- Now it's time to study the asymptotics of

$$U_n(t) = \sum_{i=1}^{n} \int_0^t H_i(u) dM_i(u)$$

- Are we there yet?
  - Wait a second, what do we mean by asymptotics?
  - that means "what if sample size gets big...," i.e., $n \to \infty$
  - what n, i.e., the order of n?
  - $n^{-1}U_n$? $n^{-1/2}U_n$?

- Mode of convergence: $X_n \to X$
  - convergence in probablity $X_n \to_P X$: for any $\epsilon > 0$

$$\Pr\{\omega : |X_n(\omega) - X(\omega)| > \epsilon\} \to 0$$

  - almost surely $X_n \to X, a.s.$:

$$\Pr\{\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\} = 1$$

  - convergence in distribution $X_n \to_{\mathscr{D}} X$

$$\Pr\{\omega : X_n(\omega) \leq x\} \to \Pr\{\omega : X(\omega) \leq x\}$$

    at continuity points of $F_X(x)$

- Alternative view on convergence in distribution
  - any $F_n(\cdot)$ defines a measure: $\mathscr{P}_n = \mathscr{P}_{F_n}$
  - $\sigma\{(-\infty, x]; x \in R\}$ defines $\mathscr{B}(R)$
  - $X_n \rightarrow_{\mathscr{D}} X \Leftrightarrow \mathscr{P}_n(B) \rightarrow \mathscr{P}(B)$ for any $B \in \mathscr{B}$
- Weak convergence of stochastic processes
  - Γ: metric space of $X(\cdot)$
  - $\mathscr{S}_0$: σ-algebra generated by all open sets
  - $\mathscr{P}_n(A) \rightarrow \mathscr{P}(A)$ for any $A \in \mathscr{S}_0$

- Example
  - $D[0,\tau]$: cadlag functions
  - $d(x,y)$: Skorohod metric for $x(t), y(t) \in D[0,\tau]$
    1. $r(\cdot)$: strictly increasing time-transformation with $r(0) = 0$, $r(\tau) = \tau$
    2. $d(x,y) = \inf\{\epsilon; \sup_t |x(t) - y[r(t)]| \leq \epsilon, \sup_t |r(t) - t| \leq \epsilon\}$
    3. neighborhood: $U(x;\delta) = \{y(t); d(x,y) < \delta\}$
    4. open set A: for any $x \in A$, there exists $U(x;\delta) \subset A$
  - $\mathscr{S}$: $\sigma-$algebra generated by all open sets
  - $X_n(\cdot) \Rightarrow X(\cdot)$ iff $\mathscr{P}_n(A) \to \mathscr{P}(A)$ for any $A \in \mathscr{S}$

- Properties on weak convergence
  1. Continuous mapping theorem: if $X_n \Rightarrow X$, then $g(X_n) \rightarrow_{\mathscr{D}} g(X)$,
     - $\text{sup}_t\{X_n(t)\} \rightarrow_{\mathscr{D}} \text{sup}_t\{X(t)\}$
     - $(X_n(t_1), \ldots, X_n(t_k)) \rightarrow_{\mathscr{D}} (X(t_1), \ldots, X(t_k))$ for any $(t_1, \ldots, t_k)$
     - $\int_0^\tau H(t)dX_n(t) \rightarrow_{\mathscr{D}} \int_0^\tau H(t)dX(t)$
  2. Slutsky's theorem: if $H_n \rightarrow_{\mathscr{D}} H$, i.e., $\text{Pr}\{\text{sup}\,|H_n(t) - H(t)| > \epsilon\} \to 0$:
     - $H_nX_n \Rightarrow HX$
     - $H_n + X_n \Rightarrow H + X$

- A theorem on weak convergence of cadlag processes
  - X_n and X are stochastic processes with cadlag sample paths on $(D[0,\tau],\mathscr{S})$, such that
    1. $(X_n(t_1),\ldots,X_n(t_k)) \to_{\mathscr{D}} (X(t_1),\ldots,X(t_k))$ for any $(t_1,\ldots,t_k)$
    2. X_n is tight

    then $X_n \Rightarrow X$ on $(D[0,\tau],\mathscr{S})$

- Tightness
  - X_n is said to be tight, if there exists a compact set K such that $\mathscr{P}_n(K) > 1-\epsilon$ for any $\epsilon > 0$.
  - Stone's sufficient condition: there exists $d > 0$ such that for any $\epsilon, \eta > 0$ and $0 \le s, t \le \tau$

$$\limsup_{n\to\infty} \Pr\{\sup_{|s-t|<d} |X_n(s) - X_n(t)| > \epsilon\} < \eta$$

- What is the asymptotic (limiting) distribution of

$$U_n(t) = \sum_{i=1}^{n} \int_0^t H_i(u) dM_i(u)$$

- Answer: Martingale Central Limit Theorem (MCLT)
  - $U_n(t) \Rightarrow U$ on $(D[0,\tau], \mathscr{S})$
  - U is a time-transformed Brownian motion
    1. $EU(t) = 0$
    2. $\text{var}\{U(t)\} = \lim_{n\to\infty} < U_n, U_n > (t)$
    3. independent increment: $U(s)$ and $U(t) - U(s)$ are independent for $s \leq t$

- Assumptions
  - convergent variance: $< U_n, U_n > (t) \to v(t)$
  - smoothness (no spikes): $< U_{n,\epsilon}, U_{n,\epsilon} > (\tau) \to 0$ for any $\epsilon > 0$, where

$$U_{n,\epsilon}(t) = \sum_{i=1}^{n} \int_0^t H_i(u) I(|H_i(u)| \geq \epsilon) dM_i(u)$$

- Two more lemmae (Tsiatis, 1981):

1. Z_i are random variables such that $Eg(Z_i)^2 < \infty$ where $g(\cdot)$ is continuous, then

$$\sup_t \left| \sum_{i=1}^{n} \frac{g(Z_i)I(X_i \geq t)}{n} - E[g(Z)I(X \geq t)] \right| \rightarrow_P 0$$

2. $Y_n(t) \rightarrow_P Y(t)$, $Z_n(t) \rightarrow_P Z(t)$ and $Q_n(t) \rightarrow_P Q(t)$. $Q_n(t)$ is positive and increasing. $Q(\tau)$ is bounded. $f(y, z)$ is continuous. Then

$$\sup_t \left| \sum_{i=1}^{n} \int_0^t f(Y_n, Z_n) dQ_n - \int_0^t f(Y, Z) dQ \right| \rightarrow_P 0$$

- Asymptotics of Nelson-Aalen estimator
  - $$U_n(t) = \sqrt{n}[\widehat{\Lambda}(t) - \Lambda(t)] = \sum_{i=1}^{n} \int_0^t \frac{\sqrt{n}}{Y(u)} dM_i(u)$$

    for $t < \tau = \sup_t\{t; \Pr(X \geq t) > 0\}$
  - $$\begin{aligned} < U_n, U_n > (t) &= \sum_i \int_0^t \frac{n}{Y(u)^2} Y_i(t)\lambda(u)du \\ &= \int_0^t \frac{n}{Y(u)^2} Y(u) \lambda(u)du \\ &= \int_0^t \frac{\lambda(u)du}{Y(u)/n} \longrightarrow v(t) = \int_0^t \frac{\lambda(u)dt}{EY_1(u)} \end{aligned}$$
  - $$\begin{aligned} < U_{n,\epsilon}, U_{n,\epsilon} > (t) &= \int_0^t \frac{n}{Y(u)} I\left(\frac{\sqrt{n}}{Y(u)} \geq \epsilon\right) \lambda(u)du \\ &= \int_0^t \frac{1}{Y(u)/n} I\left(\frac{n^{-1/2}}{Y(u)/n} \geq \epsilon\right) \lambda(u)du \longrightarrow 0 \end{aligned}$$

- Back-of-the-envelop calculations on

$$U_n(t) = \sum_{i=1}^{n} \int_0^t H_i(u) dM_i(u)$$

1. $\text{cov}\{dM_i(t), dM_j(t) \mid \mathscr{F}_{t-}\} = 0$
2. $\text{var}\{dM_i(t) \mid \mathscr{F}_{t-}\} = Y_i(t)\lambda(t)dt$

$$\int_0^t \sum_{i=1}^{n} H_i(u)^2 Y_i(u)\lambda(u)du \to v(t)$$

3. Asymptotic theory

- From $\widehat{\Lambda}(t)$ to $\widehat{S}(t) = \exp\{-\widehat{\Lambda}(t)\}$
  - functional transformation: $\phi(\cdot) = \exp(\cdot)$
  - motivation for functional Delta method
  - usual Delta method: $\widehat{\theta}_n \rightarrow_P \theta$, $f(\widehat{\theta}_n) - f(\theta) \doteq f'(\theta)(\widehat{\theta}_n - \theta)$, $\theta \in \Theta$ and $\dim(\Theta) < \infty$
  - parameter estimates: $\widehat{\Lambda}(t) \in D[0, \tau]$ with $\dim(D) = \infty$
- Differentiation of $\phi(\cdot)$ with respect to an infinite-dimensional parameter F
  - Gâteaux differentiability
  - Compact differentiability
  - Fréchet differentiability

- Gâteaux differentiability at F in the direction of G:

$$d\phi(F;G) = \lim_{\rho\to 0+} \frac{\phi[F+\rho(G-F)] - \phi(F)}{\rho} = \frac{d}{d\rho}\phi[F+\rho(G-F)]_{\rho=0}$$

  - example 1: mean functional
    * $\phi(F) = \mu_F = \int u dF(u)$
    * $d\phi(F;G) = \mu_G - \mu_F$
  - example 2: variance functional
    * $\phi(F) = \mathrm{var}_F(X) = \int\int_{(u,v)} \frac{(u-v)^2}{2} dF(u)dF(v)$
    * $d\phi(F;G) = \phi(G) - \phi(F) + (\mu_G - \mu_F)^2$

- Compact differentiability
  - for any $\rho_n \to 0$ and any $G_n \to G \in D[0, \tau]$,

$$d\phi(F; G) = \lim_{n\to\infty} \frac{\phi(F + \rho_n G_n) - \phi(F)}{\rho_n}$$

- Fréchet differentiability
  - $d\phi(F; G)$ satisfies that

$$\lim_{\|\rho\|\to 0} \frac{\|\phi(F + G) - \phi(F) - d\phi(F; G)\|}{\|\rho\|} = 0$$

- Functional Delta method
  - Suppose that $n^{1/2}(U_n - U) \Rightarrow Z$. If $\phi(\cdot)$ is compact differentiable, then

$$n^{1/2}[\phi(U_n) - \phi(U)] \Rightarrow d\phi(U; Z),$$

    and $n^{1/2}[\phi(U_n) - \phi(U)] \doteq d\phi(U; n^{1/2}(U_n - U))$
  - Ref: ABGK, pp 109-111

- Examples
  - $n^{1/2}(\log \widehat{\Lambda} - \log \Lambda) \doteq 1/\Lambda \cdot n^{1/2}(\widehat{\Lambda} - \Lambda)$
  - $n^{1/2}[\exp(-\widehat{\Lambda}) - \exp(-\Lambda)] \doteq \exp(-\Lambda) \cdot n^{1/2}(\widehat{\Lambda} - \Lambda)$

- Asymptotics on the Kaplan-Meier estimator
  - Kaplan-Meier estimator

$$\widehat{S}(t) = \prod_{X_i \le t, \Delta_i = 1} \left[1 - \frac{1}{Y(X_i)}\right] = \prod_{s \le t} \left[1 - \frac{\Delta N(s)}{Y(s)}\right]$$

  - Consistency

1. an equation (Fleming & Harrington, p. 97): let

$$Z(t) = \frac{\widehat{S}(t) - S(t)}{S(t)} = \int_0^t \frac{\widehat{S}(u-)}{S(u)} \frac{dM(u)}{Y(u)}$$

2. Lenglart's inequality

$$\Pr\left\{\sup_{0 \le t \le \tau} \left[\int_0^t H(u)dM(u)\right]^2 \ge \epsilon\right\} \le \frac{\eta}{\epsilon} + \Pr\left\{\int_0^\tau H(u)^2 d < M, M > (u) \ge \eta\right\}$$

3.

$$\Pr\left\{\sup_{0\le t\le\tau} Z(t)^2 \ge \epsilon\right\} \le \frac{\eta}{\epsilon} + \Pr\left\{\int_0^\tau \left[\frac{\widehat{S}(u-)}{S(u)}\right]^2 \frac{d\Lambda(u)}{Y(u)} \ge \eta\right\}$$

$$\le \frac{\eta}{\epsilon} + \Pr\left\{\frac{\Lambda(\tau)}{S(\tau)^2 Y(\tau)} \ge \eta\right\}$$

4. $n \to \infty$, $Y(\tau) \to \infty$, for all $\eta > 0$

$$\Pr\left\{\frac{\Lambda(\tau)}{S(\tau)^2 Y(\tau)} \ge \eta\right\} \to 0$$

5. $Z(\cdot) \to_P 0 \Rightarrow \widehat{S}(\cdot) \to_P S(t)$

– Asymptotic normality

1. Let

$$U_n(t) = n^{1/2}Z(t) = \sum_{i=1}^{n}\int_0^t \frac{-\widehat{S}(u-)}{S(u)}\frac{n^{1/2}}{Y(u)}dM_i(u)$$

2.

$$< U_n, U_n > (t) \doteq \int_0^t \left[\frac{\widehat{S}(u-)}{S(u)}\right]^2 \frac{n}{Y(u)}d\Lambda(u)$$
$$\to_P \int_0^t \frac{d\Lambda(u)}{EI(X \geq u)}$$

3.

$$< U_{n,\epsilon}, U_{n,\epsilon} > (t) \doteq \int_0^t \left[\frac{\widehat{S}(u-)}{S(u)}\right]^2 \frac{n}{Y(u)} I\left[\frac{\widehat{S}(u-)}{S(u)}\frac{n^{1/2}}{Y(u)} \geq \epsilon\right] d\Lambda(u)$$
$$\to_P 0$$

4. Martingale CLT

$$n^{1/2}Z(t) \Rightarrow U(t)$$

where var$\{U(t)\} = \int_0^t d\Lambda(u)/EI(X \geq u)$.

5.

$$n^{1/2}[\widehat{S}(t) - S(t)] \Rightarrow S(t)U(t)$$

6.

$$\begin{aligned}
\widehat{\text{var}}\{\widehat{S}(t)\} &= [\widehat{S}(t)]^2\widehat{\text{var}}[U(t)] \\
&= n^{-1}[\widehat{S}(t)]^2 \int_0^t \left[\frac{\widehat{S}(u-)}{\widehat{S}(u)}\right]^2 \frac{n}{Y(u)}[1 - \Delta\widehat{\Lambda}(u))]d\widehat{\Lambda}(u) \\
&= [\widehat{S}(t)]^2 \int_0^t \left[\frac{\prod_{v<u}\{1 - \Delta N(v)/Y(v)\}}{\prod_{v\leq u}\{1 - \Delta N(v)/Y(v)\}}\right]^2 \frac{1}{Y(u)}\left[1 - \frac{\Delta N(u)}{Y(u)}\right]\frac{dN(u)}{Y(u)} \\
&= [\widehat{S}(t)]^2 \int_0^t \left[\frac{1}{\{1 - \Delta N(u)/Y(u)\}}\right]^2 \left[1 - \frac{\Delta N(u)}{Y(u)}\right]\frac{dN(u)}{Y(u)^2} \\
&= [\widehat{S}(t)]^2 \int_0^t \frac{dN(u)}{[Y(u) - \Delta N(u)]Y(u)}
\end{aligned}$$

This is the Greenwood's formula

- Confidence intervals/bands for $\Lambda(\cdot)$
  - pointwise confidence interval at $0 \leq t \leq \tau$ such that

$$\Pr\{C_l(t) \leq \Lambda(t) \leq C_u(t)\} = 0.95$$

  - equal precision confidence band: find lower bound of $C_l(t)$ and upper bound of $C_u(t)$ such that

$$\Pr\{C_l(t) \leq \Lambda(t) \leq C_u(t) \text{ for all } 0 \leq t \leq \tau\} = 0.95$$

- Method 1: direct method
  - pointwise confidence interval
    1. $\text{asyvar}[\widehat{\Lambda}(t)] = E\int_0^t dN(u)/Y(u)^2$
    2. $\widehat{\sigma}(t)^2 = \widehat{\text{var}}[\widehat{\Lambda}(t)] = \sum_{i=1}^n \int_0^t dN_i(u)/Y(u)^2$
    3. 95% confidence interval at t_0: $\widehat{\Lambda}(t) \pm 1.96\widehat{\sigma}(t)$
  - confidence bands when the process is standardized
    1. $\sup_{0\le t\le \tau} |n^{1/2}[\widehat{\Lambda}(t) - \Lambda(t)]/v(\tau)| \to \sup_{0\le t\le 1}|W(t)|$
    2. $W(t)$ is standard Brownian motion (Wiener Process):

$$\Pr\{\sup_{0\le t\le 1}|W(t)| \le c\} = \frac{4}{\pi}\sum_{k=0}^{\infty}\frac{(-1)^k}{2k+1}e^{-\frac{\pi^2(2k+1)^2}{8c^2}} \approx 0.95$$

    3. 95% confidence bands: $\widehat{\Lambda}(t) \pm c\sqrt{\widehat{v}(\tau)}$

- Method 2: bootstrap method
  - compute centered bootstrap samples
    1. resample n subjects with replacement from $\{(X_i, \Delta_i), i = 1, 2, \ldots, n\}$
    2. compute $\widehat{\Lambda}^*(t)$ from the bootstrap sample and compute $n^{1/2}[\widehat{\Lambda}^*(t) - \widehat{\Lambda}(t)]$
  - pointwise confidence intervals: find 2.5%-tile (c_l) and 97.5%-tile (c_u) of the centered samples and compute $[\widehat{\Lambda}(t) + c_l/\sqrt{n}, \widehat{\Lambda}(t) + c_u/\sqrt{n}]$
  - confidence bands: compute the sup of the absolute values from centered samples and find the 95%-tile (c). compute $[\widehat{\Lambda}(t) - c/\sqrt{n}, \widehat{\Lambda}(t) + c/\sqrt{n}]$

- Method 3: simulation method
  - a simple example: $X_1, X_2, \ldots, X_n \sim F(x)$ with mean μ and unknown variance σ^2
    1. confidence interval for μ: $\bar{X} \pm 1.96\sigma/\sqrt{n}$, because
    $$\sqrt{n}(\bar{X} - \mu) \to_D \mathcal{N}(0, \sigma^2)$$
    2. alternative view: $\sqrt{n}(\bar{X} - \mu) = \sum_{i=1}^{n}(X_i - \mu)/\sqrt{n}$
    3. suppose X_i are known. then $X_i - \mu$ can be approximated by $X_i G_i$ where $G_i \sim \mathcal{N}(0, 1)$
    4. then $E\sum_{i=1}^{n} X_i G_i/\sqrt{n} = 0$ and $\text{var}\{\sum_{i=1}^{n} X_i G_i/\sqrt{n}\} = \sigma^2$, i.e.
    $$\sqrt{n}(\bar{X} - \mu) = \sum_{i=1}^{n} X_i G_i/\sqrt{n}$$
    5. simulate $\{G_i\}$ to construct similar sampling distribution

- $U_n(t) = \sum_{i=1}^n \int_0^t H_i(u) dM_i(u)$
  1. $\tilde{U}_n(t) = \sum_{i=1}^n \int_0^t H_i(u) G_i dN_i(u)$
  2. $E\tilde{U}_n(t) = 0$ and $\text{var}\{\tilde{U}_n(t)\} = \text{var}\{U_n(t)\}$
  3. $\text{cov}\{\tilde{U}_n(s), \tilde{U}_n(t)\} = \text{cov}\{U_n(s), U_n(t)\}$
- generate simulation samples
  1. simulate n iid copies of $G_i \sim \mathcal{N}(0, 1)$
  2. compute $\tilde{U}_n(t)$
- pointwise confidence interval: find the 2.5%-tile (c_l) and 97.5%-tile (c_u) of $\tilde{U}_n$ from the simulated samples, then $\Pr\{c_l \leq EU_n(t) \leq c_u\} \approx 0.95$
- confidence bands: find the 95%-tile (c) of $\sup_{0 \leq t \leq \tau} |\tilde{U}_n(t)|$ from the simulated samples, then $\Pr\{\sup_{0 \leq t \leq \tau} |U_n(t)| \leq c\} \approx 0.95$