Stochastic Analysis

Billy Sumners

January 27, 2019

Chapter 1

Introduction and Preliminaries

1.1 Probability Theory

1.1.1 Random Variables

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, and $(E, \mathscr{E})$ another measure space. An *E-valued random variable*, or simply a *random variable* is a measurable function $X : (\Omega, \mathscr{F}) \to (E, \mathscr{E})$. Most commonly, we will take $(E, \mathscr{E}) = (\mathbb{R}, \mathscr{B}\mathbb{R})$ and call X a *real random variable*.

The *law* or *distribution* of a random variable X is the probability measure $\mathbb{P}_X$ on $(E, \mathscr{E})$, also denoted $\mathscr{L}_X$, defined by

$$\mathbb{P}_X[A] := \mathbb{P}[X \in A] = \mathbb{P}[\{\omega \in \Omega \colon X(\omega) \in A\}].$$

The σ-*algebra generated by* X is defined by

$$\sigma(X) := \{X^{-1}(A) \colon A \in \mathscr{E}\}.$$

It is usually interpreted as the "information" known to X.

An *E-valued stochastic process* is a parameterized family $X = (X_\alpha \colon \alpha \in A)$ of E-valued random variables X_α on $(\Omega, \mathscr{F}, \mathbb{P})$. It also generates a σ-algebra defined by

$$\sigma(X) := \sigma\left(\bigcup_{\alpha \in A} \sigma(X_\alpha)\right).$$

1.1.2 Independence

Events $A, B \in \mathscr{F}$ are *independent* if

$$\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B]. \tag{1.1}$$

Two σ-algebras $\mathscr{A}, \mathscr{B} \subseteq \mathscr{F}$ are *independent* if (1.1) holds for all $A \in \mathscr{A}$ and $B \in \mathscr{B}$. To check this, it suffices to check (1.1) for all $A \in \Pi_{\mathscr{A}}, B \in \Pi_{\mathscr{B}}$, where $\Pi_{\mathscr{A}}$ and $\Pi_{\mathscr{B}}$ are π-systems generating $\mathscr{A}$ and $\mathscr{B}$ respectively.

Two random variables/stochastic processes X,Y are *independent* if $\sigma(X)$ and $\sigma(Y)$ are independent.

The above defintions can be generalized to the situation in which we have an arbitrary family $(\mathscr{A}_\lambda)_{\lambda\in\Lambda}$ of σ-algebras. In this case, the $\mathscr{A}_\lambda$ are *independent* if

$$\mathbb{P}[A_{\lambda_1}\cap\cdots\cap A_{\lambda_n}] = \mathbb{P}[A_{\lambda_1}]\cdots\mathbb{P}[A_{\lambda_n}] \tag{1.2}$$

For all distinct $\lambda_1,\ldots,\lambda_n\in\Lambda$, and all $A_{\lambda_i}\in\mathscr{A}_{\lambda_i}$.

If X and Y are independent and integrable real random variables (integrable meaning $\mathbb{E}[|X|]<\infty$ and $\mathbb{E}[|Y|]<\infty$, then

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]. \tag{1.3}$$

This can be shown by approximating by simple functions.

If X,Y are independent $(E_1,\mathscr{E}_1)$-valued random variables, and we are given measurable functions $f,g:(E_1,\mathscr{E}_1)\to(E_2,\mathscr{E}_2)$, then $f(X)$ and $g(Y)$ are independent. Actually, this is quite easy to show: let $A,B\in\mathscr{E}_2$. Then

$$\begin{aligned}\mathbb{P}[\{f(X)\in A\}\cap\{g(Y)\in B\}] &= \mathbb{P}[\{X\in f^{-1}(A)\}\cap\{Y\in g^{-1}(B)\}]\\ &= \mathbb{P}[X\in f^{-1}(A)]\mathbb{P}[Y\in g^{-1}(B)]\\ &= \mathbb{P}[f(X)\in A]\mathbb{P}[g(Y)\in B].\end{aligned} \tag{1.4}$$

As a corollary, if $f(X)$ and $g(Y)$ are integrable (this is the situation if, e.g., f and g are bounded) and real, then

$$\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]. \tag{1.5}$$

Also, since

$$\mathbb{P}[\{X\in A\}\cap\{Y\in B\}] = \mathbb{E}[\mathbb{1}_A(X)\mathbb{1}_B(Y)] \tag{1.6}$$

for all $A,B\in\mathscr{E}_1$, we can check independence of X and Y by proving (1.5) for all bounded measurable $f:(E_1,\mathscr{E}_1)\to(\mathbb{R},\mathscr{B}(\mathbb{R}))$.

1.1.3 Characteristic Functions

Given a real random variable X, its *characteristic function* is the complex-valued function $\varphi_X(\theta):=\mathbb{E}[e^{i\theta X}]$ for $\theta\in\mathbb{R}$. The characteristic function always exists since $\left|e^{i\theta X}\right|=1$. More precisely, the characteristic function of X is the inverse Fourier transform of the probability measure $\mathbb{P}_X$. Fourier inversion then gives us that $\mathbb{P}_X$ is recoverable from its characteristic function φ_X. In particular, $\varphi_X=\varphi_Y$ implies $X\overset{\mathscr{D}}{=}Y$. More explicit methods of calculating $\mathbb{P}_X$ given φ_X are available in the literature, in the form of numerous *inversion formulae*.

We can also define the *characteristic function* of an $\mathbb{R}^d$-valued random variable X by $\varphi_X(\theta):=\mathbb{E}[e^{i\theta\cdot X}]$ for $\theta\in\mathbb{R}^d$. The above theory also holds in this situation.

Real random variables $X_1,\ldots,X_n$ are independent if and only if

$$\mathbb{E}[e^{i\theta_1X_1+\cdots\theta_nX_n}] = \mathbb{E}[e^{i\theta_1X_1}]\cdots\mathbb{E}[e^{i\theta_nX_n}] \tag{1.7}$$

for all $\theta_1,\ldots,\theta_n\in\mathbb{R}$.

1.1.4 The Normal Distribution

The *normal distribution* with *mean* $\mu \in \mathbb{R}$ and *variance* $\sigma^2 \geq 0$ is the probability measure $\mathscr{N}(\mu, \sigma^2)$ on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ defined by

$$\mathscr{N}(\mu, 0) := \delta_0 \qquad \text{if } \sigma^2 = 0; \tag{1.8}$$

$$\mathscr{N}(\mu, \sigma^2)(\mathrm{d}x) := \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,\mathrm{d}x \qquad \text{if } \sigma^2 > 0. \tag{1.9}$$

A real random variable distributed normally is called a *Gaussian random variable*.

Proposition 1. *Let $X \sim \mathscr{N}(\mu, \sigma^2)$. Then the following hold.*

(1) *For all $a \in \mathbb{R}$, $X + a \sim \mathscr{N}(\mu + a, \sigma^2)$.*

(2) *For all $\alpha \in \mathbb{R}$, $\alpha X \sim \mathscr{N}(\alpha\mu, (\alpha\sigma)^2)$.*

(3) *The characteristic function of X is*

$$\varphi_X(\theta) = e^{\mathrm{i}\theta\mu - \frac{1}{2}\theta^2\sigma^2}. \tag{1.10}$$

Furthermore, if $X_1 \sim \mathscr{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathscr{N}(\mu_2, \sigma_2^2)$ are independent, then $X_1 + X_2 \sim \mathscr{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

Proof. We do the proofs only in the case $\sigma^2 > 0$. For (1) and (2), we use the fact that the set $\{(-\infty, t] : t \in \mathbb{R}\}$ of half-infinite intervals generates the Borel σ-algebra $\mathscr{B}(\mathbb{R})$.

(1) Let $t \in \mathbb{R}$. We then calculate

$$\begin{aligned}
\mathbb{P}[X + a \leq t] &= \mathbb{P}[X \leq t - a] \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{t-a} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{t} e^{-\frac{(y-(\mu+a))^2}{2\sigma^2}}\,\mathrm{d}x \quad \text{where } y = x + a \\
&= \mathscr{N}(\mu + a, \sigma^2)((-\infty, t]).
\end{aligned} \tag{1.11}$$

(2) We similarly calculate

$$\begin{aligned}
\mathbb{P}[\alpha X \leq t] &= \mathbb{P}\left[X \leq \frac{t}{\alpha}\right] \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\frac{t}{\alpha}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{t} e^{-\frac{(\frac{y}{\alpha}-\mu)^2}{2\sigma^2}}\,\mathrm{d}x \quad \text{where } y = \alpha x \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{t} e^{-\frac{(y-\alpha\mu)^2}{2(\alpha\sigma)^2}}\,\mathrm{d}x \\
&= \mathscr{N}(\alpha\mu, (\alpha\sigma)^2)((-\infty, t]).
\end{aligned} \tag{1.12}$$

(3) Suppose first that $\mu = 0$ and $\sigma^2 = 1$, so that X has density

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \tag{1.13}$$

Note that

$$f'(x) = \frac{1}{\sqrt{2\pi}} \frac{d}{dx} e^{-\frac{x^2}{2}} = -\frac{x}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} = -xf(x). \tag{1.14}$$

We integrate by parts to find

$$\int_{\mathbb{R}} f'(x) e^{i\theta x}\, dx = -\int_{\mathbb{R}} i\theta f(x) e^{i\theta x}\, dx = -i\theta \varphi_X(\theta) \tag{1.15}$$

and also

$$\int_{\mathbb{R}} x f(x) e^{i\theta x}\, dx = -i \frac{d}{d\theta} \int_{\mathbb{R}} f(x) e^{i\theta x}\, dx = -i\varphi_X'(\theta). \tag{1.16}$$

Integrating (1.14) against $e^{i\theta x}$, we obtain the ODE

$$-i\theta\varphi_X(\theta) - i\varphi_X'(\theta) = 0, \tag{1.17}$$

which we solve to find $\varphi_X(\theta) = e^{-\frac{\theta^2}{2}}$.

Next, consider the situation in which $\mu \in \mathbb{R}$ and $\sigma^2 > 0$ are arbitrary. Note then that $\frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$, so our above calculation implies

$$\mathbb{E}[e^{i\theta\left(\frac{X-\mu}{\sigma}\right)}] = e^{-\frac{\theta^2}{2}}. \tag{1.18}$$

However,

$$\mathbb{E}[e^{i\theta\left(\frac{X-\mu}{\sigma}\right)}] = e^{-i\frac{\theta}{\sigma}\mu} \mathbb{E}[e^{i\frac{\theta}{\sigma}X}] \tag{1.19}$$

Making the transformation $\theta \mapsto \sigma\theta$, this implies

$$\varphi_X(\theta) = e^{i\theta\mu - \frac{1}{2}\theta^2\sigma^2}, \tag{1.20}$$

as required.

For our final statement, we make use of (1.7) to find

$$\varphi_{X_1+X_2}(\theta) = e^{i\theta\mu_1 - \frac{1}{2}\theta^2\sigma_1^2} e^{i\theta\mu_2 - \frac{1}{2}\theta^2\sigma_2^2} = e^{i\theta(\mu_1+\mu_2) - \frac{1}{2}\theta^2(\sigma_1^2+\sigma_2^2)}. \tag{1.21}$$

So $X_1 + X_2$ has the characteristic function of an $\mathcal{N}(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$ real random variable, which implies $X_1+X_2 \sim \mathcal{N}(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$ since characteristic functions determine the distribution of a random variable.

□

Proposition 2. *Let $X \sim \mathcal{N}(\mu, \sigma^2)$. Then*

(1) $\mathbb{E}[X] = \mu$;

(2) $\mathbb{E}[(X-\mu)^2] = \sigma^2$.

A real stochastic process $X = (X_\alpha \colon \alpha \in A)$ is *Gaussian* if $\sum_{i=1}^n \theta_i X_{\alpha_i}$ is a Gaussian random variable for all $\theta_i \in \mathbb{R}$ and $\alpha_i \in A$. If (X,Y) is a Gaussian pair, then X and Y are independent if and only if $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.

1.2 Brownian Motion

An $\mathbb{R}$-valued stochastic process $B = (B\colon t \geq 0)$ is a *Brownian motion* if

(B1) It has independent increments. That is, the random variables $B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_N} - B_{t_{N-1}}$ are independent for all $0 \leq t_1 \leq \cdots \leq t_N$.

(B2) The map $t \mapsto B_t(\omega)$ is continuous a.s.

(B3) $B_0 = 0$ and $B_t - B_s \sim \mathscr{N}(0, t-s)$ for all $0 \leq s \leq t$.

This definition of Brownian motion is not universal, although all definitions are equivalent. One variant is the *Gaussian variant*: a stochastic process $(B_t : t \geq 0)$ is a Brownian motion if it is a Gaussian process, $\mathbb{E}[B_t] = 0$ for all $t \geq 0$, $\mathbb{E}[B_s B_t] = s \wedge t$ for all $s, t \geq 0$, and the map $t \mapsto B_t$ is continuous a.s. A second variant is the so-called *Bachelier variant*, where (B3) is replaced by (B3'): for all $s, t \geq 0$ and $\Delta > 0$, we have $B_{t+\Delta} + B_{s+\Delta} \overset{\mathscr{D}}{=} B_t + B_s$. The variant in Evans' lecture notes doesn't require (B2), since one can prove that if B is a stochastic process satisfying (B1) and (B2), there exists another stochastic process $\widetilde{B}$ such that $\widetilde{B_t} = B_t$ a.s. for all $t \geq 0$, and $t \mapsto \widetilde{B_t}$ is continuous a.s.

Our variant will replace (B1) by (B1'): $B_t - B_s$ is independent of $\sigma(B_r : 0 \leq r \leq s)$ for all $t \geq s \geq 0$. Let's show that (B1) and (B1') are actually equivalent. Suppose (B1) holds, and let $t \geq s \geq 0$. Choose $0 \leq r_1 \leq \cdots \leq r_N \leq s$. For Borel sets $A_1, \ldots, A_N, B \in \mathscr{B}(\mathbb{R})$, we have

$$\begin{aligned}
&\mathbb{P}[B_{r_1} \in A_1, B_{r_2} - B_{r_1} \in A_2, \ldots, B_{r_N} - B_{r_{N-1}} \in A_N, B_t - B_s \in B] \\
&\quad = \mathbb{P}[B_{r_1} \in A_1, B_{r_2} - B_{r_1} \in A_2, \ldots, B_{r_N} - B_{r_{N-1}} \in A_N, B_s - B_{r_N} \in \mathbb{R}, B_t - B_s \in B],
\end{aligned} \tag{1.22}$$

so we can assume, without loss of generality, that $r_N = s$. Let $f, g\colon \mathbb{R}^N \to \mathbb{R}$ be bounded and Borel measurable functions. Define $\widetilde{f}\colon \mathbb{R} \to \mathbb{R}$ by

$$\widetilde{f}(y_1, \ldots, y_N) := f(y_1, y_1 + y_2, \ldots, y_1 + \cdots y_N). \tag{1.23}$$

Then $f(x_1, \ldots, x_N) = \widetilde{f}(x_1, x_2 - x_1, \ldots, x_N - x_{N-1})$. By independent increments,

$$\begin{aligned}
\mathbb{E}[f(B_{r_1}, \ldots, B_{r_N}) g(B_t - B_s)] &= \mathbb{E}[\widetilde{f}(B_{r_1}, B_{r_2} - B_{r_1}, \ldots, B_{r_N} - B_{r_{N-1}}) g(B_t - B_s)] \\
&= \mathbb{E}[\widetilde{f}(B_{r_1}, B_{r_2} - B_{r_1}, \ldots, B_{r_N} - B_{r_{N-1}})] \mathbb{E}[g(B_t - B_s)] \\
&= \mathbb{E}[f(B_{r_1}, \ldots, B_{r_N})] \mathbb{E}[g(B_t - B_s)].
\end{aligned} \tag{1.24}$$