- (ICS Example 4.3) Kidney Infection Data:
  - From 38 kidney patients using portable dialysis equipment.
  - Recorded the times of infection from tie time of insertion of the catheter (thin plastic tube) for first and second infections.

    " A nephrostomy tube is a catheter (thin plastic tube) that is inserted through your skin and into your kidney. The nephrostomy tube is placed to drain urine from your kidney into a collecting bag outside your body. You may need one tube for each kidney." from Googling
  - After the occurrence or censoring of the first infection, sufficient time was allowed for the infection to be cured before the second insertion was allowed.
  - Covariates: Sex, age and disease types

Table 1
Recurrence data and frailty estimates

Patient number	Recurrence times	Event types	Age	Sex	Disease type	Frailty estimate
1	8, 16	1, 1	28	1	3	2.3
2	23, 13	1, 0	48	2	0	1.9
3	22, 28	1, 1	32	1	3	1.2
4	447, 318	1, 1	31–32	2	3	.5
5	30, 12	1, 1	10	1	3	1.5
6	24, 245	1, 1	16–17	2	3	1.1
7	7, 9	1, 1	51	1	0	3.0
8	511, 30	1, 1	55–56	2	0	.5
9	53, 196	1, 1	69	2	1	.7
10	15, 154	1, 1	51–52	1	0	.4
11	7, 333	1, 1	44	2	1	.6
12	141, 8	1, 0	34	2	3	1.2
13	96, 38	1, 1	35	2	1	1.4
14	149, 70	0, 0	42	2	1	.4
15	536, 25	1, 0	17	2	3	.4
16	17, 4	1, 0	60	1	1	1.1
17	185, 177	1, 1	60	2	3	.8
18	292, 114	1, 1	43–44	2	3	.8
19	22, 159	0, 0	53	2	0	.5
20	15, 108	1, 0	44	2	3	1.3
21	152, 562	1, 1	46–47	1	2	.2
22	402, 24	1, 0	30	2	3	.6
23	13, 66	1, 1	62–63	2	1	1.7
24	39, 46	1, 0	42–43	2	1	1.0
25	12, 40	1, 1	43	1	1	.7
26	113, 201	0, 1	57–58	2	1	.5
27	132, 156	1, 1	10	2	0	1.1
28	34, 30	1, 1	52	2	1	1.8
29	2, 25	1, 1	53	1	0	1.5
30	130, 26	1, 1	54	2	0	1.5
31	27, 58	1, 1	56	2	1	1.7
32	5, 43	0, 1	50–51	2	1	1.3
33	152, 30	1, 1	57	2	2	2.9
34	190, 5	1, 0	44–45	2	0	.7
35	119, 8	1, 1	22	2	3	2.2
36	54, 16	0, 0	42	2	3	.7
37	6, 78	0, 1	52	2	2	2.1
38	63, 8	1, 0	60	1	2	1.2

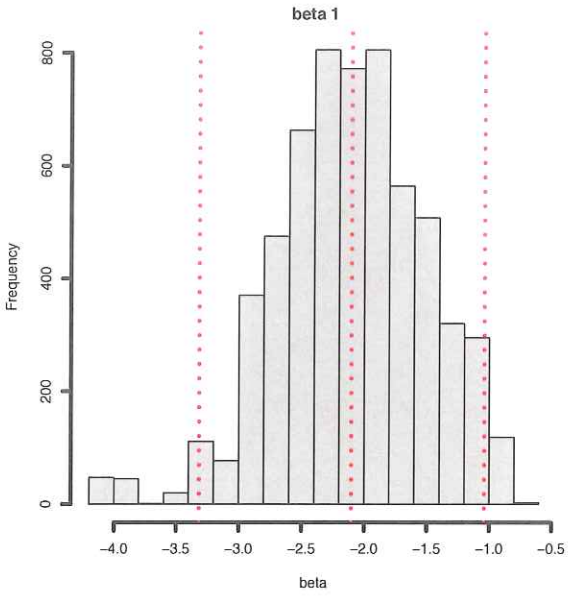
gence occurs. Estimates of β together with standard errors are:

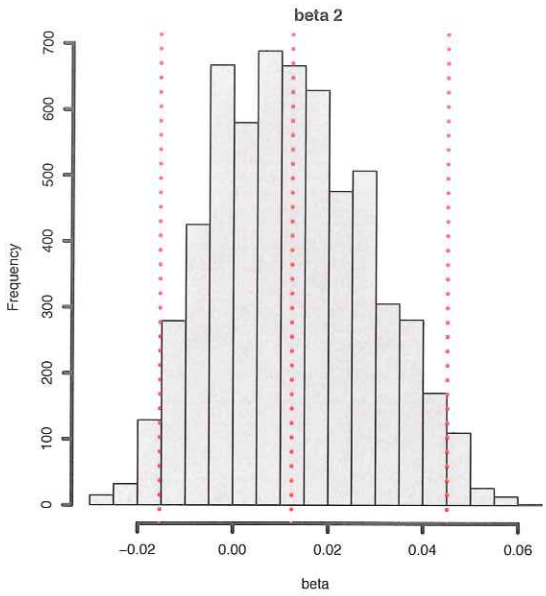
Variable	*Age*	*Sex*	*GN*	*AN*	*PKD*
Regression coefficient estimate	.0063	−1.7947	.2062	.4099	−1.2961
Standard error	.0134	.4337	.4840	.4937	.7120

Estimates of frailties are listed in the last column of Table 1.

The estimate of σ^2 is .3821. In general, the effect of the prior distribution on frailty terms is to shrink estimates toward the origin, thereby biasing the estimate of σ^2. Nevertheless the frailty estimates given in Table 1 appear very reasonable. The only regression coefficient that is significantly large compared to its standard error is that of the sex variable, indicating a lower infection rate for female patients.

- (ICS Example 4.3 – Model II) Weibull Baseline Hazard Model with Multiplicative Gamma Frailties

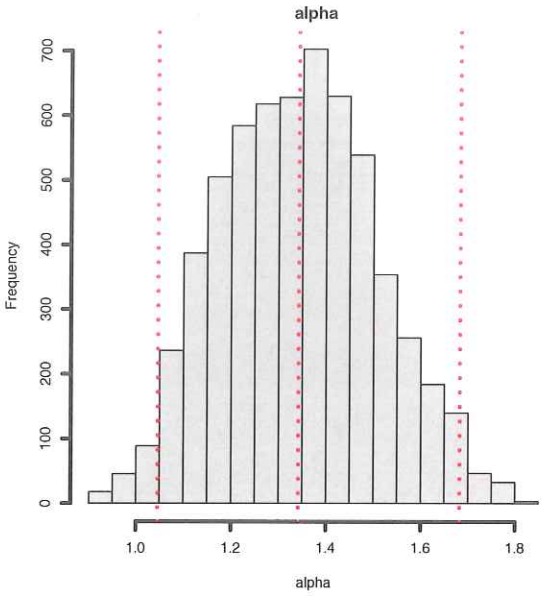




$\beta_1(\hat{\beta}_1 = -2.104\ (0.597))$

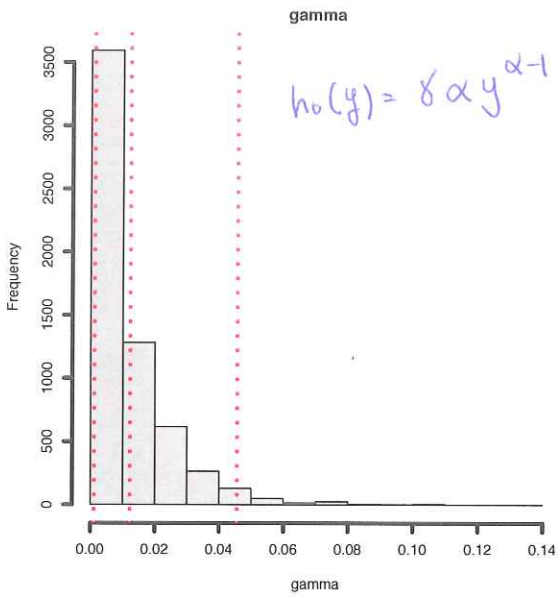
$\beta_2(\hat{\beta}_2 = 0.012\ (0.016))$

- (ICS Example 4.3 – Model II) Weibull Baseline Hazard Model with Multiplicative Gamma Frailties



$\alpha(\hat{\alpha} = 1.342\ (0.164))$

1.278 (0.190)



$h_0(y) = \gamma \alpha y^{\alpha-1}$

$\gamma(\hat{\gamma} = 0.012\ (0.013))$

0.016 (0.015)

$\omega_i \approx 1$

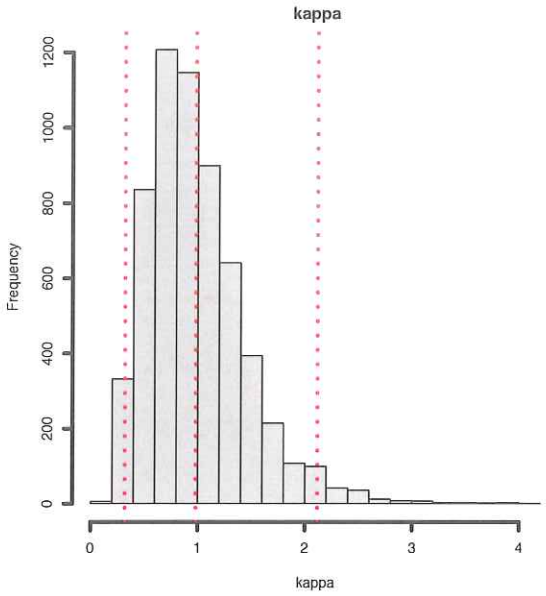
$\omega_i < 1$ h smaller

$\omega_i > 1$ h larger

$h_0(y) \cdot e^{x\beta}\ \omega_i$

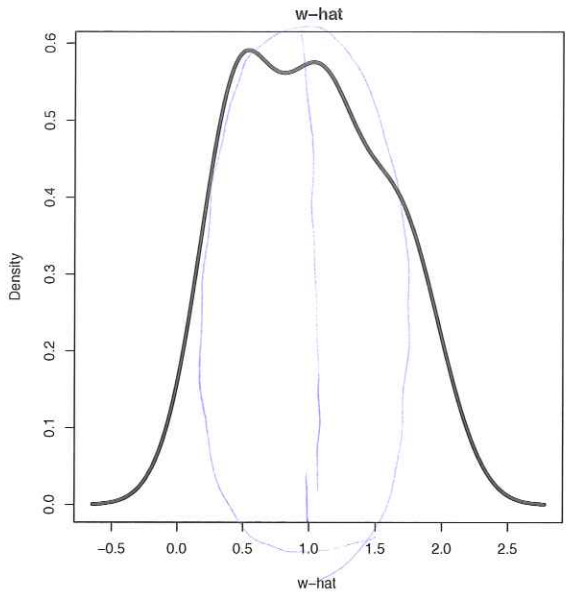
- (ICS Example 4.3 – Model II) Weibull Baseline Hazard Model with Multiplicative Gamma Frailties

$\hat{\omega}_i$



$\kappa(\hat{\kappa} = 0.983\ (0.464))$

0.585 (0.307)



Kernel density estimate of $\hat{w}_i$

- (Additive hazard Weibull model formulation) Reformulate the hazard function

$$h(y_{ij} \mid \mathbf{X}_{ij}, w_i) = \gamma\alpha y_{ij}^{\alpha-1} w_i \exp(\mathbf{X}'_{ij}\beta) = \xi_{ij}\alpha y_{ij}^{\alpha-1},$$

where $\xi_{ij} = \gamma w_i \exp(\mathbf{X}'_{ij}\beta)$. Then let

$$\log(\xi_{ij}) = \zeta + \mathbf{X}'_{ij}\beta + b_i.$$

Weibull baseline hazard with multiplicative log-normal frailty

- ⋆⋆ $b_i \overset{iid}{\sim} \mathsf{N}(0, \kappa^{-1})$ and $\kappa \sim \mathsf{Gamma}(\phi_1, \phi_2)$
- ⋆⋆ $\alpha \sim \mathsf{Gamma}(a_1, a_2)$
- ⋆⋆ $\zeta \sim \mathsf{N}(\bar{\zeta}, \omega^2)$ and $\beta \sim \mathsf{N}_p(\bar{\beta}, \Sigma)$.

- Gamma process prior for $H_0(t)$ (ICS Section 4.1.2)

$$H_0(t) \sim \text{GP}(c_0 H^\star, c_0).$$

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

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

    where $h_j = H_0(s_j) - H_0(s_{j-1})$ and $\alpha_{0,j} = c_0 H^\star(s_j)$.

  - ⋆⋆ Specify priors for w_i, $\eta = \kappa^{-1}$ and β.

- Piecewise exponential model for $h_0(t)$ (ICS Section 4.1.3)

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

  By letting $I_j = (s_{j-1}, s_j]$ for $j = 1, \ldots, J$, we have

$$h_0(t) = \lambda_j, \qquad \text{for} t \in I_j.$$

  ⋆⋆ Before, we had

$$\lambda_j \overset{indep}{\sim} \text{Gamma}(a_j, b_j)$$

  ⋆⋆ Specify priors for w_i, $\eta = \kappa^{-1}$ and β.

- ICS discusses two priors for λ_j to correlate the λ_j's in adjacent intervals.

★★ **Prior 1** Given $(\lambda_1, \ldots, \lambda_{k-1})$, let

$$\lambda_j \mid \lambda_1, \ldots, \lambda_{j-1} \sim \text{Gamma}(\alpha_j, \frac{\alpha_j}{\lambda_{j-1}}), \quad j = 1, \ldots, J,$$

where $\lambda_0 = 1$, and $\text{E}(\lambda_j \mid \lambda_1, \ldots, \lambda_{j-1}) = \lambda_{j-1}$.

- α_j controls the amount of smoothness
- If $\alpha_j \to 0$, λ_j and λ_{j-1} are independent.
- If $\alpha_j \to \infty$, $\lambda_j = \lambda_{j-1}$.

$\text{Var}(\lambda_j \mid \text{---}) = \frac{\alpha_j}{\frac{\alpha_j^2}{\lambda_{j-1}^2}} = \frac{\lambda_{j-1}^2}{\alpha_j}$

★★ **Prior 2** Let $\log(\lambda_j) = \xi_j$ and consider $\xi_j \mid \xi_{j-1} \sim \text{N}(\xi_{j-1}, \tau^2)$ with $\xi_0 = 0$.

Comments on J

- Large J → - make the model more nonparametric
  - fewer observations per interval
  ⇒ unstable estimates of λ
- Small J → possibly not flexible enough
  ⇒ poor fit

• Assume that given all parameters, the responses (y_{1i}, y_{2i}) are conditionally independent

$$\begin{aligned} h_{1i}(y_1 \mid \mathbf{u}_i, \alpha_1, \beta_1) &= \exp(\mathbf{u}_i'\alpha_1 + \mathbf{X}_i'\beta_1 + b_i)h_{01}(y), \\ h_{2i}(y_2 \mid \mathbf{u}_i, \alpha_2, \beta_2) &= \exp(\mathbf{u}_i'\alpha_2 + \mathbf{X}_i'\beta_2 + b_i)h_{02}(y). \end{aligned}$$

★★ $\mathbf{u}_i$: 16-dim vector, $u_{ij} = 1$ if patient i is from site j and $u_{ij'} = 0$ for $j' \neq j$

★★ α_1 and α_2: each a 16-dim vector. $\Rightarrow (\alpha_{1j}, \alpha_{2j})$ are bivariate center-specific effects and a center effect is not the same for both responses.

★★ β_1 and β_2 modulate the relationship between the covariate vector and the response pairs.

★★ b_i are patient-specific log-frailties.

- Further specify models for priors of α_1, α_2, β_1, β_2, b_i, h_{01} and h_{02}.
- More? Please check ICS 4.3

- Required that time of origin and end-point be clearly defined

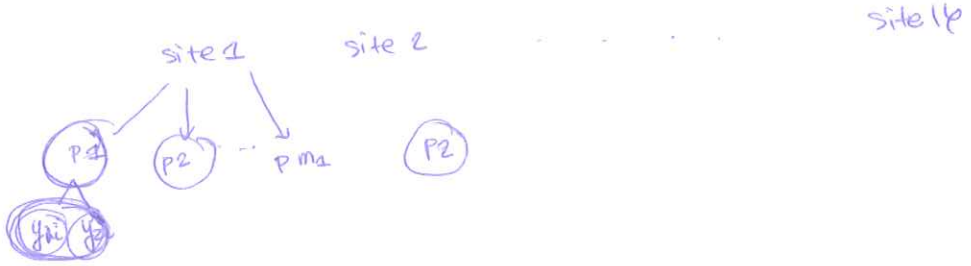
2. Time origin: pre-defined, comparable across subjects

   e.g.: date of randomization in clinical trial, date of enrollment in observational study, date of birth

   Typically not the same calendar date

3. Event:

   - e.g.: death, disease incidence, relapse
   - Usually assume at most one failure per individual.
   - If more than one failure per individual, we need advanced methods.
   - If more than one failure type, we need competing risks methods.



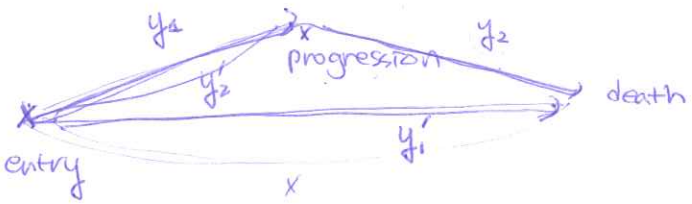
- Frailty models can be used for more complicate settings — e.g. multilevel multivariate data (ICS 4.3)

↻ Colorectal Cancer Data (Gustafson, 1997)

- ⋆⋆ Data from a clinical trial of chemotherapies for advanced cases of colorectal cancer.
- ⋆⋆ The trial was conducted at 16 North American clinical sites.
- ⋆⋆ From each patient, the bivariate failure time pairs are collected, (y_{1i}, y_{2i}).
- ⋆⋆ One of 6 different treatment arms including the standard treatment is given to patients.
- ⋆⋆ 4 other baseline covariates are also recorded $(\mathbf{X}_i)$.

y_1 : time from entry in the trial to disease progression

y_2 : time from disease progression to death



- If more than one failure per individual, we need advanced methods.

$\Rightarrow$ We may introduce frailties to jointly model multivariate survival times from a subject.

- If more than one failure type, we need competing risks methods.

  $\star\star$ Competing risks are said to be present when a patient is **at risk of more than one mutually exclusive event**, such as death from different causes, and the occurrence of one of these will prevent any other event from ever happening – Gichangi & Vach (2005)

  $\star\star$ is a variation of the competing risks problem in which a terminal event censors a non-terminal event, but not vice versa, called **semi-competing risk** - Fine et al.

  e.g. death censors progression, but progression dose not censor death.

AMS 276
Lecture 6: Cure Rate Models

Fall 2016

- **Revisit Example 1.2** Melanoma
  - ⋆⋆ Data from phase III clinical trial by the Eastern Cooperative Oncology Group (ECOG) — **E1684**
  - ⋆⋆ Two-arm clinical trial: high-dose interferon (IFN) vs observation.
  - ⋆⋆ Response: overall survival (time from randomization until death)
  - ⋆⋆ Covariates: age (standardized, X_1), gender (X_2), and performance status (X_3: fully active, other)
  - ⋆⋆ A total of $n = 284$ observations (after deleting missing observations)
  - ⋆⋆ Several years earlier, a similar melanoma study with the same patient population was conducted by ECOG – **E1673**.
  - ⋆⋆ **E1673** serves as the historical data for analysis of **E1684**.

• The Kaplan-Meier survival curve shows a long and stable plateau with heavy censoring at the tail $\Rightarrow$ empirical evidence of a cured fraction.

$S(\infty) = 0$

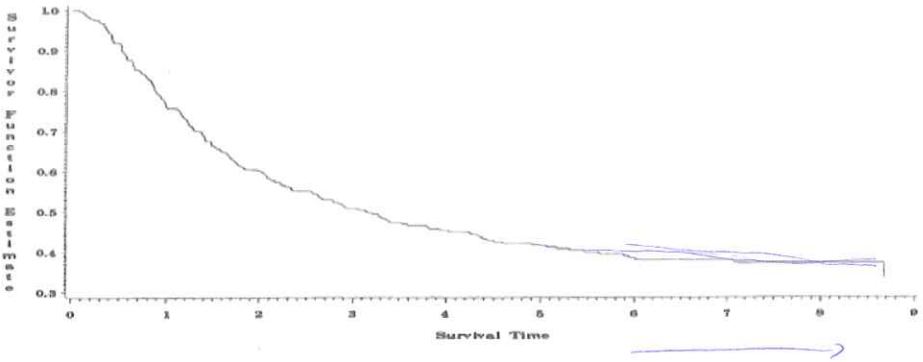


Figure 1. Kaplan–Meier Plot for E1684 Data.

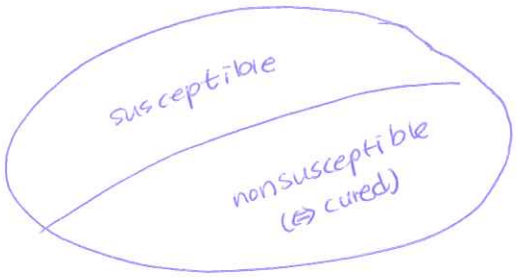
♣ **Cure Rate Models:** Models for right-censored survival data for populations with a surviving (cure) fraction.

- Usually assume that if complete follow-up were possible for all subjects, each would eventually experience the event of interest.

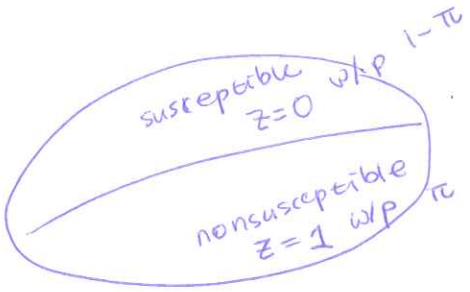
- However, in some situations some of subjects who do not experience the event at the end of the observation period are actually cured.

- Cured? meaning that even after an extended follow-up, no further events are observed.

  e.g. Consider an example of patients with tonsil cancer treated using radiation therapy. Cure occurs if the radiation kills all the cancer cells. There is a time window within which most or all of the recurrences are expected to occur. A patient without any recurrence beyond this window can usually be considered as being cured.



- Not all the subject are susceptible. $\Rightarrow$ The use of standard survival analysis for such data may not be appropriate
- **Objectives:** To study the cure rate (incidence) and survival distribution (latency) and the effect of any covariate.
  - How covariates influence the cure rate — primary interest.
  - How covariates relate to when the event happens.
- Two approaches:
  - Standard cure rate model (ICS chapter 5.1 & paper by Sy and Taylor (2000))
  - Alternative model (ICS chapter 5.2-5.4 & paper by ICS (1999))



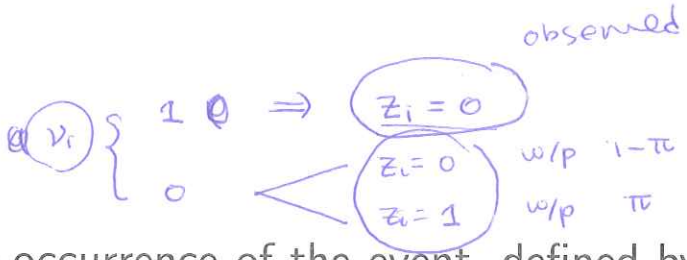
♣ Standard Cure Rate Model:

- In the standard cure model, the population is a **mixture** of susceptible and nonsusceptible (cured) individuals.

  $\Rightarrow$ Possible convenient choice is a mixture model with two components!

- Let Z be the indicator that the individual will eventually ($Z = 0$) or never ($Z = 1$) experience the event, with $\pi = P(Z = 1)$.

  A fraction π of the population is "cured" $\Rightarrow (1 - \pi)$: not cured.



- Let Y denote the time to occurrence of the event, defined by when $Z = 0$, with density $f(y \mid Z = 0)$ and survival function $S^*(y)$.
  - For a censored individual, Z is not observed.
  - For an observed individual, Z is observed to be 0.

- The survival function for the entire population is

$$S_1(y) = \pi + (1 - \pi)S^*(y).$$

cured; not cured

If cured, $S(y) = 1$ for any $y > 0$

$$S_1(y) = P(Y > y \mid Z = 1)\,P(Z = 1) + P(Y > y \mid Z = 0)\,P(Z = 0)$$

$$= 1 \times \pi + S^*(y) \cdot (1 - \pi)$$

- Covariate effects?

  ⋆⋆ Logistic regression model for π

$$\pi_i = \pi(\mathbf{X}_i) = \frac{\exp(\widetilde{\mathbf{X}}_i'\alpha)}{\exp(\widetilde{\mathbf{X}}_i'\alpha) + 1},$$

  where the covariate vector $\widetilde{\mathbf{X}} = (1, \mathbf{X}')'$ includes the intercept.

  ⋆⋆ Cox proportional hazards model for $S^*(y)$

$$h(y \mid \mathbf{X}_i, \beta) = h_0(y)\exp(\mathbf{X}_i'\beta),$$

  where $h_0(y)$ is the conditional baseline hazard function.

  ⋆⋆ Through α and β, the model separates the covariates' effects on the incidence and the latency.

π S^*

$S = e^{-H}$

$H = -\log S$

$h = -\frac{d\log(S)}{dy}$

$S_1(y) = \pi + (1-\pi)\cdot S^*(y)$

$= \pi + (1-\pi)\cdot e^{-H_0(y)e^{X'\beta}}$

$h_1(y) = -\frac{d\log(S_1(y))}{dy} = \frac{-\frac{dS_1(y)}{dy}}{S_1(y)} = \frac{-(-h_0(y)e^{X'\beta})\,e^{-H_0(y)e^{X'\beta}}\cdot(1-\pi)}{\pi + (1-\pi)e^{-H_0(y)e^{X'\beta}}}$

$\Rightarrow$ no longer proportional

$$Z_i \overset{\text{indep}}{\sim} \text{Ber}\left(\pi_i = \frac{e^{\alpha' \tilde{x}_i}}{e^{\alpha' \tilde{x}_i} + 1}\right)$$

- The likelihood?

$$\mathcal{L}(\alpha, \beta \mid \text{data}) = \prod_{i=1}^{n} \pi_i^{z_i} (1 - \pi_i)^{1 - z_i}$$

$$\prod_{i:\, z_i = 0} \left\{ h_0(y_i) e^{x_i' \beta} \right\}^{\nu_i} \cdot e^{-H_0(y_i) e^{x_i' \beta}}$$

$$f(y) = h(y) S(y)$$

$$S(y)$$

- Use the E-M algorithm to find the MLE (for details, read Sy and Taylor (2000))

- Bayesian Standard Cure Rate Model

  ⋆⋆ Develop priors for α, β and $h_0(\cdot)$.

  ⋆⋆ ICS Chapter 5.1: This approach is attractive but has some drawbacks.

  If the covariates are modeled through π via a binomial regression model,

  ↻ It cannot have a proportional hazards structure.

  ↻ The standard cure rate model yields improper posterior distributions for many types of noninformative improper priors, including the uniform prior for the regression coefficient.

  e.g. (Th 5.2.2) If an improper uniform prior for α, the posterior is improper.

  we assume