① Finish PC Bayesian Today

② G.P Bayesian

③ Time - dependent covariate

* 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; for $t \in (s_{j-1}, s_j]$

Stage II III IV Age

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

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.

** Use a piecewise constant hazard model with the following priors;

** Let $\boldsymbol{\beta} \sim \mathrm{N}_4(\bar{\boldsymbol{\beta}}, \Sigma)$.

** Let $\lambda_j \overset{iid}{\sim} \mathrm{Gamma}(\alpha_0, \lambda_0)$. $j=1,\ldots,J$

* Revisit [Example: Male Laryngeal Cancer Patients]

```
> ###  set up hyperparameters
> hyper <- NULL
>
> ### Be \sim N_p(Beta_bar, Sig)
> ## fit the frequentist Cox to set hyperparamters
> coxph.fit <- coxph(Surv(time, delta) ~ as.factor(stage) + age,
method="breslow", data=larynx)
> hyper$Beta_bar <- as.matrix(coxph.fit$coefficient)
> hyper$Beta_bar[4] <- hyper$Beta_bar[4]*sd(larynx$age)
##  since I will standardize age
> hyper$Sig <- diag(2.0, p)
> hyper$Inv_Sig <- solve(hyper$Sig)
>
> ##  lambda_j \iid \Ga(a0, lam0)
> hyper$a0 <- 0.1
> hyper$lam0 <- 0.1
>
```

∗ Revisit [Example: Male Laryngeal Cancer Patients] How did I set intervals, $(s_{j-1}, s_j]$?

∗∗ Used empirical quantiles.

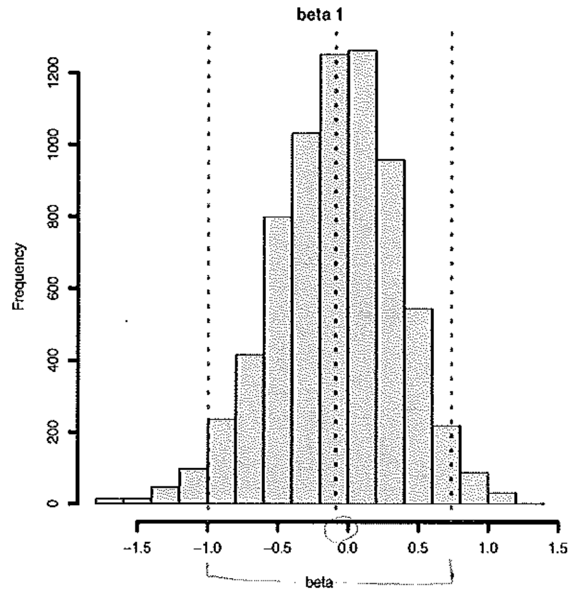
```
> Int_dat <- c(min(larynx$time)-0.01, quantile(larynx$time,
seq(0, 1.0, by=0.025))+0.01)  ## time grid
> Int_dat <- cbind(c(0,Int_dat[-length(Int_dat)]), Int_dat)
## (starting time, end time]
>
> J <- nrow(Int_dat)  ## number of intervals
```

s_{j-1} s_j

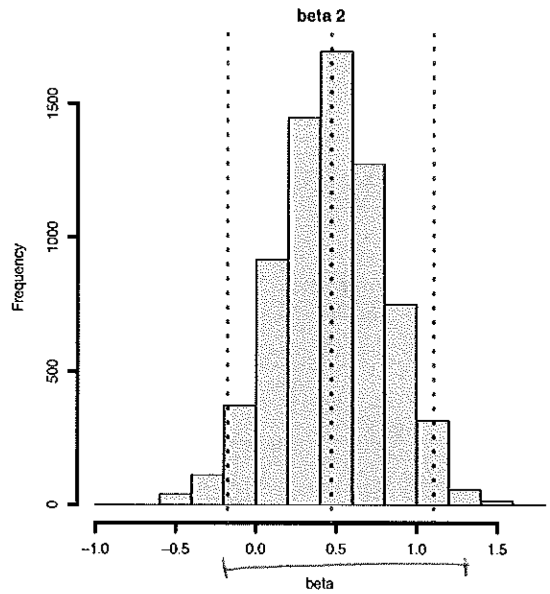
- Piecewise Constant Hazard Model

Stage III

Stage II



$\beta_1(\hat{\beta}_1 = -0.087)$

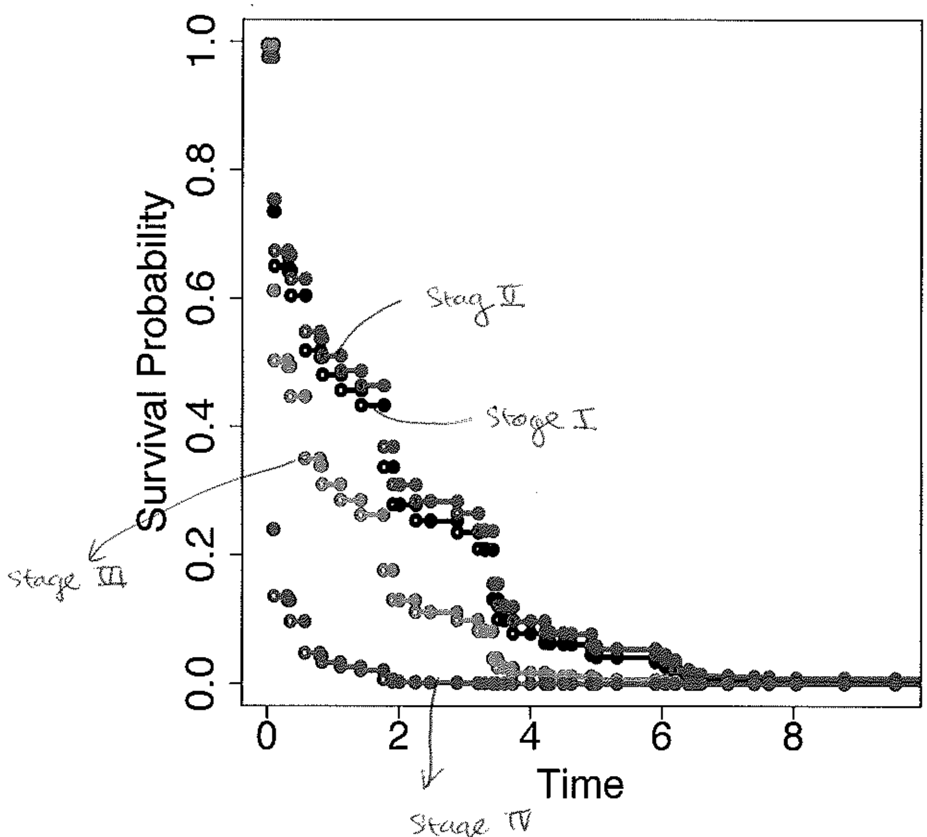


$\beta_2(\hat{\beta}_2 = 0.470)$

$h(x) = h_0\, e^{x\beta}$

- Piecewise Constant Hazard Model

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



♠ Models Using a Gamma Process (ICS Chapter 3.2)

- Recall $Z \sim \text{Gamma}(\theta, \lambda)$ with the pdf

$$f(z \mid \theta, \lambda) = \frac{\lambda^\theta}{\Gamma(\theta)} z^{\theta-1} e^{-\lambda z}, \text{ for } z > 0.$$

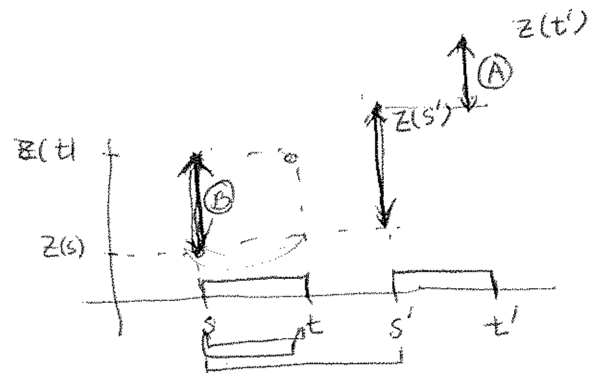
Then, $\mathrm{E}(Z) = \theta/\lambda$ and $\mathrm{Var}(Z) = \theta/\lambda^2$.

- Use the **gamma process** prior for the baseline cumulative hazard function $H_0(t)$.

- The gamma process: the most commonly used nonparametric prior process for the Cox model.

- We will *briefly* cover the gamma process.

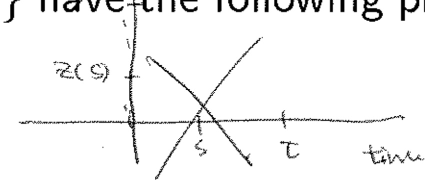
- Read the paper by Kalbflesch (1978) for more details



$\star\star$ Let $\alpha(t)$, $t \geq 0$ be an increasing left continuous function with $\alpha(0) = 0$.

$\star\star$ Let a stochastic process $\{Z(t), t > 0\}$ have the following properties.

S(0) = 1



- $Z(0) = 0$
- $Z(t)$ has independent increments in disjoint intervals, and
- for $t > s$, $Z(t) - Z(s) \sim \text{Gamma}(c(\alpha(t) - \alpha(s)), c)$

$\Rightarrow$ The process is called a gamma process, $Z(t) \sim \mathcal{GP}(c\alpha(t), c)$.

- What does this mean?

$\star\star$ It is a stochastic process with independent increments and the increments have the gamma distribution. $t > s$

$$\begin{aligned} E(Z(t) - Z(s)) &= c(\alpha(t) - \alpha(s))/c = \alpha(t) - \alpha(s), \\ \mathrm{Var}(Z(t) - Z(s)) &= c(\alpha(t) - \alpha(s))/c^2 = (\alpha(t) - \alpha(s))/c \end{aligned}$$

- $\alpha(t)$ is the mean of the process.
- c is a weight or confidence parameter about the mean.

$c \uparrow \Rightarrow \mathrm{Var}(Z(t) - Z(s)) \downarrow \Rightarrow Z(t) \approx \alpha(t)$

$c \downarrow \Rightarrow \mathrm{Var}(Z(t) - Z(s)) \uparrow \Rightarrow Z(t)$

- Facts! $H_0(t)$

  ⋆⋆ The sample path of the gamma process is almost surely strictly increasing purely discontinuous function of t ($\Rightarrow$ can be used as a prior for $H_0(t)$)

  ⋆⋆ The probability that any preassigned value of t is a jump is zero.

  ⋆⋆ *The above is great. Why?* Observed survival times can occur with positive probability even though the survival times have a continuous distribution.

  ⋆⋆ (side note) Connection to the Dirichlet process! (alternative definition of the DP) – read Ferguson (1973).

- Gamma Process as a Prior for Cumulative Hazard

  ⋆⋆ Consider the gamma process as a prior for $H_0(t)$.

$$H_0 \sim \mathcal{GP}(c_0 H^\star, c_0),$$

  where $H^\star(t)$ is an increasing function and $c_0 > 0$.

  ⋆⋆ Need to specify $H^\star$ and c_0.

$h(t) = \lambda$

$H(t) = \lambda t$

- Examples of $H^\star(t)$

$\star\star$ *Example 1:* $H^\star(t) = \gamma_0 t$ where γ_0 is a specified hyperparameter.

$E(H_\delta(t))$

Observe $\cancel{E(Z(t))} = \gamma_0 t$

$\Rightarrow$ noisy exponential baseline hazard function.

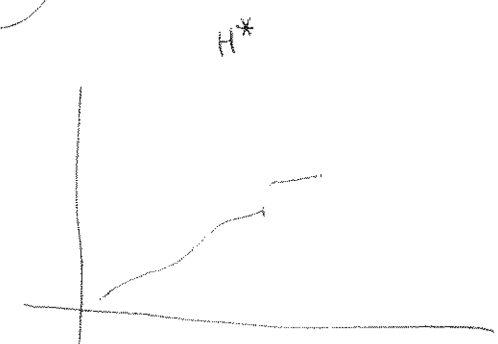
$\star\star$ *Example 2:* $H^\star(t) = \underline{\eta_0 t^{\kappa_0}}$ where η_0 and κ_0 are specified hyperparameters.

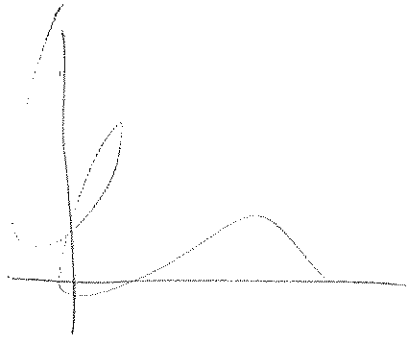
$E(H_0(t))$ $\qquad$ $h(t) = \eta_0 \kappa_0 t^{\kappa_0 - 1}$

$H(t) = \eta_0 t^{\kappa_0}$

Observe $\cancel{E(Z(t))} = \eta_0 t^{\kappa_0}$

$\Rightarrow$ noisy Weibull baseline hazard function.





- Gamma Process with Grouped-Data Likelihood

  ⋆⋆ 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, $(\underset{\uparrow s_0}{0}, s_1]$, $(s_1, s_2]$,..., $(s_{J-1}, s_J]$.

  ⋆⋆ One choice: use the distinct survival times as our $s_1, \ldots, s_{J-1}$ after ordering.

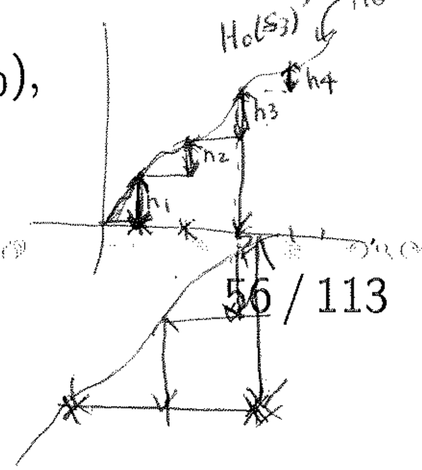
  ⋆⋆ Let h_j denote the increment in H_0 in interval j, that is,

$$h_j = H_0(s_j) - H_0(s_{j-1}), \quad j = 1, \ldots, J.$$

Assuming $H_0 \sim \mathcal{GP}(c_0 H^\star, c_0)$, we know h_j are **independent** and $h_1 + h_2 + h_3$

$$h_j \overset{indep}{\sim} \text{Gamma}(\alpha_{0,j} - \alpha_{0,j-1}, c_0),$$

where $\alpha_{0,j} = c_0 H^\star(s_j)$, $j = 1, \ldots, J$.



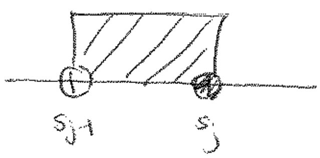
$$\alpha_{0j} - \alpha_{0,j-1} = c_0 H^*(s_j) - c_0 H^*(s_{j-1})$$
$$= c_0 (H^*(s_j) - H^*(s_{j-1}))$$

$I_j = (s_{j-1}, s_j]$

- Consider one subject and find $P(y_i \in I_j \mid \mathbf{X}_i, \mathbf{h}, \beta)$;

$$P(y_i \in I_j \mid \mathbf{X}_i, \mathbf{h}, \beta) = P(s_{j-1} < y_i \leq s_j \mid \mathbf{X}_i, \mathbf{h}, \beta)$$

$$= P(y_i \leq s_j \mid x_i, h, \beta) - P(y_i \leq s_{j-1} \mid x_i, h, \beta)$$



$$= P(y_i > s_{j-1} \mid x_i, h, \beta) - P(y_i > s_j \mid x_i, h, \beta)$$

$$= e^{-\underbrace{H_0(s_{j-1})} e^{x_i\beta}} - e^{-H_0(s_j) e^{x_i\beta}}$$

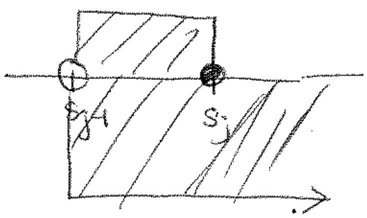
$S(t \mid x) = e^{-H(t \mid x)}$

$= e^{-\underbrace{H_0(t)}\underbrace{e^{x\beta}}}$

$$= e^{-(\sum_{g=1}^{j-1} h_g) e^{x_i\beta}} - e^{-(\sum_{g=1}^{j} h_g) \cdot e^{x_i\beta}}$$

$$= e^{-(\sum_{g=1}^{j-1} h_g) e^{x_i\beta}} \left(1 - e^{-h_j e^{x_i\beta}} \right)$$

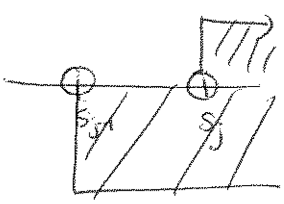
- How about $P(\underbrace{y_i \in I_j}_{} \mid \underbrace{y_i > s_{j-1}}_{}, \mathbf{X}_i, \mathbf{h}, \boldsymbol{\beta})$?



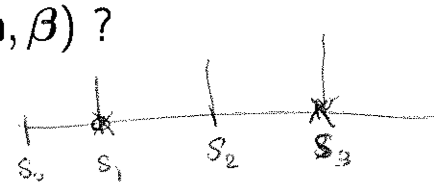
$$= \frac{P(y_i \in I_j \mid x_i, h, \beta)}{P(y_i > s_{j-1} \mid x_i, h, \beta)}$$

$$= \frac{e^{-e^{x_i'\beta} \cancel{\sum_{g=1}^{j-1} h_g}} \left(1 - e^{-e^{x_i'\beta} h_j}\right)}{e^{-e^{x_i'\beta} \cancel{\sum_{g=1}^{j-1} h_g}}} = \underbrace{1 - e^{-e^{x_i'\beta} h_j}}_{}$$

- How about $P(y_i > s_j \mid y_i > s_{j-1}, \mathbf{X}_i, \mathbf{h}, \boldsymbol{\beta})$?



$$= e^{-e^{x_i'\beta} h_j}$$



$$\Rightarrow P(y_i \in I_j \mid x_i, h, \beta) = \prod_{g=1}^{j-1} P(y_i > s_g \mid y_i > s_{g-1}, x_i, h, \beta)$$

$$\times P(y_i \in I_j \mid y_i > s_{j-1}, x_i, h, \beta)$$

* For intervals $1, \ldots, j-1$, subject i is at risk but does not fail

* For interval j, subject i is at risk and fails.

y_i, ν_i

- Gamma Process with Grouped-Data Likelihood (contd)

  ⋆⋆ We express the observed data $D = (\mathbf{X}, \mathcal{R}_j, \mathcal{D}_j, j = 1, \ldots, J)$ where

  - $\mathcal{R}_j$: the risk set of interval j

    s_{j-1}

  - $\mathcal{D}_j$: the failure set of interval j

  ⋆⋆ We will write down the likelihood using $\mathcal{R}_j$ and $\mathcal{D}_j$.

- We will express the likelihood

$$\mathcal{L}(\boldsymbol{\beta}, \mathbf{h} \mid D) \propto \prod_{j=1}^{J} G_j, \qquad H_0$$

where

$$G_j = \underbrace{\exp\{-h_j \sum_{k \in \mathcal{R}_j - \mathcal{D}_j} \exp(\mathbf{X}_k' \beta)\}}_{\text{at risk and not fail}} \prod_{\ell \in \mathcal{D}_j} \underbrace{[1 - \exp\{-h_j \exp(\mathbf{X}_\ell' \beta)\}]}_{\text{at risk and fail}}.$$

- Returning to the prior,

  ⋆⋆ Observe that H_0 enters the likelihood only through the h_j's.

  ⋆⋆ We need to specify priors for h_j and $\boldsymbol{\beta}$.

  ⋆⋆ Recall that h_j are **independent** and

$$h_j \overset{indep}{\sim} \text{Gamma}(\alpha_{0,j} - \alpha_{0,j-1}, c_0),$$

  where $\alpha_{0,j} = c_0 H^\star(s_j)$.

  $\Rightarrow$ We need to specify $H^\star(t)$ and c_0 (or we can consider $c_0 \sim \pi$).

  ⋆⋆ We may use $\boldsymbol{\beta} \sim \text{N}_p(\bar{\boldsymbol{\beta}}, \Sigma)$ or other priors including $\pi(\boldsymbol{\beta}) \propto 1$.

- $h_j \overset{indep}{\sim} \text{Gamma}(\alpha_{0j} - \alpha_{0,j-1},\ c_0)$

$$H^*(t) = \eta_0 t^{\kappa_0}$$

$$\alpha_{0,j} - \alpha_{0,j-1} = \underline{c_0 \eta_0}\left(s_j^{\kappa_0} - s_{j-1}^{\kappa_0} \right)$$

- $\beta \sim N_p(\bar{\beta}, \Sigma)$

$$\pi(h, \beta \mid data) \propto \prod_{j=1}^{J} \left\{ G_j \cdot \overbrace{h_j^{(\alpha_{0j} - \alpha_{0,j-1}) - 1} \exp(-c_0 \cdot h_j)}^{\pi(h_j)} \right\} \cdot \underbrace{\exp\left(-\frac{1}{2}(\beta - \bar{\beta})' \Sigma^{-1} (\beta - \bar{\beta})\right)}_{\pi(\beta)}$$

$$= \prod_{j=1}^{J} \left\{ \exp\left(-h_j \sum_{k \in R_j - D_j} \exp(X_k' \beta)\right) \cdot \prod_{l \in D_j} \left(1 - \exp(-h_j \exp(X_l' \beta))\right) \times h_j^{(\alpha_{0j} - \alpha_{0j-1}) - 1} \exp(-c_0 h_j) \right\}$$

$$\times \exp\left(-\frac{1}{2}(\beta - \bar{\beta})' \Sigma^{-1} (\beta - \bar{\beta})\right)$$

① Update β

$$\pi(\beta \mid h, \text{data}) \propto \prod_{j=1}^{J} \exp\left(-h_j \sum_{k \in R_j - D_j} \exp(X_k' \beta)\right) \cdot \prod_{l \in D_j} \left(1 - \exp(-h_j \exp(X_l' \beta))\right)$$

$$\times \exp\left(-\frac{1}{2}(\beta - \bar{\beta})' \Sigma^{-1} (\beta - \bar{\beta})\right)$$

② Update h_j, $j = 1, \ldots, J$

$$\pi(h_j \mid \beta, \text{Data}) \propto \exp\left(-h_j \sum_{k \in R_j - D_j} \exp(X_k' \beta)\right) \cdot \prod_{l \in D_j} \left(1 - \exp(-h_j \exp(X_l' \beta))\right)$$

$$\times h_j^{(\alpha_{0j} - \alpha_{0j-1}) - 1} \exp(-c_0 h_j)$$

- Joint posterior.

- Full conditionals.

* Revisit [Example: Male Laryngeal Cancer Patients] Hyperparameters

```
> ###  set up hyperparameters
> hyper <- NULL
>
> ### Be \sim N_p(Beta_bar, Sig)
> ## fit the frequentist Cox to set hyperparamters
> coxph.fit <- coxph(Surv(time, delta) ~ as.factor(stage) + age,
method="breslow", data=larynx)
> hyper$Beta_bar <- as.matrix(coxph.fit$coefficient)
> hyper$Beta_bar[4] <- hyper$Beta_bar[4]*sd(larynx$age)
##  since I will standardize age
> hyper$Sig <- diag(2, p)
> hyper$Inv_Sig <- solve(hyper$Sig)
>
> ##  h_j \iid \GP(c0 H*, c0)
> hyper$eta0 <- 0.1
> hyper$k0 <- 1.5  ## shape for weibull
> hyper$c0 <- 1 ## prior belief
>
> hyper$a_diff <- hyper$c0*hyper$eta0
*(Int_dat$end^hyper$k0 -Int_dat$start^hyper$k0)
>
```

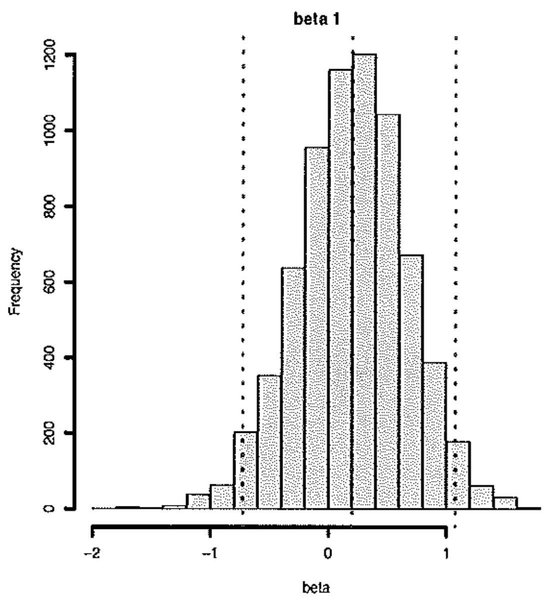
$\eta_0 t^{\kappa_0} = H^*(t)$

* Revisit [Example: Male Laryngeal Cancer Patients] How did I set intervals, $(s_{j-1}, s_j]$?

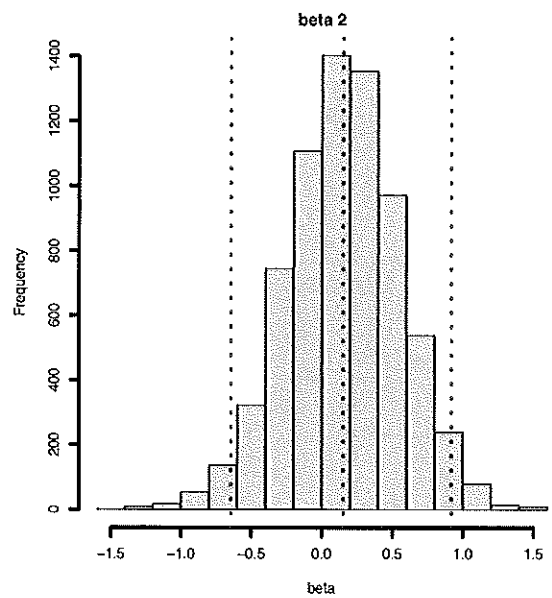
** Used distinct times.

```
> Int_dat <- sort(unique(larynx$time))  ## time grid
> Int_dat <- cbind(c(0,Int_dat[-length(Int_dat)]), Int_dat)
## (starting time, end time]
>
> J <- nrow(Int_dat)  ## number of intervals
> rownames(Int_dat) <- (1:J)
```

- Gamma Process Model



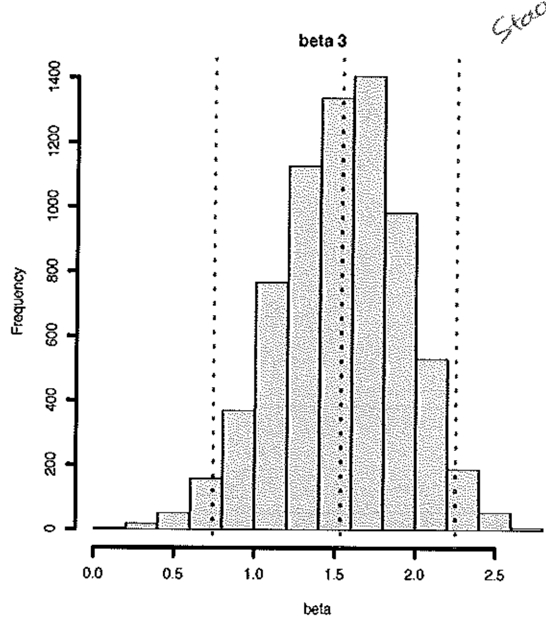
$\beta_1(\hat{\beta}_1 = 0.201)$



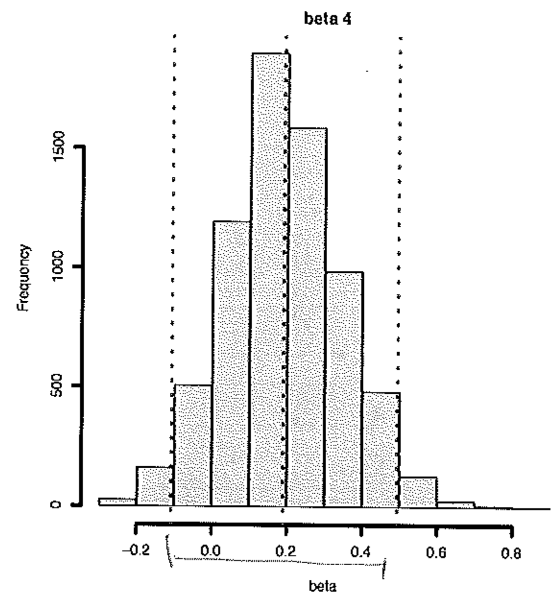
$\beta_2(\hat{\beta}_2 = 0.153)$

- Piecewise Constant Hazard Model

Stage IV



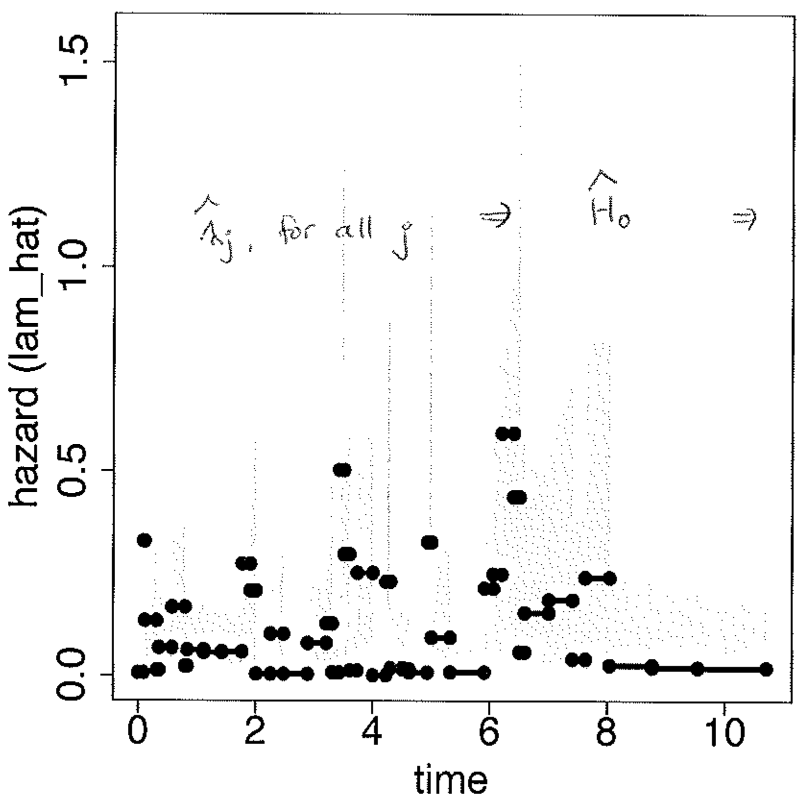
$\beta_3(\hat{\beta}_3 = 1.535)$



$\beta_4(\hat{\beta}_4 = 0.191)$

- Piecewise Constant Hazard Model

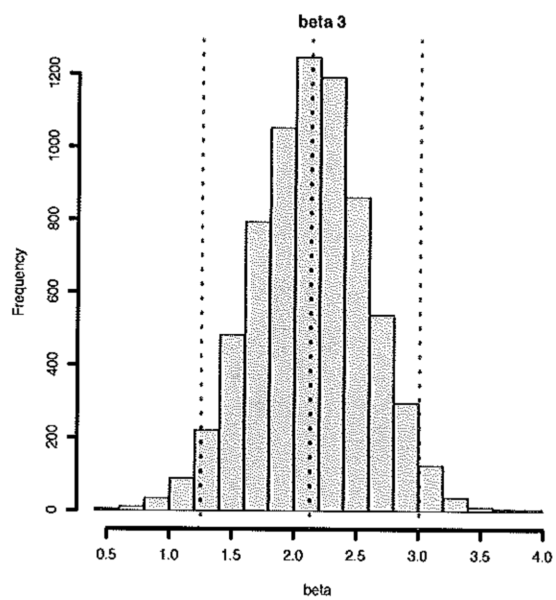
⋆ Posterior mean of λ_j with their 95% credible intervals



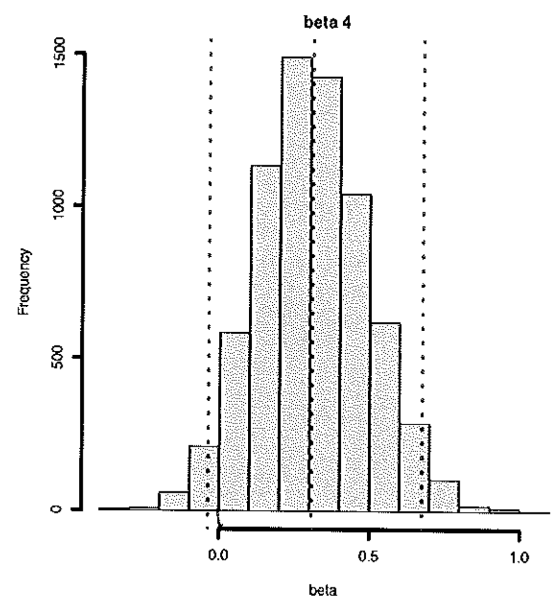
$\hat{\lambda}_j$, for all j $\Rightarrow$ $\hat{H}_0$ $\Rightarrow$ $\hat{S}_0 = e^{-\hat{H}_0}$ $\hat{S}(\cdot \mid X)$

$\hat{\beta}$

- Gamma Process Model



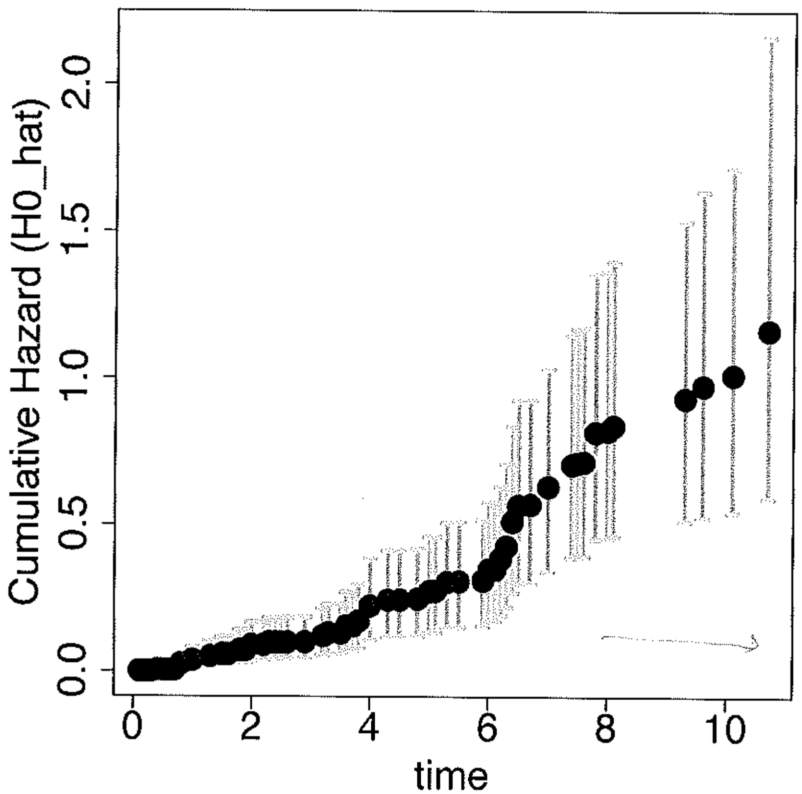
$\beta_3(\hat{\beta}_3 = 2.219)$



$\beta_4(\hat{\beta}_4 = 0.307)$

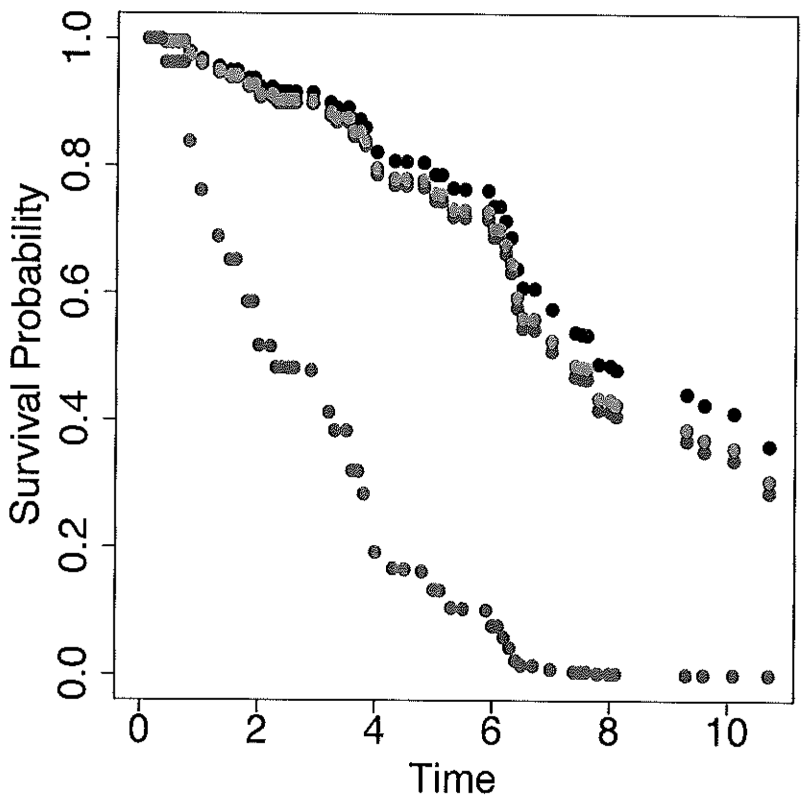
- Gamma Process Model

⋆ Posterior mean of H_0 with their 95% credible intervals



- Gamma Process Model

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



- Gamma Process Model (check with Kaplan-Meier estimates)