Plan for the lecture

- Brief summary of the Nelson-Aalen estimator
- Do two groups have different hazard rates?
  - formulated as a hypothesis test
- Using a transformation to find an alternative confidence interval

Multiplicative intensity model

- ⋆ Multiplicative intensity model

$$\lambda(t) = \alpha(t)Y(t)$$

  - $Y(t)$: predictable process

- ⋆ Frequent situation leading to a multiplicative intensity model
  - n individuals
  - each individual has the same hazard rate $\alpha(t)$
  - may have truncation and/or censoring
  - $Y(t)$: number of individuals at risk just before time t

Nelson–Aalen estimator

- $N(t)$: counting process with $\lambda(t) = \alpha(t)Y(t)$
- Notation:
  - $J(t) = \mathbb{I}(Y(t) > 0)$
  - $A(t) = \int_0^t \alpha(t)dt$
  - $A^\star(t) = \int_0^t J(t)\alpha(t)dt$
- Nelson–Aalen estimator for $A(t)$:

$$\widehat{A}(t) = \int_0^t \frac{J(s)}{Y(s)} dN(s) = \sum_{j: T_j \leq t} \frac{1}{Y(T_j)}$$

- Recall: $E[\widehat{A}(t) - A^\star(t)] = 0$

Nelson–Aalen estimator

- $\star$ $N(t)$: counting process with $\lambda(t) = \alpha(t)Y(t)$
- $\star$ Notation:
  - $J(t) = \mathbb{I}(Y(t) > 0)$
  - $A(t) = \int_0^t \alpha(t)dt$
  - $A^\star(t) = \int_0^t J(t)\alpha(t)dt$
- $\star$ Nelson–Aalen estimator for $A(t)$:

$$\widehat{A}(t) = \int_0^t \frac{J(s)}{Y(s)} dN(s) = \sum_{j:T_j \le t} \frac{1}{Y(T_j)}$$

- $\star$ Recall: $E[\widehat{A}(t) - A^\star(t)] = 0$
- $\star$ Using the optional variation process we found

$$\mathrm{Var}[\widehat{A}(t) - A^\star(t)] = E\left[\int_0^t \frac{J(s)}{Y^2(s)} dN(s)\right]$$

Nelson–Aalen estimator

- $\star$ $N(t)$: counting process with $\lambda(t) = \alpha(t)Y(t)$
- $\star$ Notation:
  - $J(t) = \mathbb{I}(Y(t) > 0)$
  - $A(t) = \int_0^t \alpha(t)dt$
  - $A^\star(t) = \int_0^t J(t)\alpha(t)dt$
- $\star$ Nelson–Aalen estimator for $A(t)$:

$$\widehat{A}(t) = \int_0^t \frac{J(s)}{Y(s)} dN(s) = \sum_{j:T_j \leq t} \frac{1}{Y(T_j)}$$

- $\star$ Recall: $E[\widehat{A}(t) - A^\star(t)] = 0$
- $\star$ Using the optional variation process we found

$$\mathrm{Var}[\widehat{A}(t) - A^\star(t)] = E\left[\int_0^t \frac{J(s)}{Y^2(s)} dN(s)\right]$$

- $\star$ So an unbiased estimator for $\mathrm{Var}[\widehat{A}(t) - A^\star(t)]$ is

$$\widehat{\sigma}^2(t) = \int_0^t \frac{J(s)}{Y^2(s)} dN(s) = \sum_{j:T_j \leq t} \frac{1}{Y^2(T_j^2)}$$

Nelson–Aalen with tied observations

- Recall
$$\widehat{A}(t) = \int_0^t \frac{J(s)}{Y(s)} dN(s), \widehat{\sigma}^2(t) = \int_0^t \frac{J(s)}{Y(s)^2} dN(s)$$
- Why tied observations:
  - (i) events happens in continuous time, but we observe ties because of rounding
  - (ii) events happens in discrete time

- Denote event times by $T_1 < T_2 < \ldots$ and multiplicities $d_1, d_2, \ldots$
- Assuming (i) we get
$$\widehat{A}(t) = \sum_{j: T_j \leq t} \left[\sum_{\ell=0}^{d_j-1} \frac{1}{Y(T_j) - \ell} \right], \widehat{\sigma}^2(t) = \sum_{j: T_j \leq t} \left[\sum_{\ell=0}^{d_j-1} \frac{1}{(Y(T_j) - \ell)^2} \right]$$
- Assuming (ii) we get (we haven't discussed the reason for $\widehat{\sigma}^2(t)$)
$$\widehat{A}(t) = \sum_{j: T_j \leq t} \frac{d_j}{Y(T_j)}, \widehat{\sigma}^2(t) = \sum_{j: T_j \leq t} \frac{(Y(T_j) - d_j) d_j}{Y(T_j)^3}$$

Large sample properties for $\widehat{A}(t)$

- Assume:
  - n individuals
  - each individual has the same hazard rate $\alpha(t)$
  - may have truncation and/or censoring
  - $Y(t)$ is number of individuals at risk just before time t
- Multiplicative intensity process: $\lambda(t) = \alpha(t)Y(t)$
- Assume also

$$\frac{Y(t)}{n} \to y(t) > 0 \text{ when } n \to \infty$$

- Then Rebolledo's theorem gives that

$$\sqrt{n}(\widehat{A}(t) - A^\star(t))$$

  converges to a Gaussian martingale $U(t)$ with

$$\langle U \rangle(t) = \int_0^t \frac{\alpha(s)}{y(s)} ds$$

- Thus, for large n: $\widehat{A}(t) \approx N(A(t), \sigma^2(t))$

Stochastic integral

- Stochastic integral

$$\mathbb{I}(t) = \int_0^t H(s)dM(s) = \lim_{n\to\infty} \sum_{k=1}^{n} H_k(M_k - M_{k-1})$$

- Consequences of the definitions
  - $\mathbb{I}(t)$ is a mean zero martingale
  - $\langle \int HdM \rangle = \int H^2 d\langle M \rangle$
  - $[\int HdM] = \int H^2 d[M]$
  - $\langle \int H_1 dM_1, \int H_2 dM_2 \rangle = \int H_1 H_2 d\langle M_1, M_2 \rangle$
  - $[\int H_1 dM_1, \int H_2 dM_2] = \int H_1 H_2 d[M_1, M_2]$

Counting processes

- ⋆ $N(t)$: A counting process, $N(t)$ adapted to $\{\mathcal{F}_t\}$
  - right continuous
  - jumps of size one
  - constant between jumps
- ⋆ Recall informal definition of intensity process

$$\lambda(t)dt = P(dN(t) = 1|\mathcal{F}_{t-}) = \mathrm{E}[dN(t)|\mathcal{F}_{t-}]$$

- ⋆ Precise definition of intensity process by Doob-Meyer

$$N(t) = \Lambda(t) + M(t)$$

- ⋆ When $\Lambda(t)$ is absolutely continuous

$$\Lambda(t) = \int_0^t \lambda(s)ds$$

- ⋆ Variation processes

$$\langle M\rangle(t) = \Lambda(t) = \int_0^t \lambda(s)ds \quad \text{and} \quad [M](t) = N(t)$$