Random Walk, Martingales and Markov processes

1 Introduction

Evans discusses conditional expectations in §H of chapter 2 of his lecture notes [1]. The same topics can be found in § 2.1 of [4] where form the definitions of section 2 are taken. The definition of martingale follows instead § I of chapter 2 of [1]. A nice mathematical presentation of martingales in the case of countable state space is given in §11 of chapter 1 of [3]. The solution of the master equation for the Poisson process can be also found in §3.8.3 of [2].

2 Some definitions

Definition 2.1. *Let* $(\Omega\,, \mathcal{F}_n\,, \mathrm{P})$ *be a probability space. A (discrete time) filtration is an increasing sequence* $\boldsymbol{\mathcal{F}} := \{\mathcal{F}_k\}_{k=0}^n$ *of* σ*-algebras* $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \dots \mathcal{F}_n$*. The quadruple* $(\Omega\,, \mathcal{F}_n\,, \boldsymbol{\mathcal{F}}\,, \mathrm{P})$ *is called a filtered probability space*

Definition 2.2. *A stochastic process is just a collection of random variables* $\{\boldsymbol{\xi}_t, t \geq 0\}$*, indexed by a time parameter* t *discrete or continuous.*

Definition 2.3. *Let* $(\Omega\,, \mathcal{F}_n\,, \boldsymbol{\mathcal{F}}\,, \mathrm{P})$ *be a filtered probability space. A stochastic process* $\{\boldsymbol{\xi}_i\}_{i=1}^n$ *is called* $\mathcal{F}_n$*-adapted* $\boldsymbol{\xi}_n$ *if is* $\mathcal{F}_n$*-measurable for every* n*, and is called* $\mathcal{F}_n$*-predictable if* $\boldsymbol{\xi}_n$ *is* $\mathcal{F}_{n-1}$*-measurable for every* n*.*

Definition 2.4. *Let* $(\Omega\,, \mathcal{F}\,, \mathrm{P})$ *be a probability space and* $\{\boldsymbol{\xi}_n\}_n$ *be a stochastic process. The filtration generated by* $\{\boldsymbol{\xi}_n\}_n$ *is defined as* $\mathcal{F}_n^{\boldsymbol{\xi}} = \sigma\left(\boldsymbol{\xi}_1\,, \dots\,, \boldsymbol{\xi}_n\right)$ *and the process is* $\mathcal{F}_n^{\boldsymbol{\xi}}$*-adapted by construction.*

3 Conditional Expectation, Heuristics

Let $\boldsymbol{\xi}$ an integrable random variable

$$\mathrm{E} \parallel \boldsymbol{\xi} \parallel < \infty \tag{3.1}$$

on the probability space $(\Omega\,, \mathcal{F}\,, \mathrm{P})$. Let then $\mathcal{G}$ a finite dimensional σ-algebra generated by the atomic decomposition (or partition) $\{A_i\}_{i=1}$ of Ω i.e.

$$\Omega = \bigcup_{i=1}^{n} A_i \tag{3.2}$$

In such a case we define the *conditional expectation* of $\boldsymbol{\xi}$ with respect to $\mathcal{G}$ as the random variable

$$\mathrm{E}\,\{\boldsymbol{\xi}|\mathcal{G}\} = \sum_{i=1}^{n} \frac{\mathrm{E}\,\{\boldsymbol{\xi}\chi_{A_i}\}}{\mathrm{P}(A_i)} \chi_{A_i}\,(\omega) \tag{3.3}$$

where $\omega \in \Omega$ and

$$\chi_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases} \tag{3.4}$$

Note that if the discrete random variable $\chi_A(\omega)$ is independent of $\boldsymbol{\xi}$ the definition implies immediately

$$\mathrm{E}\left\{\boldsymbol{\xi}|\mathcal{G}\right\} = \mathrm{E}\boldsymbol{\xi} \sum_{i=1}^{n} \frac{\mathrm{E}\left\{\chi_{A_i}\right\}}{\mathrm{P}(A_i)} \chi_{A_i}\left(\omega\right) \equiv \mathrm{E}\boldsymbol{\xi} \sum_{i=1}^{n} \chi_{A_i}\left(\omega\right) \tag{3.5}$$

whence

$$\mathrm{E}\left\{\boldsymbol{\xi}|\mathcal{G}\right\} = \mathrm{E}\boldsymbol{\xi} \tag{3.6}$$

by virtue of

$$\sum_{i=1}^{n} \chi_{A_i}\left(\omega\right) = \chi_{\Omega}\left(\omega\right) = 1 \tag{3.7}$$

The latter equation just states that $\{A_i\}_{i=1}$ is a partition of Ω and that $\chi_\Omega(\omega)$ reduces to the trivial random variable equal to the unity whenever it is sampled i.e. $\forall\, \omega \,\in\, \Omega$.

Definition 3.1. *Let $(\Omega\,, \mathcal{F}\,, \mathrm{P})$ be a probability space and suppose $\mathcal{F}'$ is a σ-algebra, $\mathcal{F}' \subseteq \mathcal{F}$. If*

$$\boldsymbol{\xi}\colon \omega \mapsto \mathbb{R}^d \tag{3.8}$$

is an integrable random variable, we define

$$\boldsymbol{\xi}' := \mathrm{E}\left\{\boldsymbol{\xi}|\mathcal{F}'\right\} \tag{3.9}$$

to be any random variable on Ω such that

i $\boldsymbol{\xi}'$ is $\mathcal{F}'$-measurable;

ii For all $F' \in \mathcal{F}'$ the identity

$$\mathrm{E}_{\mathrm{P}}\left\{\chi_{F'}\boldsymbol{\xi}\right\} = \mathrm{E}_{\mathrm{P}}\left\{\chi_{F'}\boldsymbol{\xi}'\right\} \tag{3.10}$$

holds true.

If $\mathcal{F}' = \sigma\left(\boldsymbol{\eta}\right)$ i.e. is the σ-algebra generated by $\boldsymbol{\eta}$ we will write

$$\boldsymbol{\xi}' := \mathrm{E}\left\{\boldsymbol{\xi}|\mathcal{F}'\right\} \equiv \mathrm{E}\left\{\boldsymbol{\xi}|\boldsymbol{\eta}\right\} \tag{3.11}$$

For more details please read § H of chapter 2 and/or § 2.1 of [4].

4 Martingales

Definition 4.1. *Let $(\Omega\,, \mathcal{F}_n\,, \boldsymbol{\mathcal{F}}\,, \mathrm{P})$ be a filtered probability space and $\{\boldsymbol{\xi}_i\}_{i=1}^{n}$ a $\mathcal{F}_n$-adapted stochastic process such that*

$$\mathrm{E}\left|\xi_i\right| \,<\, \infty \qquad \forall i \tag{4.1}$$

If

$$\xi_k = \mathrm{E}\left\{\xi_j|\mathcal{F}_k\right\} \qquad a.s \qquad \forall\, j \,\geq\, k$$

holds true we say that $\{\xi_i\}_{i=1}^{n}$ is a (discrete) martingale.

More generally we will consider stochastic processes $\{\boldsymbol{\xi}_t, t \in T\}$ with T being a subset or coinciding with either $\mathbb{R}_+$ or $\mathbb{N}$. Furthermore we can posit that for any fixed $t \in T$ the random variable $\boldsymbol{\xi}_t$ takes values in a *state space* $\mathbb{S}$ which may be finite (as for the random walk), countable or (a subset of) $\mathbb{R}^d$. In such a case the general definition is

Definition 4.2. *Let $\boldsymbol{\mathcal{F}}_t$ be a filtration of the probability space $(\Omega\,, \mathcal{F}_t\,, \mathrm{P})$ and let $\{\boldsymbol{\xi}_t\,, t \in T\}$ an $\mathbb{S}$-valued stochastic process adapted to $\boldsymbol{\mathcal{F}}_t$ satisfying*

$$\mathrm{E}\; \| \,\boldsymbol{\xi}_t \,\| < \,\infty \tag{4.2}$$

for all $t \in T$. We say that $\{\boldsymbol{\xi}_t\,, t \in T\}$ is a $\boldsymbol{\mathcal{F}}_t$-martingale if

$$\mathrm{E}\,\{\boldsymbol{\xi}_t|\mathcal{F}_s\} = \boldsymbol{\xi}_s \qquad \forall\, t \,\geq\, s \,\in\, T \tag{4.3}$$

If instead

$$\mathrm{E}\,\{\boldsymbol{\xi}_t|\mathcal{F}_s\} \leq \boldsymbol{\xi}_s \qquad \forall\, t \,\geq\, s \,\in\, T \tag{4.4}$$

we say that $\{\boldsymbol{\xi}_t\,, t \in T\}$ is a $\boldsymbol{\mathcal{F}}_t$-super-martingale. Finally, if

$$\mathrm{E}\,\{\boldsymbol{\xi}_t|\mathcal{F}_s\} \geq \boldsymbol{\xi}_s \qquad \forall\, t \,\geq\, s \,\in\, T \tag{4.5}$$

we say that $\{\boldsymbol{\xi}_t\,, t \in T\}$ is a $\boldsymbol{\mathcal{F}}_t$-sub-martingale.

An important consequence of the martingale property is the conservation of the expectation value. Namely we must have

$$\mathrm{E}\mathrm{E}\,\{\boldsymbol{\xi}_t|\mathcal{F}_s\} = \mathrm{E}\boldsymbol{\xi}_s \tag{4.6}$$

but also from the definition of conditional expectation

$$\mathrm{E}\mathrm{E}\,\{\boldsymbol{\xi}_t|\mathcal{F}_s\} = \mathrm{E}\boldsymbol{\xi}_t \tag{4.7}$$

Hence for any $t\,, s$

$$\mathrm{E}\boldsymbol{\xi}_t = \mathrm{E}\boldsymbol{\xi}_s \tag{4.8}$$

5 Random Walk as Martingale

We defined the random walk as

$$\Xi_n = \sum_{i=1}^{n} \xi_i \tag{5.1}$$

with $\{\xi_i\}_{i=1}^N$ i.i.d. random variables with $\xi_i \stackrel{d}{=} \xi$ for all i. Furthermore

$$\xi \colon \Omega \mapsto \{-x, x\} \tag{5.2}$$

then

$$|\Xi_n| \,\leq\, n \qquad \forall\, n = 1, \ldots\, N \tag{5.3}$$

We have

$$\mathrm{E}\,\{\Xi_n\,|\,\Xi_n\} = \Xi_n \tag{5.4}$$

by definition of conditional expectation. We have

$$\mathrm{E}\{\Xi_n\,|\,\Xi_{n-1}\} = \mathrm{E}\{\Xi_{n-1}+\xi_n\,|\,\Xi_{n-1}\} = \mathrm{E}\{\Xi_{n-1}\,|\,\Xi_{n-1}\} + \mathrm{E}\{\xi_n\,|\Xi_{n-1}\} = \Xi_{n-1} + \mathrm{E}\{\xi_n\,|\Xi_{n-1}\} \tag{5.5}$$

By definition of conditional expectation

$$\mathrm{E}\{\xi_n\,|\Xi_{n-1}\} = \mathrm{E}\,\xi_n = (2\,p-1)\,x \tag{5.6}$$

if $\mathrm{P}(\xi = x) = p$ as ξ_n is independent of Ξ_{n-1}. Repeating for arbitrary $k \leq n$

$$\mathrm{E}\{\Xi_n\,|\,\Xi_k\} = \Xi_k + \sum_{i=k+1}^{n} \mathrm{E}\{\xi_i\,|\,\Xi_k\} = \Xi_k + (n-k)\,(2\,p-1)\,x \tag{5.7}$$

We verified that $\{\Xi_n\,, 1 \leq n \leq N\}$ is

- a sub-martingale if $p\; >\; 1/2$;
- a martingale if $p = 1/2$;
- a super-martingale if $p\; <\; 1/2$.

From Ξ_n it is always possible to construct a martingale by subtracting its *compensator*:

$$\tilde{\Xi}_n = \Xi_n - A_n \tag{5.8}$$

In the case of the random walk

$$A_n = \sum_{i=1}^{n} E\xi_i = n\,\mathrm{E}\xi = n\,(2\,p-1)\,x \tag{5.9}$$

It is straightforward to verify that

$$\tilde{\Xi}_n = \sum_{i=1}^{n} \tilde{\xi}_i \tag{5.10}$$

is specified by the sum of i.i.d. random variables with zero average. Hence

$$\mathrm{E}\left\{\tilde{\Xi}_n\,|\,\tilde{\Xi}_k\right\} = \tilde{\Xi}_k + \sum_{i=k+1}^{n} \mathrm{E}\left\{\tilde{\xi}_i\,|\,\tilde{\Xi}_k\right\} = \tilde{\Xi}_k \tag{5.11}$$

which proves that $\{\Xi_n\,, 1 \leq n \leq N\}$ is a martingale.

6 Markov process

Let us consider a stochastic process $\{\xi_n\,, n \in \mathbb{N}\}$ valued to a countable state space $\mathbb{S} \subseteq \mathbb{Z}$:

$$\xi_n \colon \Omega \times \mathbb{N} \mapsto \mathbb{S} \tag{6.1}$$

We suppose that the evolution law for its probability distribution generalizes the form we found for the random walk

$$\mathrm{P}_{n+1}(i) = \sum_{k \in \mathbb{S}} \mathrm{P}(i, n+1|k, n)\,\mathrm{P}_n(k) \tag{6.2}$$

or equivalently

$$\mathrm{P}_{n+1}(i) - \mathrm{P}_n(i) = \sum_{k \in \mathbb{S}} [\mathrm{P}(i, n+1|k, n) - \delta_{i\,k}]\, \mathrm{P}_n(k) \tag{6.3}$$

Using the normalization condition

$$\sum_{k \in \mathbb{S}} \mathrm{P}(k, n+1|i, n) = 1 \tag{6.4}$$

we can couch (6.3) into the form

$$\mathrm{P}_{n+1}(i) - \mathrm{P}_n(i) = \sum_{k \in \mathbb{S}} [\mathrm{P}(i, n+1|k, n)\, \mathrm{P}_n(k) - \mathrm{P}(k, n+1|i, n)\, \mathrm{P}_n(i)] \tag{6.5}$$

The master equation (6.2) states that at any time step we can reconstruct the probability of the stochastic process at the ensuing step if we know its "present" distribution. A more pictorial description is that the "future" depends only upon the "present" but not upon the "past". Such a property is the distinguishing feature of Markov processes.

7 Continuous limit

Up to now we considered a unit time step. We may instead introduce a time unit τ and rescale probabilities

$$\mathrm{P}_n(m) = \tilde{\mathrm{P}}_{n\,\tau}(m) \tag{7.1}$$

The aim is to study the limit

$$\tau \downarrow 0 \qquad \& \qquad t = n\,\tau \in \mathbb{R}_+ \tag{7.2}$$

After rescaling we couch (6.5) into the form

$$\tilde{\mathrm{P}}_{t+\tau}(i) - \tilde{\mathrm{P}}_t(i) = \sum_{k \in \mathbb{S}} [\tilde{\mathrm{P}}(i, t+\tau|k, t)\, \tilde{\mathrm{P}}_t(k) - \tilde{\mathrm{P}}(k, t+\tau|i, t)\, \tilde{\mathrm{P}}_t(i)] \tag{7.3}$$

The expansion in Taylor series

$$\tilde{\mathrm{P}}_{t+\tau}(i) = \tilde{\mathrm{P}}_t(i) + \tau\, \partial_t \tilde{\mathrm{P}}_t(i) + O(\tau^2) \tag{7.4}$$

and

$$\tilde{\mathrm{P}}(i, t+\tau|k, t) = \delta_{i\,k} + \tau\, \mathrm{K}_t(i|k) + O\left(\tau^2\right) \tag{7.5}$$

yields

$$\partial_t \tilde{\mathrm{P}}_t(i) = \sum_{k \in \mathbb{S}} [\mathrm{K}_t(i|k)\, \tilde{\mathrm{P}}_t(k) - \mathrm{K}_t(k|i)\, \tilde{\mathrm{P}}_t(i)] + O(\tau) \tag{7.6}$$

Thus in the limit $\tau \downarrow 0$ we are left with

$$\partial_t \tilde{\mathrm{P}}_t(i) = \sum_{k \in \mathbb{S}} [\mathrm{K}_t(i|k)\, \tilde{\mathrm{P}}_t(k) - \mathrm{K}_t(k|i)\, \tilde{\mathrm{P}}_t(i)] \tag{7.7}$$

Probability conservation now requires

$$\sum_{i \in \mathbb{S}} \partial_t \tilde{\mathrm{P}}_t(i) = \partial_t \sum_{i \in \mathbb{S}} \tilde{\mathrm{P}}_t(i) = 0 \tag{7.8}$$

entailing

$$\sum_{i\,,k\in\mathbb{S}}[\mathrm{K}_t(i|k)\,\tilde{\mathrm{P}}_t(k) - \mathrm{K}_t(k|i)\,\tilde{\mathrm{P}}_t(i)] = 0 \tag{7.9}$$

which is satisfied identically. Two observations are in order

- the diagonal component $\mathrm{K}_t(i|k)$ of the transition rate $\mathrm{K}_t(\cdot)$ does not contribute to (7.7).
- The condition

$$\sum_{k\in\mathbb{S}}\mathrm{K}_t(k|i) = 0 \tag{7.10}$$

is a sufficient condition for (7.9) to hold true. It is also guarantees to leading order in $O(\tau)$ that

$$1 = \sum_{i\in\mathbb{S}}\tilde{\mathrm{P}}(i, t+\tau|k, t) = \sum_{i\in\mathbb{S}}\left[\delta_{i\,k} + \tau\,\mathrm{K}_t(i|k) + O\left(\tau^2\right)\right] = 1 + O(\tau^2) \tag{7.11}$$

By virtue of the first observation, it is not restrictive to assume that (7.10) always holds true. In such a case we can write

$$\partial_t\mathrm{P}_t(i) = \sum_{k\in\mathbb{S}}[\mathrm{K}_t(i|k)\,\mathrm{P}_t(k) \tag{7.12a}$$

$$\sum_{k\in\mathbb{S}}\mathrm{K}_t(k|i) = 0 \tag{7.12b}$$

8 Poisson process

We now make a special choice for the transition rates in (7.12) and set for some $\gamma \in \mathbb{R}_+$

$$\mathrm{K}_t(i|k) = \gamma\,\delta_{i\,,k+1} - \gamma\,\delta_{k\,,i} \tag{8.1}$$

The resulting equation is

$$\partial_t\mathrm{P}(i\,,t) = \gamma\,\mathrm{P}(i-1\,,t) - \gamma\,\mathrm{P}(i\,,t)$$

This is the evolution for a process that can make (or not make) jumps only towards the right of its current position. If we assume that the initial distribution

$$\mathrm{P}(i,0) = \mathrm{P}_o(i)$$

has support on $\mathbb{N}$ then the process will stay there for any further time. The equation can be solved exactly by computing the characteristic function

$$\check{\mathrm{P}}(u,t) := \sum_{k=0}^{\infty} e^{\imath\,k\,u}\,\mathrm{P}(k,t)$$

Namely, it is straightforward to see that $\check{\mathrm{P}}(u,t)$ satisfies:

$$\partial_t\check{\mathrm{P}}(u,t) = \gamma\,(e^{\imath\,u} - 1)\,\check{\mathrm{P}}(u,t)$$

The solution for the initial condition $\check{\mathrm{P}}(u,0) = \check{\mathrm{P}}_o(u)$

$$\check{\mathrm{P}}(u,t) = e^{\gamma\, t\,(e^{\imath\, u}-1)}\check{\mathrm{P}}_o(u)$$

If we specialize for an initial condition

$$\mathrm{P}_o(i) = \delta_{i\,0}$$

(i.e. we assume that the process starts from the origin) we obtain

$$\check{\mathrm{P}}(u,t) = e^{\gamma\, t\,(e^{\imath\, u}-1)}$$

In order to infer the probability distribution associated to the characteristic function we can write

$$\check{\mathrm{P}}(u,t) = e^{\gamma\, t\, e^{\imath\, u}} e^{-\gamma\, t} = e^{-\gamma\, t} \sum_{j=0}^{\infty} \frac{(\gamma\, t)^j}{\Gamma\,(j+1)} e^{\imath\, u\, j}$$

which implies that $\check{\mathrm{P}}(u,t)$ is the characteristic function of the Poisson process, with probability distribution:

$$\mathrm{P}(j\,,t) = \frac{(\gamma\, t)^j}{\Gamma\,(j+1)} e^{-\gamma\, t}$$

References

[1] L. C. Evans. An Introduction to Stochastic Differential Equations. Lecture notes, UC Berkeley, 2006.

[2] C. W. Gardiner. *Handbook of stochastic methods for physics, chemistry and the natural sciences*, volume 13 of *Springer series in synergetics*. Springer, 2 edition, 1994.

[3] A. N. Shiryaev. *Probability*, volume 95 of *Graduate texts in mathematics*. Springer, 2 edition, 1996.

[4] R. van Handel. Stochastic Calculus and Stochastic Control. Lecture notes, Caltech, 2007.