Basic Theory and Some Applications of Martingales

*ACM SIGMETRICS/Performance 2012 Tutorial*

Richard A. Hayden

rh@doc.ic.ac.uk

Department of Computing
Imperial College London

June 11th 2012
What is a martingale?

According to Oxford English Dictionary:

- **Horse riding**: A strap or arrangement of straps fastened at one end to the noseband, bit, or reins of a horse and at the other to its girth, in order to prevent it from rearing or throwing its head back, or to strengthen the action of the bit;

- **Nautical**: A stay which holds down the jib boom of a square-rigged ship, running from the boom to the dolphin-striker;

- **Gambling**: Any of various gambling systems in which a losing player repeatedly doubles or otherwise increases a stake such that any win would cover losses accrued from preceding bets.
Martingales as gambling models of *fair games*

- $X_n$ is the amount of money gambler has at time $n \geq 0$ and $X_0 = w$ is their initial wealth;
- At each time $n$, gambler wagers some amount $W_n \leq X_n$ of his money ($W_n$ may be random and can depend on the past);
- Round decided by a fair coin toss, independent of previous rounds: w.p. $1/2$, loses the wager and w.p. $1/2$, wins it;
- Game stops if gambler goes broke or they can choose to stop playing at any time (‘winning’ whatever money they’re left with).

- In terms of i.i.d. random variables $P_n$ that equal 1 w.p. $1/2$ and $-1$ w.p. $1/2$, we can write:

$$X_n = w + \sum_{m=1}^{n} P_m W_m$$

So then:

$$\mathbb{E}[X_{n+1}|(X_m)_{m \leq n}] = X_n + \mathbb{E}[P_{n+1}W_{n+1}|(X_m)_{m \leq n}]$$

$$= X_n + \mathbb{E}[P_{n+1}]\mathbb{E}[W_{n+1}|(X_m)_{m \leq n}] = X_n$$

**Our first martingale**
What is the best strategy?

A strategy here consists of:

- The amount to wager at each step, the $W_n \leq X_n$;
- The (random) ‘stopping time’ $T$ at which the gambler chooses to stop playing (if they haven’t already gone broke by that time), for example, could have $T := \min\{n \geq 0 : X_n \geq d\}$ for some $d > 0$.

Writing $S := \min\{n \geq 0 : X_n = 0\}$, we are then interested in the quantity $\mathbb{E}[X_{T \wedge S}]$, that is, the average return of the strategy when the game ends. We assume also that $\mathbb{E}[T \wedge S] < \infty$.

Write $R := T \wedge S$ and define $X^R_n := 1_{\{R \leq n\}} X_R + 1_{\{R \geq n\}} X_n$ as the stopped process, which also enjoys a martingale property:

$$
\mathbb{E}[X^R_{n+1} | (X_m)_{m \leq n}] = \mathbb{E}[1_{\{R \leq n\}} X_R | (X_m)_{m \leq n}] + \mathbb{E}[1_{\{R > n\}} X_{n+1} | (X_m)_{m \leq n}]
$$

$$
= \mathbb{E}[1_{\{R \leq n\}} X_R | (X_m)_{m \leq n}] + 1_{\{R > n\}} \mathbb{E}[X_{n+1} | (X_m)_{m \leq n}]
$$

$$
= 1_{\{R \leq n\}} X_R + 1_{\{R > n\}} X_n = X^R_n
$$

Thus, in particular, $\mathbb{E}[X^R_n] = w$ for all $n \geq 0$ and, furthermore, note that $X^R_n \to X_R$ w.p. 1, so, taking expectations,\(^1\) we have $w \to \mathbb{E}[X_R]$, that is, $\mathbb{E}[X_R] = w$, independent of the strategy.

\(^1\) Formally, since $|X^R_n| \leq 2^R w$, we may apply the dominated convergence theorem.
Definition: Martingale

A sequence of $\mathbb{R}$-valued random variables $(X_n)_{n \geq 0}$ is a martingale with respect to a filtration $(\mathcal{F}_n)_{n \geq 0}$ if, for all $n \geq 0$:

- $\mathbb{E}[|X_n|] < \infty$;
- $\mathbb{E}[X_{n+1} | \mathcal{F}_n] = X_n$ w.p. 1 ($\iff \mathbb{E}[X_{n+k} | \mathcal{F}_n] = X_n$ w.p. 1, $k > 0$).

In continuous time, $\mathbb{E}[|X_t|] < \infty$ and $\mathbb{E}[X_{s+t} | \mathcal{F}_s] = X_s$ w.p. 1 for all $s, t \in \mathbb{R}_+$.

Theorem: Optional stopping

Let $(X_t)_{t \geq 0}$ be a discrete- or continuous-time martingale with respect to $(\mathcal{F}_t)_{t \geq 0}$ and suppose that $S \leq T$ are $\mathbb{R}_+$-valued random variables such that:

- $S, T$ are $\mathcal{F}_t$-stopping times, that is, it is possible to say whether $T \leq t$ and $S \leq t$ given the information up to time $t$, that is, given $\mathcal{F}_t$;
- $T < \infty$ w.p. 1;
- For all $t \geq 0$, $|X_t^T| < C$ and $\mathbb{E}[C] < \infty$; or $T \leq K < \infty$ w.p. 1.

Then $\mathbb{E}[X_T | \mathcal{F}_S] = X_S$ w.p. 1.
Martingales and the optional stopping theorem

**Theorem: Optional stopping**

Let \((X_t)_{t \geq 0}\) be a discrete- or continuous-time martingale wrt. \((\mathcal{F}_t)_{t \geq 0}\) and suppose that \(S \leq T\) are \(\mathbb{R}_+\)-valued random variables such that:

- \(S, T\) are \(\mathcal{F}_t\)-stopping times, that is, it is possible to say whether \(T \leq t\) and \(S \leq t\) given the information up to time \(t\), that is, given \(\mathcal{F}_t\);
- \(T < \infty\) w.p. 1;
- For all \(t \geq 0\), \(|X_t^T| < C\) and \(\mathbb{E}[C] < \infty\); or \(T \leq K < \infty\) w.p. 1.

Then \(\mathbb{E}[X_T | \mathcal{F}_S] = X_S\) w.p. 1

Proof for the case \(S, T\) bounded above by constant \(N\):

\[
\mathbb{E}[X_T - X_S | \mathcal{F}_S] = \sum_{j=0}^{N} \mathbb{E}[1_{\{S=j\}}(X_T - X_j) | \mathcal{F}_j]
\]

\[
= \sum_{j=0}^{N} \sum_{k=j}^{N} 1_{\{S=j\}} \mathbb{E}[(X_{k+1}^T - X_k^T) | \mathcal{F}_j] \quad \text{since } X_{N+1}^T = X_T, \ X_j^T = X_j
\]

\[
= \sum_{j=0}^{N} \sum_{k=j}^{N} 1_{\{S=j\}} \mathbb{E}[\mathbb{E}[(X_{k+1}^T - X_k^T) | \mathcal{F}_k] | \mathcal{F}_j] = 0
\]

=0 by martingale property
Martingales and the optional stopping theorem

**Theorem: Optional stopping**

Let \( (X_t)_{t \geq 0} \) be a discrete- or continuous-time martingale wrt. \( (\mathcal{F}_t)_{t \geq 0} \) and suppose that \( S \leq T \) are \( \mathbb{R}_+ \)-valued random variables such that:

- \( S, T \) are \( \mathcal{F}_t \)-stopping times, that is, it is possible to say whether \( T \leq t \) and \( S \leq t \) given the information up to time \( t \), that is, given \( \mathcal{F}_t \);
- \( T < \infty \) w.p. 1;
- For all \( t \geq 0 \), \( |X_t^T| < C \) and \( \mathbb{E}[C] < \infty \); or \( T \leq K < \infty \) w.p. 1.

Then \( \mathbb{E}[X_T|\mathcal{F}_S] = X_S \) w.p. 1

**Proof sketch for continuous time:**

- Define \( D_n := \{k2^{-n} : k \in \mathbb{Z}_+\} \), \( T_n := \inf\{q \in D_n : q > T\} \) and \( S_n := \inf\{q \in D_n : q > S\} \) so \( T_n \downarrow T \) and \( S_n \downarrow S \)
- Consider the discrete-time martingales \( (X^n_k)_{k \geq 0} \), \( X^n_k := X_{k2^{-n}} \) with filtrations \( (\mathcal{F}_k^n)_{k \geq 0} \), \( \mathcal{F}_k^n := \mathcal{F}_{k2^{-n}} \)
- Discrete-time optional stopping gives \( \mathbb{E}[X^n_{T_n}|\mathcal{F}^n_{S_n}] = X^n_{S_n} \)
- Result then follows from an \( n \to \infty \) argument
- **Subtle point:** Also require the right-continuity assumptions: \( X_t = \lim_{s \downarrow t} X_s \) and \( \mathcal{F}_t = \bigcap_{u > t} \mathcal{F}_u \) for all \( t \geq 0 \). In continuous time, we consider only càdlàg martingales, i.e. living in function space \( D \).

\( D \) is the space of all \( \text{continue à droite, limite à gauche} \) (right-continuous with left limits) functions.
Martingales and the optional stopping theorem

**Theorem: Optional stopping**

Let \((X_t)_{t \geq 0}\) be a discrete- or continuous-time martingale wrt. \((\mathcal{F}_t)_{t \geq 0}\) and suppose that \(S \leq T\) are \(\mathbb{R}_+\)-valued random variables such that:

- \(S, T\) are \(\mathcal{F}_t\)-stopping times, that is, it is possible to say whether \(T \leq t\) and \(S \leq t\) given the information up to time \(t\), that is, given \(\mathcal{F}_t\);
- \(T < \infty\) w.p. 1;
- For all \(t \geq 0\), \(|X_t^T| < C\) and \(\mathbb{E}[C] < \infty\); or \(T \leq K < \infty\) w.p. 1.

Then \(\mathbb{E}[X_T|\mathcal{F}_S] = X_S\) w.p. 1

- Martingale property is crucial, why is simply \(\mathbb{E}[X_t] = \mathbb{E}[X_0]\) for all \(t > 0\) not enough?
  - Let \(X_0\) be Bernoulli(0.5) and then define \(X_{n+1} = 1 - X_n\) for \(n > 0\). Then if \(T := \min\{n \geq 0 : X_n = 1\}\), \(\mathbb{E}[X_T] = 1\).
  - Consider the betting process again and assume now that we can go into arbitrary debt with a maximum wager per turn of \(W\). Consider stopping at \(T := \min\{n \geq 0 : X_n \geq d\}\) for \(d > w\). But then \(\mathbb{E}[X_T] \geq d > w\), what’s going on?
    - \(X_n\) is still a martingale and \(T\) is still a stopping time. Therefore one of the second two conditions must be violated, at least we must have \(\mathbb{E}[T] = \infty\) (i.e. you are likely to get lost in debt and remain there for long periods of time).²

² Your bank or loan shark is unlikely to support this as a gambling strategy for too long.
Application: the least variable PH distribution is Erlang

Theorem: Erlang distribution minimises the coefficient of variation over \( n \) phases

Let \((X_t)_{t \geq 0}\) be a CTMC on \(\{0, \ldots, n\}\) with \(X_0 = k\) and \(T_0 := \inf\{t \geq 0 : X_t = 0\}\) its absorption time. Assume further (WLOG) that the states are ordered such that the expected time to absorption from each is non-decreasing. Then:

\[
\frac{\sqrt{\text{Var}[T_0]}}{\mathbb{E}[T_0]} \geq \frac{1}{\sqrt{n}}
\]

- Define \(Y_t := h(X_t) + \min(t, T_0) - h(X_0)\) where \(h(i)\) is the expected time to absorption starting in state \(i\). Claim: \((Y_t)_{t \geq 0}\) is a martingale wrt. \((X_t)_{t \geq 0}\)

- To see this, note that
  \[
  \mathbb{E}[Y_{s+t} - Y_s | (X_u)_{u \leq s}] = \mathbb{E}[Y_{s+t} - Y_s | X_s] = \sum_{i \neq 0} 1\{X_s = i\} \mathbb{E}[Y_{s+t} - Y_s | X_s = i]
  \]
- Also, \(\mathbb{E}[Y_{s+t} - Y_s | X_s = i] = \mathbb{E}[h(X_{s+t}) - \min(t, T_0 - s) - h(i) | X_s = i]\)
- This is the same as \(\mathbb{E}[Y_t | X_0 = i]\) since \((Y_t)_{t \geq 0}\) is Markov wrt. \((X_t)_{t \geq 0}\) with infinitesimal generator:
  \[
  \lim_{\delta t \to 0} (1/\delta t) \mathbb{E}[Y_{t+\delta t} - Y_t | X_t = i] = \delta t \sum_{j \neq i} \lambda_i p_{ij} h(j) + (1 - \delta t \lambda_i) h(i) + \delta t - h(i)
  \]
- Also, \(h(i) = 1/\lambda_i + \sum_{j \neq i} p_{ij} h(j)\), so the generator is identically zero and the claim is proved

Application: the least variable PH distribution is Erlang

Theorem: Erlang distribution minimises the coefficient of variation over \( n \) phases

Let \((X_t)_{t \geq 0}\) be a CTMC on \(\{0, \ldots, n\}\) with \(X_0 = k\) and \(T_0 := \inf\{t \geq 0 : X_t = 0\}\) its absorption time. Assume further (WLoG) that the states are ordered such that the expected time to absorption from each is non-decreasing. Then:

\[
\sqrt{\frac{\text{Var}[T_0]}{\mathbb{E}[T_0]}} \geq \frac{1}{\sqrt{n}}
\]

- \(Y_t := h(X_t) + \min(t, T_0) - h(X_0)\) is a martingale wrt. \((X_t)_{t \geq 0}\)
- We want to compute \(\text{Var}[T_0] = \text{Var}[Y_{T_0}] = \mathbb{E}[Y_{T_0}^2]\)
- For \(N \geq 0\), define refining partition of \(\mathbb{R}_+\): \(0 = t_0^N < t_1^N < \ldots\), then:\[^3\]
  \[
  \mathbb{E}[Y_{T_0}^2] = \lim_{N \to \infty} \mathbb{E}\left[\sum_{k=1}^{N} 1\{t_{k-1}^N < T_0\} (Y_{t_k^N}^2 - Y_{t_{k-1}^N}^2)\right] = \lim_{N \to \infty} \mathbb{E}\left[\sum_{k=1}^{N} 1\{t_{k-1}^N < T_0\} \mathbb{E}[Y_{t_k^N}^2 - Y_{t_{k-1}^N}^2 | (X_s)_{s \leq t_{k-1}^N}]\right]
  \]
- Note that: \(Y_{t_k^N}^2 - Y_{t_{k-1}^N}^2 = (Y_{t_k^N} - Y_{t_{k-1}^N})^2 + 2Y_{t_{k-1}^N} (Y_{t_k^N} - Y_{t_{k-1}^N})\), so
  \[
  \mathbb{E}[Y_{t_k^N}^2 - Y_{t_{k-1}^N}^2 | (X_s)_{s \leq t_{k-1}^N}] = \mathbb{E}[(Y_{t_k^N} - Y_{t_{k-1}^N})^2 | (X_s)_{s \leq t_{k-1}^N}]
  \]
- So: \(\mathbb{E}[Y_{T_0}^2] = \mathbb{E}\left[\lim_{N \to \infty} \sum_{k=1}^{N} 1\{t_{k-1}^N < T_0\} (Y_{t_k^N} - Y_{t_{k-1}^N})^2\right] = \mathbb{E}\left[\sum_{s \leq T_0} (h(X_s) - h(X_{s-}))^2\right] = \mathbb{E}\left[\sum_{s} (h(X_s) - h(X_{s-}))^2\right]
  \]


Application: the least variable PH distribution is Erlang

Theorem: Erlang distribution minimises the coefficient of variation over \( n \) phases

Let \((X_t)_{t \geq 0}\) be a CTMC on \( \{0, \ldots, n\} \) with \( X_0 = k \) and \( T_0 := \inf\{ t \geq 0 : X_t = 0 \} \) its absorption time. Assume further (WLoG) that the states are ordered such that the expected time to absorption from each is non-decreasing. Then:

\[
\frac{\sqrt{\text{Var}[T_0]}}{\mathbb{E}[T_0]} \geq \frac{1}{\sqrt{n}}
\]

- \( \text{Var}[T_0] = \mathbb{E} \left[ \sum_s (h(X_s) - h(X_{s-}))^2 \right] \)
- Recall that the states are ordered such that \( h(i) \) is non-decreasing
- Then we claim that for any absorbing path \( k = i_0, i_1, \ldots, i_L = 0 \),
  \[
  \sum_{u=1}^{L} [h(i_u) - h(i_{u-1})]^2 \geq \sum_{i=1}^{k} [h(i) - h(i - 1)]^2
  \]
- To see this note that:
  \[
  \left[ \sum_{i \in I_u} [h(i) - h(i - 1)] \right]^2 = [h(i_u) - h(i_{u-1})]^2
  \]
  where: \( I_u := \{ j : \min(i_u, i_{u-1}) < j \leq \max(i_u, i_{u-1}) \} \). Also the terms of the sum are non-negative, so
  \[
  [h(i_u) - h(i_{u-1})]^2 \geq \sum_{i \in I_u} [h(i) - h(i - 1)]^2
  \]
- Then since the absorbing path must hit \( k \) and 0, \( \bigcup_{u=1}^{L} I_u \) must contain \( \{1, \ldots, k\} \) and the claim follows.

Application: the least variable PH distribution is Erlang

**Theorem:** Erlang distribution minimises the coefficient of variation over \( n \) phases

Let \((X_t)_{t \geq 0}\) be a CTMC on \( \{0, \ldots, n\} \) with \( X_0 = k \) and \( T_0 := \inf\{ t \geq 0 : X_t = 0 \} \) its absorption time. Assume further (WLoG) that the states are ordered such that the expected time to absorption from each is non-decreasing. Then:

\[
\frac{\sqrt{\text{Var}[T_0]}}{\mathbb{E}[T_0]} \geq \frac{1}{\sqrt{n}}
\]

- \( \text{Var}[T_0] \geq \sum_{i=1}^{k} [h(i) - h(i-1)]^2 \)
- Then, by Cauchy–Schwarz:
  \[
  \text{Var}[T_0] \geq \left[ \sum_{i=1}^{k} [h(i) - h(i-1)] \right]^2 \frac{1}{k}
  \]
- Finally, \( \text{Var}[T_0] \geq \frac{h^2(k)}{n} = \frac{\mathbb{E}^2[T_0]}{n} \)
- Extension to possibly random initial conditions is straightforward by conditioning and using this result

---

3 Cauchy–Schwarz for the usual inner product on \( \mathbb{R}^n \):
\[
(\sum_{i=1}^{n} x_i y_i)^2 \leq (\sum_{i=1}^{n} x_i^2)(\sum_{i=1}^{n} y_i^2)
\]

Maximal inequality

Optional stopping theorem has immediate implications for the sample-path behaviour of a martingale:

**Theorem: Maximal inequality**

Let \((X_t)_{t \geq 0}\) be a discrete- or continuous-time non-negative martingale and define \(X_t^* := \sup_{0 \leq s \leq t} X_s\) and \(X_\infty^* := \sup_{0 \leq s < \infty} X_s\). Then \(\mathbb{P}\{X_\infty^* \geq c\} \leq \mathbb{E}[X_0] / c\).

**Proof.**

- Define \(T := \min\{t \geq 0 : X_t \geq c\}\), then \(T \land t\) is a bounded stopping time for any \(t \geq 0\).
- So we may apply optional stopping to obtain \(\mathbb{E}[X_{T \land t}] = \mathbb{E}[X_0]\).
- Also \(\mathbb{E}[X_{T \land t}] \geq \mathbb{E}[1_{\{T \leq t\}} c] = c\mathbb{P}\{T \leq t\} = c\mathbb{P}\{X_t^* \geq c\}\), so \(\mathbb{P}\{X_t^* \geq c\} \leq \mathbb{E}[X_0] / c\).
- The result then follows by taking the limit \(t \to \infty^4\).

---

4 This time by the **monotone convergence theorem**.
Upcrossing inequality

**Theorem: Upcrossing inequality**

Let \((X_n)_{n \geq 0}\) be a martingale and let \(a < b\) be reals. Define \(N_n(a, b]\) to be the number of *upcrossings* by \((X_m)_{m \leq n}\) of the interval \((a, b]\) by time \(n\). Then \(\mathbb{E}[N_n(a, b)] \leq \mathbb{E}[(a - X_n) \vee 0]/(b - a)\).

An *upcrossing* is a minimal path beginning in \([0, a]\) and ending in \((b, \infty)\).
Theorem: Upcrossing inequality

Let \((X_n)_{n \geq 0}\) be a martingale and let \(a < b\) be reals. Define \(N_n(a, b)\) to be the number of upcrossings by \((X_m)_{m \leq n}\) of the interval \((a, b]\) by time \(n\). Then \(\mathbb{E}[N_n(a, b)] \leq \mathbb{E}[(a - X_n) \vee 0]/(b - a)\).

\[
\begin{align*}
S_1 & := \min\{n \geq 0 : X_n \leq a\} \\
S_m & := \min\{n \geq T_{m-1} : X_n \leq a\} \\
T_m & := \min\{n \geq S_m : X_n > b\}
\end{align*}
\]
**Upcrossing inequality**

**Theorem: Upcrossing inequality**

Let \((X_n)_{n \geq 0}\) be a martingale and let \(a < b\) be reals. Define \(N_n(a, b]\) to be the number of upcrossings by \((X_m)_{m \leq n}\) of the interval \((a, b]\) by time \(n\). Then 

\[
\mathbb{E}[N_n(a, b)] \leq \mathbb{E}[ (a - X_n) \lor 0] / (b - a).
\]

\[
Z_n := \begin{cases} 
0 : & n \leq S_1 \text{ or } T_m < n \leq S_{m+1} \\
1 : & S_m < n \leq T_m \end{cases}
\]

<table>
<thead>
<tr>
<th>(X_n)</th>
<th>(Z_n = 0)</th>
<th>(Z_n = 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(X_n \leq a)</td>
<td>(Z_{n+1} = 1)</td>
<td>(Z_{n+1} = 1)</td>
</tr>
<tr>
<td>(a &lt; X_n \leq b)</td>
<td>(Z_{n+1} = 0)</td>
<td>(Z_{n+1} = 1)</td>
</tr>
<tr>
<td>(X_n &gt; b)</td>
<td>(Z_{n+1} = 0)</td>
<td>(Z_{n+1} = 0)</td>
</tr>
</tbody>
</table>
Theorem: Upcrossing inequality

Let \((X_n)_{n \geq 0}\) be a martingale and let \(a < b\) be reals. Define \(N_n(a, b]\) to be the number of upcrossings by \((X_m)_{m \leq n}\) of the interval \((a, b]\) by time \(n\). Then \(\mathbb{E}[N_n(a, b)] \leq \mathbb{E}[(a - X_n) \lor 0]/(b - a)\).

\[
Z_n := \begin{cases} 
0 : n \leq S_1 \text{ or } T_m < n \leq S_{m+1} \\
1 : S_m < n \leq T_m 
\end{cases}
\]

\((Z_n)_{n \geq 0}\) is predictable
Upcrossing inequality

**Theorem: Upcrossing inequality**

Let \((X_n)_{n \geq 0}\) be a martingale and let \(a < b\) be reals. Define \(N_n(a, b]\) to be the number of upcrossings by \((X_m)_{m \leq n}\) of the interval \((a, b]\) by time \(n\). Then \(\mathbb{E}[N_n(a, b)] \leq \mathbb{E}[(a - X_n) \vee 0] / (b - a)\).

\[
(b - a)N_n(a, b] \leq \sum_{k=1}^{n} Z_k (X_k - X_{k-1}) + (a - X_n) \vee 0
\]
Upcrossing inequality

**Theorem: Upcrossing inequality**

Let \((X_n)_{n \geq 0}\) be a martingale and let \(a < b\) be reals. Define \(N_n(a, b)\) to be the number of upcrossings by \((X_m)_{m \leq n}\) of the interval \((a, b]\) by time \(n\). Then \(\mathbb{E}[N_n(a, b)] \leq \mathbb{E}[(a - X_n) \lor 0]/(b - a)\).

\[
(b - a)N_n(a, b) \leq \sum_{k=1}^{n} Z_k(X_k - X_{k-1}) + (a - X_n) \lor 0
\]

\[
(b - a)\mathbb{E}[N_n(a, b)|(X_m)_{m \leq n-1}] \leq \mathbb{E}[(a - X_n) \lor 0|(X_m)_{m \leq n-1}]
\]

\[
+ \mathbb{E}[Z_n(X_n - X_{n-1})|(X_m)_{m \leq n-1}] + \sum_{k=1}^{n-1} \mathbb{E}[Z_k(X_k - X_{k-1})|(X_m)_{m \leq n-1}]
\]

\[
0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1
\]

\[
S_1 \quad T_1 \quad S_2 \quad T_2 \quad S_3 \quad T_3 \quad S_4 \quad T_4 \quad S_5
\]
Upcrossing inequality

**Theorem: Upcrossing inequality**

Let \((X_n)_{n \geq 0}\) be a martingale and let \(a < b\) be reals. Define \(N_n(a, b]\) to be the number of upcrossings by \((X_m)_{m \leq n}\) of the interval \((a, b]\) by time \(n\). Then

\[
\mathbb{E}[N_n(a, b)] \leq \mathbb{E}[(a - X_n) \lor 0]/(b - a).
\]

\[
(b - a)\mathbb{E}[N_n(a, b)|(X_m)_{m \leq n-1}] \leq \mathbb{E}[(a - X_n) \lor 0|(X_m)_{m \leq n-1}]
\]

\[
+ Z_n\mathbb{E}[(X_n - X_{n-1})|(X_m)_{m \leq n-1}] + \sum_{k=1}^{n-1} \mathbb{E}[Z_k(X_k - X_{k-1})|(X_m)_{m \leq n-1}]
\]

\[
(b - a)\mathbb{E}[N_n(a, b)|(X_m)_{m \leq n-1}] \leq \mathbb{E}[(a - X_n) \lor 0|(X_m)_{m \leq n-1}]
\]

\[
+ \sum_{k=1}^{n-1} \mathbb{E}[Z_k(X_k - X_{k-1})|(X_m)_{m \leq n-1}]
\]
Upcrossing inequality

**Theorem: Upcrossing inequality**

Let $(X_n)_{n \geq 0}$ be a martingale and let $a < b$ be reals. Define $N_n(a, b]$ to be the number of upcrossings by $(X_m)_{m \leq n}$ of the interval $(a, b]$ by time $n$. Then $\mathbb{E}[N_n(a, b)] \leq \mathbb{E}[(a - X_n) \lor 0]/(b - a)$.

\[
(b - a)\mathbb{E}[N_n(a, b]|(X_m)_{m \leq n-2}] \leq \mathbb{E}[(a - X_n) \lor 0|(X_m)_{m \leq n-2}]
+ \mathbb{E}[Z_{n-1}(X_{n-1} - X_{n-2})|(X_m)_{m \leq n-2}] + \sum_{k=1}^{n-2} \mathbb{E}[Z_k(X_k - X_{k-1})|(X_m)_{m \leq n-2}]
\]

\[
(b - a)\mathbb{E}[N_n(a, b]|(X_m)_{m \leq n-2}] \leq \mathbb{E}[(a - X_n) \lor 0|(X_m)_{m \leq n-2}]
+ \sum_{k=1}^{n-2} \mathbb{E}[Z_k(X_k - X_{k-1})|(X_m)_{m \leq n-2}]
\]
Upcrossing inequality

Theorem: Upcrossing inequality

Let \( (X_n)_{n \geq 0} \) be a martingale and let \( a < b \) be reals. Define \( N_n(a, b] \) to be the number of upcrossings by \( (X_m)_{m \leq n} \) of the interval \((a, b]\) by time \( n \). Then

\[
E[N_n(a, b)] \leq E[(a - X_n) \vee 0]/(b - a).
\]

\[
(b - a)E[N_n(a, b]|(X_m)_{m \leq n-2}] \leq E[(a - X_n) \vee 0|(X_m)_{m \leq n-2}]
\]

\[
+ \sum_{k=1}^{n-2} E[Z_k(X_k - X_{k-1})|(X_m)_{m \leq n-2}]
\]

\[
\ldots
\]

\[
(b - a)E[N_n(a, b)] \leq E[(a - X_n) \vee 0]
\]
Martingale convergence theorem

**Theorem: Martingale convergence**

Let \((X_n)_{n \geq 0}\) be a non-negative martingale, or a martingale with \(\mathbb{E}[|X_n|] \leq K\) independent of \(n\). Then w.p. 1, \(X_n \to X\) for some finite random variable \(X\).

\(a\) Extends to continuous time under càdlàg assumptions.

**Proof.**

- Total number of crossings: \(\lim_{n \to \infty} N_n(a, b)\) (limit exists, possibly \(\infty\)).

- \(\mathbb{E}[\lim_{n \to \infty} N_n(a, b)] = \lim_{n \to \infty} \mathbb{E}[N_n(a, b)],\)\(^5\) which is finite by *upcrossing inequality* since: \(\mathbb{E}[N_n(a, b)] \leq \mathbb{E}[(a - X_n) \lor 0] \leq \mathbb{E}[|a - X_n|] \leq \mathbb{E}[|X_n|] + a\)
  which is bounded by \(\mathbb{E}[X_0] + a\) in the non-negative case, or \(K + a\) otherwise.

- Probability of *any rational interval* being upcrossed infinitely many times is 0 (countable additivity).

- \(X_n\) must thus converge, w.p. 1, to a (possibly infinite) limit (if \(X_n\) doesn’t diverge, its oscillations must be ‘damped’ over time).

- \(\mathbb{E}[\lim_{n \to \infty} |X_n|] \leq \liminf_{n \to \infty} \mathbb{E}[|X_n|],\)\(^6\) so \(\lim_{n \to \infty} X_n\) is finite w.p. 1.

\(^5\) Monotone convergence theorem again.

\(^6\) By *Fatou’s Lemma*. 
Application: the PASTA property

**Theorem**

Let \((N_t)_{t\geq0}\) be a stochastic process (e.g. queue length) and \((A_t)_{t\geq0}\) a rate \(\lambda > 0\) Poisson process. Let \(\{N_t \in B\}\) be an event of interest and write \(U_t := 1_{\{N_t \in B\}}\).

- **Time avg.:** \(V_t := \frac{1}{t} \int_0^t U_s \, ds\)
- **Arrivals avg.:** \(Y_t / A_t\) for \(Y_t := \int_0^t U_s \, dA_s\)

Assume for each \(t \geq 0\), \(\{A_{t+u} - A_t : u \geq 0\}\) and \(\{U_s : 0 \leq s \leq t\}\) are independent. Then \(V_t \to V(\infty)\) w.p. 1 \(\Rightarrow Y_t / A_t \to V(\infty)\) w.p. 1.

- Define \(R_t := Y_t - \lambda t V_t\). **Claim:** \((R_t)_{t\geq0}\) is a martingale wrt. \((R_t)_{t\geq0}\)

- \(\mathbb{E}[R_{t+s} - R_s | (R_u)_{u\leq s}] = \mathbb{E} \left[ \mathbb{E} \left[ Y_{t+s} - Y_s - \lambda \int_s^{s+t} U_u \, d\lambda | (A_u, U_u)_{u\leq s} \right] | (R_u)_{u\leq s} \right]\)

- For \(n \geq 0\), approximate \(Y_{t+s} - Y_s = \int_s^{s+t} U_u \, dA_u\) by \(\tilde{Y}_n := \sum_{k=0}^{n-1} U_{s+kt/n} [A_{s+(k+1)t/n} - A_{s+kt/n}]\)

- Then \(\mathbb{E}[\tilde{Y}_n | (A_u, U_u)_{u\leq s}] = \mathbb{E} \left[ \sum_{k=0}^{n-1} \frac{t}{n} U_{s+kt/n} | (A_u, U_u)_{u\leq s} \right] \lambda\)

- Taking the limit \(n \to \infty\) gives the claim\(^7\)

---

\(^7\) Dominated convergence again.

Application: the PASTA property

Theorem

Let \((N_t)_{t \geq 0}\) be a stochastic process (e.g. queue length) and \((A_t)_{t \geq 0}\) a rate \(\lambda > 0\) Poisson process. Let \(\{N_t \in B\}\) be an event of interest and write \(U_t := 1_{\{N_t \in B\}}\).

**Time avg.:** \(V_t := \frac{1}{t} \int_0^t U_s \, ds\)  
**Arrivals avg.:** \(Y_t/A_t\) for \(Y_t := \int_0^t U_s - dA_s\)

Assume for each \(t \geq 0\), \(\{A_{t+u} - A_t : u \geq 0\}\) and \(\{U_s : 0 \leq s \leq t\}\) are independent. Then \(V_t \to V(\infty)\) w.p. 1 ⇒ \(Y_t/A_t \to V(\infty)\) w.p. 1.

\[ R_t := Y_t - \lambda t V_t \text{ is a martingale wrt. } (R_t)_{t \geq 0} \]

Then for any \(h > 0\), \((B_n)_{n \geq 0} := (R_{nh})_{n \geq 0}\) is a discrete-time martingale

\[ \text{Write } Z_n := \sum_{k=1}^n \frac{1}{k} (B_k - B_{k-1}), \text{ so } (Z_n)_{n \geq 0} \text{ is a martingale wrt. } (B_n)_{n \geq 0} \]

Also

\[ \mathbb{E}[(B_k - B_{k-1})^2] \leq \mathbb{E} \left[ \left( \int_{(k-1)h}^{kh} U_s - dA_s \right)^2 \right] + \lambda^2 \mathbb{E} \left[ \left( \int_{(k-1)h}^{kh} U_s \, ds \right)^2 \right] \leq \lambda h + 2\lambda^2 h^2 \]

So \(\mathbb{E}[Z_n^2] = \sum_{k=1}^n \frac{1}{k^2} \mathbb{E}[(B_k - B_{k-1})^2] + (\text{cross terms}) \text{ of the form } \frac{1}{u} \frac{1}{v} \mathbb{E}[(B_u - B_{u-1})(B_v - B_{v-1})] \text{ where } u < v \)

Then \(\mathbb{E}[\mathbb{E}[(B_u - B_{u-1})(B_v - B_{v-1})|(B_s)_{s \leq v-1}]] = 0\), so

\[ \mathbb{E}[Z_n^2] = \sum_{k=1}^\infty \frac{1}{k^2} \mathbb{E}[(B_k - B_{k-1})^2] =: C < \infty \]

Then martingale convergence theorem gives \(Z_n \to Z\) w.p. 1

*Kronecker’s lemma* \(^7\) gives \(\frac{R_{nh}}{n} \to 0\) w.p. 1, so \(\frac{R_t}{t} \to 0\) as \(h\) is arbitrary

Finally, \(\frac{R_t}{t} = \frac{Y_t}{A_t} \frac{A_t}{t} - \lambda V_t\) and result follows from Poisson SLLN

\(^7\) For \(0 < b_1 < b_2 < \ldots\), suppose the sum \(\sum_{k=1}^\infty \frac{x_k}{b_k}\) converges to a finite limit. Then \(\frac{1}{b_n} \sum_{k=1}^n x_k \to 0\) as \(n \to \infty\).

Part 2: The martingale method for proving heavy-traffic limits of Markovian queues
Point processes

- A point process \((N_t)_{t\geq 0}\) is a non-decreasing, non-negative-integer-valued stochastic process in \(D\) (right continuous with lefthand limits) with \(N_0 = 0\).

- For example, Poisson processes, renewal processes, arrival and departure processes of a queue . . .

- We consider only non-explosive processes: \(N_t < \infty\) w.p. 1 for all \(t \in [0, \infty)\).

- Most queueing models can be constructed as simple (to write down) combinations of such processes.
M/M/N queue in terms of Poisson processes

**Ingredients:** given mutually-independent $Q_0$ and Poisson processes:

$$(A_t)_{t \geq 0} \text{ (rate } \lambda) \quad \{(S^k_t)_{t \geq 0} : 1 \leq k \leq N\} \text{ (rate } \mu)$$

How might we construct $(Q_t)_{t \geq 0}$ the queue-length process of an M/M/N queue with arrivals at rate $\lambda$ and services at rate $\mu$?

$$Q_t = Q_0 + A_t - \sum_{k=1}^{N} \int_0^t 1\{Q_u - k \geq 1\} \ dS^k_u$$

where, since we are working with unit-jump point processes:

$$\int_0^t 1\{Q_u - k \geq 1\} \ dS^k_u = \sum_{i=1}^{\infty} 1\{t \geq T^k_i\} 1\{Q_{T^k_i} - k \geq 1\}$$

for $T^k_0 := 0, T^k_1, T^k_2, \ldots$ the jump times of $S^k_t$
Where do martingales come into the picture?

- Representations like this in terms of Poisson processes are very useful for spawning martingales.
- For example, independent increments gives:

\[
\mathbb{E}[A_{s+t} - \lambda(s + t)|(A_u)_{u \leq s}] = A_s + \mathbb{E}[A_{s+t} - A_s|(A_u)_{u \leq s}] - \lambda(s + t)
\]

\[
= A_s + \lambda t - \lambda(s + t) = A_s - \lambda s
\]

so \( M_t^A := A_t - \lambda t \) is a martingale.

- This lets us decompose:

\[
Q_t = Q_0 + M_t^A + \lambda t - \sum_{k=1}^{N} \int_0^t \mathbf{1}_{\{Q_u \geq k\}} \, dS^k_u
\]

- What about the term \( \sum_{k=1}^{N} \int_0^t \mathbf{1}_{\{Q_u \geq k\}} \, dS^k_u \)?
- Any guesses?

\[
M_t^S = \sum_{k=1}^{N} \int_0^t \mathbf{1}_{\{Q_u \geq k\}} \, dS^k_u - \mu \sum_{k=1}^{N} \int_0^t \mathbf{1}_{\{Q_u \geq k\}} \, du
\]

- We’ll prove that this is a martingale shortly . . . but with respect to which filtration? \( \mathcal{F}_t := (Q(0), (A_s)_{s \leq t}, (S^k_s)_{s \leq t} : 1 \leq k \leq N) \)
The Doob–Meyer decomposition theorem

**Definition: Sub/super-martingale**

A discrete- or continuous-time sequence of \( \mathbb{R} \)-valued random variables \((X_t)_{t \geq 0}\) is a sub-martingale (resp. super-martingale) wrt. an adapted filtration \((\mathcal{F}_t)_{t \geq 0}\) if:

- \( \mathbb{E}[|X_t|] < \infty \) for all \( t \geq 0 \);
- \( \mathbb{E}[X_{s+t} | \mathcal{F}_s] \geq (\leq) X_s \) w.p. 1 for all \( s, t \geq 0 \).

Note: any increasing process is a sub-martingale, e.g. point processes

**Theorem: Doob–Meyer decomposition for non-negative sub-martingales**

Let \((X_t)_{t \geq 0}\) be a non-negative sub-martingale wrt. \((\mathcal{F}_t)_{t \geq 0}\). Then there exists a unique (w.p. 1) non-negative, non-decreasing \( \mathcal{F}_t \)-predictable\(^a\) process \((C_t)_{t \geq 0}\) (called the compensator of \( X_t \)) with \( C_0 = 0 \) and \( \mathbb{E}[C_t] < \infty \) for each \( t \geq 0 \), such that \( M_t := X_t - C_t \) is an \( \mathcal{F}_t \)-martingale.

\(^a\) Meaning \( C_t \) is knowable just prior to time \( t \), that is, with the information in \( \mathcal{F}_{t^-} := \bigcup_{s < t} \mathcal{F}_s \).

For example:

- \( \lambda t \) is the compensator of \( A_t \)
- \( \mu \sum_{k=1}^N \int_0^t 1_{\{Q_u \geq k\}} \, du \) is the compensator of \( \sum_{k=1}^N \int_0^t 1_{\{Q_u^- \geq k\}} \, dS_u^k \) (still to be proven)
The Doob–Meyer decomposition theorem

**Theorem:** Doob–Meyer decomposition for non-negative sub-martingales

Let \((X_t)_{t \geq 0}\) be a non-negative sub-martingale wrt. \((\mathcal{F}_t)_{t \geq 0}\). Then there exists a unique (w.p. 1) non-negative, non-decreasing \(\mathcal{F}_t\)-predictable\(^a\) process \((C_t)_{t \geq 0}\) (called the *compensator* of \(X_t\)) with \(C_0 = 0\) and \(E[C_t] < \infty\) for each \(t \geq 0\), such that \(M_t := X_t - C_t\) is an \(\mathcal{F}_t\)-martingale.

\(^a\) Meaning \(C_t\) is knowable just prior to time \(t\), that is, with the information in \(\mathcal{F}_{t^-} := \bigcup_{s < t} \mathcal{F}_s\).

In discrete time, the decomposition is straightforward:

- Write \(X_{n+1} = X_0 + \sum_{i=0}^{n} (X_{i+1} - X_i)\)
- Then: \(X_{n+1} = X_0 + \sum_{i=0}^{n} (X_{i+1} - E[X_{i+1} | \mathcal{F}_i]) + \sum_{i=0}^{n} (E[X_{i+1} | \mathcal{F}_i] - X_i)\)
  
  \[\underbrace{:= M_{n+1}}_{\text{:= } M_{n+1}} + \underbrace{:= C_{n+1}}_{\text{:= } C_{n+1}}\]
- \(C_n\) is non-negative and non-decreasing by the sub-martingale property of \(X_n\). Predictability is clear since \(E[X_{n+1} | \mathcal{F}_n]\) is known at time \(n\)
- \(E[M_{n+1} | \mathcal{F}_n] = X_0 + \sum_{i=0}^{n-1} (X_{i+1} - E[X_{i+1} | \mathcal{F}_i]) = M_n\) so \(M_n\) is a martingale
The Doob–Meyer decomposition theorem

**Theorem: Doob–Meyer decomposition for non-negative sub-martingales**

Let \((X_t)_{t \geq 0}\) be a non-negative sub-martingale wrt. \((F_t)_{t \geq 0}\). Then there exists a unique (w.p. 1) non-negative, non-decreasing \(F_t\)-predictable\(^a\) process \((C_t)_{t \geq 0}\) (called the *compensator* of \(X_t\)) with \(C_0 = 0\) and \(E[C_t] < \infty\) for each \(t \geq 0\), such that \(M_t := X_t - C_t\) is an \(F_t\)-martingale.

\(^a\) Meaning \(C_t\) is knowable just prior to time \(t\), that is, with the information in \(F_{t-} := \cup_{s < t} F_s\).

- To prove uniqueness, let \(X_n = M^{(1)}_n + C^{(1)}_n\) and \(X_n = M^{(2)}_n + C^{(2)}_n\) be two such decompositions wrt. the same filtration \(F_n\)
- So \((M^{(1)}_n - M^{(2)}_n) + (C^{(1)}_n - C^{(2)}_n) = 0\) for all \(n \geq 0\), where \(M^{(1)}_n - M^{(2)}_n\) is an \(F_n\)-martingale and \(C^{(1)}_n - C^{(2)}_n\) is \(F_n\)-predictable
- Now \(E[C^{(1)}_{n+1} - C^{(2)}_{n+1} | F_n] = C^{(1)}_{n+1} - C^{(2)}_{n+1}\) by predictability
- and \(E[M^{(1)}_{n+1} - M^{(2)}_{n+1} | F_n] = M^{(1)}_{n+1} - M^{(2)}_{n+1}\) by martingale property
- So also \((M^{(1)}_n - M^{(2)}_n) + (C^{(1)}_{n+1} - C^{(2)}_{n+1}) = 0\) for all \(n \geq 0\)
- Thus \((C^{(1)}_{n+1} - C^{(2)}_{n+1}) = 0\) so since \(C^{(1)}_0 = C^{(2)}_0\), we have \(C^{(1)}_n = C^{(2)}_n\) for all \(n \geq 0\)

Harder in continuous time.
The Doob–Meyer decomposition theorem

**Theorem: Doob–Meyer decomposition for non-negative sub-martingales**

Let \((X_t)_{t \geq 0}\) be a non-negative sub-martingale wrt. \((\mathcal{F}_t)_{t \geq 0}\). Then there exists a unique (w.p. 1) non-negative, non-decreasing \(\mathcal{F}_t\)-predictable process \((C_t)_{t \geq 0}\) (called the *compensator* of \(X_t\)) with \(C_0 = 0\) and \(\mathbb{E}[C_t] < \infty\) for each \(t \geq 0\), such that \(M_t := X_t - C_t\) is an \(\mathcal{F}_t\)-martingale.

\(a\) Meaning \(C_t\) is knowable just prior to time \(t\), that is, with the information in \(\mathcal{F}_{t^-} := \bigcup_{s < t} \mathcal{F}_s\).

- This decomposition can also be extended to *finite variation* processes \(X_t\), where the variation of \(X_t\), \(V^X_t := \sup P_t \sum_{i=1}^M |X_{t_i}^M - X_{t_{i-1}^M}| < \infty\) for all \(t \geq 0\). The sup is over \(P_t := \{(t_i^M)_{i=1}^M : (t_i^M)_{i=0}^M\ is\ a\ partition\ of\ [0, t]\}\)

- Any process of finite variation can be expressed as the difference of two increasing non-negative processes, e.g. \(X_t = V^X_t - (V^X_t - X_t)\)

- This gives a version of the decomposition for finite variation processes:

**Theorem: Doob–Meyer decomposition for finite variation processes**

Let \((X_t)_{t \geq 0}\) be of finite variation with \(\mathbb{E}[V^X_t] < \infty\) and \(\mathbb{E}[|X_t|] < \infty\) for \(t \geq 0\); and let \((\mathcal{F}_t)_{t \geq 0}\) be an adapted filtration. Then there exists a unique (w.p. 1) \(\mathcal{F}_t\)-predictable process \((C_t)_{t \geq 0}\) (called the *compensator* of \(X_t\)) with \(C_0 = 0\) and \(\mathbb{E}[C_t] < \infty\) for each \(t \geq 0\), such that \(M_t := X_t - C_t\) is an \(\mathcal{F}_t\)-martingale.
Integrals of predictable processes wrt. bounded variation martingales are martingales

To show: $\mu \sum_{k=1}^{N} \int_{0}^{t} 1\{Q_u \geq k\} du$ is the compensator of $\sum_{k=1}^{N} \int_{0}^{t} 1\{Q_u \geq k\} dS_u^k$, set $X_t = S_t^k - \mu t$ and $C_t = 1\{Q_t \geq k\}$ for each $k$ in:

Theorem: A version of the martingale integration theorem

Let $(X_t)_{t \geq 0}$ be a finite variation martingale wrt. a filtration $(\mathcal{F}_t)_{t \geq 0}$ and let $(C_t)_{t \geq 0}$ be an $\mathcal{F}_t$-predictable process such that $\mathbb{E}\left[\int_{0}^{t} |C_s| dX_s\right] < \infty$ for all $t \geq 0$. Then the process $\int_{0}^{t} C_s dX_s$ is an $\mathcal{F}_t$-martingale.

- We prove it for simple predictable processes $C_t := C^{(0)} + \sum_{i=1}^{n-1} C^{(i)} 1\{t \in (T_i, T_{i+1}]\}$ where the $0 = T_0 \leq T_1 \leq \ldots \leq T_n$ are $\mathcal{F}_t$-stopping times and the $C^{(i)}$ are known given $\mathcal{F}_{T_i}$

- Then $\int_{0}^{t} C_s dX_s = \sum_{i=1}^{n-1} C^{(i)} (X_{t \wedge T_{i+1}} - X_{t \wedge T_i})$

- Now $\mathbb{E}\left[\int_{0}^{s+t} C_u dX_u | \mathcal{F}_s\right] = \mathbb{E}\left[\int_{0}^{s} C_u dX_u + \int_{s}^{s+t} C_u dX_u | \mathcal{F}_s\right] = \int_{0}^{s} C_u dX_u + \mathbb{E}\left[\int_{s}^{s+t} C_u dX_u | \mathcal{F}_s\right]$

- So need to show $\mathbb{E}\left[\int_{s}^{s+t} C_u dX_u | \mathcal{F}_s\right] = 0$
Integrals of predictable processes wrt. bounded variation martingales are martingales

To show: $\mu \sum_{k=1}^{N} \int_{0}^{t} 1\{Q_u \geq k\} \, du$ is the compensator of $\sum_{k=1}^{N} \int_{0}^{t} 1\{Q_u \geq k\} \, dS^k_u$, set $X_t = S^k_t - \mu t$ and $C_t = 1\{Q_t \geq k\}$ for each $k$ in:

Theorem: A version of the martingale integration theorem

Let $(X_t)_{t \geq 0}$ be a finite variation martingale wrt. a filtration $(\mathcal{F}_t)_{t \geq 0}$ and let $(C_t)_{t \geq 0}$ be an $\mathcal{F}_t$-predictable process such that $\mathbb{E} \left[ \int_{0}^{t} |C_s| \, |dX_s| \right] < \infty$ for all $t \geq 0$. Then the process $\int_{0}^{t} C_s \, dX_s$ is an $\mathcal{F}_t$-martingale.

- So need to show $\mathbb{E} \left[ \int_{s}^{s+t} C_u \, dX_u | \mathcal{F}_s \right] = 0$
- Now $\mathbb{E} \left[ \int_{s}^{s+t} C_u \, dX_u | \mathcal{F}_s \right] = \mathbb{E} \left[ \sum_{i=1}^{n-1} 1\{T_i \leq s+t\} C^{(i)} (X_{s\vee((s+t)\wedge T_{i+1})} - X_{s\vee((s+t)\wedge T_{i})}) | \mathcal{F}_s \right]$
- Furthermore, for each term:

\[
\mathbb{E} \left[ 1\{T_i \leq s+t\} C^{(i)} (X_{s\vee((s+t)\wedge T_{i+1})} - X_{s\vee((s+t)\wedge T_{i})}) | \mathcal{F}_s \right] = \\
\mathbb{E} \left[ \mathbb{E} \left[ 1\{T_i \leq s+t\} C^{(i)} (X_{s\vee((s+t)\wedge T_{i+1})} - X_{s\vee((s+t)\wedge T_{i})}) | \mathcal{F}_{s\vee((s+t)\wedge T_{i})} \right] | \mathcal{F}_s \right] = \\
\mathbb{E} \left[ 1\{T_i \leq s+t\} C^{(i)} \mathbb{E} \left[ (X_{s\vee((s+t)\wedge T_{i+1})} - X_{s\vee((s+t)\wedge T_{i})}) | \mathcal{F}_{s\vee((s+t)\wedge T_{i})} \right] | \mathcal{F}_s \right] = 0
\]

=0 by optional stopping
Integrals of predictable processes wrt. bounded variation martingales are martingales

To show: $\mu \sum_{k=1}^{N} \int_{0}^{t} 1\{Q_{u-} \geq k\} \, du$ is the compensator of $\sum_{k=1}^{N} \int_{0}^{t} 1\{Q_{u-} \geq k\} \, dS_{u}^{k}$, set $X_{t} = S_{t}^{k} - \mu t$ and $C_{t} = 1\{Q_{t-} \geq k\}$ for each $k$ in:

**Theorem: A version of the martingale integration theorem**

Let $(X_{t})_{t \geq 0}$ be a finite variation martingale wrt. a filtration $(\mathcal{F}_{t})_{t \geq 0}$ and let $(C_{t})_{t \geq 0}$ be an $\mathcal{F}_{t}$-predictable process such that $\mathbb{E} \left[ \int_{0}^{t} |C_{s}| \, |dX_{s}| \right] < \infty$ for all $t \geq 0$. Then the process $\int_{0}^{t} C_{s} \, dX_{s}$ is an $\mathcal{F}_{t}$-martingale.

- Extends to general predictable processes by intricate and technical limiting arguments
- However, note that $1\{Q_{t-} \geq k\}$ is simple predictable. Consider the sequence of stopping times at which $Q_{t}$ crosses $k$ in each direction
Doob–Meyer (semi-martingale) decomposition of queue

\[ Q_t = Q_0 + A_t - \sum_{k=1}^{N} \int_{0}^{t} 1_{\{Q_u \geq k\}} \, dS_u^k \]

\[ Q_t = Q_0 + M_t^A - M_t^S + \lambda t - \mu \int_{0}^{t} (Q_u \wedge N) \, du \]

where:

\[ M_t^A = A_t - \lambda t \]

\[ M_t^S = \sum_{k=1}^{N} \int_{0}^{t} 1_{\{Q_u \geq k\}} \, dS_u^k - \mu \int_{0}^{t} (Q_u \wedge N) \, du \]

Equation ‘looks like’ an ODE, up to the martingales \( M_t^A \) and \( M_t^S \)

Under an appropriate scaling, can be shown that the limiting martingales become Brownian motions giving an SDE
Heavy-traffic scaling regime

- Sequence of models indexed by $N$, the number of servers
- Each server operates at a fixed rate $\mu$
- Arrival rate $\lambda_N = \mu N - \mu \beta \sqrt{N}$ for some $\beta \in \mathbb{R}$
- Then $\rho_N := \frac{\lambda_N}{N \mu} = 1 - \frac{\beta}{\sqrt{N}} \to 1$
- The $\sqrt{N}$ term puts us in the so-called quality and efficiency driven (QED) regime\(^3\) — limiting probability of customer delay in $(0, 1)$

Then we have:

$$Q_t^N = Q_0^N + M_{t}^{A,N} - M_{t}^{S,N} + \lambda_N t - \mu \int_0^t (Q_u^N \wedge N) \, du$$

where:

$$M_{t}^{A,N} = A_{t}^N - \lambda_N t$$

$$M_{t}^{S,N} = \sum_{k=1}^{N} \int_0^t \mathbf{1}_{\{Q_u^N \geq k\}} \, dS_k^u - \mu \int_0^t (Q_u^N \wedge N) \, du$$

are martingales

---

Heavy-traffic scaling regime

- Define the rescaled processes: \( \bar{Q}_t^N := \frac{Q_t^N - N}{\sqrt{N}}, \bar{M}_t^{A,N} := \frac{M_t^{A,N}}{\sqrt{N}} \) and
  \( \bar{M}_t^{S,N} := \frac{M_t^{S,N}}{\sqrt{N}} \)

- Then we have:

\[
\bar{Q}_t^N = \frac{Q_0^N - N}{\sqrt{N}} + \bar{M}_t^{A,N} - \bar{M}_t^{S,N} + \frac{\lambda N t - N \mu t}{\sqrt{N}} - \mu \int_0^t \frac{Q_u^N \wedge N - N}{\sqrt{N}} du
\]

  \[
  = \bar{Q}_0^N + \bar{M}_t^{A,N} - \bar{M}_t^{S,N} - \mu \beta t - \mu \int_0^t (\bar{Q}_u^N \wedge 0) du
  \]

- We will see that by virtue of their martingale structure
  \( (\bar{M}_t^{A,N}, \bar{M}_t^{S,N}) \Rightarrow (\sqrt{\mu} B_t^{(1)}, \sqrt{\mu} B_t^{(2)}) \) in \( D^2 \) as \( N \to \infty \) where \( B_t^{(1)} \) and \( B_t^{(2)} \)
  are independent standard Brownian motions

- Then by the continuous mapping theorem, if \( \bar{Q}_0^N \Rightarrow X_0 \) in \( \mathbb{R} \) as \( N \to \infty \),
  we will have that \( \bar{Q}_t^N \Rightarrow X_t \) in \( D \) as \( N \to \infty \), where:

\[
X_t = X_0 + \sqrt{\mu} B_t^{(1)} - \sqrt{\mu} B_t^{(2)} - \mu \beta t - \mu \int_0^t (X_u \wedge 0) du
\]

\[
\overset{d}{=} X_0 + \sqrt{2 \mu} B_t - \mu \beta t - \mu \int_0^t (X_u \wedge 0) du
\]
Martingale functional central limit theorem (FCLT)

- For \((X_t)_{t\geq 0}\) and \((Y_t)_{t\geq 0}\), define, where it exists, the quadratic covariation process: 
  \[ [X_t, Y_t] := \sup_{P_t} \sum_{i=1}^{M}(X_t^M - X_{t_{i-1}}^M)(Y_t^M - Y_{t_{i-1}}^M), \]
  where the sup is again over all partitions \(P_t\) of \([0, t]\). 
  Quadratic variation \([X_t] := [X_t, X_t]\)

- If \([X_t, Y_t]\) satisfies the conditions of Doob–Meyer, can define predictable quadratic covariation \(\langle X_t, Y_t \rangle\) as the unique compensator of \([X_t, Y_t]\) (wrt. to a filtration \((\mathcal{F}_t)_{t\geq 0}\)). 
  Predictable quadratic variation \(\langle X_t \rangle := \langle X_t, X_t \rangle\)

Theorem: Lévy’s characterisation of multi-dimensional Brownian motion

Let \((M_t)_{t\geq 0} = (M_1^t, \ldots, M_k^t)\) be a martingale wrt. a filtration \((\mathcal{F}_t)_{t\geq 0}\) which has continuous sample paths in \(\mathbb{R}^k\), \(M_0 = 0\) and \(M_t\) is square integrable, i.e. 
\(\mathbb{E}[\|M_t\|^2] < \infty\) for all \(t \geq 0\). Let \(C = (c_{ij})\) be a \(k\times k\) real covariance matrix (non-negative-definite and symmetric), such that the predictable quadratic covariation processes \(\langle M_i^t, M_j^t \rangle\) (wrt. \((\mathcal{F}_t)_{t\geq 0}\)) exist for each \(1 \leq i, j \leq k\), and for each \(t \geq 0\):

\[ \langle M_i^t, M_j^t \rangle = c_{ij}t \]

Then \(M_t\) is a \(k\)-dimensional Brownian motion with mean 0 and covariance matrix \(Ct\).
Martingale functional central limit theorem (FCLT)

- For \((X_t)_t \geq 0\) and \((Y_t)_t \geq 0\), define, where it exists, the quadratic covariation process: 
\[ [X_t, Y_t] := \sup_{P_t} \sum_{i=1}^{M} (X_{t_i} - X_{t_{i-1}})(Y_{t_i} - Y_{t_{i-1}}), \]
where the sup is again over all partitions \(P_t\) of \([0, t]\). Quadratic variation 
\[ [X_t] := [X_t, X_t] \]
- If \([X_t, Y_t]\) satisfies the conditions of Doob–Meyer, can define predictable quadratic covariation 
\[ \langle X_t, Y_t \rangle \]
as the unique compensator of \([X_t, Y_t]\) (wrt. to a filtration \((\mathcal{F}_t)_{t \geq 0}\)). Predictable quadratic variation 
\[ \langle X_t \rangle := \langle X_t, X_t \rangle \]

Theorem: A version of the martingale FCLT

For each \(N \geq 1\), let \((M_t^N)_{t \geq 0} = (M_{t,1}^N, \ldots, M_{t,k}^N)\) be a martingale in \(D^k\) wrt. a filtration \((\mathcal{F}_t^N)_{t \geq 0}\) such that \(M_0^N = 0\) and \(M_t^N\) is square integrable, i.e. \(\mathbb{E}[\|M_t^N\|^2] < \infty\) for all \(t \geq 0\) and \(N \geq 1\). Let \(C = (c_{ij})\) be a \(k \times k\) real covariance matrix (non-negative-definite and symmetric), such that the predictable quadratic covariation processes \(\langle M_t^N, i, M_t^N, j \rangle\) (wrt. \((\mathcal{F}_t^N)_{t \geq 0}\)) exist for each \(1 \leq i, j \leq k\), and for each \(t \geq 0\):
\[ \langle M_t^N, i, M_t^N, j \rangle \Rightarrow c_{ij}t \quad \text{in } \mathbb{R} \text{ as } N \to \infty \]

Assume further that for each \(t \geq 0\) the expected values of the maximum jump in \([0, t]\) of \(\langle M_t^N, i, M_t^N, j \rangle\) and the maximum squared jump of \(M_t^N\) tend to zero as \(N \to \infty\). Then \(M_t^N \Rightarrow B_t\) in \(D^k\) as \(N \to \infty\) where \((B_t)_{t \geq 0}\) is a \(k\)-dimensional Brownian motion with mean \(0\) and covariance matrix \(C_t\).

Martingale functional central limit theorem (FCLT)

- For \((X_t)_{t \geq 0}\) and \((Y_t)_{t \geq 0}\), define, where it exists, the quadratic covariation process: 
  \[ [X_t, Y_t] := \sup_{P_t} \sum_{i=1}^{M} (X_{t_i}^M - X_{t_{i-1}}^M)(Y_{t_i}^M - Y_{t_{i-1}}^M), \]
  where the sup is again over all partitions \(P_t\) of \([0, t]\). Quadratic variation \([X_t] := [X_t, X_t]\)

- If \([X_t, Y_t]\) satisfies the conditions of Doob–Meyer, can define predictable quadratic covariation \(\langle X_t, Y_t \rangle\) as the unique compensator of \([X_t, Y_t]\) (wrt. to a filtration \((\mathcal{F}_t)_{t \geq 0}\)). Predictable quadratic variation \(\langle X_t \rangle := \langle X_t, X_t \rangle\)

- Interested in \((\bar{M}_{t}^{A,N}, \bar{M}_{t}^{S,N})\) and so \(\langle \bar{M}_{t}^{A,N}, \bar{M}_{t}^{S,N} \rangle, \langle \bar{M}_{t}^{A,N} \rangle, \langle \bar{M}_{t}^{S,N} \rangle\)

- We will show:
  - \(\langle \bar{M}_{t}^{A,N}, \bar{M}_{t}^{S,N} \rangle = 0\) (w.p. 1), so that the marginal BMs are independent
  - \(\langle \bar{M}_{t}^{A,N} \rangle \Rightarrow \mu t\) and \(\langle \bar{M}_{t}^{S,N} \rangle \Rightarrow \mu t\) so that each marginal BM has infinitesimal variance \(\mu\)

- Note that \([X_t + Y_t, U_t + V_t] = [X_t, U_t] + [X_t, V_t] + [Y_t, U_t] + [Y_t, V_t]\)

- Also, if \(X_t\) is continuous and \(Y_t\) of finite variation, we have:
  \[
  [X_t, Y_t] \leq \sup_{P_t} \max_{i} |X_{t_i}^M - X_{t_{i-1}}^M| \sup_{P_t} \sum_{i=1}^{M} |Y_{t_i}^M - Y_{t_{i-1}}^M| \\
  \leq \sup_{P_t} \max_{i} |X_{t_i}^M - X_{t_{i-1}}^M| V_t^Y \\
  = 0 \quad \text{since } X_t \text{ is (uniformly) continuous on } [0, t]
  
  Further, if \(X_t\) and \(Y_t\) are both pure jump: 
  \([X_t, Y_t] = \sum_{s \leq t} \Delta X_s \Delta Y_s\)
Showing $\langle \bar{M}_t^{A,N}, \bar{M}_t^{S,N} \rangle = 0$

This is straightforward since:

- $\bar{M}_t^{A,N} = \frac{1}{\sqrt{N}} (A_t^N - \lambda N t)$
- $\bar{M}_t^{S,N} = \frac{1}{\sqrt{N}} (\sum_{k=1}^{N} \int_0^t 1_{\{Q_u^N \geq k\}} dS^k_u - \mu \int_0^t (Q_u^N \land N) \, du)$

- So

$$[\bar{M}_t^{A,N}, \bar{M}_t^{S,N}] = \frac{1}{N} \left[ A_t^N, \sum_{k=1}^{N} \int_0^t 1_{\{Q_u^N \geq k\}} dS^k_u \right]$$

$$\quad - \frac{1}{N} \left[ A_t^N, \mu \int_0^t (Q_u^N \land N) \, du \right]$$

$$\quad - \frac{1}{N} \left[ \lambda N t, \sum_{k=1}^{N} \int_0^t 1_{\{Q_u^N \geq k\}} dS^k_u \right]$$

$$\quad + \frac{1}{N} \left[ \lambda N t, \mu \int_0^t (Q_u^N \land N) \, du \right]$$

- The last three terms are zero (w.p. 1) since they are quadratic covariations of finite variation (w.p. 1) processes and continuous ones
- The first term is zero (w.p. 1) since $A_t^N$ and $S^k_t$ are independent and thus share no jumps w.p. 1
- So finally also $\langle \bar{M}_t^{A,N}, \bar{M}_t^{S,N} \rangle = 0$
Showing $\langle \tilde{M}_t^{A,N} \rangle \Rightarrow \mu t$

This is also straightforward since:

$\triangleright$ $\tilde{M}_t^{A,N} = \frac{1}{\sqrt{N}} (A_t^N - \lambda_N t)$

$\triangleright$ So

$$[\tilde{M}_t^{A,N}] = \frac{1}{N} [A_t^N, A_t^N] - \frac{2}{N} [A_t^N, \lambda_N t] + \frac{1}{N} [\lambda_N t, \lambda_N t]$$

$\triangleright$ The last two terms are zero (w.p. 1) since they are quadratic covariations of finite variation (w.p. 1) processes and continuous ones

$\triangleright$ $[A_t^N, A_t^N] = A_t^N$ (w.p. 1) since $A_t^N$ is a non-decreasing unit-jump process

$\triangleright$ So $[\tilde{M}_t^{A,N}] = \frac{1}{N} A_t^N$ and thus $\langle \tilde{M}_t^{A,N} \rangle = \frac{\lambda_N t}{N}$

$\triangleright$ Recall that $\lambda_N = \mu N - \mu \beta \sqrt{N}$ and the result follows
Showing $\langle \bar{M}_t^{S,N} \rangle \Rightarrow \mu t$

This is less straightforward:

- $\bar{M}_t^{S,N} = \frac{1}{\sqrt{N}} (\sum_{k=1}^{N} \int_0^t \mathbf{1}_{\{Q_u^N \geq k\}} \, dS_u^k - \mu \int_0^t (Q_u^N \wedge N) \, du)$

- So

$$[ar{M}_t^{S,N}] = \frac{1}{N} \left[ \sum_{k=1}^{N} \int_0^t \mathbf{1}_{\{Q_u^N \geq k\}} \, dS_u^k, \sum_{k=1}^{N} \int_0^t \mathbf{1}_{\{Q_u^N \geq k\}} \, dS_u^k \right]$$

$$- \frac{2}{N} \left[ \sum_{k=1}^{N} \int_0^t \mathbf{1}_{\{Q_u^N \geq k\}} \, dS_u^k, \mu \int_0^t (Q_u^N \wedge N) \, du \right]$$

$$+ \frac{1}{N} \left[ \mu \int_0^t (Q_u^N \wedge N) \, du, \mu \int_0^t (Q_u^N \wedge N) \, du \right]$$

- As before, the last two terms are zero (w.p. 1)

- And

$$[\sum_{k=1}^{N} \int_0^t \mathbf{1}_{\{Q_u^N \geq k\}} \, dS_u^k, \sum_{k=1}^{N} \int_0^t \mathbf{1}_{\{Q_u^N \geq k\}} \, dS_u^k] = \sum_{k=1}^{N} \int_0^t \mathbf{1}_{\{Q_u^N \geq k\}} \, dS_u^k$$

- So $[\bar{M}_t^{S,N}] = \frac{1}{N} \sum_{k=1}^{N} \int_0^t \mathbf{1}_{\{Q_u^N \geq k\}} \, dS_u^k$ and thus

$$\langle \bar{M}_t^{S,N} \rangle = \frac{1}{N} \mu \int_0^t (Q_u^N \wedge N) \, du$$

- So it remains to show that $\frac{1}{N} \mu \int_0^t (Q_u^N \wedge N) \, du \Rightarrow \mu t$
Showing \( \frac{1}{N} \mu \int_0^t (Q^N_u \wedge N) \, du \Rightarrow \mu t \)

- Would follow by the continuous mapping theorem from \( \frac{Q^N_t \wedge N}{N} \Rightarrow 1 \) in \( D \)
  where 1 is the identically-one process, which is implied by
  \[ \frac{Q^N_t - N}{\sqrt{N}} = \frac{\bar{Q}^N_t}{\sqrt{N}} \Rightarrow 0 \text{ in } D \]

**Definition: Stochastic boundedness in \( D^k \)**

A sequence of stochastic processes \( (X^N_t)_{t \geq 0} \) for \( N \geq 1 \) is **stochastically bounded** in \( D^k \) if the sequence of random variables \( \sup_{t \in [0,T]} \|X^N_t\| \) in \( \mathbb{R} \) is **tight** for each \( T > 0 \), that is, if for any \( \epsilon > 0 \), there exists some compact subset \( K \) of \( \mathbb{R} \), such that \( \mathbb{P}\{\sup_{t \in [0,T]} \|X^N_t\| \in K\} > 1 - \epsilon \) for all \( n \geq 1 \).

**Theorem: Functional weak law of large numbers from stochastic boundedness**

Let \( (X^N_t)_{t \geq 0} \) be a sequence of stochastically-bounded stochastic processes in \( D^k \) and let \( \alpha_N \to \infty \) as \( N \to \infty \) be real numbers. Then \( \frac{X^N_t}{\alpha_N} \Rightarrow 0 \) in \( D^k \).

So it remains to show that \( \bar{Q}^N_t \) is stochastically bounded in \( D \)
Showing that $\bar{Q}_t^N$ is stochastically bounded

- Recall: $\bar{Q}_t^N = \bar{Q}_0^N + \bar{M}_t^{A,N} - \bar{M}_t^{S,N} - \mu \beta t - \mu \int_0^t (\bar{Q}_u^N \land 0) \, du$
- Stochastic boundedness is preserved under the summation and integration above (e.g. Lemmas 5.4, 5.5 in [5])
- It then suffices to show stochastic boundedness of $\bar{Q}_0^N$, $\bar{M}_t^{A,N}$ and $\bar{M}_t^{S,N}$
- Stochastic boundedness of $\bar{Q}_0^N$ by assuming that $\mathbb{E}[\bar{Q}_0^N] < \infty$ for $N \geq 1$

**Theorem: Lenglart–Rebolledo inequality**

Suppose that $(M_t)_{t \geq 0}$ is a square-integrable martingale wrt. $(\mathcal{F}_t)_{t \geq 0}$ such that its predictable quadratic variation $\langle M_t \rangle$ exists. Then for all $c > 0$ and $d > 0$:

$$\mathbb{P}\{\sup_{t \in [0,T]} |M_t| > c\} \leq \frac{d}{c^2} + \mathbb{P}\{\langle M_T \rangle > d\}$$

(See e.g. pg. 66 in [6] for a proof)

**Theorem: Stochastic boundedness of square-integrable martingales**

Suppose that $(M_{t}^N)_{t \geq 0}$ for $N \geq 1$ is a sequence of square-integrable martingales wrt. $(\mathcal{F}_{t}^N)_{t \geq 0}$ such that the predictable quadratic variations $\langle M_{t}^N \rangle$ exist. Then if the sequence of real random variables $\langle M_{T}^N \rangle$ is tight for each $T > 0$, the sequence $(M_{t}^N)_{t \geq 0}$ of stochastic processes is stochastically bounded in $D$.

Suffices to show tightness of $\langle \bar{M}_{T}^{A,N} \rangle = \frac{\lambda N T}{N}$ and $\langle \bar{M}_{T}^{S,N} \rangle = \frac{1}{N} \mu \int_0^T (Q_u^N \land N) \, du$


Tightness of $\langle \bar{M}_{T}^{A,N} \rangle = \frac{\lambda_{N} T}{N}$ and $\langle \bar{M}_{T}^{S,N} \rangle = \frac{1}{N} \mu \int_{0}^{T} (Q_{u}^{N} \wedge N) \, du$

- Tightness of $\langle \bar{M}_{T}^{A,N} \rangle = \frac{\lambda_{N} T}{N}$ is immediate since $\lambda_{N} = \mu N - \mu \beta \sqrt{N}$ and so $|\langle \bar{M}_{T}^{A,N} \rangle| \leq \mu T (1 + |\beta|)$

- For $\langle \bar{M}_{T}^{S,N} \rangle = \frac{1}{N} \mu \int_{0}^{T} (Q_{u}^{N} \wedge N) \, du$, apply the crude inequality:

  $$\frac{1}{N} \mu \int_{0}^{T} (Q_{u}^{N} \wedge N) \, du \leq \frac{1}{N} \mu T (Q_{0}^{N} + A_{T}^{N})$$

- Recall that by assumption $\bar{Q}_{0}^{N} = \frac{Q_{0}^{N} - N}{\sqrt{N}} \Rightarrow X_{0}$, so $\frac{Q_{0}^{N} - N}{N} \Rightarrow 0$ and $\frac{Q_{0}^{N}}{N} \Rightarrow 1$. Then $\frac{Q_{0}^{N}}{N}$ is thus tight

- Finally $A_{T}^{N} \overset{d}{=} A_{T}^{N} \overset{d}{=} A_{T}^{N}$ for a rate-1 Poisson process $(A_{t})_{t \geq 0}$ and by the strong law of large numbers for the Poisson process $\frac{A_{T}^{N} \overline{\lambda}_{N}}{T \lambda_{N}} \overset{d}{=} 1 \text{ w.p. 1 as } N \to \infty$ so $\frac{A_{T}^{N} \overline{\lambda}_{N}}{N - \beta \sqrt{N}} \overset{d}{=} \frac{1}{1 - \beta \sqrt{N}} A_{T}^{N} \overset{d}{=} \mu T$ and thus $\frac{A_{T}^{N}}{N} \overset{d}{=} \mu T \text{ w.p. 1 as } N \to \infty$ and $\frac{A_{T}^{N}}{N}$ is thus tight
General martingale recipe for heavy-traffic limits of Markovian queues

- Establish semi-martingale decomposition of queueing process of interest
  - Construct the process in terms of integral equations with respect to Poisson processes
  - Find the Doob–Meyer decomposition of each term using the martingale integration theorem
- Fix an appropriate heavy-traffic scaling regime and obtain the semi-martingale decompositions of the sequence of queueing processes
- Identify the predictable quadratic covariation processes of the martingales
- Prove convergence of the predictable quadratic covariations as required by the martingale FCLT
  - In some cases, this is trivial
  - In others, an auxiliary fluid limit is required
    - The fluid limit is obtained by showing a stochastic boundedness property of the queueing process
    - And this usually follows from tightness of the martingales’ predictable quadratic variations
- Finally, apply the continuous mapping theorem to transmit convergence of the scaled martingales to that of the queueing process
More interesting example

Allow $K$ customer classes with service rates $\mu_j$ ($1 \leq j \leq K$) and arrival rate of each class $p_j \lambda$

- $(A_t)_{t \geq 0}$ (rate $\lambda$)
- $\{(S^{j,l}_t)_{t \geq 0} : 1 \leq j \leq K, 1 \leq l \leq N\}$ (rate $\mu_j$)
- $\{\alpha_j : j \geq 1\}$ i.i.d. $\alpha_j = 1$ w.p. $p_1$, $\alpha_j = 2$ w.p. $p_2$
- $Q_t$ is the total number of customers in the queue (not in service)
- $\hat{Q}^j_t$ is the number of class-$j$ customers in service

\[
D^j_t = \sum_{l=1}^{N} \left. \int_{0}^{t} 1\{\hat{Q}^j_s \geq l\} 1\{Q_s > 0\} \, dS^{j,l}_s \right|_{t=0} \quad D_t = \sum_{j=1}^{K} D^j_t
\]

Customers who left queue and entered service in $[0, t]$ due to termination of service of a $j$-customer

\[
B_t = D_t + \left. \int_{0}^{t} 1\{\sum_{j=1}^{K} \hat{Q}^j_s < N\} \, dA_s \right|_{t=0}
\]

Total customers who started service in $[0, t]$

More interesting example

Allow $K$ customer classes with service rates $\mu_j$ ($1 \leq j \leq K$) and arrival rate of each class $p_j \lambda$

- $(A_t)_{t \geq 0}$ (rate $\lambda$)
- $\{(S_{t,l}^j)_{t \geq 0} : 1 \leq j \leq K, 1 \leq l \leq N\}$ (rate $\mu_j$)
- $\{\alpha_j : j \geq 1\}$ i.i.d. $\alpha_j = 1$ w.p. $p_1$, $\alpha_j = 2$ w.p. $p_2$
- $Q_t$ is the total number of customers in the queue (not in service)
- $\hat{Q}_t^j$ is the number of class-$j$ customers in service

$$Q_t = Q_0 + \int_0^t \mathbf{1}\{\sum_{j=1}^K \hat{Q}_s^j = N\} \, dA_s - D_t$$

More interesting example

Allow $K$ customer classes with service rates $\mu_j \ (1 \leq j \leq K)$ and arrival rate of each class $p_j \lambda$

- $(A_t)_{t\geq0}$ (rate $\lambda$)
- $\{(S^{j,l}_t)_{t\geq0} : 1 \leq j \leq K, 1 \leq l \leq N\}$ (rate $\mu_j$)
- $\{\alpha_j : j \geq 1\}$ i.i.d. $\alpha_j = 1$ w.p. $p_1$, $\alpha_j = 2$ w.p. $p_2$
- $Q_t$ is the total number of customers in the queue (not in service)
- $\hat{Q}_t^j$ is the number of class-$j$ customers in service

$$\hat{Q}_t^j = \hat{Q}_0^j + \int_0^t \left[ \mathbf{1}_{\{\sum_{j=1}^K \hat{Q}_s^j < N\}} \mathbf{1}_{\{\alpha_{B_s} = j\}} \right] dA_s$$

Class-$j$ customers who go straight into service

$$+ \sum_{j=1}^K \int_0^t \mathbf{1}_{\{\alpha_{B_s} = j\}} dD_s^j$$

Class-$j$ customers who go into service after a completion

$$- \sum_{l=1}^N \int_0^t \mathbf{1}_{\{\hat{Q}_s^j \geq l\}} dS_s^{j,l}$$

Class-$j$ customers leaving due to their own completion

More interesting example

Allow $K$ customer classes with service rates $\mu_j$ ($1 \leq j \leq K$) and arrival rate of each class $p_j \lambda$

- $(A_t)_{t \geq 0}$ (rate $\lambda$)
- $\{(S_{t_i}^j)_{t \geq 0} : 1 \leq j \leq K, 1 \leq i \leq N\}$ (rate $\mu_j$)
- $\{\alpha_j : j \geq 1\}$ i.i.d. $\alpha_j = 1$ w.p. $p_1$, $\alpha_j = 2$ w.p. $p_2$
- $Q_t$ is the total number of customers in the queue (not in service)
- $\hat{Q}_t^j$ is the number of class-$j$ customers in service

$$Q_t = Q_0 + \lambda \int_0^t 1_{\{\sum_{j=1}^K \hat{Q}_s^j = N\}} \, ds - \sum_{j=1}^K \mu_j \int_0^t 1_{\{Q_s > 0\}} \hat{Q}_s^j \, ds + M_t$$

More interesting example

Allow $K$ customer classes with service rates $\mu_j$ ($1 \leq j \leq K$) and arrival rate of each class $p_j \lambda$

$\begin{itemize}
\item (\At)_{t \geq 0}$ (rate $\lambda$)
\item $\{(\St^j_\cdot)^{\cdot} : 1 \leq j \leq K, 1 \leq l \leq N\}$ (rate $\mu_j$)
\item $\{\alpha_j : j \geq 1\}$ i.i.d. $\alpha_j = 1$ w.p. $p_1$, $\alpha_j = 2$ w.p. $p_2$
\item $Q_t$ is the total number of customers in the queue (not in service)
\item $\hat{Q}_t^j$ is the number of class-$j$ customers in service
\end{itemize}

$$
\hat{Q}_t^j = \hat{Q}_0^j + p_j \lambda \int_0^t \mathbf{1}_{\left\{\sum_{j=1}^K \hat{Q}_s^j < N\right\}} ds \\
+ p_j \sum_{j=1}^K \mu_j \int_0^t \mathbf{1}_{\{Q_s > 0\}} \hat{Q}_s^j ds \\
- \mu_j \int_0^t \hat{Q}_s^j ds + \hat{M}_t^j
$$


Thank you!