

Step 2: when Inference:

treat ① other competing failure
② lost to follow / withdrawl

as censored.

Note: before use PH Model, first need to assess PH Assump.

Independent Assumption

- metioned in CH1. Indep Assump:

$$h(t \mid \underset{\text{group}}{G}, \underset{\text{Censored}}{Ce}) = h(t \mid G, \underset{\text{Non-Censored}}{NCe})$$

$\Rightarrow$ treat other competing risks as censoring, which assumes their h(t) for getting the factor we assess is the same as those NCe ones.

- **How to prove if they are indep?**

  Can never determine
  it is counterfactual

- **How to proceed with assump not satisfied?**

  strategy 1: use clinical / biological / other backgrounds to assess Indep.

  strategy 2: Include common risk factors for competing risks (conditional Indep.)

Strategy 3: Sensitivity Analysis

- step 1: make worst cases
  (e.g. ① all censored will die from this event ② all censored will live to the end of study)
- step 2: fit $h(t,X)$ to each case
- step 3: compare $\hat{HR}$ between those extreme cases with the Indep. one)

results:

worst cases
- X meaningfully different → at most a small bias when assuming Indep.
- meaningfully different → only extreme of bias but not actual bias is determined.

Main point:

- No method to directly assess indep. assumption.
- Typical analysis assumes Indep. Assump. is satisfied.

Method 2: The Lunn-McNeil (LM)

Idea: use a single Model with Dummy variables to apply on several competing risks

v.s Method 1: give more flexibility.

General Stratified Cox LM Model:

$g = 1, 2, \cdots, C$

$$h_g^*(t, X) = h_{0g}^*(t) \times \exp[\beta_1 X_1 + \cdots + \beta_p X_p$$
$$+ \delta_{21} D_2 X_1 + \cdots + \delta_{2p} D_2 X_p$$
$$+ \delta_{31} D_3 X_1 + \cdots + \delta_{3p} D_3 X_p$$
$$+ \cdots$$
$$+ \delta_{C1} D_C X_1 + \cdots + \delta_{Cp} D_C X_p]$$

where $h_{0g}^*(t)$ diff for each g.

$D_1\ D_2 \cdots D_C$ is dummy variables.

Note we only use $C-1$ dummy variable in the equation.

Since when $g=1$, we can set $D_2, \cdots D_C = 0$

it also can have interaction terms

$X_i X_j$.

General Stratified Cox LMalt Model

$g = 1, \cdots, c$

$h_g'(t, X) = h_{0g}'(t) \times$

$$\exp[\delta_{11}' D_1 X_1 + \cdots + \delta_{1p}' D_1 X_p + \delta_{21}' D_2 X_1 + \cdots + \delta_{2p}' D_2 X_p + \delta_{31}' D_3 X_1 + \cdots + \delta_{3p}' D_3 X_p + \cdots + \delta_{c1}' D_c X_1 + \cdots + \delta_{cp}' D_c X_p]$$

Note here we use all of

$D_1 D_2 \cdots D_c$

LM vs LMalt:

- HR, test stats, interval estimates are the same
- coefficients are diff

- LM_{alt}: output can be provided directly from coefficient δ

  LM: indirectly: $(\delta + \beta)$

Unstratified LM model (LM_u)

$$h^*(t,X) = h_0^*(t) \times \exp[\gamma_2 D_2 + \gamma_3 D_3 + \cdots + \gamma_p D_p + \delta_{21} D_2 X_1 + \delta_{22} D_2 X_2 + \cdots + \delta_{2p} D_2 X_p + \cdots + \delta_{c1} D_c X_1 + \cdots + \delta_{cp} D_c X_p]$$

Note it is unstratified version

↳ $h_0^*(t)$: same for all data

LM_u vs. LM: Need to check PH Assump.

$\boxed{\text{3 plots}}$

KM:

- Indep Assump: treat other risk factors as censoring & $h(t|G, C_e) = h(t, |G \cap C_e)$
- informative for etiology

CIC: Cumulative Incidence Curv

- X Indep Assump
- Informative for treatment utility in cost-effectiveness analyze.
- vs. KM: for each competing risk, CIC uses $\hat{S}(t)$ over all risk factors, while KM only uses $\hat{S}(t)$ for that specific factor.

For how to compute CIC curve please see P448

and how to use the definition and math Method which similar with $h(t)$ and PH model. please see P451

CPC: Conditional probability Curves

$$CPC_c = P(T_c \le t \mid T \ge t)$$

$$CPC_c = CIC_c / (1 - CIC_c)$$