What to do today (2022/03/15)?

Part IV. Advanced Topics

- **Part IV.1 Counting Process Formulation**
  - **Part IV.1.1 Theoretical Preparation**
    - *IV.1.1A Basic concepts*
    - *IV.1.1B An introduction to stochastic process*
    - **IV.1.1C Counting process and martingale: the key results**
  - **Part IV.1.2 Counting Process Formulation in LIDA and Applications: Revisits to KM, logrank, Cox PH**
    - **IV.1.2A Formulation**
    - **IV.1.2B Revisit to Logrank Test**
    - *IV.1.2C Revisit to Cox's Partial Likelihood Approach*
    - *IV.1.2D Revisit to Kaplan-Meier Estimator*
    - *IV.1.2E Others*
- *Part IV.2 Selected Recent Topics in LIDA*
- *Part IV.3 Beyond Lifetime Data Analysis*

Part IV.1.1C Martingale: An Introduction

Martingale Definition
Consider a stochastic process $X(\cdot) = \{X(\omega; t) : \omega \in \Omega, t \geq 0\}$ on $(\Omega, \mathcal{F}, \mathcal{P})$, adapted to a filtration $\{\mathcal{F}_t : t \geq 0\}$.

Suppose $X(\cdot)$ is right-continuous left-hand limits. It is a martingale wrt the filtration $\{\mathcal{F}_t\}$ if

- (a) $\forall t$, $E|X(t)| < \infty$ (integrable)
- (b) $\forall t$, $E\big[X(t+s)|\mathcal{F}_t\big] = X(t)$ a.s. for $s \geq 0$.

- *"fair game" in gambling*: $E(X(t)) = E(X(0))$
- *Sub (Super) Martingale:* "$=$" in (b) of a martingale definition is replaced by $\geq$ ($\leq$)

Part IV.1.1C Martingale: An Introduction

- *Predictable.* stochastic process $X(\cdot)$ is predictable wrt $\{\mathcal{H}_t\}$ if $X(t)$ is determined by $\mathcal{H}_{t-} = \sigma(X(u) : 0 \leq u < t)$.

  Processes with left-continuous sample path are predictable.

- *Stopping Time.* r.v. τ is a stopping time wrt a filtration $\{\mathcal{F}_t\}$ if $\{\tau \leq t\} \in \mathcal{F}_t, \forall t$.

- *Stopped Process.* Suppose $X(\cdot)$ is a stochastic process adapted to $\{\mathcal{F}_t\}$, and τ a stopping time wrt $\{\mathcal{F}_t\}$. We call $X^\tau(\cdot)$ a stopped process by τ: $X^\tau(t) = X(t \wedge \tau)$.

Part IV.1.1C Martingale: An Introduction

- *Notion of "Localization" in Theory*.
  - A sequence of stopping times, non-decreasing $\{\tau_n\}$ is called "localizing sequence" if $P(\tau_n \geq t) \to 1$ as $n \to \infty$ for $t \in \mathcal{T}$.
  - We say $X(\cdot)$ has a certain property locally if there is a localizing sequence $\{\tau_n\}$ such that, $\forall n$, $I(\tau_n > 0)X^{\tau_n}(\cdot)$ has the property.

e.g. $X(\cdot)$ is locally bounded.

e.g. $X(\cdot)$ is a local martingale.

Part IV.1.1C Martingale: An Introduction

Important Martingale Results

- Provided $\{M(t) : t \in \mathcal{T}\}$ is a martingale wrt $\{\mathcal{F}_t\}$ and $M(0) = 0$,
  - $E(M(t)) = 0, \forall t \in \mathcal{T}$
  - $Cov(M(t+s) - M(t), M(t)) = 0$ and $Cov(M(t), M(t+s)) = Var(M(t))$, $\forall t \in \mathcal{T}, s > 0$

- Provided $\{M_1(t) : t \in \mathcal{T}\}$ and $\{M_2(t) : t \in \mathcal{T}\}$ are martingales wrt $\{\mathcal{F}_t\}$, $M(\cdot) = aM_1(\cdot) + bM_2(\cdot)$ is a martingale.

- Provided $\{M(t) : t \in \mathcal{T}\}$ is a martingale wrt $\{\mathcal{F}_t\}$, $M^2(\cdot)$ is a submartingale wrt $\{\mathcal{F}_t\}$.

Part IV.1.1C Martingale: An Introduction

Doob-Meyer Decomposition. If $X(\cdot)$ is a submartingale adapted to $\{\mathcal{F}_t : t \geq 0\}$, there exists a unique, right-continuous, non-decreasing, predictable process $A(\cdot)$ with $A(0) = 0$ and $M(\cdot) = X(\cdot) - A(\cdot)$ is a martingale wrt $\{\mathcal{F}_t : t \geq 0\}$. The process $A(\cdot)$ is called the *compensator* for $X(\cdot)$.

Example. $N(\cdot)$ is a Poisson process with intensity function ρ. (i) $N(\cdot)$ is a submartingale. (ii) Its compensator $A(\cdot)$ is $A(t) = \int_0^t \rho(u)du$, $t \geq 0$.

This result holds for any counting process $N(\cdot)$.

Part IV.1.1C Martingale: An Introduction

- If $M(.)$ is martingale and (locally) square integrable, M^2 is a (local) sub-martingale and has a compensator, denoted by $< M, M > (.)$ such that $M^2(.)- < M, M > (.)$ is a martingale.

  The process $< M, M > (.)$ is called the predictable quadratic variation of $M(.)$.

  - If $E[M(t)] = 0$, then $var[M(t)] = E[M^2(t)] = E[< M, M > (t)]$.
  - If the zero mean martingale $M = N - A$ satisfies $E[M^2(t)] < \infty$ and $A(.)$ is continuous, then $< M, M > (.) = A(.)$, a.s.

Part IV.1.1C Martingale: An Introduction

- $M_1, M_2, \ldots$ are zero mean martingales defined on the same filtration.
  - If both M_1 and M_2 are (locally) square integrable, then there is a right continuous predictable process $< M_1, M_2 >$ such that $< M_1, M_2 > (0) = 0$, $E[\big| < M_1, M_2 > \big|] < \infty$, and $M_1(.)M_2(.) - < M_1, M_2 > (.)$ is a martingale.
  - If $< M_1, M_2 >= 0$, a.s., M_1 and M_2 are called orthogonal.
  - If $M_j = N_j - A_j$, where N_j is a counting process and A_j its (continuous) compensator, then if N_i and N_j don't jump at the same ime, $< M_i, M_j > (.) = 0$, a.s.

Suppose that $N(.)$ is a counting process with $E[N(t)] < \infty$, that $A(.)$ is the compensator for $N(.)$ and $H(.)$ a bunded, predictable process. Define $Q(.)$ by

$$Q(t) = \int_0^t H(s)dM(s), \text{ where } M(.) = N(.) - A(.).$$

- $Q(.)$ is a zero-mean martingale.
- $var[Q(t)] = E[Q^2(t)]$
- $Q^2(.)$ is a sub-martingale, and thus there is $< Q, Q > (.)$ such that $Q^2(.)- < Q, Q > (.)$ is a martingale. Hence, $var[Q(t)] = E[< Q, Q > (t)]$
- If $A(.)$ is continuous, $< Q, Q > (t) = \int_0^t H^2(s)dA(s)$, a.s.
- If M_1, M_2 are martingales, H_1, H_2 bounded predictable processes, and Q_1, Q_2 are defined in the same way as $Q(.)$, the predictable quadratic covariance process $< Q_1, Q_2 >$ satisfies $< Q_1, Q_2 > (t) = \int_0^t H_1(s)H_2(s)d < M_1, M_2 > (s)$.

Local Properties:

- M is a local martingale (sub-martingale) if $\exists$ a localizing sequence $\{\tau_n\}$ such that for each n, $M_n(.) = \{M(t) : 0 \leq t \leq \tau_n\}$ is a martingale (sub-martingale).

- If as above but we also have $\sup_t E[M_n^2(t)] < \infty$ for all n, M is called a locally square-integrable martingale.

- **Extended Doob-Meyer Decomposition**: If $X(.)$ is a local sub-martingale, $\exists$ a right continuous, nondecreasing predictable process $A(.)$ such that $X(.) - A(.)$ is a local martingale.

By the extended Doob-Meyer Decomposition,

- If $N(.)$ is a counting process (not necessaryly $E[N(t)] < \infty$), then $N(.)$ is a local sub-martingale. Thus, $\exists$ a unique $A(.)$ (nondecreasing, right-continuous, predictable) such that $M(.) = N(.) - A(.)$ is a local martingale.

- If $A(.)$ is locally bounded, then $M(.)$ above is a local square integrable martingale.

- If $H(.)$ is locally-bounded and predictable, and $M(.)$ is a local martingale, then $Q = \int HdM$ is a locally square integrable martingale.

Part IV.1.1C Martingale: An Introduction

Martingale Central Limit Theorem (CLT)
For $i = 1, \ldots, n$, let $N_{in}(.)$ be counting process, $A_{in}(.)$ its continuous compensator, $H_{in}(.)$ locally bounded predictable process. Define

$$M_{in} = N_{in} - A_{in}, \quad U_{in} = \int H_{in} dM_{in}, \quad U_{in,\epsilon} = \int H^*_{in,\epsilon} dM_{in},$$

$$H^*_{in,\epsilon} = H_{in} \mathbf{1}[|H_{in}| \geq \epsilon], \quad \bar{U}_n = \sum_i U_{in}, \quad , \bar{U}_{n,\epsilon} = \sum_i U_{in,\epsilon}$$

Suppose that

(a) $< \bar{U}_n, \bar{U}_n > (t) \to \alpha(t)$ in prob, $\forall t \geq 0$ and some function $\alpha(.)$.

(b) $< \bar{U}_{n,\epsilon}, \bar{U}_{n,\epsilon} > (t) \to 0$ in prob, $\forall t \geq 0$ and $\forall \epsilon \geq 0$.

Then $\bar{U}_n \to U = \int f dW$ weakly, where $W(.)$ is a Wiener process and $\int_0^t f^2(s) ds = \alpha(t)$.

Part IV.1.1C Martingale: An Introduction

That is, the limiting process $U(.)$ is a zero-mean Gaussian process with independent increments and $var[U(t)] = \alpha(t)$.
By the definitions of N_{in}, A_{in}, and H_{in},

$$< \bar{U}_n, \bar{U}_n > (t) = \sum_{i=1}^{n} \int_0^t H_{in}^2(s) dA_{in}(s),$$

$$< \bar{U}_{n,\epsilon}, \bar{U}_{n,\epsilon} > (t) = \sum_{i=1}^{n} \int_0^t H_{in,\epsilon}^{*2}(s) dA_{in}(s).$$

Part IV.1.2 Counting Process Formulation in LIDA and Applications: Revisits to KM, Logrank, Cox PH

Part IV.1.2A Formulation

- Consider T, a lifetime with hazard function $h(\cdot)$.
- Let C be a censoring time, $U = T \wedge C$ and $\delta = 1, 0$ for $U = T, C$, respectively.
- Define $N(t) = I(U \leq t, \delta = 1)$ and $Y(t) = I(U \geq t)$.

If $T \perp\!\!\!\perp C$ and $T \sim f(\cdot)$, $C \sim g(\cdot)$,

$\Longrightarrow E\big[N(t)\big] = E\big[A(t)\big]$ with $A(t) = \int_0^t h(x)Y(x)dx$.

$$E\big[N(t)\big] = P(T \leq t, T \leq C) = \int_0^t \int_{y \geq x} f(x)g(y)dydx$$

$$E\big[A(t)\big] = \int_0^t h(x)E\big[Y(x)\big]dx = \int_0^t h(x)P(U \geq x)dx$$

Part IV.1.2A Formulation

- Consider T, a lifetime with hazard function $h(\cdot)$.
- Let C be a censoring time, $U = T \wedge C$ and $\delta = 1, 0$ for $U = T, C$, respectively.
- Define $N(t) = I(U \leq t, \delta = 1)$ and $Y(t) = I(U \geq t)$.

If $T \perp\!\!\!\perp C$ and $T \sim f(\cdot)$, $C \sim g(\cdot)$,

$\Longrightarrow$
$A(\cdot)$ is the continuous compensator of $N(\cdot)$ wrt $\{\mathcal{H}_t\}$, and denote $M(\cdot) = N(\cdot) - A(\cdot)$.

$$E\Big[N(u + \Delta u) - N(u)\Big| Y(u)\Big] \approx Y(u)h(u)\Delta u$$

Part IV.1.2A Formulation

Recall *right-censored* lifetimes $\{(U_i, \delta_i) : i = 1, \ldots, n\}$ from n indpt individuals.

Define $N_i(t) = I(U_i \leq t, \delta_i = 1)$ and $Y_i(t) = I(U_i \geq t)$.

Let $N_{\cdot}(t) = \sum_{i=1}^{n} N_i(t)$, number of failures observed till t, and $Y_{\cdot}(t) = \sum_{i=1}^{n} Y_i(t)$, number of individuals at risk at time t.

Denote $dN_i(t) = \begin{cases} 1 \times dt & \textit{when } t = U_i,\ \delta_i = 1 \\ 0 & \textit{otherwise} \end{cases}$

Then $N_i(t) = \int_0^t dN_i(v)$ for $t \geq 0$.

Part IV.1.2A Formulation

Let $A_i(t) = \int_0^t Y_i(x)h(x)dx$ and $M_i(t) = N_i(t) - A_i(t)$.

Thus $N_\cdot(t) = \int_0^t Y_\cdot(v)h(v)dv + M_\cdot(t)$.

Part IV.1.2B Revisit to Logrank Test

Logrank Test.

- Consider T_1, T_0 event times in the treatment, control group, with hazard functions $h_1(\cdot), h_0(\cdot)$. Assume $T_1 \perp\!\!\!\perp T_0$ and indepdent censoring time C.
- $H_0 : h_1(\cdot) = h_0(\cdot)$ vs H_1 : *otherwise*.
- Data are
  $\{(U_{1i}, \delta_{1i}) : i = 1, \ldots, n_1\} \bigcup \{U_{0j}, \delta_{0j}) : j = 1, \ldots, n_0\}$
- Test statistic

$$Z = \frac{O - E}{\sqrt{V}} \to N(0, 1) \text{ in distn}$$

under H_0 as $n \to \infty$ with $n = n_1 + n_0$ and $n_1/n \to p \in (0, 1)$.

to verify the asymptotic normality?

Part IV.1.2B Revisit to Logrank Test

Introducing $N_{1i}(t) = I(U_{1i} \leq t, \delta_{1i} = 1)$ and $Y_{1i}(t) = I(U_{1i} \geq t)$, $N_{0j}(t) = I(U_{0j} \leq t, \delta_{0j} = 1)$ and $Y_{0j}(t) = I(U_{0j} \geq t)$, and $N_{1.}(t), Y_{1.}(t), N_{0.}(t), Y_{0.}(t)$.

With $0 < V_1 < \ldots < V_K < \infty$ all the distinct observed failure times,

$$O = N_{1.}(\infty) = \sum_{i=1}^{n_1} \int_0^\infty dN_{1i}(t) = \sum_k d_{1k}$$

$E = \sum_{k=1}^K E_k$ with

$$E_k = d_k \frac{N_{k1}}{N_k} = \int_{V_{k-1}}^{V_k} \frac{Y_{1.}(v)}{Y_{1.}(v) + Y_{0.}(v)} d[N_{1.}(v) + N_{0.}(v)]$$

Part IV.1.2B Revisit to Logrank Test

$$
\begin{aligned}
O - E &= \int_0^\infty dN_{1.}(t) - \int_0^\infty \frac{Y_{1.}(t)}{Y_{1.}(t) + Y_{0.}(t)} d[N_{1.}(t) + N_{0.}(t)] \\
&= \int_0^\infty [1 - H_n(t)] dN_{1.}(t) - \int_0^\infty H_n(t) dN_{0.}(t)
\end{aligned}
$$

Further, introducing $M_{li}(t) = N_{li}(t) - A_{li}(t)$ and
$A_{li} = \int_0^t h_l(v) Y_{li}(v) dv$ for $i = 1, \ldots, n_l$, $l = 1, 0$.
Under H_0, $h_1(\cdot) = h_0(\cdot)$, and

$$
\begin{aligned}
O - E &= \int_0^\infty [1 - H_n(t)] dM_{1.}(t) - \int_0^\infty H_n(t) dM_{0.}(t) \\
&= \sum_i \int_0^\infty [1 - H_n(t)] dM_{1i}(t) - \sum_j \int_0^\infty H_n(t) dM_{0j}(t),
\end{aligned}
$$

denoted by $U(\infty)$.

Part IV.1.2B Revisit to Logrank Test

(a) Let $U_n(t) = \frac{1}{\sqrt{n}} U(t)$:

$$
\begin{aligned}
< U_n, U_n > (t) &= \sum_i \int_0^t \frac{1}{n}[1 - H_n(v)]^2 dA_{1i}(v) + \sum_j \int_0^t \frac{1}{n} H_n(v)^2 dA_{0j}(v) \\
&= \int_0^t \frac{1}{n}[1 - H_n(v)]^2 h_1(v) Y_{1.}(v) dv + \int_0^t \frac{1}{n} H_n(v)^2 h_0(v) Y_{0.}(v) dv
\end{aligned}
$$

converges to $\alpha(t)$ in prob, because

(i) $Y_{l.}(v)/n_l \to P(U_l \geq v)$;

(ii) $H_n(v) \to pG_1(v)/[pG_1(u) + (1-p)G_0(u)]$.

Part IV.1.2B Revisit to Logrank Test

(b) Let $U_{n,\epsilon}(t)$ be

$$\sum_i \int_0^t \frac{1}{\sqrt{n}}[1-H_n(t)]I(|H_n^{**}(v)| \geq \epsilon)dM_{1i}(t) - \sum_j \int_0^t \frac{1}{\sqrt{n}}H_n(t)I(|H_n^{*}(v)| \geq \epsilon)dM_{0j}(t)$$

$< U_{n,\epsilon}, U_{n,\epsilon} > (t)$ is

$$\begin{aligned} & \sum_i \int_0^t \frac{1}{n}[1-H_n(v)]^2 I(|H_n^{**}(v)| \geq \epsilon) Y_{1i}(v) h_1(v) dv \\ + \quad & \sum_j \int_0^t \frac{1}{n} H_n(v)^2 I(|H_n^{*}(v)| \geq \epsilon) Y_{0j}(v) h_0(v) dv \end{aligned}$$

converges to 0 in prob, because $I(|H_n^*(v)| \geq \epsilon)$ and

$I(|H_n^{**}(v)| \geq \epsilon)$ converge to 0 in prob.

Part IV.1.2B Revisit to Logrank Test

By the Martingale CLT,

$$U_n(t) \to Gaussian\big(0, \alpha(t)\big)$$

in distn as $n \to \infty$. Thus

$$\frac{1}{\sqrt{n\alpha(\infty)}} U(\infty) \to N(0, 1).$$

in distn as $n \to \infty$.

Part IV.1.2B Revisit to Logrank Test

Recall that $O - E = U(\infty)$ is

$$
\begin{aligned}
& \int_0^\infty \frac{Y_{0.}(v)}{Y_{1.}(v) + Y_{0.}(v)} dN_{1.}(v) - \int_0^\infty \frac{Y_{1.}(v)}{Y_{1.}(v) + Y_{0.}(v)} dN_{0.}(v) \\
= \; & \int_0^\infty \frac{Y_{0.}(v) Y_{1.}(v)}{Y_{1.}(v) + Y_{0.}(v)} \left[\frac{dN_{1.}(v)}{Y_{1.}(v)} - \frac{dN_{0.}(v)}{Y_{0.}(v)} \right]
\end{aligned}
$$

Part IV.1.2B Revisit to Logrank Test

Remarks

- Nelson-Aalen Estimator for the cumulative hazards $H_l(\cdot)$:

$$\hat{H}_1(t) = \int_0^t \frac{1}{Y_{1.}(v)} dN_{1.}(v), \quad \hat{H}_0(t) = \int_0^t \frac{1}{Y_{0.}(v)} dN_{0.}(v)$$

(Nelson, 1969; Aalen, 1972)

$\Longrightarrow \hat{S}_l(t) = \exp\{-\hat{H}_l(t)\}$,

an estimator for the survivor function $S_l(\cdot)$: not the same as KM estimator but asymptotically equivalent, with higher finite sample efficiency.

Part IV.1.2B Revisit to Logrank Test

- A Class of Rank Tests: $H_0 : H_0(\cdot) = H_1(\cdot)$

$$\int_0^\infty W_n(t)[d\hat{H}_1(t) - d\hat{H}_0(t)],$$

with $W_n(\cdot)$ predictable, non-negative and order of n.

  - e.g. $W_n(t) = \frac{Y_{0.}(t)Y_{1.}(t)}{Y_{1.}(t)+Y_{0.}(t)} \Longrightarrow$ *Logrank Test*
  - e.g. $W_n(t) = \frac{Y_{0.}(t)Y_{1.}(t)}{n} \Longrightarrow$ *Gehan-Wilcoxon Test*
  - e.g. $W_n(t) = \hat{S}^-(t)\frac{Y_{0.}(t)Y_{1.}(t)}{Y_{1.}(t)+Y_{0.}(t)} \Longrightarrow$ *Prentice-Wicoxon Test* ($\hat{S}^-(t)$ the left-cont's version of SM estm)
  - e.g. $W_n(t) = [\hat{S}^-(t)]^\alpha\frac{Y_{0.}(t)Y_{1.}(t)}{Y_{1.}(t)+Y_{0.}(t)} \Longrightarrow$ *Fleming-Hurrington Test*

Part IV.1.2B Revisit to Logrank Test

- A Class of Rank Tests: $H_0 : H_0(\cdot) = H_1(\cdot)$

$$\int_0^{\infty} W_n^*(t) \frac{Y_{0.}(t) Y_{1.}(t)}{Y_{1.}(t) + Y_{0.}(t)} [d\hat{H}_1(t) - d\hat{H}_0(t)],$$

with $W_n^*(\cdot)$ predictable, bounded, and non-negative. *How to choose $W_n^*(\cdot)$ to achieve a powerful test?*

  - If $H_1 : \log\left[h_1(t)/h_0(t)\right] = \alpha g(t)$, optimal weight?

e.g. Lagakos and Schoenfeld (1984).

What to study next?

Part IV. Advanced Topics

- **Part IV.1 Counting Process Formulation** (Revisits to KM estm, Logrank test, and Cox PH model)
  - *Part IV.1.1 Theoretical Preparation*
  - **Part IV.1.2 Counting Process Formulation in LIDA and Applications: Revisits to KM, Logrank, Cox PH**
    - *Part IV.1.2A Formulation*
    - *Part IV.1.2B Revisit to Logrank Test*
    - **Part IV.1.2C Revisit to Cox's Partial Likelihood Approach**
    - **Part IV.1.2D Revisit to Kaplan-Meier Estimator**
    - *Part IV.1.2E Others*

- *Part IV.2 Selected Recent Topics in LIDA*
- *Part IV.3 Beyond Lifetime Data Analysis*