THE COIN TOSS

JOHN THICKSTUN

While studying probability, I have often wished I had a concrete example of a martingale to work with. Below is an (almost[1]) complete construction, fully worked out, of the classic martingale of symmetric bets on successive coin tosses. This is not the way a probabilist thinks about stochastic processes and martingales, which explains why a worked out example like this–while pedagogically useful–does not typically appear in the literature. After completing the construction, I make some remarks on why probabilists do not think in these terms.

Let $S = \{H, T\}$ be a two element set with members H and T. We will operate on the space of outcomes $\Omega = S^{\mathbb{N}}$. This is an indexed set with with $\omega_n \in S$ for each $\omega \in \Omega$, $n \in \mathbb{N}$. The idea is that the n^{th} index of an outcome ω models a bit of information at time n: for example the result of a coin flip. Let $\Omega_n = S^n$ and $\mathcal{F}_n = \otimes_{k=1}^n \mathcal{P}(S) = \mathcal{P}(\Omega_n)$. Note that $\mathcal{F}_n$ is a σ-algebra. Furthermore, there is a canonical injection of $\mathcal{F}_n$ into $\mathcal{P}(\Omega)$. Define the projection operator $\Pi : \Omega \to \Omega_n$ where $\Pi(\omega)$ is the unique $\omega^n \in \Omega_n$ such that $\omega_k = \omega_k^n$, $1 \leq k \leq n$. Intuitively, Π "forgets" what happens after time n. The injection from $\mathcal{F}_n$ into $\mathcal{P}(\Omega)$ is then defined by the pre-image of Π. We will identify $\mathcal{F}_n$ with $\Pi^{-1}(\mathcal{F}_n)$, giving us $\mathcal{F}_n \subset \mathcal{P}(\Omega)$.

It is easy to check that $\mathcal{F}_n$ (understood as a subset of $\mathcal{P}(\Omega)$) is a σ-algebra (making Π^{-1} an injective homomorphism). Moreover $\mathcal{F}_m \subset \mathcal{F}_n$ for $m < n$. Therefore $\mathcal{F}_n$ is a filtration $\mathbb{F}$. We now proceed to construct a probability measure on Ω. Define P_k on $\mathcal{P}(S)$ by $P_k(\emptyset) = 0$, $P_k\{H\} = P_k\{T\} = 1/2$, and $P_k\{H, T\} = 1$. It is trivial to check that P_k is a probability measure. Now define $P_n : \Omega_n \to \mathbb{R}$ by $P_n(\omega) = \prod_{k=1}^{\infty} P_k(\omega_n)$. Then each P_n is a probability measure, the measures P_n are Kolmogorov consistent, and by the Kolmogorov extension theorem P_n extends to a probability measure P on the product σ-algebra $\mathcal{F} = \otimes_{n\in\mathbb{N}} \mathcal{F}_n \subset \mathcal{P}(\Omega)$. Thus we have a probability space $(\Omega, \mathcal{F}, P)$.

Let X_k be a random variable on S defined by $X_k(\omega) = \mathbb{1}_{\omega=H} - \mathbb{1}_{\omega=T}$. We intend for X_k to represent the value of a one dollar bet on the outcome of a coin toss. Consider the sum S_n on Ω_n defined by $S_n = \sum_{k=1}^n X_k$. Then S_n represents the cumulative outcome of our bets on X_k. Note that if $M \subset \Omega_n$, $S_n^{-1}(M) \in \mathcal{F}_n$ (trivially, since these are power sets). Therefore S_n is adapted to $\mathbb{F}$. In fact, it is not hard to check that $\mathbb{F}$ is the natural filtration of S_n. Furthermore, S_n is a martingale. To see this, observe that $\mathbb{E}|S_n| \leq \sum_{k=1}^{\infty} |X_k| = n < \infty$.

[1]We appeal to the Kolmogorov extension theorem, which is non-constructive, to establish a probability measure on an infinite state space.

Moreover, if $m < n$, then

$$\mathbb{E}(S_n|\mathcal{F}_m) = \mathbb{E}\left(\sum_{k=1}^{n} X_k|\mathcal{F}_m\right) = \sum_{k=1}^{m} \mathbb{E}(X_k|\mathcal{F}_m) + \sum_{k=m+1}^{n} \mathbb{E}(X_k|\mathcal{F}_m)$$

$$= \mathbb{E}\left(\sum_{k=1}^{m} X_k|\mathcal{F}_m\right) + \sum_{k=m+1}^{n} \mathbb{E}X_k = \mathbb{E}(S_m|\mathcal{F}_m) + 0 = S_m.$$

It may be helpful to step back at this point and meditate on why processes are not usually presented in this explicit, constructive way. Unlike our toy example, the set Ω is typically understood to be a very complex space, perhaps even the configuration state of the universe at points in time and space. Then $\mathcal{F}$ in our coin toss example would be taken to be the σ-algebra generated by events (understood as subsets of $\mathcal{P}(\Omega)$ via Π) that influence the tosses: the vigor of the toss, the height, the wind speed, etc..

Let's flesh out the technical details of this new model. We will introduce an intermediary r.v. $C_n : \Omega \to S$ which maps all the combined physical influences to the outcome of the coin toss. When we say that $\mathcal{F}$ is the events that influence the tosses, what we really mean is that C_n is $\mathcal{F}$-measurable. The outcome of our bet X_k then becomes a function of the coin toss, rather than a function on the probability space. Of course, we can repeat this process ad-nauseum, building ever more random variables into our model and pushing the ultimate source of uncertainty–the probability outcomes–further and further away.

Suppose we construct our model in such a way that random variables capture all the relevant information, and all the relationships we are able to describe: in our example this is the outcome of each coin toss, the outcome of each bet, and the running total of the bets. Then by definition we have nothing to say about the space of outcomes and events $(\Omega, \mathcal{F})$. All that matters is that our r.v.'s are $\mathcal{F}$-measurable. Probabilists therefore find it convenient to dim the lights on $(\Omega, \mathcal{F})$. Instead of writing $X(\omega)$ we simply write X. Instead of constructing an explicit filtration $\mathcal{F}$ we just take the smallest one that makes X measurable: $\sigma(X)$.