observed survival times are distince

- Partial likelihoods for distinct-event time data

- We will express the data in V and W to find the partial likelihood.

(V, U)

- Set-up

  ** Data: $(y_i, \nu_i, \mathbf{X}_i)$, $i = 1, \ldots, n$ (n individuals)

  ** Absolutely continuous failure time distribution

  ** Assume noninformative censoring

  ** d distinct event times (d observed failures) and $n - d$ right censored survival times.

  ** $t_0(= 0) < t_1 < t_2 < \ldots < t_d < t_{d+1}(= \infty)$: the distinct ordered event times (no ties between the event times)

  ** Let (j) be the label for individual failing at t_j. Note that $y_{(j)} = t_j$.

- Set-up (contd)
  - ** Covariates for d failures, $\mathbf{X}_{(j)}$, $j = 1, \ldots, d$
  - ** Censorship times in $[t_j, t_{j+1})$: $(t_{j1}, \ldots, t_{jm_j})$ with corresponding covariates $\mathbf{X}_{(j1)}, \ldots, \mathbf{X}_{(jm_j)}$.
  - ** Let (j, ℓ) be the label for individual censored at $t_{j,\ell}$. Note that $y_{(j)} = t_j$.
- Now we divide the data into sets

$$(V_1, U_1, W_1, \underbrace{V_2, U_2, W_2}, \ldots, W_d, V_{d+1}, U_{d+1}),$$

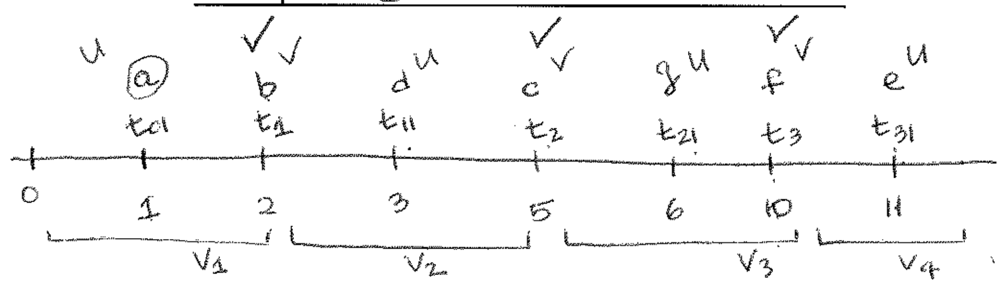
where

- ** $V_j = \{t_{j-1,1}, \ldots, t_{j-1,m_{j-1}}, t_j\}$: tells us time information of death and censoring in $(t_{j-1}, t_j]$.
- ** $U_j = \{(j, \ell), \ell = 1, \ldots, m_j\}$: tells us ~~time information of death and censoring~~ who ~~has died or~~ was censored in $(t_{j-1}, t_j]$.
- ** $W_j = \{(j)\}$ tells us who died at time t_j in the sample.

- Example:

covariate	X_a	X_b	X_c	X_d	X_e	X_f	X_g
id	a	b	c	d	e	f	g
y_i	1	2	5	3	11	10	6
ν_i	0	1	1	0	0	1	0

$d = 3$



$V_1 = \{1, 2\}$ $V_2 = \{3, 5\}$ $V_3 = \{6, 10\}$ $V_4 = \{11\}$

$U_1 = \{a\}$ $U_2 = \{d\}$ $U_3 = \{g\}$ $U_4 = \{e\}$

$W_1 = \{b\}$ $W_2 = \{c\}$ $W_3 = \{f\}$

- Goal: Build a likelihood on a subset of the full dataset
  ⋆⋆ Carrying most of the information about β

  ⋆⋆ Ignore information on nuisance parameters $\{h_0(t) : t \geq 0\}$.
- Generate a likelihood of $\{W_1, \ldots, W_d\}$.

  $h(t \mid x) = \underbrace{h_0(t)} \cdot e^{x\beta'}$
- Justification
  ⋆⋆ Timing of events $\{t_1, \ldots, t_d\}$ can be explained by $h_0(t)$.

  ⋆⋆ Censoring times and censored labels can be ignored if we assume *noninformative censorship* (independent censoring)

  $\Leftrightarrow$ Censoring mechanism does not depend on the risk of the failure of interest.

- Further

W_j

** $Q_j = (V_1, U_1, W_1, \ldots, V_{j-1}, U_{j-1}, W_{j-1}, V_j, U_j)$

** $\mathcal{F}(t_j) = (Q_j, \mathbf{X})$ denotes the information up to time t_j.

U, W

$\Rightarrow$ tells who has died or was censored prior to time t_j, when they died or was censored, together with their covariates. V

$\Rightarrow$ tells the individuals at risk at time t_j and their covariate values.

$\Rightarrow$ tells that a death occurs at interval $[t_j, t_j + \Delta t)$.

$\Rightarrow$ does *not* tell the individual who was observed to die among those at risk at time t_j.

** Define risk set R_j as the set of all individuals who are event-free and uncensored at time just prior to t_j. That is, $R_j = \{i \mid y_i \geq t_j\}$.

- Recap!

- Let $t_0 < t_1 < \ldots < t_d < t_{d+1}$ denote the values of observed survival times along the time axis.

- Recall $Q_j = (V_1, U_1, W_1, \ldots, V_{j-1}, U_{j-1}, W_{j-1}, V_j, U_j)$ and $\mathcal{F}(t_j) = (Q_j, \mathbf{X})$.

- The data can be expressed as

$\{y_i, \nu_i, x_i\}$ equiv.

$$(\mathcal{F}(t_1), W_1, \mathcal{F}(t_2), W_2, \ldots, \mathcal{F}(t_d), W_d, \mathcal{F}(t_{d+1})).$$

- The likelihood becomes

$$P(\mathcal{F}(t_1)) \times P(W_1 \mid \mathcal{F}(t_1)) \times P(\mathcal{F}(t_2) \mid \mathcal{F}(t_1), W_1) \cdots$$

- Partial likelihood is only part of the likelihood of the observed data!

$$\mathrm{PL} = \prod_{j=1}^{d} P(W_j = (j) \mid \mathcal{F}(t_j)).$$

Observe we are conditioning on R_j defined by $\mathcal{F}(t_j)$ and we know there is only one failure event at time t_j.

- Risk set at time t_j: $R_j = \{i \mid y_i \geq t_j\}$.
- Consider $P(W_j = (j) \mid \mathcal{F}(t_j))$. What does this imply?

  Among the subjects in R_j (by $\mathcal{F}(t_j)$), what is the probability that the observed death happened to subject (j) (who is actually observed to die at time t_j) rather than the other subjects?

P(individual (j) dies at time t_j | one death at t_j, R_j: a set of subjects alive at t_j)

- Let's work more on $P(W_j = (j) \mid \mathcal{F}(t_j))$.

↻ Let $[t_j, t_j + \Delta t)$ sufficiently small so that at most one death can occur within the interval.

$$
\begin{aligned}
P(W_j = (j) \mid \mathcal{F}(t_j)) &= \frac{\lim_{\Delta t \to 0^+} P(y_{(j)} \in [t_j, t_j + \Delta t) \mid \mathcal{F}(t_j))/\Delta t}{\lim_{\Delta t \to 0^+} \sum_{k \in R_j} P(y_k \in [t_j, t_j + \Delta t) \mid \mathcal{F}(t_j))/\Delta t} \\
&= \frac{h(t_j \mid \mathbf{X}_{(j)})}{\sum_{k \in R_j} h(t_j \mid \mathbf{X}_k)} = \frac{\cancel{h_0(t_j)} \cdot e^{X_{(j)}\beta}}{\sum_{k \in R_j} \cancel{h_0(t_j)} e^{X_k \beta}} \\
&= \frac{\exp(\beta' \mathbf{X}_{(j)})}{\sum_{k \in R_j} \exp(\beta' \mathbf{X}_k)}.
\end{aligned}
$$

- So the partial likelihood becomes

$$\begin{aligned}
\mathrm{PL}(\beta) &= \prod_{j=1}^{d} P(W_j = (j) \mid \mathcal{F}(t_j)) \\
&= \prod_{j=1}^{d} \frac{h(t_j \mid \mathbf{X}_{(j)})}{\sum_{k \in R_j} h(t_j \mid \mathbf{X}_k)} \\
&= \prod_{j=1}^{d} \frac{\exp(\beta' \mathbf{X}_{(j)})}{\sum_{k \in R_j} \exp(\beta' \mathbf{X}_k)} \\
&= \prod_{i=1}^{n} \left\{ \frac{\exp(\beta' \mathbf{X}_{(j)})}{\sum_{k \in \underline{R_i}} \exp(\beta' \mathbf{X}_k)} \right\}^{\nu_i}.
\end{aligned}$$

Risk set at y_i

- Partial likelihoods for distinct-event time data

$$p(\mathbf{y}, \boldsymbol{\nu} \mid \boldsymbol{\beta}, \mathbf{X}) = \prod_{j=1}^{d} \frac{\exp\left(\boldsymbol{\beta}'\mathbf{X}_{(j)}\right)}{\sum_{k \in R_j} \exp\left(\boldsymbol{\beta}'\mathbf{X}_k\right)}$$

⋆⋆ The numerator: information from the individuals who experience the event (individuals with censored survival times do not contribute)

⋆⋆ The denominator: information from all individuals who have not experienced the event (including some individuals who will be censored later)

⋆⋆ The likelihood depends only on the ranking of the event times. $\Rightarrow$ inferences about $\boldsymbol{\beta}$ depend only on the rank order of the survival times.

- Partial likelihoods for distinct-event time data (contd)

$$p(\mathbf{y}, \boldsymbol{\nu} \mid \boldsymbol{\beta}, \mathbf{X}) = \prod_{j=1}^{d} \frac{\exp\left(\boldsymbol{\beta}'\mathbf{X}_{(j)}\right)}{\sum_{k \in R_j} \exp\left(\boldsymbol{\beta}'\mathbf{X}_{j}\right)}$$

⋆⋆ Treat as a usual likelihood so inference is carried out by usual means ⇒ obtain the partial maximum likelihood estimate of $\boldsymbol{\beta}$ (numerical methods are needed)

⋆⋆ Specifying the baseline hazard $h_0(t)$ is unnecessary.

⋆⋆ This ignores the part of the likelihood that records information between failure times. In other words, the interval between events does not inform the PL function.

⋆⋆ Inference based on the partial likelihood has many of the properties of inference based on the full likelihood function, including consistency and asymptotic normality.

∗ [Example: K-M 1.8 Death Times of Male Laryngeal Cancer Patients (page 9)]

Kardaun (1983) reports data on 90 males diagnosed with cancer of the larynx during the period 1970–1978 at a Dutch hospital. The followings are recorded;

- ⋆⋆ Survival – the intervals (in years) between first treatment and either death or the end of the study (January 1, 1983).
- ⋆⋆ Covariates – the patient's age at the time of diagnosis, the year of diagnosis, and the stage of the patient's cancer.

  The four stages of disease in the study were recorded; the four groups are Stage I with 33 patients, Stage II with 17 patients, Stage III with 27 patients and Stage IV with 13 patients.

∗ Revisit [Example: Male Laryngeal Cancer Patients (KM Ex 8.2)]

We use the proportional hazards model using the main effects of age and stage for this data;

Stage II 0, 1 Stage III IV Age

$$h(t \mid \mathbf{X}_i) = h_0(t)\exp(\beta'\mathbf{X}_i) = h_0(t)\exp(\beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4})$$

where X_k, $k = 1, 2, 3$ are the indicators of stage II, III and IV disease, respectively, and X_4 is the age of the patient.

```
> library(KMsurv)  # To get the datasets in K-M
> library(survival) # R functions
>
> data(larynx)
>
> coxph.fit <- coxph(Surv(time, delta) ~ as.factor(stage) + age,
method="breslow", data=larynx)
```

∗∗ For "method", you can choose one from "efron", "breslow" and "exact" (different in how to handle partial likelihood when ties are present).

* Revisit [Example: Male Laryngeal Cancer Patients]

```
> coxph.fit
Call:
coxph(formula = Surv(time, delta) ~ as.factor(stage) + age, data = larynx,
    method = "breslow")


                      coef  exp(coef) se(coef)    z        p
as.factor(stage)2 0.1386 β̂1  1.1486   0.4623 0.30    0.764    0.76
as.factor(stage)3 0.6383 β̂2  1.8934   0.3561 1.79    0.073    0.073
as.factor(stage)4 1.6931 β̂3  5.4361   0.4222 4.01  6.1e-05    0.00 ✓
age               0.0189 β̂4  1.0191   0.0143 1.33    0.185    0.185

Likelihood ratio test=18.1  on 4 df, p=0.0012
n= 90, number of events= 50
>
```

* Revisit [Example: Male Laryngeal Cancer Patients]

  ** The information matrix is the negative of the matrix of second derivatives of the log likelihood.

  ** For large samples, the (partial) MLE has a $p-$variate normal distribution with mean $\boldsymbol{\beta}$ and variance-covariance estimated by $I^{-1}(\widehat{\boldsymbol{\beta}})$

  ** Now can do a test based on the asymptotic normality of the (partial) MLE (Wald's test, Likelihood ratio test, Score test...)

```
> va <- coxph.fit$var
# I^(-1), estimated cov matrix of the estimates
> va
            [,1]         [,2]          [,3]          [,4]
[1,] 0.213726421 0.0683008943  0.0689498211  0.0008015350
[2,] 0.068300894 0.1267932600  0.0682084308  0.0003144479
[3,] 0.068949821 0.0682084308  0.1782595628 -0.0003990833
[4,] 0.000801535 0.0003144479 -0.0003990833  0.0002030920
>
```

∗ Revisit [Example: Male Laryngeal Cancer Patients] We use the proportional hazards model using the main effects of age and stage for this data;

$$h(t \mid \mathbf{X}_i) = h_0(t)\exp(\beta'\mathbf{X}_i) = h_0(t)\exp(\beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4})$$

where X_k, $k = 1, 2, 3$ are the indicators of stage II, III and IV disease, respectively, and X_4 is the age of the patient.

- Interpretation:
  - A 95% confidence interval for the risk of death for patients in Stage IV relative to the risk of death for patients in Sta~~te~~ge I would be

    $$(\exp(\underline{1.6931} - 1.96 \times \underline{0.4222}), \exp(1.6931 + 1.96 \times 0.4222)) = (2.39, 12.44)$$

    ⇔ With approximately 95% confidence, $\exp(\beta_3)$ will lie between 2.38 and 12.44.

$\underline{h_0(t)}\, e^{X_4\beta_4}$ ← Stage I

$h(t) =$

• How does **X** affect the hazard function under the proportional hazards regression model?

↻ The conditional survival function of an individual with covariate vector **X** can be expressed in terms of a baseline survival function $S_0(t)$

$$S_0(t) = \exp(-\int_0^t h_0(u)du) = \exp(-H_0(t)),$$
$$S(t \mid \mathbf{X}) = \left(S_0(t)\right)^{\exp(\beta'\mathbf{X})}.$$

⇒ By fitting a Cox model, one can readily interpret the multiplicative effect of **X** on the hazard.

⇒ However, $H_0(t)$ (or $h_0(t)$) is required to determine **X**'s effect on $S(t \mid \mathbf{X})$.

$$S(t \mid x) = \exp\left(-\int_0^t h(u \mid x)\, du\right)$$
$$= \exp\left(-\int_0^t h_0(u) \cdot e^{x\beta}\, du\right)$$
$$= \left(\exp\left(-\int_0^t h_0(u)\, du\right)\right)^{e^{x\beta}}$$
$$= \left(S_0(t)\right)^{e^{x\beta}}$$

- Estimation of the Survival Function

$\star$ Although covariate effects are of primary interest, investigators are often interested in the survival function.

$\star$ Strategy: Estimate $H_0(t)$, $S_0(t)$ $\Rightarrow$ $S(t|X) = \{S_0(t)\}^{\exp(X_i'\beta)}$.

$\star$ Let

$$\widehat{H}_0(t) = \sum_{t_j \le t} \frac{m_j}{\sum_{i \in R_j} \exp\left(\sum_{k=1}^{p} \hat{\beta}_k X_{ik}\right)},$$ KM p 258

where m_j is the number of deaths at time t_j.

$\star$ Thus, the baseline survival function, $S_0(t) = \exp(-H_0(t))$ is given by

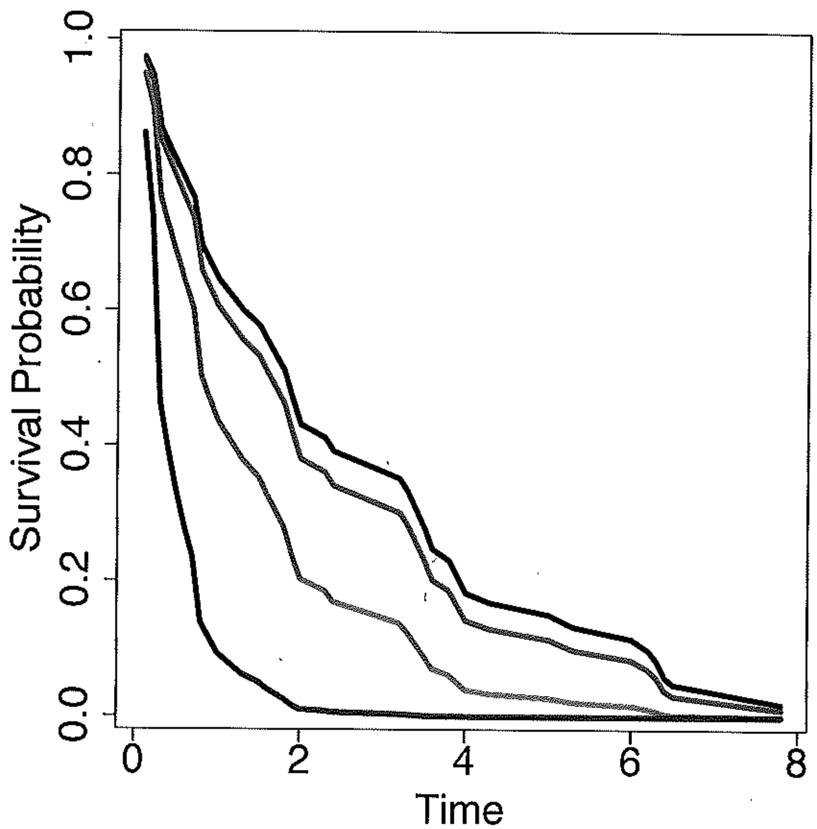
$$\widehat{S}_0(t) = \exp(-\widehat{H}_0(t)).$$

* Revisit [Example: Male Laryngeal Cancer Patients]

```
my.survfit.object <- survfit(coxph.fit)
## to obtain the baseline survival function

S0 <- (summary(my.survfit.object))$surv  ## S0

S1 <- (S0)^(exp(coxph.fit$coefficients[4]*60))  stage I
S2 <- (S0)^(exp(coxph.fit$coefficients[1] + coxph.fit$coefficients[4]*60)) II
S3 <- (S0)^(exp(coxph.fit$coefficients[2] + coxph.fit$coefficients[4]*60)) III
S4 <- (S0)^(exp(coxph.fit$coefficients[3] + coxph.fit$coefficients[4]*60))
                                                                           IV
t.grid <- (summary(my.survfit.object))$time


pdf("S-Cox.pdf")
par(mar=c(4.5, 4.5, 2.1, 2.1))
plot(t.grid, S1, type="l", lwd=4, ylab="Survival Probability",
xlab="Time", cex.axis=2, cex.lab=2)
lines(t.grid, S2, lwd=4, col=2)
lines(t.grid, S3, lwd=4, col=3)
lines(t.grid, S4, lwd=4, col=4)
dev.off()
```

∗ Revisit [Example: Male Laryngeal Cancer Patients]

⋆ For each cancer stage at age 60, the survival function is



- **Bayesian proportional hazards model**

$$
\begin{aligned}
h(t \mid \mathbf{X}) &= h_0(t)\exp(\beta'\mathbf{X}) \\
\Rightarrow H(t \mid \mathbf{X}) &= \int_0^t h(u \mid \mathbf{X})du \\
&= \int_0^t h_0(u)\exp(\beta'\mathbf{X})du \\
&= H_0(t)\exp(\beta'\mathbf{X})
\end{aligned}
$$

- We use all the information contained in data. $\Rightarrow$ No partial likelihood!

- We can assume a parametric model for $h_0(t)$. e.g. exponential, Weibull. $\alpha\gamma t^{\alpha-1} = h_0(t)$ $h_0(t) = \gamma$

$\Rightarrow$ So the model becomes fully parametric.

$\mathcal{L}_i = \{f(y_i)\}^{\nu_i}\{S(y_i)\}^{1-\nu_i}$

$= \{h(y_i)\}^{\nu_i} S(y_i)$

- As a specific example, suppose we assume $h_0(t) = \alpha\gamma t^{\alpha-1}$ (that is, the Weibull for the baseline.)

  ⋆⋆ Place priors on α, γ and $\boldsymbol{\beta}$.

  ⋆⋆ We know this model is equivalent to the AFT model with $W_i \overset{iid}{\sim}$ V.

$h(t \mid \mathbf{X}) = h_0(t) \cdot e^{X\beta}$

- The likelihood function becomes

$$\begin{aligned}
\mathcal{L}(\alpha, \gamma, \boldsymbol{\beta} \mid \mathcal{D}) &= \prod_{i=1}^{n} \{\underbrace{h_0(y_i)\exp(\mathbf{X}_i'\boldsymbol{\beta})}_{h(y_i \mid x_i)}\}^{\nu_i} \underbrace{\{S_0(y_i)\}^{\exp(\mathbf{X}_i'\boldsymbol{\beta})}}_{S(y_i \mid x_i) = e^{-H(y_i \mid x_i)}} \\
&= \prod_{i=1}^{n} \{h_0(y_i)\exp(\mathbf{X}_i'\boldsymbol{\beta})\}^{\nu_i} \underbrace{\exp\{-\sum_{i=1}^{n} \exp(\mathbf{X}_i'\boldsymbol{\beta}) H_0(y_i)\}}
\end{aligned}$$

- Abrams *et al.*(1996) – Bayesian parametric proportional hazards model with a lot more about prior specification.

$X_i \overset{indep}{\sim} Ga(a_i, 1)$, $i=1,\dots,n$

$\left(\frac{X_1}{\sum X_i}, \qquad \right) \sim Dir(a_1, \dots, a_n)$

- Bayesian proportional hazards model.

$$
\begin{aligned}
h(t \mid \mathbf{X}) &= h_0(t)\exp(\beta'\mathbf{X}) \\
\Rightarrow H(t \mid \mathbf{X}) &= \int_0^t h(u \mid \mathbf{X})du \\
&= \int_0^t h_0(u)\exp(\beta'\mathbf{X})du \\
&= H_0(t)\exp(\beta'\mathbf{X})
\end{aligned}
$$

piecewise constant

Gamma Process

- Assume nonparametric prior processes for $h_0(t)$ or $H_0(t)$
- Allows a more general modeling strategy with fewer assumptions

• Examples of nonparametric prior processes for $h_0(t)$ or $H_0(t)$ (ICS Chapter 3)

⋆⋆ Piecewise constant hazard model (3.1) for $h_0(t)$

⋆⋆ Gamma process model (3.2) and correlated Gamma process (3.6) for $H_0(t)$

⋆⋆ Beta process model (3.5) for $H(t)$

⋆⋆ Dirichlet process model (3.7) for $S(t)$

• Many of them are beyond our course material so let's focus on piecewise constant hazard model and briefly on the Gamma process model.

♠ Piecewise constant hazard model (ICS Chapter 3.1)

⋆⋆ Recall we have

$$h(t \mid \mathbf{X}) = h_0(t)\exp(\beta'\mathbf{X})$$

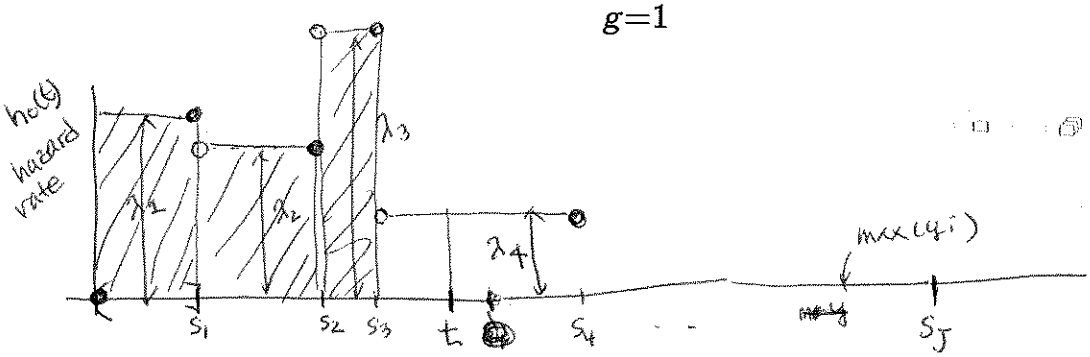
⋆⋆ Construct a finite partition of the time axis, $0 < s_1 < s_2 < \ldots < s_J$ with $s_J > \max(y_i)$.

$\Rightarrow$ we have the J intervals, $(0, s_1]$, $(s_1, s_2]$,..., $(s_{J-1}, s_J]$.

⋆⋆ In interval j, we assume a constant baseline hazard.
For $t \in I_j = (s_{j-1}, s_j]$,

$$\begin{aligned} h_0(t) &= \lambda_j \\ \Rightarrow H_0(t) &= \sum_{g=1}^{j-1} \lambda_g(s_g - s_{g-1}) + \lambda_j(t - s_{j-1}). \end{aligned}$$

$H_0(t)$



$H_0(t) = (s_1 - 0) \times \lambda_1 + (s_2 - s_1) \times \lambda_2 + (s_3 - s_2)\lambda_3 + (t - s_3)\lambda_4$

- We need to specify priors for $\boldsymbol{\beta}$ and $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_J)'$.

$\star\star$ May consider $\boldsymbol{\beta} \sim \mathrm{N}_p$ or $\pi(\boldsymbol{\beta}) \propto 1$

$\star\star$ Two examples of prior for $\boldsymbol{\lambda}$.

- ↺ Independent gamma prior, $\lambda_j \overset{indep}{\sim} \mathrm{Gamma}(\alpha_{0j}, \beta_{0j})$
- ↺ Let $\psi_j = \log(\lambda_j)$ and use a correlated prior $\boldsymbol{\psi} \sim \mathrm{N}_J(\boldsymbol{\psi}_0, \Sigma)$.

- Comments on the piecewise constant hazard model

  ⋆⋆ If $J = 1$, the model reduces to a parametric exponential distribution with failure rate parameter $\lambda \equiv \lambda_1$.

  ⋆⋆ This semiparametric model is called a *piecewise exponential model*

  ⋆⋆ Accommodate various shapes of the baseline hazard over the intervals.

  ⋆⋆ Simple and useful. Serves as the benchmark for comparison

- Shall we write down the likelihood?

** Define censoring indicator ν_i as

$$\nu_i = \begin{cases} 1 & \text{if subject } i \text{ failed (observed survival time),} \\ 0 & \text{otherwise.} \end{cases}$$

** Define δ_{ij} as

$$\delta_{ij} = \begin{cases} 1 & \text{if subject } i \text{ failed or censored in interval } j \\ 0 & \text{otherwise} \end{cases}$$

** The likelihood function of $(\boldsymbol{\beta}, \boldsymbol{\lambda})$ is

For patient i,

$$\mathcal{L}_i = \left(h(y_i \mid x_i, \beta, \lambda) \right)^{\nu_i} \cdot \exp\left(- H(y_i \mid x_i, \beta, \lambda) \right)$$

$$= \left(\prod_{j=1}^{J} (\lambda_j)^{\delta_{ij}} \cdot e^{x_i' \beta} \right)^{\nu_i}$$

$$\cdot \exp\left(- \sum_j \delta_{ij} \left(\sum_{g=1}^{j-1} \lambda_g (s_g - s_{g-1}) + \lambda_j (y_i - s_{j-1}) \right) \right)$$

$$\mathcal{L}(\beta, \lambda) = \prod_{i=1}^{n} \mathcal{L}_i$$

For patient i, we have $y_i, x_i, \nu_i, \delta_{ij}$.

$$\mathcal{L}_i = \left(h(y_i \mid x_i, \beta, \lambda)\right)^{\nu_i} \exp\left(-H(y_i \mid x_i, \beta, \lambda)\right)$$

$$= \left(\prod_{j=1}^{J} (\lambda_j)^{\delta_{ij}} \times e^{\beta' x_i} \right)^{\nu_i} \times \exp\left(-\sum_{j=1}^{J} \delta_{ij}\left(\sum_{g=1}^{j-1} \lambda_g (s_g - s_{g-1}) + \lambda_j (y_i - s_j)\right) \times e^{\beta' x_i} \right)$$

For all patients $(i=1, \dots, n)$

$$\mathcal{L} = \prod_{i=1}^{n} \mathcal{L}_i$$

$$= \prod_{i=1}^{n} \left\{ \left(\prod_{j=1}^{J} (\lambda_j)^{\delta_{ij}} \cdot e^{\beta' x_i} \right)^{\nu_i} \cdot \exp\left(-\sum_{j=1}^{J} \delta_{ij} \left(\sum_{g=1}^{j-1} \lambda_g (s_g - s_{g-1}) + \lambda_j (y_i - s_j)\right) e^{\beta' x_i} \right) \right\}$$

The joint posterior is

$$P(\beta, \lambda \mid \text{Data}) = \mathcal{L}(\beta, \lambda) \cdot \pi(\beta)\, \pi(\lambda)$$

$\Rightarrow$ <u>Full conditionals</u>

① Update β

Let $$H_{0i} = \sum_{j=1}^{J} \delta_{ij} \left(\sum_{g=1}^{j-1} \lambda_g (s_g - s_{g-1}) + \lambda_j (y_i - s_j) \right)$$

$$p(\beta \mid \lambda, \text{data}) \propto \prod_{i=1}^{n} e^{\beta' x_i \cdot \nu_i} \cdot \exp\left(-H_{0i} \cdot e^{\beta' x_i}\right) \cdot \pi(\beta)$$

<u>Suppose</u> $\pi(\beta) = N_p(\bar{\beta}, \Sigma)$

$$\log\left(p(\beta \mid \lambda, \text{data})\right) \propto \sum_{i=1}^{n}\left(\nu_i \cdot \beta' x_i - H_{0i} \cdot e^{\beta' x_i}\right) - \frac{1}{2}(\beta - \bar{\beta})' \Sigma^{-1} (\beta - \bar{\beta})$$

(2) Update λ_j, $j = 1, \ldots, J$.

one λ_j at a time.

$$p(\lambda_j \mid \lambda_{-j}, \beta, \text{data}) \propto \prod_{i=1}^{n} \left(\prod_{j=1}^{J} (\lambda_j)^{\delta_{ij} \cdot \nu_i}\right) \cdot \exp\left(-\underbrace{H_{0i}}_{\text{function of } \lambda} \cdot e^{\beta' x_i}\right) \pi(\lambda_j)$$

Suppose $\lambda_j \sim Ga(\alpha_{0j}, \lambda_{0j})$

$$\log\left(p(\lambda_j \mid \lambda_{-j}, \beta, \text{data})\right) \propto \sum_{\substack{i=1 \\ y_i > s_{j-1}}}^{n} \left\{ \nu_i \left(\sum_{j=1} \delta_{ij} \log(\lambda_j)\right) - H_{0i} e^{\beta' x_i} \right\}$$

$$+ (\alpha_{0j} - 1) \cdot \log(\lambda_j) - \lambda_{0j} \cdot \lambda_j$$