Frailty Model for Parametric Analysis of Bivariate Right Censored Data with Competing Risks

Ankit Sinha ,
Supervisor : Dr.Sudipta Das

July 12, 2024

Introduction

- **Definition:**
  - **Bivariate right censored data:** This is basically a bivariate right censored data along with censoring indicators.
  - **Competing risks:** Different potential events that prevent, or affect the occurrence of the primary event of interest.
- **Objective:** To analyze bivariate right censored data using shared frailty models.
- **Key Concept:** Frailty captures unobserved heterogeneity within a population.

Bivariate Failure Time Data with Competing Risks (1)

- **Observed Data:** (X_1, X_2, J_1, J_2)
  - X_1 and X_2 denote the two Censoring times.
  - $J_k = 0$ if observation is censored for $k = 1, 2$.
- **Sub-distribution Function:**

$$F_j^{(k)}(t_k; \theta) = P(T_k \leq t_k, J_k = j; \theta) = \int_0^{t_k} f_j^{(k)}(u; \theta)\, du$$

  - $f_j^{(k)}(u; \theta)$: sub-density function for the k-th individual.

Bivariate Failure Time Data with Competing Risks (2)

- **Cumulative Distribution Function:**

$$F^{(k)}(t_k;\theta) = P(T_k \leq t_k;\theta) = \sum_{j=1}^{L_k} F_j^{(k)}(t_k;\theta)$$

- **Survival Function:**

$$S^{(k)}(t_k;\theta) = P(T_k > t_k;\theta) = 1 - F^{(k)}(t_k;\theta)$$

Bivariate Failure Time Data with Competing Risks (3)

- **Joint Sub-distribution Function:**

$$F_{j_1 j_2}(t_1, t_2; \theta) = P(T_1 \leq t_1, T_2 \leq t_2, J_1 = j_1, J_2 = j_2; \theta)$$

- **Cause-specific Hazard Function:**

$$\lambda_j^{(k)}(t_k; \theta) = \lim_{h \to 0^+} \frac{P(t_k < T_k \leq t_k + h, J_k = j; \theta)}{h} = \frac{f_j^{(k)}(t_k; \theta)}{S_{-j}^{(k)}(t_k; \theta)}$$

- **Frailty Variables:**

$$\epsilon^{(k)} = (\epsilon_1^{(k)}, \ldots, \epsilon_{L_k}^{(k)})$$

  - $\theta = (\lambda, \phi)$
- **Multiplicative Model for Hazard Function:**

$$\lambda_j^{(k)}(t_k; \lambda \mid \epsilon^{(k)}) = \lambda_{0j}^{(k)}(t_k; \lambda)\epsilon^{(k)}$$

  - $\lambda_{0j}^{(k)}(t_k; \lambda)$: cause-specific baseline hazard function.
- **Cumulative Cause-specific Baseline Hazard Function:**

$$H_{0j}^{(k)}(t_k; \lambda) = \int_0^{t_k} \lambda_{0j}^{(k)}(u; \lambda)\, du$$

Shared Gamma Frailty Model

- **Frailty Variable:** Assumes a Gamma distribution with parameters $\frac{1}{\alpha^2}$ and $\frac{1}{\alpha^2}$.
- **Density Function:**

$$g(\nu;\alpha) = \frac{1}{\Gamma\left(\frac{1}{\alpha^2}\right)\alpha^{\frac{2}{\alpha^2}}}\nu^{\frac{1}{\alpha^2}-1}e^{-\frac{\nu}{\alpha^2}},$$

  where $\nu > 0$ and $\alpha > 0$
- **Applications:** E.g., modeling cancer occurrence times in a father and son due to shared genetic/environmental factors.

Maximum Likelihood Estimation (1)

- **Observed Data:** (X_1, X_2, J_1, J_2)
  - $J_k = 0$ indicates right-censoring for the k-th individual $(k = 1, 2)$.
  - $x_1 > 0$ and $x_2 > 0$ are the observed times.
  - j_1 and j_2 are the causes of failure $(j_k = 0, 1, \ldots, L_k)$.

Maximum Likelihood Estimation (2)

- **Likelihood Contributions:**
  1. **Case 1:** $j_1 = 1, \ldots, L_1$ **and** $j_2 = 1, \ldots, L_2$

$$L(x_1, x_2, j_1, j_2, \nu, \alpha) \propto f_{j_1 j_2}(x_1, x_2, \nu, \alpha)$$

  2. **Case 2:** $j_1 = 0$ **and** $j_2 = 1, \ldots, L_2$

$$L(x_1, x_2, j_1, j_2, \nu, \alpha) \propto f_{j_2}^{(2)}(x_2, \nu, \alpha) - \int_0^{x_1} \sum_{j=1}^{L_1} f_{jj_2}(u_1, u_2, \nu, \alpha) du_1$$

- **Likelihood Contributions:**
  1. **Case 3:** $j_1 = 1, \ldots, L_1$ **and** $j_2 = 0$

$$L(x_1, x_2, j_1, j_2, \nu, \alpha) \propto f_{j_1}^{(1)}(x_1, \nu, \alpha) - \int_0^{x_2} \sum_{j=1}^{L_2} f_{j_1 j}(u_1, u_2, \nu, \alpha)\, du_2$$

  2. **Case 4:** $j_1 = 0$ **and** $j_2 = 0$

$$L(x_1, x_2, j_1, j_2, \nu, \alpha) \propto S(x_1, x_2, \nu, \alpha)$$

Simulation (1)

- **Objective:** Generate 100 samples of bivariate current status data.
- **Weibull Baseline Hazard Function:**

$$h(t) = \alpha_1 \beta_1 t^{\alpha_1 - 1} e^{-(\beta_1 t)^{\alpha_1}}$$

  - Shape parameter: α_1
  - Scale parameter: β_1^{-1}

- **Algorithm Steps:**
  1. Generate a frailty variable ϵ from a Gamma distribution with parameter σ^{-2}.
  2. Generate two random variables T_1 and T_2 from Weibull distributions:

$$T_1 \sim \text{Weibull}\left(\alpha_1, \frac{1}{\beta_1 \epsilon}\right), \quad T_2 \sim \text{Weibull}\left(\alpha_2, \frac{1}{\beta_2 \epsilon}\right)$$

    - α_1 and α_2: shape parameters
    - $\frac{1}{\beta_1 \epsilon}$ and $\frac{1}{\beta_2 \epsilon}$: scale parameters

Simulation (3)

- **Algorithm Steps (continued):**
  3. Generate two censoring times X_1 and X_2 from exponential distributions with parameter θ.
  4. Determine censoring occurrence:

$$j_1 = \begin{cases} 1, & \text{if } T_1 \leq X_1 \\ 0, & \text{otherwise} \end{cases} \quad j_2 = \begin{cases} 1, & \text{if } T_2 \leq X_2 \\ 0, & \text{otherwise} \end{cases}$$

- Repeat the process until 100 samples are obtained.

Results

Simulation Parameters:

- $\alpha_1 = 0.45$, $\alpha_2 = 2.57$, $\beta_1 = \beta_2 = 5$, $\sigma = 0.5$

Parameter Estimates:

Parameter	Estimate	Error in Estimation (Absolute Value)
α_1	1.2658	0.8158
α_2	4.2358	1.6658
β_1	4.5256	0.4744
β_2	8.2145	3.2145
σ	0.145	0.355

Conclusion

- **Key Findings:**
  - Shared frailty models effectively capture unobserved heterogeneity.
  - The shared Gamma frailty model is robust for understanding dependencies in bivariate failure time data with competing risks.
- **Future Research:**
  - Explore non-Gamma frailty distributions.
  - Incorporate more complex dependency structures.

References

1. L. Duchateau and P. Janssen. The Frailty Model. New York: Springer Verlag, 2008.
2. Biswadeep Ghosh, Anup Dewanji, and Sudipta Das. Technical Report No. ASU/2024/01, Applied Statistics Unit, Indian Statistical Institute, Kolkata, 2024.
3. A. Wienke. Frailty models in survival analysis. Chapman and Hall/CRC, 1st edition, 2010.

Questions & Answers

Thank you for your attention!
Please feel free to ask any questions.