3. Counting Processes and Martingales

- This lecture's topics:
  - Stochastic processes
  - Definitions
  - Filtration
  - Intensity process
  - Martingales
- Text: TG Chapter 2, FH Chapter 1-2

Background

- We have used counting processes to represent one-sample estimators
  - demonstrated asymptotic unbiasedness of N-A estimator
  - derived N-A variance estimator
  - demonstrated asymptotic unbiasedness of variance estimator
- Note: we have yet to derive the limiting distribution of $n^{1/2}\{\widehat{\Lambda}(t) - \Lambda(t)\}$
- In deriving properties of $\widehat{\Lambda}(t)$, conditioning played an essential role
- An entire body of theoretical research has been built around such conditioning techniques
  - Martingale theory
- In fact, a Martingale Central Limit Theorem (MCLT) has been developed, whereby convergence to a Normal is automatic, under some mild conditions

Definitions: Algebra

- To begin our study of Martingales, we first establish some definitions and basic properties
- *Probability space* $(\Omega, \mathcal{F}, \mathcal{P})$ where an abstract space Ω is a set of all possible outcomes, $\mathcal{F}$ is a σ-algebra and $\mathcal{P}$ is a set function (measure)
  - let $\mathcal{A}$ be a collection of subsets (of outcomes) from Ω
  - if $E \in \mathcal{A}$ implies complement $\overline{E} \in \mathcal{A}$ and if $E_1 \in \mathcal{A}$ and $E_2 \in \mathcal{A}$ implies $E_1 \cup E_2 \in \mathcal{A}$, then $\mathcal{A}$ is an *algebra*.
  - if $E \in \mathcal{A}$ implies $\overline{E} \in \mathcal{A}$, and if $E_j \in \mathcal{A}$ for $(j = 1, 2, 3, \ldots)$ implies $E_1 \cup E_2 \cup \ldots \in \mathcal{A}$, then $\mathcal{A}$ is an *σ-algebra*.
  - a σ-algebra is a collection of events, closed under countable unions and intersections

Definitions: Stochastic Process

- *Stochastic process*: a collection of random variables $X = \{X(t);\ t \in \mathcal{T}\}$ defined on the same probability space
  - frequently, $\mathcal{T} = [0, \infty)$
  - another common choice: $\mathcal{T} = (0, \tau_*)$, where $P(X_i > \tau_*) > 0$ for $i = 1, \ldots n$
- *Path*: realization of a stochastic process
- *Counting process*: stochastic process for which each path is a non-decreasing, piece-wise constant, cadlag, step-function with increments of size 1
  - Typically, $N(0) = 0$ and $N(t) < \infty$ for all $t \in \mathcal{T}$

Definitions: Filtration

- *Filtration (history)*: an increasing family of σ-algebras $\{\mathcal{F}_t : t \geq 0\}$

- A filtration is increasing if $s \leq t$ implies $\mathcal{F}_s \subset \mathcal{F}_t$; i.e., if $A \in \mathcal{F}_s$ implies $A \in \mathcal{F}_t$
- *Adapted*: a stochastic process X is adapted to $\mathcal{F}_t$ if $X(t)$ is $\mathcal{F}_t$-measurable for all t
  - Essentially, a quantity is measurable if meaningful probability statements can be made about it
  - In particular, if $X(t)$ is adapted to $\mathcal{F}_t$ then $E[X(t)|\mathcal{F}_t] = X(t)$
- Any process is adapted to its own history
- Often in survival analysis, it is convenient to define $\mathcal{F}_t$ to represent the history of X
  - e.g., $\mathcal{F}_t = \sigma\{X(s);\ 0 \leq s \leq t\}$
  - i.e., $\mathcal{F}_t$ contains all data generated by X over $(0, t]$
  - $\mathcal{F}_t$ contains all the information available to an observer who has been watching each fan from the time it was put in service until it has run for t hours (including any fan failures or censorings in $[0, t]$)
- A filtration we will often use:

$$\mathcal{F}_t = \sigma\{N_i(s), Y_i(s+);\ s \in (0, t]\ i = 1, \ldots n\}$$

Conditional Expectation

- If a random variate X is $\mathcal{F}$-measurable, and if $\mathcal{G} \subset \mathcal{F}$, then:
  - $E[X|\mathcal{F}] = X$
  - $E[aX|\mathcal{F}] = aX$, where a is a constant
  - $E[XY|\mathcal{F}] = XE[Y|\mathcal{F}]$
  - $E[X|\mathcal{G}]$ is $\mathcal{G}$-measurable
  - for all events $B \in \mathcal{G}$, $E[XI(B)] = E[E[X|\mathcal{G}]I(B)]$

Properties of Stochastic Processes

- A stochastic process is ...
  - *integrable* if

$$\sup_{t \in \mathcal{T}} E[|X(t)|] < \infty$$

  - *square integrable* if

$$\sup_{t \in \mathcal{T}} E[X(t)^2] < \infty$$

  - *uniformly bounded* if

$$P\left\{\sup_{t \in \mathcal{T}} |X(t)| < c\right\} = 1$$

Intensity Process

- An intensity process $A(t)$,
  - corresponds to a counting process; e.g., $N(t)$
  - is defined with respect to a filtration; say $\mathcal{F}_t$
- Set $A(t) = \int_0^t dA(s)$, where

$$dA(t) \;=\; E[dN(t)|\mathcal{F}_{t^-}] = Y(t)\lambda(t)dt$$

- More specifically,

$$dA(t) \;=\; \lim_{dt \downarrow 0} E[N(t^- + dt) - N(t^-)|\mathcal{F}_{t^-}]$$

  where $\mathcal{F}_{t^-}$ contains information on $(0, t)$

- In the settings of interest to us, we assume that the probability of > 1 event in $[t, t + dt)$ is negligible; i.e.,

$$\lim_{dt \downarrow 0} P\{N(t^- + dt) - N(t^-) > 1|\mathcal{F}_{t^-}\} \;=\; o(dt^2)$$

- As such, an equivalent definition of the intensity process is given by,

$$dA(t) \;=\; \lim_{dt \downarrow 0} P\{N(t^- + dt) - N(t^-) = 1|\mathcal{F}_{t^-}\}$$

Martingale: Definition

- A right-continuous stochastic process $X = \{X(t) :\ t \geq 0\}$ is a martingale with respect to the filtration $\{\mathcal{F}_t : t \geq 0\}$ if
  (i) X is adapted to $\{\mathcal{F}_t : t \geq 0\}$
  (ii) $E[|X(t)|] < \infty$ for all $t < \infty$
  (iii) $E[X(t+s)|\mathcal{F}_t] = X(t)$ a.s. for all $t \geq 0$ and $s \geq 0$
- Note: X will be a sub-martingale if we replace condition (iii) with
  (iii)* $E[X(t+s)|\mathcal{F}_t] \geq X(t)$
- Note: X is a super-martingale if we replace condition (iii) with
  (iii)** $E[X(t+s)|\mathcal{F}_t] \leq X(t)$
- A martingale is a pure random noise process
  - has conditional mean 0, given the history
  - conditionally centered process
  - fluctuates about the mean randomly over t
- e.g., random walk; gambling (fair game)
- Define $dX(t)$ to be the martingale increment in X over $[t, t + dt)$; i.e.,
$$dX(t) \ = X(t^- + dt) - X(t^-)$$
- Preceding properties imply that $E[dX(t)|\mathcal{F}_{t^-}] = 0$
- We demonstrate that X, an $\mathcal{F}_t$ martingale, has uncorrelated increments
  - recall: $E[X(t)|\mathcal{F}_s] = X(s)$ for $s < t$
  - now, for $s < t$, $E[X(s)\{X(t) - X(s)\}]$
$$\begin{aligned} &= E\left[E[X(s)\{X(t) - X(s)\}|\mathcal{F}_s\right] \\ &= E\left[X(s)E[\{X(t) - X(s)\}|\mathcal{F}_s\right] \\ &= E\left[X(s)\{E[X(t)|\mathcal{F}_s] - E[X(s)|\mathcal{F}_s]\}\right] \\ &= 0 \end{aligned}$$

- Consider the usual counting process for univariate survival, $N(t) = I(X \leq t, \Delta = 1)$
- Set $M(t) = N(t) - A(t)$, where

$$\begin{aligned} A(t) &= \int_0^t dA(s) \\ dA(t) &= Y(t)\lambda(t)dt \\ &= E[dN(t)|\mathcal{F}_{t^-}] \end{aligned}$$

- We refer to the integrated intensity process, $A(t)$, as the compensator of $N(t)$ - centering the process
- We now demonstrate that $A(t)$ is in fact the compensator of $N(t)$.

Centering Increments

- Suppose that failure times are subject to independent (right) censoring, then
  - pertinent counting process: $N(t)$
  - filtration: $\mathcal{F}_t = \sigma\{N_i(s), Y_i(s+);\ i = 1, \ldots, n; s \in (0, t]\}$
- The compensator increments are given by,

$$\begin{aligned} E[dN_i(t)|\mathcal{F}_{t^-}] &= \Pr[dN_i(t) = 1|\mathcal{F}_{t^-}] \\ &= \Pr[dN_i(t) = 1|Y(t)] \\ &= Y_i(t) \Pr[t \leq T_i < t + dt | t \leq T_i, t \leq C_i] \\ &= Y_i(t) \Pr[t \leq T_i < t + dt | t \leq T_i] \\ &= Y_i(t) dA(t) \end{aligned}$$

- $M = N - A$? i.e., $\mathrm{E}[N_i(t)] = \mathrm{E}[A(t)]$.

Predictable Process

A stochastic process H is predictable w.r.t. the filtration $\mathcal{F}_t$ if for each t, the value of $H(t)$ is a function of (or is specified by) $\mathcal{F}_{t^-}$.

- H is predictable if its value at time t is fixed just prior to t, i.e. the behavior of H at t is determined by its behavior on $[0, t)$.
- Left-continuous processes are predictable, e.g., $Y(t)$.
- Any deterministic function is predictable, e.g., $S(t)$, $\lambda(t)$.
- $E[H(t)|\mathcal{F}_{t^-}] = H(t)$

Stochastic Integral

- Suppose that M is an $\mathcal{F}$-martingale
  - the process

$$Z(t) = \int_0^t H(s)dM(s)$$

    is a stochastic integral with respect to $M(t)$

- **<u>Claim</u>**: If H is predictable with respect to the filtration, $\mathcal{F}$, and if M is an $\mathcal{F}$ martingale, then $Z(t)$ defined above is an $\mathcal{F}$ martingale

- **<u>Proof</u>**:

We now show that $E[Z(t) - Z(s)|\mathcal{F}_s] = 0$

$$\begin{aligned} E[Z(s)|\mathcal{F}_s] &= E\left[\int_0^s H(u)dM(u)|\mathcal{F}_s\right] \\ &= \int_0^s E[H(u)dM(u)|\mathcal{F}_s] \\ &= \int_0^s H(u)dM(u) \\ &= Z(s) \end{aligned}$$

$$\begin{aligned} E[Z(t)|\mathcal{F}_s] &= E\left[\int_0^t H(u)dM(u)\middle|\mathcal{F}_s\right] \\ &= \int_0^t E[H(u)dM(u)|\mathcal{F}_s] \\ &= Z(s) + \int_s^t E[H(u)dM(u)|\mathcal{F}_s] \end{aligned}$$

- As before, we will apply conditioning; this time, to a conditional quantity
- First, we consider conditional expectations
- Iterating conditional expectations,

$$\begin{aligned} E[H(u)dM(u)|\mathcal{F}_s] &= E[E[H(u)dM(u)|\mathcal{F}_s, \mathcal{F}_{u^-}]|\mathcal{F}_s] \\ &= E[E[H(u)dM(u)|\mathcal{F}_{u^-}]|\mathcal{F}_s] \\ &= E[H(u)E[dM(u)|\mathcal{F}_{u^-}]|\mathcal{F}_s] \\ &= 0 \end{aligned}$$

- Therefore, we have

$$E[Z(t)|\mathcal{F}_s] \;=\; Z(s)$$

  such that $E[Z(t) - Z(s)|\mathcal{F}_s] = 0$ as required.

- Conclusion: stochastic integrals with respect to martingales are, themselves, martingales

- **Iterated Conditional Expectations**:

  – Assume (without loss of generality) that X, Y and Z are continuous

  – **Claim**: $E[X|Y] = E[E(X|Y,Z)|Y]$

  – **Proof**:

$$\begin{aligned}
E[X|Y,Z] &= \int_{\mathcal{X}} xf(x|y,z)dx \\
E[E(X|Y,Z)|Y] &= \int_{\mathcal{Z}}\int_{\mathcal{X}} xf(x|y,z)dxf(z|y)dz \\
&= \int_{\mathcal{Z}}\int_{\mathcal{X}} xf(x|y,z)f(z|y)dx\ dz \\
&= \int_{\mathcal{X}} x\int_{\mathcal{Z}} f(x|y,z)f(z|y)dz\ dx \\
&= \int_{\mathcal{X}} x\int_{\mathcal{Z}} \frac{f(x,y,z)}{f(y,z)}\frac{f(y,z)}{f(y)}dz\ dx \\
&= \int_{\mathcal{X}} x\int_{\mathcal{Z}} f(x,z|y)dz\ dx \\
&= \int_{\mathcal{X}} xf(x|y)dx \\
&= E[X|Y]
\end{aligned}$$

Key Martingales Properties

- We can express the (key) martingale property (iii) in terms of the increments
- Claim: $E[X(t) - X(s)|\mathcal{F}_s] = 0 \Longleftrightarrow E[dX(t)|\mathcal{F}_{t^-}] = 0$
- Proof:
  - assume that the LHS holds; then,

    set $t = u^- + du$ and $s = u^-$, to obtain

$$\begin{aligned} E[dX(u)|\mathcal{F}_{u^-}] &= E[X(u^- + du) - X(u^-)|\mathcal{F}_{u^-}] \\ &= E[X(t) - X(s)|\mathcal{F}_s] \\ &= 0 \end{aligned}$$

  - now, assume that the RHS holds; then,

$$\begin{aligned} E[X(t) - X(s)|\mathcal{F}_s] &= E\left[\int_s^t dX(u)\middle| \mathcal{F}_s\right] \\ &= \int_s^t E[dX(u)|\mathcal{F}_s] \\ &= \int_s^t E\left[E[dX(u)|\mathcal{F}_{u^-}, \mathcal{F}_s]|\mathcal{F}_s\right] \\ &= \int_s^t E\left[E[dX(u)|\mathcal{F}_{u^-}]|\mathcal{F}_s\right] \\ &= 0 \end{aligned}$$

- **Claim**: If M is an $\mathcal{F}$-martingale and H is predictable, then $Z = \int_0^t HdM$ is an $\mathcal{F}$-martingale

- **Proof**:

$$\begin{aligned} dZ(t) &= H(t)dM(t) \\ E[dZ(t)|\mathcal{F}_{t^-}] &= E[H(t)dM(t)|\mathcal{F}_{t^-}] \\ &= H(t)E[dM(t)|\mathcal{F}_{t^-}] \\ &= 0 \end{aligned}$$

More Properties of Stochastic Integrals

- Suppose that M is an $\mathcal{F}$ martingale and H is $\mathcal{F}$ predictable.
  - $Z = \int HdM$ is also an $\mathcal{F}$ martingale
- Properties of $Z = \int HdM$,

  (i) $E[Z(t)] = 0$

  (ii) $\mathrm{corr}[Z(t) - Z(s), Z(s)] = 0$

  (iii) $V\{Z(t)\} = \ldots$
- We have already proved (i) and (ii)
- We now derive (iii)

Variance of Stochastic Integral

- **Claim**: If M is an $\mathcal{F}$ martingale, H is $\mathcal{F}$ predictable and $Z = \int HdM$, then

$$\begin{aligned} V\{Z(t)\} &= E\left[\int_0^t H^2(s)dA(s)\right] && (1) \\ &= E\left[\int_0^t H^2(s)dN(s)\right] && (2) \end{aligned}$$

- **Proof**:

  We break the $(0, t]$ interval up into m non-overlapping subintervals of equal length, t/m

  $(0, t_1]$, $(t_1, t_2] \ldots, (t_{m-1}, t]$

$$\begin{aligned} V\{Z(t)\} &= E[Z(t)^2] \\ &= E\left[\left\{\int_0^t H(s)dM(s)\right\}^2\right] \end{aligned}$$

- Then, for $m \to \infty$,

$$\begin{aligned} V\{Z(t)\} &= \lim_{m\to\infty} E\left[\left\{\sum_{j=1}^m H(t_j)\Delta M_j\right\}^2\right] \\ &= \lim_{m\to\infty}\{\sum_{j=1}^m E\left[\{H(t_j)\Delta M_j\}^2\right] \\ &\quad +2\sum_{j=1}^m \sum_{k=j+1}^m E\left[H(t_j)\Delta M_j H(t_k)\Delta M_k\right]\} \end{aligned}$$

- Consider the cross-products in the second term ($t_j < t_k$),

$$\begin{aligned} E\left[H(t_j)\Delta M_j H(t_k)\Delta M_k\right] &= E\left[E[H(t_j)\Delta M_j H(t_k)\Delta M_k|\mathcal{F}_{k-1}]\right] \\ &= E\left[H(t_j)\Delta M_j H(t_k) E[\Delta M_k|\mathcal{F}_{k-1}]\right] \\ &= 0 \end{aligned}$$

- We can also use conditioning to compute the first term,

$$\begin{aligned} E[\{H(t_j)\Delta M_j\}^2] &= E\left[E[H(t_j)^2\Delta M_j^2|\mathcal{F}_{j-1}]\right] \\ &= E\left[H(t_j)^2 E[\Delta M_j^2|\mathcal{F}_{j-1}]\right] \\ &= E\left[H(t_j)^2 V(\Delta M_j|\mathcal{F}_{j-1})\right] \\ &= E\left[H(t_j)^2 V(\Delta N_j|\mathcal{F}_{j-1})\right] \end{aligned}$$

- Now, recall that, for $m \to \infty$,
  - $\Delta N_j | \mathcal{F}_{j-1} \sim \text{Binomial}(Y(t_j), \lambda(t_j)\Delta(t_j))$, such that
  - $E[\Delta N_j | \mathcal{F}_{j-1}] = \Delta A_j$ and
  - $V(\Delta N_j | \mathcal{F}_{j-1}) = \Delta A_j \{1 - \lambda(t_j)\Delta(t_j)\} = \Delta A_j + o(\Delta A_j)$

  provided that H is bounded.

- Then, we obtain

$$E[\{H(t_j)\Delta M_j\}^2] = E\left[H(t_j)^2(\Delta A_j + o(\Delta A_j))\right]$$

- Combining the above results,

$$\begin{aligned} V\{Z(t)\} &= \lim_{m\to\infty} E\left[\sum_{j=1}^{m} H^2(t_j)(\Delta A_j + o(\Delta A_j))\right] \\ &= E\left[\int_0^t H^2(s)dA(s)\right] \end{aligned}$$

  thus proving (1) in property (iii)

- To prove (2), we first note that $\int H^2 dM$ is an $\mathcal{F}$ martingale, since M is an $\mathcal{F}$ martingale and H^2 is predictable

- Therefore, we obtain

$$\begin{aligned} E\left[\int_0^t H^2(s)dM(s)\right] &= 0 \ \text{ such that} \\ E\left[\int_0^t H^2(s)\{dN(s) - dA(s)\}\right] &= 0 \\ E\left[\int_0^t H^2(s)dN(s)\right] &= E\left[\int_0^t H^2(s)dA(s)\right] \end{aligned}$$

  thus proving (2) in property (iii).

- In summary, we have demonstrated that, for $Z = \int HdM$,

$$\begin{aligned} V\{Z(t)\} &= E\left[\int_0^t H^2(s)dA(s)\right] \\ &= E\left[\int_0^t H^2(s)dN(s)\right] \end{aligned}$$

Application: Nelson-Aalen Estimator

- Example: Using the properties of $Z = \int HdM$, derive the variance of the Nelson-Aalen estimator, $\widehat{\Lambda} = \int Y^{-1}dN$

$$\begin{aligned} V\{\widehat{\Lambda}(t)\} &= V\{\widehat{\Lambda}(t) - \Lambda(t)\} \\ \widehat{\Lambda}(t) - \Lambda(t) &= \int_0^t Y(s)^{-1}\{dN(s) - Y(s)d\Lambda(s)\} \\ &= \int_0^t Y(s)^{-1}dM(s) = \int HdM \end{aligned}$$

where $H = Y^{-1}$.

- Since $M(t)$ is a martingale with respect to the filtration

$$\mathcal{F}_t = \sigma\{N_i(s), Y_i(s+);\ i = 1, \cdots, n; s \in (0, t]\}$$

and since $H(t) = Y(t)^{-1}$ is $\mathcal{F}(t)$ predictable, we obtain

$$\begin{aligned} V\{\widehat{\Lambda}(t)\} &= E\left[\int_0^t Y(s)^{-2}dA(s)\right] \\ &= E\left[\int_0^t Y(s)^{-1}d\Lambda(s)\right] \end{aligned}$$

- Using (2) in property (iii), we also have

$$V\{\widehat{\Lambda}(t)\} = E\left[\int_0^t Y(s)^{-2}dN(s)\right]$$

- Note that both (1) and (2) suggest the following estimator,

$$\begin{aligned}\widehat{V}\{\widehat{\Lambda}(t)\} &= \int_0^t Y(s)^{-1} d\widehat{\Lambda}(s) \\ &= \int_0^t Y(s)^{-2} dN(s)\end{aligned}$$

Doob-Meyer Decomposition

- Implication: Any counting process may be uniquely decomposed as the sum of a martingale and a predictable, right-continuous process, compensator.

- Theorem: For any non-negative right-continuous $\mathcal{F}$ sub-martingale, X, there exists a unique increasing right-continuous predictable process, A, such that $A(0) = 0$ and $M = X - A$ is an $\mathcal{F}$ martingale.
  - X is separated into systematic (predictable) and random noise (martingale) parts
  - A is known as the compensator for X, since $X - A$ is a random error process
  - the compensator of a sub-martingale can be determined via $E[dX(t)|\mathcal{F}_{t^-}]$

- Corollary: Let N be a counting process adapted to $\mathcal{F}$, with $E[N(t)] < \infty$ for all t. There exists a unique increasing right-continuous predictable process, A, such that $A(0) = 0$, $E[A(t)] < \infty$ for all t, where $N - A$ is an $\mathcal{F}$ martingale.
  - N is a sub-martingale
  - $A(t) = \int_0^t Y(s)\lambda(s)ds = \int_0^t dA(s)$, where $dA(s)$ is known as the intensity process
  - we have $E[N(t)] = E[A(t)]$, as described previously

- **Note**: in contrast to the hazard function, $\lambda(t)$, the intensity process $dA(t) = Y(t)\lambda(t)dt$ is random

Predictable Variation Process

- The *predictable variation process* of a square integrable martingale M is denoted by

$$\begin{aligned} \langle M\rangle(t) &= \int_0^t d\langle M\rangle(s) \\ d\langle M\rangle(s) &= V\{dM(s)|\mathcal{F}_{s^-}\} \end{aligned}$$

- Main interest: $\langle M\rangle$ is related to $V(M)$

- Claim: The variance of an $\mathcal{F}$ martingale M is the mean of its compensator A.

  Proof: We will prove the result by demonstrating that:

  - $M^2(t)$ is a sub-martingale
  - $M^2(t) - \langle M\rangle(t)$ is a martingale
  - $\langle M\rangle(t) = A(t)$

- since $V\{M(t)\} = E[M^2(t)]$, we will show that $V\{M(t)\} = E[A(t)]$

- Claim: The process M^2 is a sub-martingale.

Proof: for $s < t$,

$$\begin{aligned} E[M^2(t)|\mathcal{F}_s] &= E[\{M(t) - M(s)\}^2 \\ &\quad -M^2(s) + 2M(t)M(s)|\mathcal{F}_s] \\ &= E[\{M(t) - M(s)\}^2|\mathcal{F}_s] \\ &\quad -E[M^2(s)|\mathcal{F}_s] + 2E[M(t)M(s)|\mathcal{F}_s] \\ &= E[\{M(t) - M(s)\}^2|\mathcal{F}_s] + E[M^2(s)|\mathcal{F}_s] \\ &\geq M^2(s) \end{aligned}$$

- Claim: The predictable variation process of a compensated counting process is the compensator itself.

  Proof:

$$\begin{aligned} M(t) &= N(t) - A(t) \\ \langle M\rangle(t) &= \int_0^t d\langle M\rangle(s) \end{aligned}$$

$$\begin{aligned} d\langle M\rangle(s) &= V\{dM(s)|\mathcal{F}_{s^-}\} \\ &= V\{dN(s) - dA(s)|\mathcal{F}_{s^-}\} \\ &= V\{dN(s)|\mathcal{F}_{s^-}\} \end{aligned}$$

$$\begin{aligned} E[dN(s)|\mathcal{F}_{s^-}] &= dA(s) \\ V\{dN(s)|\mathcal{F}_{s^-}\} &= dA(s)\{1 - dA(s)\} \approx dA(s) \end{aligned}$$

  Integrating both sides of the last equality, we have $\Longrightarrow \langle M\rangle(t) = A(t)$.

- Claim: $M^2 - \langle M\rangle$ is an $\mathcal{F}$ martingale.

  Proof: We have already demonstrated that M^2 is a sub-martingale. Using the Doob-Meyer decomposition, we determine the compensator of M^2 through its increments.

i.e., we need to show that $E[dM^2(t)|\mathcal{F}_{t^-}] = d\langle M\rangle(t)$

$$\begin{aligned}
d\langle M\rangle(t) &\equiv V\{dM(t)|\mathcal{F}_{t^-}\} \\
&= E[\{dM(t)\}^2|\mathcal{F}_{t^-}] \\
&= E[\{M(t^- + dt) - M(t^-)\}^2|\mathcal{F}_{t^-}] \\
&= E[M^2(t^- + dt) + M^2(t^-) \\
&\quad -2M(t^- + dt)M(t^-)|\mathcal{F}(t^-)] \\
&= E[M^2(t^- + dt)|\mathcal{F}_{t^-}] + M^2(t^-) - 2M^2(t^-) \\
&= E[M^2(t^- + dt)|\mathcal{F}_{t^-}] - M^2(t^-) \\
&= E[M^2(t^- + dt) - M^2(t^-)|\mathcal{F}_{t^-}] \\
&\equiv E[dM^2(t)|\mathcal{F}_{t^-}]
\end{aligned}$$

Therefore, upon integrating, the compensator for M^2 is given by $\langle M\rangle$, meaning that $M^2 - \langle M\rangle$ is an $\mathcal{F}$ martingale.

- Thus, we have proven that

$$E[M^2(t)] = E[\langle M\rangle(t)]$$

and hence that

$$V\{M(t)\} = E[A(t)]$$

Predictable Covariation Process

- Consider a collection of n martingales w.r.t. to a common filtration $\mathcal{F}_t$.
  - Then $M(t) = \sum_{i=1}^{n} M_i(t)$ is a martingale w.r.t. $\mathcal{F}_t$.
  - The variance process for M depends on the covariation among these martingales.
  - The *predictable covariation process*, $\langle M_i, M_j\rangle$ is a compensator of M_iM_j.

- Theorem: Let M_1 and M_2 be square integrable $\mathcal{F}$ martingales. There exists a unique right-continuous predictable process, $\langle M_1, M_2\rangle$ such that $\langle M_1, M_2\rangle(0) = 0$, $E[\langle M_1, M_2\rangle(t)] < \infty$ for all t and $M_1M_2 - \langle M_1, M_2\rangle$ is an $\mathcal{F}$ martingale.

- The process $\langle M_1, M_2\rangle$ is known as the *predictable covariation process*

$$\begin{aligned}
\langle M_1, M_2\rangle(t) &= \int_0^t d\langle M_1, M_2\rangle(s) \\
d\langle M_1, M_2\rangle(s) &= \text{cov}\{dM_1(s), dM_2(s)|\mathcal{F}_{s^-}\} \\
\text{cov}\{M_1(t), M_2(t)\} &= E[\langle M_1, M_2\rangle(t)]
\end{aligned}$$

  – bivariate analog of $\langle M\rangle$, predictable variation process

- Claim: If the counting processes N_i $(i = 1, 2)$ never jump simultaneously, then the $\mathcal{F}$ martingales $M_i = N_i - A_i$ $(i = 1, 2)$ are uncorrelated.

  Proof:

$$\begin{aligned}
\text{cov}\{M_1(t), M_2(t)\} &= E[\langle M_1, M_2\rangle(t)] \\
\langle M_1, M_2\rangle(t) &= \int_0^t d\langle M_1, M_2\rangle(s) \\
d\langle M_1, M_2\rangle(s) &= \text{cov}\{dM_1(s), dM_2(s)|\mathcal{F}_{s^-}\} \\
&= E[dM_1(s)dM_2(s)|\mathcal{F}_{s^-}] \\
&= \text{cov}[dN_1(s)dN_2(s)|\mathcal{F}_{s^-}] \\
&= E[dN_1(s)dN_2(s)|\mathcal{F}_{s^-}] \\
&\quad -E[dN_1(s)|\mathcal{F}_{s^-}]E[dN_2(s)|\mathcal{F}_{s^-}] \\
&= -dA_1(s)dA_2(s) \\
&= O(ds^2)
\end{aligned}$$

  Therefore $\langle M_1, M_2\rangle(t) = 0$

- If $\langle M_i, M_j\rangle(t) = 0\ \forall t$, then M_i and M_j are called *orthogonal* martingales.

Martingale Transformation

Claim: If M is an $\mathcal{F}$ martingale and H is a bounded predictable process, and $Z = \int HdM$, then $\langle Z\rangle(t) = \int H^2 dA$.

- Proof:

$$\begin{aligned} \langle Z\rangle(t) &= \int_0^t d\langle Z\rangle(s) \\ d\langle Z\rangle(s) &= V\{dZ(s)|\mathcal{F}_{s^-}\} \\ &= V\{H(s)dM(s)|\mathcal{F}_{s^-}\} \\ &= H^2(s)V\{dM(s)|\mathcal{F}_{s^-}\} \\ &= H^2(s)d\langle M\rangle(s) \\ &= H^2(s)dA(s) \end{aligned}$$

with the desired result obtained upon integration

- Theorem: Let N_i $(i = 1, 2)$ be a counting process and let H_i $(i = 1, 2)$ be bounded $\mathcal{F}$ predictable processes. For $\mathcal{F}$ martingales $M_i = N_i - A_i$ and $Z_i = \int H_i dM_i$,

$$\begin{aligned} \langle Z_1, Z_2\rangle &= \int H_1 H_2 d\langle M_1, M_2\rangle \\ E[Z_1 Z_2] &= E\left[\int H_1 H_2 d\langle M_1, M_2\rangle\right] \end{aligned}$$

- Corresponding to the preceeding theorem, we have the following generalization ...

- Corollary: Let $U_k = \sum_{i=1}^n \int H_{ik} dM_i$ for $k = 1, 2$, where $M_i = N_i - A_i$ is an $\mathcal{F}$ martingale $(i = 1, \ldots, n)$ and H_{ik} are bounded predictable processes.

(i) U_k is a martingale

(ii) $E[U_k(t)] = 0$

(iii) covariation,

$$E[U_k(t)U_\ell(t)] = E\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\int_0^t H_{ik}(s)H_{j\ell}(s)d\langle M_i, M_j\rangle(s)\right]$$

Multivariate Counting Process

- Definition: A n-variate process, $\mathbf{N} = [N_1, \ldots, N_n]$, called a *multivariate counting process* if N_i is a counting process for $i = 1, \ldots, n$ and N_i and N_j cannot jump at the same time.

- Theorem: Let $\mathbf{N}$ be a multivariate counting process, where A_i is a continuous compensator of N_i ($i = 1, \ldots, n$), $U_k = \sum_{i=1}^{n} \int H_{ik} dM_i$ ($k = 1, \ldots, K$), with H_{ik} being bounded and predictable.

  (i) U_k is a martingale

  (ii) $E[U_k(t)] = 0$

  (iii)

$$\mathrm{cov}\{U_k(s)U_\ell(t)\} = \sum_{\mathbf{i=1}}^{\mathbf{n}} \int_0^{s\wedge t} E[H_{ik}(u)H_{i\ell}(u)dA_i(u)]$$

$$Var[U_k(t)] = \sum_{i=1}^{n} \int_0^{t} E\{H_{ik}^2(u)dA_i(u)\}$$

$$Cov\{U_k(s), U_k(t)\} = \sum_{i=1}^{n} \int_0^{s\wedge t} E\{H_{ik}^2(u)dA_i(u)\}$$

Proof: $s \geq t$

$$Cov\{U_k(s), U_l(t)\} = E\{U_k(s)U_l(t)\}$$

$$= E[E\{U_k(s)U_l(t)|\mathcal{F}_t\}] = E[E\{U_k(s)|\mathcal{F}_t\}U_l(t)] = E[U_k(t)U_l(t)]$$

$$= E\sum_{i=1}^{n}\sum_{j=1}^{n}\int_0^t H_{ik}(u)H_{jl}(u)\; d < M_i, M_j > (u)$$

$$= \sum_{i=1}^{n}\int_0^t E\{H_{ik}(u)H_{il}(u)dA_i(u).\}$$

Rebelledo's Central Limit Theorem

- Theorem (Martingale Central Limit Theorem): Let $\mathbf{N}$ be a multivariate counting process, where A_i is a continuous compensator of N_i ($i = 1, \ldots, n$), $U_k = \sum_{i=1}^n \int H_{ik} dM_i$ ($k = 1, \ldots, K$), with H_{ik} being bounded and predictable. In addition to these conditions, **if** the following hold as $n \to \infty$,

$$\langle U_k, U_\ell\rangle(t) = \sum_{i=1}^{n}\int_0^t H_{ik}(s)H_{i\ell}(s)dA_i(s) \xrightarrow{p} \sigma_{kl}(t)$$

$$\langle U_{k,\epsilon}\rangle(t) = \sum_{i=1}^{n}\int_0^t H_{ik}^2(s)I\{|H_{ik}(s)| > \epsilon\}dA_i(s) \xrightarrow{p} 0$$

for all $\epsilon > 0$, then $\mathbf{U} \xrightarrow{D} \mathbf{W}$ where $\mathbf{W} = (W_1, \ldots, W_K)^T$ is a K-variate Gaussian process with $W_k(0) = 0$, $E[W_k] = 0$ and covariance function $E[W_k(s)W_\ell(t)] = \sigma_{kl}(s \wedge t)$.

- $U_{k,\varepsilon}(t) = \sum_{i=1}^n \int_0^t H_{ik}(u)I\{|H_{ik}(u)| > \varepsilon\}dM_i(u)$ contains all the jumps of U_k larger than ε (in absolute value);
- $\sigma_{kl}(t)$ is a deterministic continuous function.
- Condition (1) requires the H_{ik} be appropriately standardized.

- Condition (2) is a Lindeberg-type condition which guaranttees that the influence of any single process is negligible in the limit.
- the condition $\langle U_{k,\epsilon}, U_{\ell,\epsilon}\rangle(t) \to 0$ implies that the paths of U_k are continuous

Summary of Basic Results

- $E\{M(t+s)|\mathcal{F}_t\} = M(t) \Rightarrow M(t)$ is martingale
- $M(t) \equiv N(t) - \int_0^t Y(u)\lambda(u)du$ is zero-mean martingale
- $M^2(t) - < M, M > (t)$ is zero-mean martingale
- $\text{Var}\{M(t)\} = E < M, M > (t)$
- $< M, M > (t) = \int_0^t Y(u)\lambda(u)du$
- $Z(t) \equiv \int_0^t H(u)dM(u)$ is zero-mean martingale
- $$\begin{aligned} < Z, Z > (t) &= \int_0^t H^2(u)d < M, M > (u) \\ &= \int_0^t H^2(u)Y(u)\lambda(u)du \end{aligned}$$
- $\text{Var}\{Z(t)\} = E\{\int_0^t H^2(u)Y(u)\lambda(u)du\}$
- $U(t) \equiv \sum_{i=1}^n \int_0^t H_i(u)dM_i(u)$ is zero-mean martingale
- $< U, U > (t) = \sum_{i=1}^n \int_0^t H_i^2(u)Y_i(u)\lambda_i(u)du$
- $\text{Var}\{U(t)\} = \sum_{i=1}^n \int_0^t E\{H_i^2(u)Y_i(u)\lambda_i(u)du\}$

Gaussian (Martingale) Process

- Consider a sequence of counting processes, $N_n(t)$, with corresponding intensity processes, $A_n(t)$, defined with respect to the sequence of filtrations, $\mathcal{F}_n(t)$
- Let H_n be a sequence of $\mathcal{F}_n$ predictable processes
- Define the $\mathcal{F}_n$ martingale process: $M_n = N_n - A_n$ and martingale transform, $Z_n = \int H_n dM_n$
- Martingale CLT (recall): If $Z_n = \int H_n dM_n$ satisfies the following conditions,

  (i) for all $t \in (0, \tau_*]$,

  $$\langle Z_n \rangle(t) \;=\; \int_0^t H_n^2(s) dA_n(s) \xrightarrow{p} \int_0^t g^2(s) ds < \infty$$

  (ii) for any $\mathcal{E} > 0$,

  $$\int_0^{\tau_*} H_n^2(s) I\{|H_n(s)| > \mathcal{E}\} dA_n(s) \;\xrightarrow{p}\; 0$$

  then Z_n converges weakly on $(0, \tau_*]$ to a zero-mean Gaussian process with independent increments and covariance function $\sigma_Z(s,t) = \int_0^{s \wedge t} g^2(u) du$

Stochastic Processes: Convergence Results

- We now consider some properties of stochastic processes.
- Various modes of convergence of Z_n can be evaluated
  - *convergence in distribution* refers to convergence of Z_n to Z at a specific time point
  - *weak convergence* of Z_n to Z refers to the behavior of Z_n as a process over all t

Stochastic Processes: Convergence in Distribution

- To state that $Z_n(t) \xrightarrow{D} Z(t)$ implies that
  - the marginal distribution of $Z_n(t_1)$ converges to the marginal distribution of $Z(t_1)$; e.g.,

$$P\{a \leq Z_n(t_1) \leq b\} \;\longrightarrow\; P\{a \leq Z(t_1) \leq b\}$$

  - for any finite set of times, $\{t_1, \ldots, t_m\}$, the joint distribution of $\{Z_n(t_1), \ldots, Z_n(t_m)\}$ converges to the joint distribution of $\{Z(t_1), \ldots, Z(t_m)\}$; e.g.,

$$\begin{aligned} & P\{a_1 \leq Z_n(t_1) \leq b_1, \ldots, a_m \leq Z_n(t_m) \leq b_m\} \\ \longrightarrow\; & P\{a_1 \leq Z(t_1) \leq b_1, \ldots, a_m \leq Z(t_m) \leq b_m\} \end{aligned}$$

Stochastic Processes: Weak Convergence

- Weak convergence of Z_n to Z, $Z_n \xrightarrow{\mathcal{W}} Z$, allows us to make statements about the entire process over $(0, \tau_*]$; e.g.,

$$P\left\{\sup_{t\in(0,\tau_*]} |Z_n(t)| \leq a\right\} \;\longrightarrow\; P\left\{\sup_{t\in(0,\tau_*]} |Z(t)| \leq a\right\}$$

- Note: the event being considered is that Z_n stays below a specific bound, a, for the entire interval of interest
- To illustrate the difference between finite and infinite-dimensional results,
  - $Z_n \xrightarrow{D} Z$ leads to sets of point-wise confidence intervals
  - $Z_n \xrightarrow{W} Z$ would be the basis of a confidence band

Properties of Gaussian Processes

- If $Z(t)$, is a zero-mean Gaussian process with independent increments and variance function $\int_0^t g^2(s)ds$, then:
  - $Z(t_1) \sim N\left(0, \int_0^{t_1} g^2(s)ds\right)$
  - for $t_1 < t_2 < \ldots < t_m$, $\{Z(t_1), \ldots, Z(t_m)\}$ follows a multivariate Normal (MVN) distribution with:

    1. mean $\mathbf{0}_m$
    2. $\text{cov}\{Z(t_j), [Z(t_{j+1}) - Z(t_j)]\} = 0$
    3. $\text{cov}\{Z(t_j), Z(t_k)\} = V\{Z(t_j)\}$ for $j < k$

- Claim: $V\{Z(t_2) - Z(t_1)\} = \int_{t_1}^{t_2} g^2(s)ds$.

  Proof:

$$\begin{aligned} V\{Z(t_2)\} = \int_0^{t_2} g^2(s)ds &= V\{Z(t_1) + [Z(t_2) - Z(t_1)]\} \\ &= V\{Z(t_1)\} + V\{Z(t_2) - Z(t_1)\} \\ &\quad +2\ \text{cov}\{Z(t_1), Z(t_2) - Z(t_1)\} \\ \int_0^{t_2} g^2(s)ds &= \int_0^{t_1} g^2(s)ds + V\{Z(t_2) - Z(t_1)\} \end{aligned}$$

  such that

$$V\{Z(t_2) - Z(t_1)\} = \int_{t_1}^{t_2} g^2(s)ds.$$

- Set $\Delta Z_j = Z(t_j) - Z(t_{j-1})$ for $j = 1, \ldots, m$, with $t_0 \equiv 0$ and $t_m \equiv t$
  - we have
$$\Delta Z_j \ \sim \ N\left(0, \int_{t_{j-1}}^{t_j} g^2(s)ds\right)$$
  - then, $\{Z(t_1), \ldots, Z(t_m)\}$ has the same joint distribution as $\{\Delta Z_1, \Delta Z_1 + \Delta Z_2, \ldots, \Delta Z_1 + \ldots + \Delta Z_m\}$
  - $\Delta Z_1 \perp \Delta Z_2 \perp \ldots \perp \Delta Z_m$

- Combining these results, we have the following:
  - set $\mathbf{t} = [t_1, \ldots, t_m]^T$
  - set $Z_n(\mathbf{t}) = [Z_n(t_1), \ldots, Z_n(t_m)]^T$
  - $Z_n(\mathbf{t}) \xrightarrow{D} Z(\mathbf{t}) \sim MVN(\mathbf{0}_m, V\{Z(\mathbf{t})\})$, where
$$V\{Z(\mathbf{t})\} \ = \ \begin{bmatrix} \sigma^2(t_1) & \sigma^2(t_1) & \ldots & \ldots & \sigma^2(t_1) \\ \sigma^2(t_1) & \sigma^2(t_2) & \ldots & \ldots & \sigma^2(t_2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sigma^2(t_1) & \sigma^2(t_2) & \ldots & \ldots & \sigma^2(t_m) \end{bmatrix}$$

- Other aspects of the Gaussian process,
  - paths are everywhere continuous
  - nowhere differentiable

Brownian Motion

- One of the most studied stochastic processes is Brownian motion (Wiener process), which has the following characteristics:
  - $W(0) = 0$ and $E[W(t)] = 0$
  - $V\{W(t)\} = t$
  - independent increments
  - stationary increments
  - Gaussian process; $W(t) \sim N(0, t)$
  - continuous sample paths: $t \longmapsto W(t)$ is continuous w.p. 1
- Regarding the independent increment structure,
  - set $0 \equiv t_0 < t_1 < \ldots < t_m \equiv t$
  - set $\Delta W_j = W(t_j) - W(t_{j-1})$, for $j = 1, \ldots, m$
  - we then have $\Delta W_1 \perp \Delta W_2 \perp \ldots \perp \Delta W_m$
- Regarding scaling property: for a constant, c,

$$\begin{aligned} cW(t) &\stackrel{D}{=} W(c^2 t) \\ W(ct) &\stackrel{D}{=} \sqrt{c}W(t) \sim N(0, ct) \end{aligned}$$

- Variance and covariance: for $s < t$,

$$\begin{aligned} \text{cov}\{W(s), W(t)\} &= E[W(s)W(t)] \\ &= E[W(s)\{W(s) + W(t) - W(s)\}] \\ &= E[W^2(s)] + E[W(s)\{W(t) - W(s)\}] \\ &= V\{W(s)\} \\ &= s \end{aligned}$$

  - generally, $\text{cov}\{W(r), W(s)\} = r \wedge s$
- Stationary increments: distribution of $\{W(t) - W(s)\}$ depends only on $(t - s)$

- For the increments of the Wiener process, consider $r < s$,

$$\begin{aligned} E[W(s) - W(r)] &= 0 \\ V\{W(s) - W(r)\} &= s - r \end{aligned}$$

- Now, set $r = t^- + dt$ and $s = t^-$ to obtain

$$\begin{aligned} E[dW(t)] &= 0 \\ V\{dW(t)\} &= dt \end{aligned}$$

Time-Transformed Brownian Motion

- $Z(t)$ (zero-mean Gaussian process with independent increments and variance function $\int_0^t g^2(s)ds$) has the same distribution as a time-transformed Brownian motion

$$Z(t) \stackrel{D}{=} W\left(\int_0^t g^2(s)ds\right) \equiv \widetilde{W}(t)$$

- i.e., $\widetilde{W}(t)$ is still a Brownian motion, but with the time scale re-defined
- We can express the time-transformed process in terms of the original process as follows,

$$\widetilde{W}(t) = \int_0^t g(s)dW(s)$$

- Mean of time-transformed Wiener process,

$$\begin{aligned} E[\widetilde{W}(t)] &= E\left[\int_0^t g(s)dW(s)\right] \\ &= \int_0^t E[g(s)dW(s)] \\ &= \int_0^t g(s)E[dW(s)] \\ &= 0 \end{aligned}$$

$$
\begin{aligned}
V\{\widetilde{W}(t)\} &= E\left[\left\{\int_0^t g(s)dW(s)\right\}^2\right] \\
&= E\left[\int_0^t \int_0^t g(u)dW(u)g(s)dW(s)\right] \\
&= \int_0^t \int_0^t g(u)g(s)E[dW(u)dW(s)] \\
&= \int_0^t g^2(s)E[\{dW(s)\}^2] \\
&= \int_0^t g^2(s)ds
\end{aligned}
$$

since $E[dW(u)dW(s)] = 0$ for $s \neq u$

.