

Scratchwork: Martingales

What is a martingale? We need to learn some conditional probability. The definition of “martingale” is very carefully done and takes a while to state.

#1 A *filtered* space $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, \mathbb{P})$ where $(\Omega, \mathcal{F}, \mathbb{P})$ is a “probability triple” and $\{\mathcal{F}_n : n \geq 0\}$ is a **filtration**. An increasing family of sub- σ -algebras of \mathcal{F} :

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}$$

$\omega \in \Omega$ is some realization (which we never directly observe) and the information available to us at time n is the sigma algebra \mathcal{F}_n .

#2 A process $X = (X_n : n \geq 0)$ is called *adapted* if for each n , X_n is \mathcal{F}_n -measurable.

#3 A process X is called a **martingale** (relative to $(\mathcal{F}_n, \mathbb{P})$) if

- X is adapted
- $\mathbb{E}(|X_n|) < \infty$ for all times $n \geq 0$
- $\mathbb{E}[X_n | \mathcal{F}_{n-1}]$ almost surely for $n \geq 1$.

Example A map $T : \Omega \rightarrow \{0, 1, 2, \dots, \infty\}$ is called a *stopping time* if

$$\{T \leq n\} = \{\omega : T(\omega) \leq n\} \in \mathcal{F}_n \text{ for all } n \leq \infty$$

And we get the Martingale Theorem. These definitions occurred *after* much exploration and much getting things wrong. We have the martingale theorem:

$$\mathbb{E}(X_{T \wedge n}) = \mathbb{E}(X_0)$$

So even though the stopping time is random, we get a very easy expectation computations. The distributions of $X_{T \wedge n}$ and X_0 might be totally different. We are saying that the values of $\mathbb{E}[X]$ are the same at both “times” $t = 0$ and $t = T \wedge n$.

In practice, our probability space will be $\Omega = \mathbb{R}$ and \mathbb{P} will be Lebesgue measure¹ dx and the sigma algebra \mathcal{F} is generated by intervals $[a, b] \subseteq \mathbb{R}$.

Our dynamical systems textbooks gives the example. Let \mathcal{A} be the partition $\{[0, \frac{1}{2}], [\frac{1}{2}, 1]\}$ and in general

$$\mathcal{A}_n = \left\{ \left[0, \frac{1}{2^n}\right], \left[\frac{1}{2^n}, \frac{2}{2^n}\right], \dots, \left[1 - \frac{1}{2^n}, 1\right] \right\}$$

Surely, we must have $\mathbb{E}[f | \mathcal{A}_n] \rightarrow \mathbb{E}[f | \mathcal{B}]$ as our measurements finer, we should be able to get the expectation we want. Right?

¹Even that is too complicated, it’s just anything we can compute with “intervals”, i.e. with a ruler. But that framing can get too inaccurate, but then we do use instead? So in actual random process theory the examples are easier the end result is the same. The modern standard is Lebesgue measure.

References

- [1] David Williams **Probability with Martingales** (Cambridge Mathematical Textbooks) Cambridge University Press, 1991 / 2011.
- [2] Manfred Einsiedler, Thomas Ward. **Ergodic Theory: with a view towards Number Theory** GTM #259 Springer, 2011.