Martingale Methods: Definitions & Examples
Karlin & Taylor, *A First Course in Stochastic Processes, pp 238–253*

MARTINGALES

We've already encountered and used martingales in this course to help study the hitting-times of Markov processes. Informally a martingale is simply a stochastic process M_t defined on some probability space $(\Omega, \mathcal{F}, \mathbf{P})$ that is "conditionally constant," *i.e.*., whose predicted value at any future time $s > t$ is the same as its present value at the time t of prediction. Formally we represent what is known at time t in the form of an increasing family of σ-algebras $\mathcal{F}_t \subset \mathcal{F}$, possibly those generated by a process $[X_s : s \leq t]$ or even by the martingale itself, $\mathcal{F}_t = \sigma([M_s : s \leq t])$, and require that $\mathbf{E}[|M_t|] < \infty$ for each t (so the conditional expectation below is well-defined) and that

$$M_t = \mathbf{E}[M_s \mid \mathcal{F}_t]$$

for each $t < s$. For discrete-time processes (like functions of the Markov chains we looked at before) it is only necessary to take $s = t + 1$, and we usually take $\mathcal{F}_t = \sigma[X_i : i \leq t]$ and write

$$M_t = \mathbf{E}[M_{t+1} \mid X_0, ..., X_t].$$

There are several "big theorems" about martingales that make them useful in studying stochastic processes:

Optional Sampling Theorem:

If τ is a *stopping time* or *Markov time*, *i.e.*., a random time that "doesn't depend on the future" (technically the requirement is that the event $[\tau \leq t]$ should be in $\mathcal{F}_t$ for each t), and if M_t is a martingale, and if either $\mathbf{E}[\tau] < \infty$ or if $\{M_t\}$ is uniformly integrable, then

$$M_t = \mathbf{E}[M_{\tau \vee t} | \mathcal{F}_t]$$

and in particular $x = \mathbf{E}[M_\tau | M_0 = x]$. More generally, if $\{\tau_n\}$ is an increasing sequence of stopping times with $\mathbf{E}[\tau_n] < \infty$ or $\{M_t\}$ uniformly integrable, then $Y_n = M_{\tau_n}$ is a martingale.

Maximal Inequalities:

If M_t is a martingale and if $t \leq \infty$ then

$$\begin{aligned}
\mathbf{P}[\sup_{s \leq t} M_s \geq \lambda] &\leq \frac{1}{\lambda} \mathbf{E}[M_t^+] \\
\mathbf{P}[\min_{s \leq t} M_s \leq -\lambda] &\leq \frac{1}{\lambda} (\mathbf{E}[M_t^+] - \mathbf{E}[M_0]) \\
\mathbf{E}[\sup_{s \leq t} |M_s|^p] &\leq \left(\frac{p}{p-1} \right)^p \sup_{s \leq t} \mathbf{E}[|M_s|^p] && (p > 1) \\
\mathbf{E}[\sup_{s \leq t} |M_s|] &\leq \frac{e}{e-1} \sup_{s \leq t} \mathbf{E}[|M_s| \log^+(|M_s|)] && (p = 1)
\end{aligned}$$

Martingale Path Regularity:

If M_t is a martingale and $a < b$ denote by $\nu_{[a,b]}^{(t)}$ the number of "upcrossings" of the interval $[a, b]$ by M_s prior to time t, the number of times it passes from below a to above b; then

$$\mathbf{E}[\nu_{[a,b]}^{(t)}] \leq \frac{\mathbf{E}[M_t^+] + |a|}{b - a}$$

and, in particular, martingale paths don't oscillate infinitely often– thus they have left and right limits at every point. This is also the key lemma to prove:

Martingale Convergence Theorems:

Let M_t be a martingale. Then:

For *any* martingale M_t, there exists an RV $M_{-\infty}$ such that
$$\lim_{t\to-\infty} M_t = M_{-\infty} \ a.s. \tag{Backwards MCT}$$
If also $\sup_{s<\infty} \mathbf{E}[M_s^+] < \infty$, then there exists an RV M_∞ such that
$$\lim_{t\to\infty} M_t = M_\infty \ a.s. \tag{Forwards MCT}$$
If also $\{|M_s|^p\}$ is uniformly integrable, then $M_\infty \in L^p$ and
$$\lim_{t\to\infty} M_t = M_\infty \text{ in } L^p. \tag{L^p}$$

Martingale Problem for Continuous-Time Markhov Chains:

Let Q_{jk}^t be a (possibly time-dependent) Markov transition matrix on a state space $\mathcal{S}$. Then an $\mathcal{S}$-valued process X_t is a Markov chain with transition matrix $Q_{jk}(t)$ if and only if, for all functions $\phi : \mathcal{S} \to \mathbb{R}$, the process

$$M_\phi(t) = \phi(X_t) - \phi(X_0) - \int_0^t \Big[\sum_{j\neq i=X_s} Q_{ij}^s \left[\phi(j) - \phi(i)\right]\Big]\, ds$$

is a martingale. Similar characterizations apply to discrete-time Markov chains and to continuous-time Markov processes with non-discrete state space $\mathcal{S}$. This is the most powerful and general way known for *constructing* Markov processes.

Doob's Martingale:

Let Y be any $\mathcal{F}$-measurable L^1 random variable and let $M_t = \mathbf{E}[Y \mid \mathcal{F}_t]$ be the best prediction of Y available at time t. Then M_t is a uniformly-integrable martingale.

To summarize, martingales are important because:

1. They have close connections with Markov processes;
2. Their expectations at stopping times are easy to compute;
3. They offer a tool for bounding the maximum and minimum of processes;
4. They offer a tool for establishing path regularity of processes;
5. They offer a tool for establishing the *a.s.* convergence of certain random sequences;
6. They are important for modeling economic and statistical time series which are, in some sense, predictions.

Examples:

1. Partial sums: $S_n = \Sigma_{i=1}^n X_i$
2. Stochastic Integral: Let X_n be an IID Bernoulli sequence with probability p; you can bet any fraction F_n you like of your (previous) fortune M_{n-1} at odds $p : 1-p$, so your new fortune is $M_{n-1}(1 - F_n(1 - X_n/p))$. If $F_n \in \sigma[X_1 \cdots X_{n-1}$, M_n is a martingale. Note that
$$M_n = M_0 + \sum_{i=1}^n F_i M_{i-1}[Y_n - Y_{n-1}]$$
for the martingale $Y_n = (S_n - np)/p$.
3. Variance of a Sum: $M_n = \big(\sum_{i=1}^n Y_i\big)^2 - n\sigma^2$, where $\mathbf{E} Y_i Y_j = \sigma^2 \delta_{ij}$
4. Radon-Nikodym Derivatives: $M_n(\omega) = \mathbf{E}[f(\omega) \mid \sigma\{(\frac{i}{2^n}, \frac{j}{2^n}]\}]$

Submartingales: $\mathbf{E}[X_t^+] < \infty$, $\mathbf{E}[X_s \mid \mathcal{F}_t] \geq X_t$, $X_t \in \mathcal{F}_t$. Jensen's inequality: if ϕ convex, then $\phi(X_t)$ is a submartingale if $\mathbf{E}[\phi(X_t)^+] < \infty$.