** Can we write the likelihood in T?

- We have $Y \mid \alpha, \lambda \sim V(\alpha, \lambda)$. $\alpha = \frac{1}{\sigma}$, $\lambda_i = \frac{\beta_0 + \beta' x_i}{\sigma}$
- Let $T = \exp(Y)$. What is the distribution of T?

  $\Rightarrow T \mid \alpha, \lambda \sim \text{Weibull}(\alpha, \lambda)$.

  ** pdf

$$\underline{f(t \mid \alpha, \lambda)} = \alpha t^{\alpha - 1} \exp(\lambda - \exp(\lambda) t^{\alpha}), \quad t > 0.$$

  ** Survival function: $\underline{S_T(t \mid \alpha, \lambda)} = \exp(-\exp(\lambda) t^{\alpha})$

- How to write down the likelihood for right-censored data in T?

$$t_i \mid \alpha, \beta_0, \beta \overset{\text{indep.}}{\sim} W\left(\alpha = \frac{1}{\sigma}, \ \lambda_i = \frac{\beta_0 + \beta' x_i}{\sigma}\right)$$

$$\Rightarrow \mathcal{L}(\beta_0, \beta, \sigma) = \prod_{i=1}^{n} \left\{ \frac{1}{\sigma} t_i^{\frac{1}{\sigma} - 1} \exp\left(\frac{\beta_0 + \beta' x_i}{\sigma} - \exp\left(\frac{\beta_0 + \beta' x_i}{\sigma}\right) t_i^{\frac{1}{\sigma}} \right) \right\}^{\nu_i}$$

$$\times \left\{ \exp\left(- \exp\left(\frac{\beta_0 + \beta' x_i}{\sigma}\right) t_i^{\frac{1}{\sigma}} \right) \right\}^{1 - \nu_i}$$

** How about the hazard function of T?

- Recall $h(t) = \alpha\gamma t^{\alpha-1}$ where $t \sim \text{Weibull}(\alpha, \gamma)$ and $\lambda = \log(\gamma)$.

$$h_0(t) = h(t \mid X=0) \overset{\alpha=\frac{1}{\sigma},\ \lambda=\frac{\beta_0}{\sigma}}{=} \left(\frac{1}{\sigma}\right) e^{\frac{\beta_0}{\sigma}} \cdot t^{\frac{1}{\sigma}-1}$$

$$h(t \mid X) \overset{\alpha=\frac{1}{\sigma},\ \lambda=\frac{\beta_0+\beta' x}{\sigma}}{=} \left(\frac{1}{\sigma}\right) e^{\frac{\beta_0+\beta' x}{\sigma}} \cdot t^{\frac{1}{\sigma}-1}$$

$$= \left(\frac{1}{\sigma}\right) e^{\frac{\beta_0}{\sigma}} t^{\frac{1}{\sigma}-1} \cdot e^{\frac{\beta' x}{\sigma}}$$

$$= h_0(t)\, e^{\frac{\beta' x}{\sigma}}$$

- We will discuss the proportional hazards property later; The Weibull distribution has both the proportional hazards and accelerated failure time properties (not true for the other distributions that we will discuss later).

Find Maximum Likelihood Estimates (MLE) of the parameters!

- MLE and covariance matrix of β_0, $\boldsymbol{\beta}$ and σ can be found numerically. Do tests on $\boldsymbol{\beta}$ or find a confidence interval based on asymptotic normality.

- R function `survreg` with option dist="weibull" provides MLE of β_0, $\boldsymbol{\beta}$ and σ.

- Be careful with the interpretation of the output!
  - ⋆⋆ Their `(Intercept)` is an estimate of our $-\beta_0$.
  - ⋆⋆ Their estimates of coefficients are estimates of our $-\beta_j$
  - ⋆⋆ Their `Log(scale)` is an estimate of our $\log(\sigma)$

∗ [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.

$\begin{cases} 1 & \text{if stage III} \\ 0 & \end{cases}$

$*$ [Example: Male Laryngeal Cancer Patients] We use the accelerated failure-time model using the main effects of age and stage for this data;

$\begin{cases} 1 & \text{if Stage II} \\ 0 & \text{ow} \end{cases}$ $\begin{cases} 1 & \text{if stage III} \\ 0 & \end{cases}$ Age

$$Y = \log(T) = -\beta_0 - \beta_1 X_1 - \beta_2 X_2 - \beta_3 X_3 - \beta_4 X_4 + \sigma W,$$

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. **Assume** $W \sim$ standard V

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

= 0 censored
= 1 observed

$(0, 10)^+$

$(0, 10)$

∗ [Example: Male Laryngeal Cancer Patients]

```
> summary(WeiFit)
```

$-3.5288 \pm 1.96 \times .9041 = (-5.3, -1.75)$

```
Call:
survreg(formula = Surv(time, delta) ~ as.factor(stage) + age,
    data = larynx, dist = "weibull")
                      Value Std. Error      z        p
(Intercept)          3.5288     0.9041  3.903 9.50e-05
as.factor(stage)2   -0.1477     0.4076 -0.362 7.17e-01
as.factor(stage)3   -0.5866     0.3199 -1.833 6.68e-02
as.factor(stage)4   -1.5441     0.3633 -4.251 2.13e-05
age                 -0.0175     0.0128 -1.367 1.72e-01
Log(scale)          -0.1223     0.1225 -0.999 3.18e-01

Scale= 0.885
```

$= -\hat{\beta}_3$ (beside -1.5441)

$\hat{\beta}_3 = 1.5441$

$h(t) = \alpha \gamma t^{\alpha - 1}$

$= \hat{\sigma}$ (beside Scale= 0.885) $\quad \hat{\alpha} = \frac{1}{\hat{\sigma}} = \frac{1}{0.885} = 1.13 \Rightarrow$

```
Weibull distribution
Loglik(model)= -141.4   Loglik(intercept only)= -151.1
Chisq= 19.37 on 4 degrees of freedom, p= 0.00066
Number of Newton-Raphson Iterations: 5
n= 90
```

∗ [Example: Male Laryngeal Cancer Patients] We use the accelerated failure-time model using the main effects of age and stage for this data; $+\beta_0 + \beta_1 X_1 +$

$$Y = \log(T) = -\beta_0 \boxed{-\beta_1} X_1 - \beta_2 X_2 - \beta_3 X_3 - \beta_4 X_4 + \sigma W,$$

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. **Assume** $W \sim$ standard V

- A positive value of β is indicative of decreased survival.
- Interpretation: $\frac{h(t|\text{IV})}{h_0(t)} = \frac{h(t\mid \text{IV})}{h(t\mid \text{I})} = e^{\frac{\hat{\beta}_3}{\sigma}} = e^{\frac{1.5411}{0.885}} = e^{1.745} = 5.73$
  - ⋆ We find that the relative risk of death (ratio of the hazards) for a Stage IV patient compared to a Stage I patient is exp(1.745) = 5.73.
  - ⋆ The acceleration factor for Stage IV disease compared to Sage I disease is exp(1.54) = 4.68, so the median lifetime for a Stage I patient is estimated to 4.68 times that of a Stage IV patient.

$\cancel{S(t\mid \text{I}) = S(4.68 \times t \mid \text{IV})}$

$S(t\mid \text{IV}) = S(4.68 \times t \mid \text{I})$

We are BAYESIAN!

$$Y = -\beta_0 - \boldsymbol{\beta}'\mathbf{X} + \sigma W$$

where $W \sim$ standardized extreme value distribution

- Parametric Bayesian: We place priors on σ and $(\beta_0, \boldsymbol{\beta})$.

  For example,

  ⋆⋆ $\sigma(= 1/\alpha) \sim \mathrm{IG}$.

  $\tilde{\boldsymbol{\beta}} = \underbrace{(\beta_0, \underset{p\text{-dim}}{\boldsymbol{\beta}})}_{(p+1)\text{-dim}} \sim \mathrm{N}_{p+1}(\bar{\boldsymbol{\beta}}, \Sigma_{\beta})$.

$W \sim$ Standard EV

$\tilde{\beta} \sim N_{p+1}(\bar{\beta}, \Sigma_\beta)$, $\sigma \sim IG(a_\sigma, b_\sigma)$

$$\tilde{x}_i = \begin{bmatrix} 1 \\ x_1 \end{bmatrix}_{(p+1)} \qquad \tilde{y}_i = \begin{cases} c_i & \text{if censored} \\ y_i & \text{if not censored} \end{cases}$$

$$\pi(\tilde{\beta}, \sigma \mid \tilde{y}, \nu) \propto \prod_{i=1}^{n} \left\{ \frac{1}{\sigma} \exp\left(\frac{y_i + \overbrace{\tilde{\beta}'\tilde{x}_i}^{= \beta_0 + \beta' x_i}}{\sigma} - \exp\left(\frac{y_i + \tilde{\beta}'\tilde{x}_i}{\sigma}\right)\right)\right\}^{\nu_i}$$

$$\times \left\{ \exp\left(-\exp\left(\frac{y_i + \tilde{\beta}'\tilde{x}_i}{\sigma}\right)\right)\right\}^{1-\nu_i}$$

$$\times \underbrace{(\sigma)^{-a_\sigma - 1} \exp\left(-\frac{b_\sigma}{\sigma}\right)}_{\propto \pi(\sigma)} \cdot \underbrace{\exp\left(-\frac{1}{2}(\tilde{\beta} - \bar{\beta})'\Sigma_\beta^{-1}(\tilde{\beta} - \bar{\beta})\right)}_{\propto \pi(\tilde{\beta})}$$

$$\otimes = \left(\frac{1}{\sigma}\right)^{\sum \nu_i} \exp\left(\sum_{i=1}^{n} \left(\nu_i \frac{y_i + \tilde{\beta}'\tilde{x}_i}{\sigma} - \exp\left(\frac{y_i + \tilde{\beta}' x_i}{\sigma}\right)\right)\right)$$

$$\times \left(\frac{1}{\sigma}\right)^{a_\sigma + 1} \exp\left(-\frac{b_\sigma}{\sigma}\right) \exp\left(-\frac{1}{2}(\tilde{\beta} - \bar{\beta})'\Sigma_\beta^{-1}(\tilde{\beta} - \bar{\beta})\right)$$

① Update $\tilde{\beta}$ $(p+1)$-dim vector $\beta_R \in \mathbb{R}$

$$\pi(\tilde{\beta} \mid \sigma, \tilde{y}, \nu) \propto \exp\left(\sum_{i=1}^{n} \left(\frac{\nu_i(y_i + \tilde{\beta}'\tilde{x}_i)}{\sigma} - \exp\left(\frac{y_i + \tilde{\beta}' x_i}{\sigma}\right)\right)\right)$$

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

② Update $\sigma > 0$

$$\pi(\sigma \mid \tilde{\beta}, \tilde{y}, \nu) \propto \left(\frac{1}{\sigma}\right)^{\sum \nu_i + a_\sigma + 1} \exp\left\{ \sum_{i=1}^{n} \left(\frac{\nu_i(y_i + \tilde{\beta}'\tilde{x}_i)}{\sigma} - \exp\left(\frac{y_i + \tilde{\beta}'\tilde{x}_i}{\sigma}\right)\right)\right\} \times \exp\left(-\frac{b_\sigma}{\sigma}\right)$$

$*$ [Example: Male Laryngeal Cancer Patients] We use the accelerated failure-time model using the main effects of age and stage for this data;

$$Y = \log(T) = -\beta_0 - \beta_1 X_1 - \beta_2 X_2 - \beta_3 X_3 - \beta_4 X_4 + \sigma W,$$

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. **Assume** $W \sim$ extreme value

$e^{\frac{\beta x}{\sigma}}$

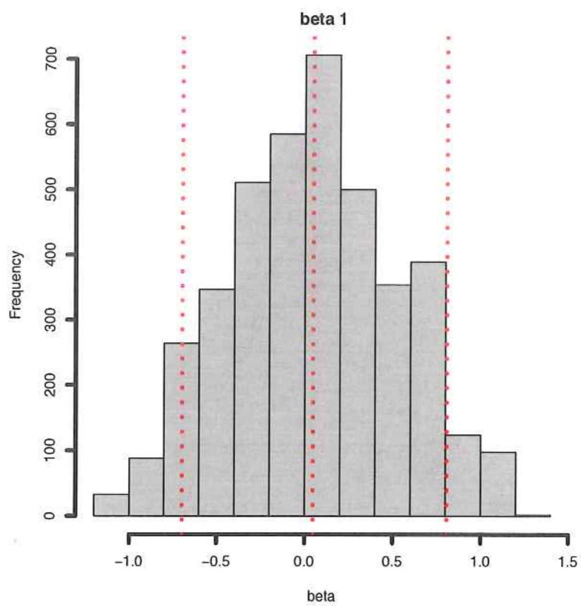
⋆⋆ Assume a priori independence for β_k, $k = 0, \ldots, 4$,

$$\beta_k \overset{iid}{\sim} \mathrm{N}(\bar{\beta}_k, 2.0^2).$$

⋆⋆ $\sigma \sim \mathrm{IG}(5, 5)$, independent of β_k

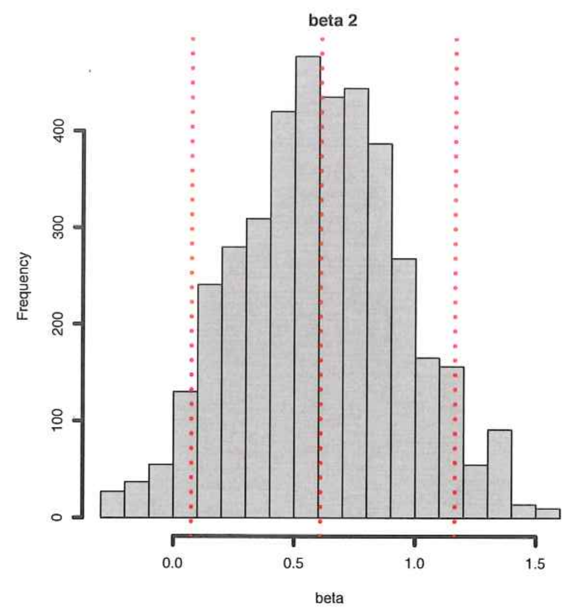
⋆⋆ Use the output from `survreg` to specify the hyperparameter values and initialize the MCMC.

- Weibull-Posterior distribution



$\beta_1(\hat{\beta}_1 = 0.048)$

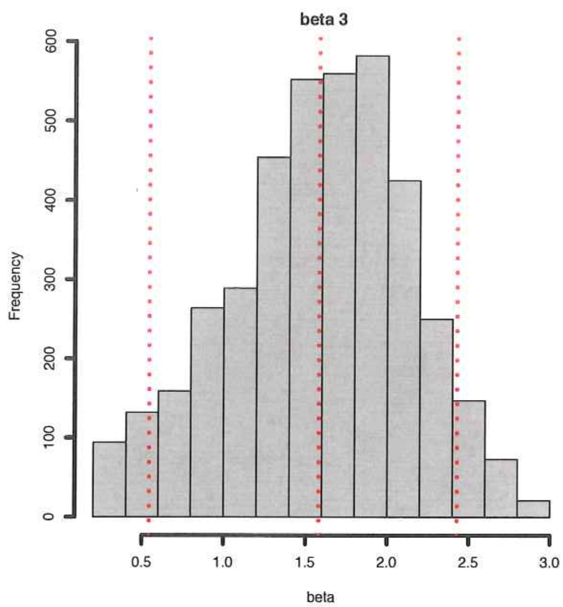
0.1477



$\beta_2(\hat{\beta}_2 = 0.608)$

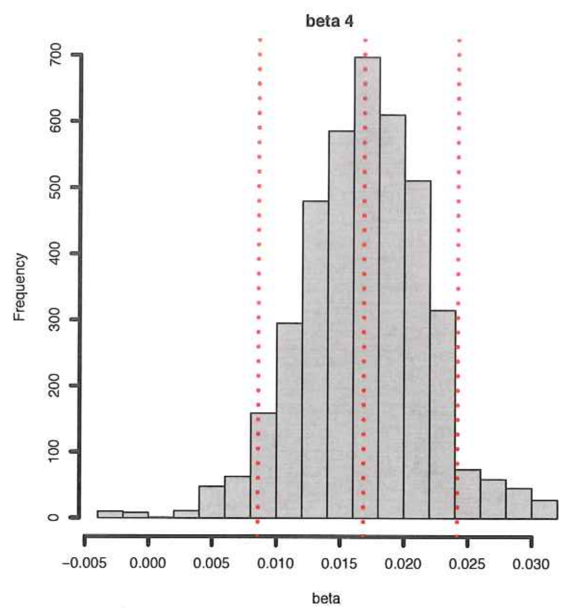
0.5866

- Weibull-Posterior distribution (contd)



$\beta_3(\hat{\beta}_3 = 1.580)$

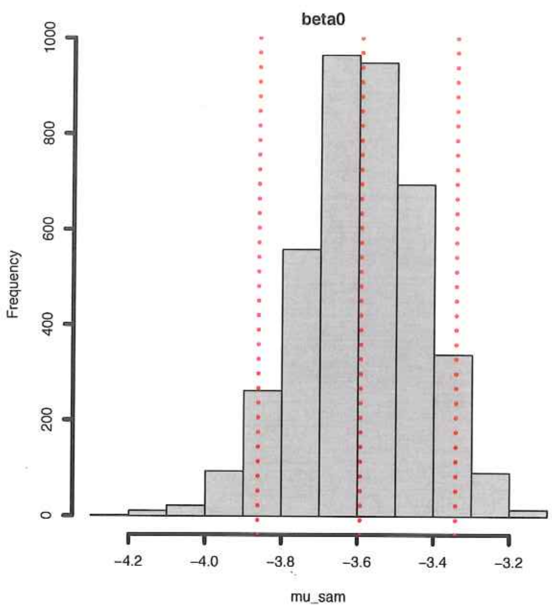
1.544



$\beta_4(\hat{\beta}_4 = 0.017)$

0.0175

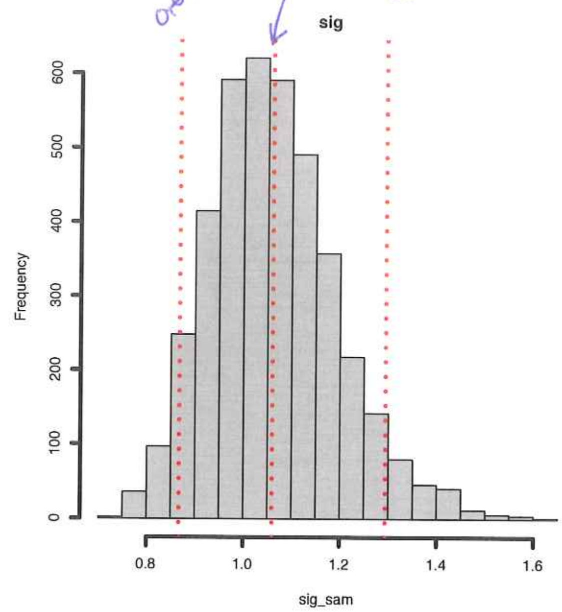
- Weibull-Posterior distribution (contd)



$\beta_0(\hat{\beta}_0 = -3.593)$

−3.5288

2.5%
~~0.025~~
posterior mean
97.5%



$\sigma(\hat{\sigma}_1 = 1.060)$

0.885

$Y_i \mid \beta_0, \beta, \sigma \sim V\left(\frac{1}{\sigma}, \frac{\beta_0 + \beta' x_i}{\sigma}\right)$

$\beta_0^* = \frac{\beta_0}{\sigma}, \quad \frac{\beta_k}{\sigma} = \beta_k^*$

Other parameterization! (ICS 2.2 & 2.3)

- We have $Y_i \mid \alpha, \lambda_i \overset{indep}{\sim} V(\alpha, \lambda_i)$.

  Equivalently, $T_i \mid \alpha, \lambda_i \overset{indep}{\sim} W(\alpha, \lambda_i)$

  ⋆⋆ We let $\lambda_i = \beta_0^\star + (\boldsymbol{\beta}^\star)'\mathbf{X}$. Then place $(\beta_0^\star, \boldsymbol{\beta}^\star) \sim \mathrm{N}_{p+1}(\bar{\boldsymbol{\beta}}^\star, \Sigma_{\beta^\star})$.

  ⋆⋆ We let $\alpha \sim \text{Gamma}(\alpha_0, \kappa_0)$.

♣ Parametric Approach 2: Standard Logistic Distribution for W

- Survival time $T \Rightarrow Y = \log(T)$
- Let $Y \mid \mu, \sigma \sim \text{Logistic}(\mu, \sigma)$.
- The distribution of $W = \frac{y-\mu}{\sigma}$?

  $f(\omega) = \frac{e^{\omega}}{(1+e^{\omega})^2}$,

  $\Rightarrow$ W follows the standardized logistic distribution. $\omega \in \mathbb{R}$
- Let's consider our $Y = -\beta_0 - \boldsymbol{\beta}'\mathbf{X} + \sigma W$ where $W \sim$ standardized logistic distribution.

  $\Rightarrow$ What is the distribution of our Y?

  $Y \mid \beta_0, \beta, \sigma \sim \text{logistic}(-\beta_0 - \beta' x, \sigma)$

- Recall! Let $Y \mid \mu, \sigma \sim \text{Logistic}(\mu, \sigma)$

$$f(y \mid \mu, \sigma) = \frac{\exp(\frac{y-\mu}{\sigma})}{\sigma(1 + \exp(\frac{y-\mu}{\sigma}))^2}, \quad -\infty < y < \infty,$$

⋆⋆ Survival function:

$$S_Y(y \mid \mu, \sigma) = \frac{1}{1 + \exp(\frac{y-\mu}{\sigma})}.$$

- Now consider our $Y = -\beta_0 - \boldsymbol{\beta}'\mathbf{X} + \sigma W$
- We know $Y \mid \boldsymbol{\beta}, \sigma \sim \text{Logistic}(\mu, \sigma)$ where

  ⋆⋆ $\mu_i = -\beta_0 - \beta' X_i$.
- How to write down the likelihood for right-censored data in Y?

$$\mathcal{L}(\beta_0, \beta, \sigma \mid \tilde{y}, \nu) = \prod_{i=1}^{n} \left\{ \frac{\exp\left(\frac{y_i + \beta_0 + \beta' x_i}{\sigma}\right)}{\sigma\left(1 + \exp\left(\frac{y_i + \beta_0 + \beta' x_i}{\sigma}\right)\right)^2} \right\}^{\nu_i}$$

$$\times \left\{ \frac{1}{1 + \exp\left(\frac{y_i + \beta_0 + \beta' x_i}{\sigma}\right)} \right\}^{1-\nu_i}$$

- Assume $Y \mid \mu, \sigma \sim \text{Logistic}(\mu, \sigma)$
- Then $T = \exp(Y) \mid \alpha, \lambda \sim \text{Log-Logistic}(\alpha, \lambda)$

$$f(t \mid \alpha, \lambda) = \frac{\lambda \alpha t^{\alpha-1}}{(1+\lambda t^{\alpha})^2}, \ 0 < t,$$

where $\alpha = 1/\sigma > 0$ and $\lambda = \exp(-\mu/\sigma)$.

⋆⋆ Survival function:

$$S_T(t \mid \alpha, \lambda) = \frac{1}{(1+\lambda t^{\alpha})}.$$

$\alpha = \frac{1}{\sigma}$

$\lambda_i = \exp(-\frac{\mu_i}{\sigma})$

$= \exp(\frac{\beta_0 + \beta' x_i}{\sigma})$

- We can write the likelihood in T. Try!

- proportional odds model

(1) $\frac{S_0(t)}{1 - S_0(t)} = \frac{\frac{1}{1+\lambda t^{\alpha}}}{1 - \frac{1}{1+\lambda t^{\alpha}}} = \frac{1}{\lambda t^{\alpha}} = \frac{1}{e^{\beta_0/\sigma} \cdot t^{1/\sigma}}$

(2) $\frac{S(t \mid x)}{1 - S(t \mid x)} = \frac{1}{e^{\frac{\beta_0 + \beta \cancel{x}}{\sigma}} t^{1/\sigma}} = \frac{1}{e^{\beta_0/\sigma} t^{1/\sigma}} \cdot \exp(-\frac{\beta' x}{\sigma})$

$= \frac{S_0(t)}{1 - S_0(t)} \cdot \exp(-\frac{\beta' x}{\sigma})$

$$\log\left(\frac{S(t|x)}{1-S(t|x)}\right) = \log\left(\frac{S_0(t)}{1-S_0(t)}\right) - \frac{\beta' x}{\sigma}$$

* [Example: Male Laryngeal Cancer Patients] We use the accelerated failure-time model using the main effects of age and stage for this data;

$$Y = \log(T) = -\beta_0 - \beta_1 X_1 - \beta_2 X_2 - \beta_3 X_3 - \beta_4 X_4 + \sigma W,$$

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. **Assume** $W \sim$ Logistic

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

∗ [Example: Male Laryngeal Cancer Patients]

```
> summary(LLogFit)

Call:
survreg(formula = Surv(time, delta) ~ as.factor(stage) + age,
    data = larynx, dist = "loglogistic")
                    Value Std. Error      z        p
(Intercept)        3.1022     0.9527  3.256 1.13e-03
as.factor(stage)2 -0.1257     0.4152 -0.303 7.62e-01
as.factor(stage)3 -0.8057     0.3539 -2.277 2.28e-02
as.factor(stage)4 -1.7661     0.4257 -4.149 3.34e-05
age               -0.0151     0.0138 -1.095 2.73e-01
Log(scale)        -0.3352     0.1202 -2.788 5.31e-03

Scale= 0.715

Log logistic distribution
Loglik(model)= -141.6   Loglik(intercept only)= -151.6
Chisq= 20.07 on 4 degrees of freedom, p= 0.00048
Number of Newton-Raphson Iterations: 4
n= 90
```

↻ We are BAYESIAN!

$$Y = -\beta_0 - \boldsymbol{\beta}'\mathbf{X} + \sigma W$$

where $W \sim$ standardized logistic distribution

- Parametric Bayesian: We place priors on σ and $(\beta_0, \boldsymbol{\beta})$.

  For example,

  ⋆⋆ $\sigma(= 1/\alpha) \sim \text{IG}$.

  ⋆⋆ $(\beta_0, \boldsymbol{\beta}) \sim \text{N}_{p+1}(\bar{\boldsymbol{\beta}}, \Sigma_\beta)$.

$*$ [Example: Male Laryngeal Cancer Patients] We use the accelerated failure-time model using the main effects of age and stage for this data;

$$Y = \log(T) = -\beta_0 - \beta_1 X_1 - \beta_2 X_2 - \beta_3 X_3 - \beta_4 X_4 + \sigma W,$$

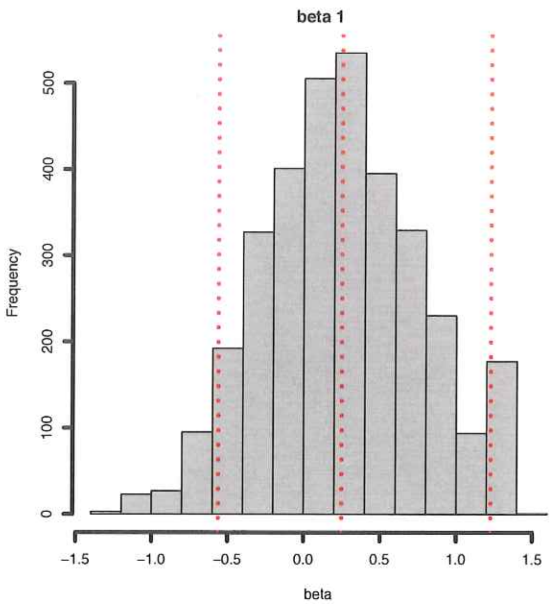
where X_k, $k = 1, 2, 3$ are the indicators of stage II, III and IV disease, respectively, and X_4 is the age of the patient. **Assume** $W \sim$ logistic

- ⋆⋆ Assume a priori independence for β_k, $k = 0, \ldots, 4$,

$$\beta_k \overset{iid}{\sim} \mathrm{N}(\bar{\beta}_k, 2.0).$$

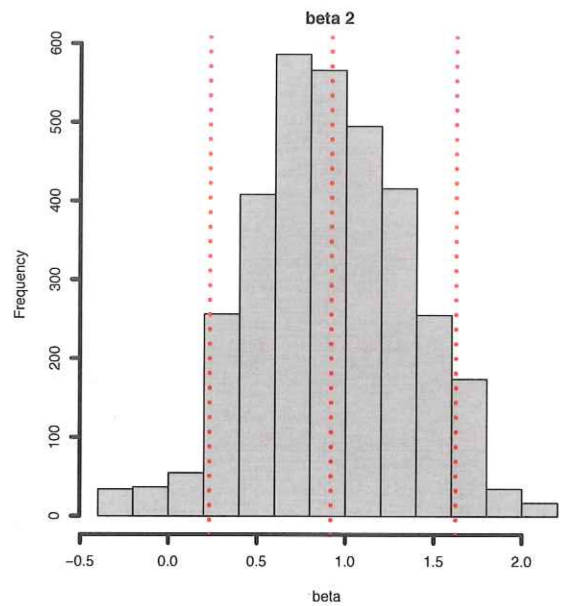
- ⋆⋆ $\sigma \sim \mathrm{IG}(5, 5)$, independent of β_k
- ⋆⋆ Use the output from `survreg` to specify the hyperparameter values and initialize the MCMC.

- Log-Logistic–Posterior distribution



$\beta_1(\hat{\beta}_1 = 0.247)$

0.1257

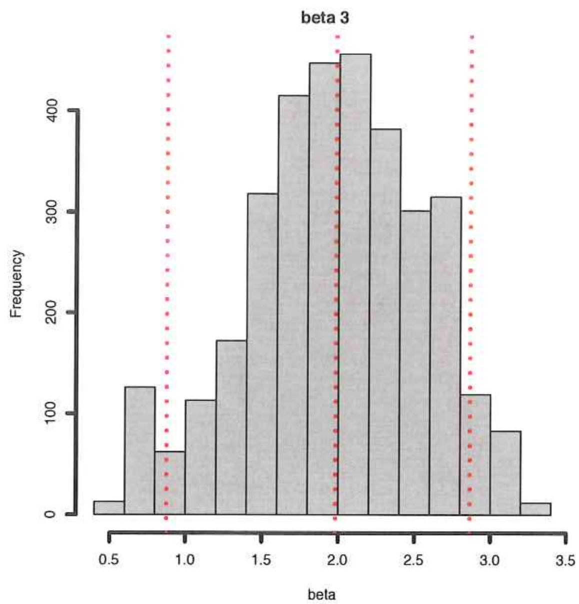


$\beta_2(\hat{\beta}_2 = 0.918)$

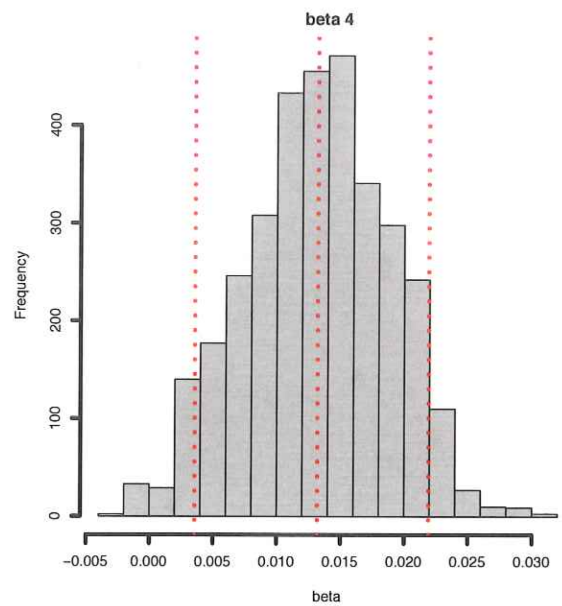
0.8057

$p(y^* \mid y) = \int p(y^* \mid \theta)\, \pi(\theta \mid y)\, d\theta$

- Log-Logistic–Posterior distribution (contd)

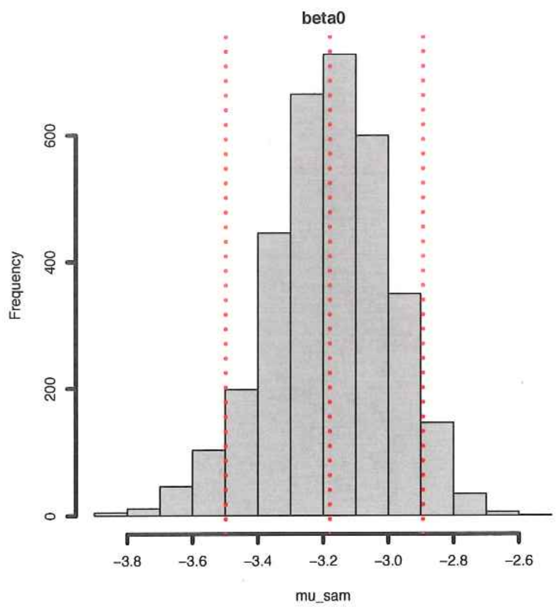


$\beta_3(\hat{\beta}_3 = 1.982)$



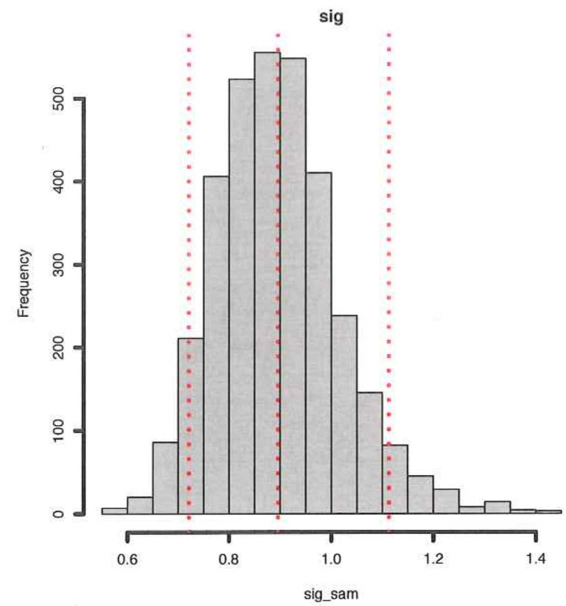
$\beta_4(\hat{\beta}_4 = 0.013)$

- Log-Logistic–Posterior distribution (contd)



$\beta_0(\hat{\beta}_0 = -3.18)$

-3.1022



$\sigma(\hat{\sigma} = 0.900)$

0.715

♣ Parametric Approach 3: Std Normal Distribution for W

- Survival time $T \Rightarrow Y = \log(T)$
- Let $Y \mid \mu, \sigma^2 \sim \mathrm{N}(\mu, \sigma^2)$. $\Rightarrow$ $T \sim LN(\mu, \sigma^2)$
- The distribution of $W = \frac{Y-\mu}{\sigma}$?

  $\Rightarrow$ W follows the standard normal distribution.

- Let's consider our $Y = -\beta_0 - \beta'\mathbf{X} + \sigma W$ where $W \sim$ standard normal distribution.

  $\Rightarrow$ What is the distribution of our Y?

  $Y \mid \beta_0, \beta, \sigma \sim N(-\beta_0 - \beta' X, \sigma^2)$

Recall!

- Let $Y \mid \mu, \sigma^2 \sim \mathrm{N}(\mu, \sigma^2)$.

$$f(y \mid \mu, \sigma) = \frac{\exp\{-\frac{1}{2}(\frac{y-\mu}{\sigma})^2\}}{\sqrt{2\pi\sigma^2}}, \quad -\infty < y < \infty.$$

⋆⋆ Survival function:

$$S_Y(y \mid \mu, \sigma^2) = 1 - \Phi(\frac{y-\mu}{\sigma}),$$

where $\Phi(\cdot)$ is the cdf of the standard normal distribution

Now consider our $Y = -\beta_0 - \boldsymbol{\beta}'\mathbf{X} + \sigma W$ and $Y = \log(T)$.

- How to write down the likelihood for right-censored data in Y?

$$\mathcal{L}(\beta_0, \beta, \sigma \mid \tilde{y}, \nu) \propto \prod_{i=1}^{n} \left\{ \phi\left(\frac{\tilde{y}_i + \beta_0 + \beta' x_i}{\sigma} \right) \right\}^{\nu_i} \times \left\{ 1 - \Phi\left(\frac{\tilde{y}_i + \beta_0 + \beta' x_i}{\sigma} \right) \right\}^{1-\nu_i}$$

- We can write the likelihood in T as well. Try!

∗ [Example: Male Laryngeal Cancer Patients] We use the accelerated failure-time model using the main effects of age and stage for this data;

$$Y = \log(T) = -\beta_0 - \beta_1 X_1 - \beta_2 X_2 - \beta_3 X_3 - \beta_4 X_4 + \sigma W,$$

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. **Assume** $W \sim \mathrm{N}(0, 1)$

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

$*$ [Example: Male Laryngeal Cancer Patients]

```
> summary(LNFit)

Call:
survreg(formula = Surv(time, delta) ~ as.factor(stage) + age,
    data = larynx, dist = "lognormal")
                   Value Std. Error     z        p
(Intercept)        3.3832     0.9356  3.62 2.99e-04
as.factor(stage)2 -0.1989     0.4423 -0.45 6.53e-01
as.factor(stage)3 -0.8995     0.3634 -2.48 1.33e-02
as.factor(stage)4 -1.8574     0.4427 -4.20 2.72e-05
age               -0.0185     0.0137 -1.35 1.77e-01
Log(scale)         0.2341     0.1065  2.20 2.79e-02
```

Scale= 1.26 = $\hat{\sigma}$

```
Log Normal distribution
Loglik(model)= -141.4   Loglik(intercept only)= -151.8
Chisq= 20.91 on 4 degrees of freedom, p= 0.00033
Number of Newton-Raphson Iterations: 4
n= 90
```

↻ We are BAYESIAN!

$$Y = -\beta_0 - \boldsymbol{\beta}'\mathbf{X} + \sigma W$$

where $W \sim$ standard Normal distribution

- Parametric Bayesian: We place priors on σ^2 and $(\beta_0, \boldsymbol{\beta})$. *For example,*
  - ⋆⋆ $\sigma^2 \sim \text{IG}$.
  - ⋆⋆ $(\beta_0, \boldsymbol{\beta}) \sim \text{N}_{p+1}(\bar{\boldsymbol{\beta}}, \Sigma_\beta)$.
- ICS 2.4
  - ⋆⋆ $\tau = 1/\sigma^2 \sim \text{Gamma}(\alpha_0/2, \lambda_0/2)$.
  - ⋆⋆ $\boxed{\pi(\beta_0, \boldsymbol{\beta}) \propto 1}$ or $\text{N}_{p+1}(\boldsymbol{\mu}_0, \tau^{-1}\Sigma_0)$.

* [Example: Male Laryngeal Cancer Patients] We use the accelerated failure-time model using the main effects of age and stage for this data;

$$Y = \log(T) = -\beta_0 - \beta_1 X_1 - \beta_2 X_2 - \beta_3 X_3 - \beta_4 X_4 + \sigma W,$$

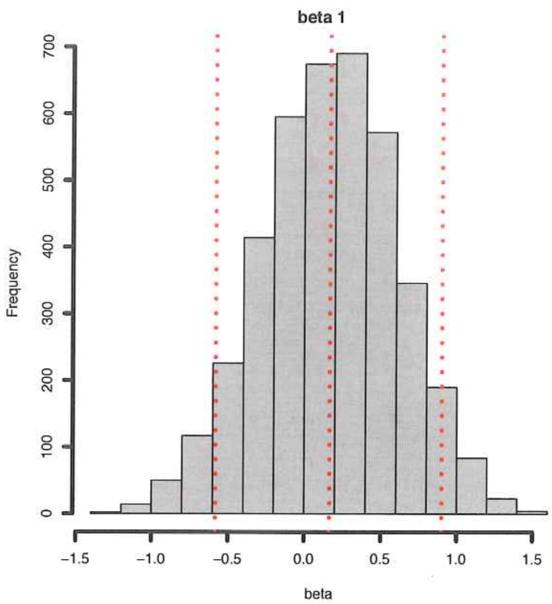
where X_k, $k = 1, 2, 3$ are the indicators of stage II, III and IV disease, respectively, and X_4 is the age of the patient. **Assume** $W \sim$ standard Normal

- ⋆⋆ Assume a priori independence for β_k, $k = 0, \ldots, 4$,

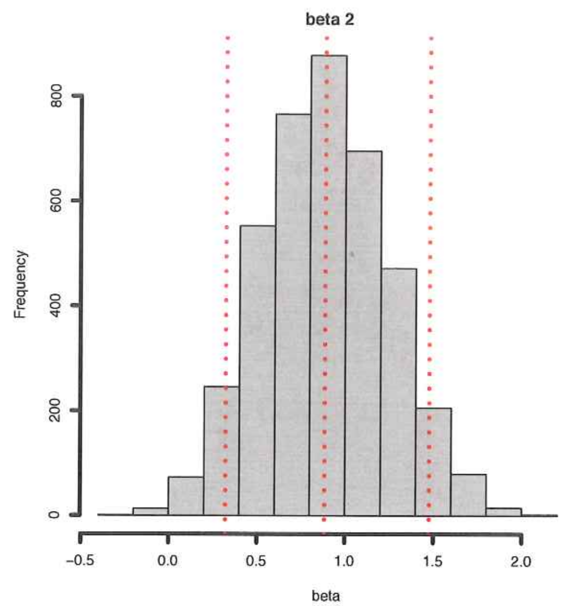
$$\beta_k \overset{iid}{\sim} \mathrm{N}(\bar{\beta}_k, 2.0).$$

- ⋆⋆ $\sigma^2 \sim$ IG$(5, 5)$, independent of β_k
- ⋆⋆ Use the output from `survreg` to specify the hyperparameter values and initialize the MCMC.

- Log-Normal–Posterior distribution

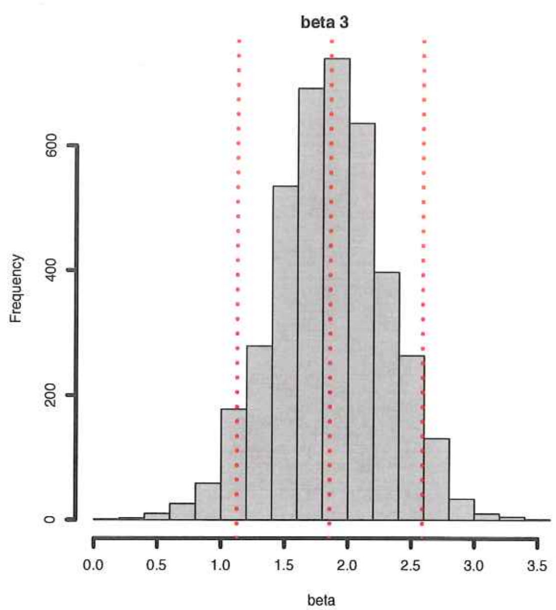


$\beta_1(\hat{\beta}_1 = 0.164)$

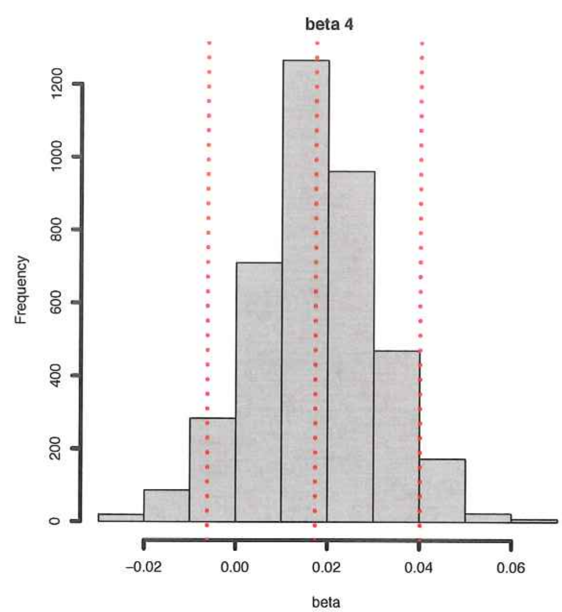


$\beta_2(\hat{\beta}_2 = 0.885)$

- Log-Normal–Posterior distribution (contd)

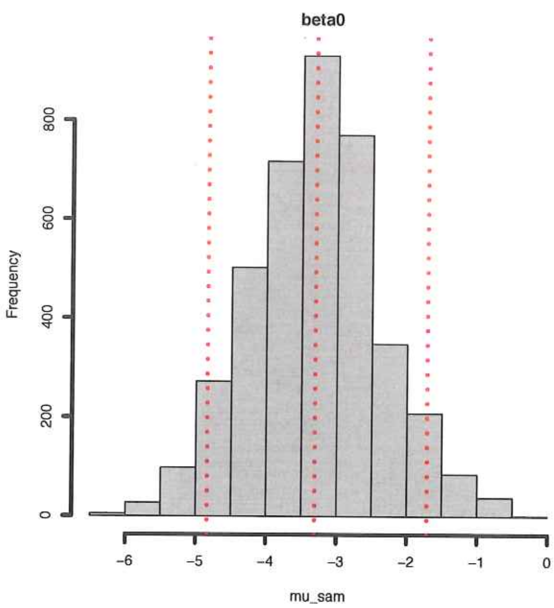


$\beta_3(\hat{\beta}_3 = 1.858)$

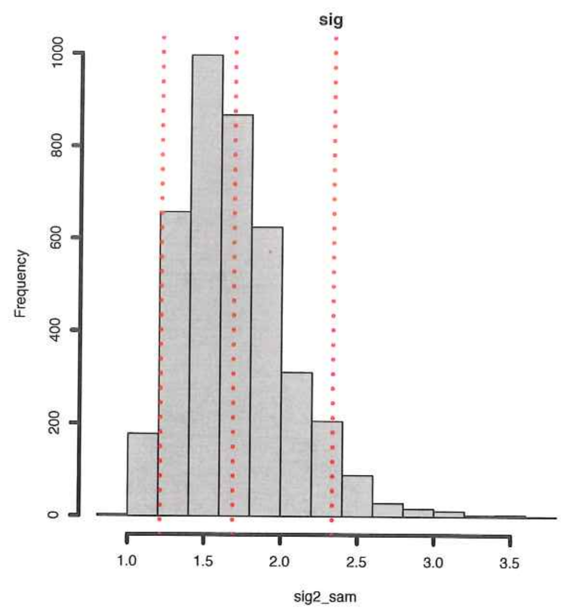


$\beta_4(\hat{\beta}_4 = 0.017)$

- Log-Normal–Posterior distribution (contd)



$\beta_0(\hat{\beta}_0 = -3.309)$



$\sigma^2(\hat{\sigma}^2 = 1.688)$