Lecture 11 Stochastic Process and Brownian Motion *

Tiejun Li

1 Axiomatic Construction of Stochastic Process

Example 1.1. *Consider the independent fair coin tossing process described by the sequence*

$$X = (X_1, X_2, \ldots, X_n, \ldots) \in \{0,1\}^{\mathbb{N}},$$

where $X_n = 0$ *or* 1 *if the* n*th output is 'Tail' (T) or 'Head' (H), respectively. Different trials are assumed to be independent and* $\mathbb{P}(X_n = 0) = \mathbb{P}(X_n = 1) = 1/2$.

Notice that for this process the number of all possible outputs is uncountable. One can not define the probability of an event through summation of the probability of each atom as the case of discrete random variables. In fact, if we define $\Omega = \{H,T\}^{\mathbb{N}}$, the probability of an *atom* $\omega = (\omega_1, \omega_2, \ldots, \omega_n, \ldots) \in \{H,T\}^{\mathbb{N}}$ is 0, i.e.

$$\mathbb{P}(X_1(\omega) = k_1, X_2(\omega) = k_2, \ldots, X_n(\omega) = k_n, \ldots) = \lim_{n\to\infty} \left(\frac{1}{2}\right)^n = 0, \quad k_j \in \{0,1\},\ j = 1,2,\ldots$$

and events like $\{X_n(\omega) = 1\}$ involve uncountably many atoms.

To set up a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ for this process, it is natural to take $\Omega = \{H,T\}^{\mathbb{N}}$ and the σ-algebra $\mathcal{F}$ as the smallest σ-algebra containing all events of the form:

$$C = \left\{\omega | \omega \in \Omega, (\omega_j)_{j=1:m} \in C_m, C_m \subset \{H,T\}^m\right\} \tag{1.1}$$

for any $m \in \mathbb{N}$, i.e. the sets whose finite time projections are specified. These sets are called *cylinder sets*, which is meaningful from the experimental observation point of view. The probability measure $\mathbb{P}$ of an event of the form (1.1) is defined to be

$$\mathbb{P}(C) = \frac{1}{2^m}|C|.$$

Denote $\mathcal{C}$ the set of cylinder sets. One can easily show that $\mathcal{C}$ is an algebra which is only closed under finite union/intersection operation. To extend the probability measure $\mathbb{P}$ from $\mathcal{C}$ to $\mathcal{F}$, we need to verify that $\mathbb{P}$ is countably additive on $\mathcal{C}$.

*School of Mathematical Sciences, Peking University, Beijing 100871, P.R. China

Lemma 1.2. *If $A_n \downarrow A$ and $A_n \in \mathcal{C}$ is non-empty, then A is non-empty.*

With Lemma 1.2, we obtain if $A_n \downarrow \emptyset$, then $\mathbb{P}(A_n) \downarrow 0$, which is equivalent to the countable additivity. From the extension theorem of measures, this probability measure $\mathbb{P}$ is well-defined on $\mathcal{F}$.

Proof of Lemma 1.2. Denote

$$A_n = \{\omega | (\omega_1, \omega_2, \ldots, \omega_{m_n}) \in C_n\}$$

where $\omega_k \in \{H, T\}$. From the non-empty condition of A_n, there exists $\omega^n \in A_n$. Consider

$$\begin{array}{cccc} \omega_1^1 & \omega_1^2 & \omega_1^3 & \cdots \\ \omega_2^1 & \omega_2^2 & \omega_2^3 & \cdots \\ \omega_3^1 & \omega_3^2 & \omega_3^3 & \cdots \\ \vdots & \vdots & \vdots & \ddots, \end{array}$$

there exist infinite superscripts n_k^1 such that $\omega_1^{n_k^1} = H$ or T always in the 1st row. Similar argument can be applied to the continued rows by an subsequence trick. Take the diagonal indices and define $n_k := n_k^k$ and $u_k := \omega_k^{n_k}$ for $k = 1, 2, \ldots$. Denote $u = (u_1, u_2, \ldots)$.

For any r, if $k \geq r$, one has $\omega_j^{n_k} = u_j$ for $1 \leq j \leq r$. For any n, if $k \geq n$, then $n_k \geq n$, and $\omega^{n_k} \in A_{n_k} \subset A_n$. So $(\omega_1^{n_k}, \omega_2^{n_k}, \ldots, \omega_{m_n}^{n_k}) \in C_n$. Take $k \geq m_n$. We get $\omega_j^{n_k} = u_j$ for $1 \leq j \leq m_n$, i.e. $u \in A_n$ for any n.

In summary, $u \in A$ and we are done. □

It is straightforward to check that for any cylinder set $F \in \{0, 1\}^{\mathbb{N}}$, the probability $\mathbb{P}(X(\omega) \in F)$ coincides with the definition we made in Example 1.1 for independent coin tossing process. We remark that the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is not uniquely defined. Another natural way is to take $\Omega = \{0, 1\}^{\mathbb{N}}$, $\mathcal{F}$ the smallest σ-algebra containing all cylinder sets in Ω, and similar probability measure $\mathbb{P}$ on $\mathcal{F}$. With this choice we have

$$X_n(\omega) = \omega_n, \quad \omega \in \Omega = \{0, 1\}^{\mathbb{N}}$$

which is called a *coordinate process* in the sense that $X_n(\omega)$ is just the nth coordinate of ω.

In general, a stochastic process is a parameterized random variables $\{X_t\}_{t \in \mathbf{T}}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in $\mathbb{R}$, the parameter set $\mathbf{T}$ can be $\mathbb{N}$, $[0, +\infty)$ or some finite interval. For any fixed $t \in \mathbf{T}$, we have a random variable

$$X_t : \Omega \to \mathbb{R} \qquad \omega \rightarrowtail X_t(\omega).$$

For any fixed $\omega \in \Omega$, we have a real-valued measurable function on $\mathbf{T}$

$$X_\cdot(\omega) : \mathbf{T} \to \mathbb{R} \qquad t \rightarrowtail X_t(\omega),$$

which is called a trajectory or sample path of X. As a bi-variate function, a stochastic process can also be viewed as a measurable function from $\Omega \times \mathbf{T}$ to $\mathbb{R}$

$$(\omega, t) \rightarrowtail X(\omega, t) := X_t(\omega),$$

with the σ-algebra in $\Omega \times \mathbf{T}$ been chosen as $\mathcal{F} \times \mathcal{T}$, and $\mathcal{T}$ is the Borel σ-algebra on $\mathbf{T}$.

The largest probability space that one can take is the infinite product space $\Omega = \mathbb{R}^{\mathbf{T}}$, i.e. Ω is the space of all real-valued functions on $\mathbf{T}$. $\mathcal{F}$ can be taken as the infinite product σ-algebra $\mathcal{B}^{\mathbf{T}}$, which is the smallest σ-algebra containing all cylinder sets

$$C = \{\omega \in \mathbb{R}^{\mathbf{T}} | (\omega(t_1), \omega(t_2), \ldots, \omega(t_k)) \in A\}, \quad A \in \mathcal{B}^k, \ t_i \in \mathbf{T} \text{ for } i = 1, \ldots, k,$$

where $\mathcal{B}, \mathcal{B}^k$ is the Borel σ-algebra on $\mathbb{R}$ and $\mathbb{R}^k$, respectively. When $\mathbf{T} = \mathbb{N}$ and X_t only takes values in $\{0, 1\}$, we are back to the setting of Example 1.1.

Finite dimensional distributions are particularly interesting for a stochastic process, since they are the ones we can really observe. Let

$$\mu_{t_1,\ldots,t_k}(F_1 \times F_2 \times \cdots \times F_k) = \mathbb{P}[X_{t_1} \in F_1, \ldots, X_{t_k} \in F_k]$$

for all $F_1, F_2, \ldots, F_k \in \mathcal{B}$. $\mu_{t_1,\ldots,t_k}$ is called the finite dimensional distributions of $\{X_t\}_{t\in\mathbf{T}}$ at the time slice $(t_1, \ldots, t_k)$, where $t_i \in \mathbf{T}$ for $i = 1, 2, \ldots, k$.

The following theorem of Kolmogorov states that an abstract probability space $(\Omega, \mathcal{F}, \mathbb{P})$ can be established for a stochastic process X by knowing its all finite dimensional distributions with suitable consistency conditions.

Theorem 1.3 (Kolmogorov's extension theorem)**.** *Assume that a family of finite dimensional distributions $\{\mu_{t_1,\ldots,t_k}\}$ satisfy the following two consistency conditions for arbitrary sets of $t_1, t_2, \ldots, t_k \in T, k \in \mathbb{N}$:*

(i) For any permutation σ of $\{1, 2, \ldots, k\}$,

$$\mu_{t_{\sigma(1)},\ldots,t_{\sigma(k)}}(F_1 \times F_2 \times \cdots \times F_k) = \mu_{t_1,\ldots,t_k}(F_{\sigma^{-1}(1)} \times F_{\sigma^{-1}(2)} \times \cdots \times F_{\sigma-1(k)}).$$

(ii) For any $m \in \mathbb{N}$,

$$\mu_{t_1,\ldots,t_k}(F_1 \times F_2 \times \cdots \times F_k) = \mu_{t_1,\ldots,t_k,t_{k+1},\ldots,t_{k+m}}(F_1 \times \cdots \times F_k \times \mathbb{R} \times \cdots \times \mathbb{R}).$$

Then there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a stochastic process $\{X_t\}_{t\in\mathbf{T}}$ such that

$$\mu_{t_1,\ldots,t_k}(F_1 \times F_2 \times \cdots \times F_m) = \mathbb{P}(X_{t_1} \in F_1, X_{t_2} \in F_2, \ldots, X_{t_m} \in F_m)$$

for any $t_1, t_2, \ldots, t_m \in \mathbf{T}$, $m \in \mathbb{N}$.

The proof the Kolmogorov extension theorem may be referred to [4, 5]. The advantage of the Kolmogorov theorem is that it is very general. The problem is that the probability space Ω is too big, so big that we can not say anything about features of paths on this space. Therefore the real challenge is to define probability measures on smaller spaces.

2 Filtration

The more we observe about a stochastic process, the more information we have at our disposal. This gives rise to a family of increasingly larger σ-algebras, which we call the *filtration* associated with the stochastic process. The filtration is the main conceptual difference between the random variables and and stochastic processes.

Definition 2.1 (Filtration)**.** *Given the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the filtration is a nondecreasing family of σ-algebras $\{\mathcal{F}_t\}_{t\geq 0}$ such that $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ for any $0 \leq s < t$.*

A stochastic process $\{X_t\}$ is called $\mathcal{F}_t$-*adapted* if X_t is $\mathcal{F}_t$-measurable, i.e. $X_t^{-1}(B) \in \mathcal{F}_t$, for any $t \geq 0$ and $B \in \mathcal{B}$. Given a stochastic process $\{X_t\}$, one can define the filtration generated by this process by: $\mathcal{F}_t^X = \sigma(X_s, s \leq t)$, which is the smallest σ-algebra such that the $\{X_s\}_{s\leq t}$ are measurable. $\mathcal{F}_t^X$ is the smallest filtration such that the process $\{X_t\}$ is adapted. The filtration $\mathcal{F}_t^X$ can be thought of as the information supplied by the process up to time t. Taking again the independent coin tossing as the example and $\Omega = \{H, T\}^{\mathbb{N}}$. In this case, $\mathbf{T} = \mathbb{N}$ and the filtration is $\{\mathcal{F}_n^X\}_{n\geq 0}$. When $n = 0$, the σ-algebra is trivial

$$\mathcal{F}_0^X = \{\emptyset, \Omega\},$$

which means that we do not know any information about the output of the coin tossing. When $n = 1$, the σ-algebra is

$$\mathcal{F}_1^X = \{\emptyset, \Omega, \{H\}, \{T\}\}$$

since the first output gives either Head or Tail and we only know this information about the first output. When $n = 2$, we have

$$\mathcal{F}_2^X = \{\emptyset, \Omega, \{H\cdot\}, \{T\cdot\}, \{\cdot H\}, \{\cdot T\}, \{HH\}, \{HT\}, \{TH\}, \{TT\}, \ldots\},$$

which contains all possible combinations of the outputs for the first two rounds of experiments. Sets like

$$\{HH\cdots T\} \quad \text{or} \quad \{HH\cdots H\}$$

are not contained in $\mathcal{F}_0^X$, $\mathcal{F}_1^X$ or $\mathcal{F}_2^X$ since the first two outputs can not tell such information. It is obvious that $\mathcal{F}_n^X$ becomes finer and finer as n increases.

3 Gaussian Process

In order to study Wiener process or Brownian motion (Brownian motion is also called Wiener process because its first rigorous mathematical foundation was established by N. Wiener in 1923 [11]), we will first introduce the Gaussian process on the continuous state space $\boldsymbol{R}$.

Definition 3.1. *A Gaussian process means that all of the finite dimensional distributions* $\mu_{t_1,\dots,t_k}$ *are Gaussian for any* $t_1, t_2, \dots, t_k \in T$.

We know that any Gaussian vector $X = (X_1, X_2, \dots, X_n)^T$ is completely determined by its first moment $m = \mathbb{E}X$ and second moment $K = \mathbb{E}(X-m)(X-m)^T$, where $m_i = \mathbb{E}X_i$ and $K_{ij} = \mathbb{E}(X_i - m_i)(X_j - m_j)^T$. The corresponding pdf is

$$p(\boldsymbol{x}) = \frac{1}{Z} e^{-\frac{1}{2}(\boldsymbol{x}-m)^T K^{-1}(\boldsymbol{x}-m)}$$

if K is invertible, where Z is a normalization constant. For the general case, we need to represent X via the characteristic function

$$\mathbb{E}e^{i\boldsymbol{\xi}\cdot X} = e^{i\boldsymbol{\xi}\cdot m - \frac{1}{2}\boldsymbol{\xi}^T K \boldsymbol{\xi}}.$$

From the above interpretation, a Gaussian process is uniquely determined by the mean function $m(t) = \mathbb{E}X_t$ and the covariance function $K(s,t) = \mathbb{E}(X_s - m(s))(X_t - m(t))$. We have $K(s,t) = K(t,s)$ by definition. If we consider the finite dimensional distribution at the time slice $(t_1, t_2, \dots, t_n)$, then $m(t)$ and $K(s,t)$ give the first moment

$$M = \big(m(t_1), m(t_2), \dots, m(t_n)\big)$$

and second moment

$$K = \begin{bmatrix} K(t_1,t_1) & K(t_1,t_2) & \cdots & K(t_1,t_n) \\ K(t_2,t_1) & K(t_2,t_2) & \cdots & K(t_2,t_n) \\ \vdots & \vdots & \ddots & \vdots \\ K(t_n,t_1) & K(t_n,t_2) & \cdots & K(t_n,t_n) \end{bmatrix}.$$

It is straightforward to observe that for any $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$ we have

$$\begin{aligned} \sum_{i,j} K(t_i,t_j) x_i x_j &= \sum_{i,j} \mathbb{E}(X_{t_i} - m(t_i))(X_{t_j} - m(t_j)) x_i x_j \\ &= \mathbb{E}\Big(\sum_i (X_{t_i} - m(t_i)) x_i \Big)^2 \geq 0. \end{aligned}$$

Thus we may view $m(t)$ as an infinite dimensional vector, and $K(s,t)$ as an infinite dimensional positive semi-definite matrix. From the characteristic function point of view, the Gaussian process X can be explained as a Gaussian random element in an infinite dimensional space $L^2(T)$ since we have at least formally in the current stage

$$\mathbb{E}e^{i(\xi,X)} = e^{i(\xi,m) - \frac{1}{2}(\xi, K\xi)},$$

where $(\xi, m) = \int_a^b \xi(t) m(t) dt$ is the inner-product in $L^2(T)$, and $(K\xi)(t) = \int_a^b K(t,s)\xi(s) ds$ is the action of the kernel function K on the function ξ. Based on the Kolmogorov's

extension theorem, we can construct a Gaussian process X from a given mean function $m(t)$ and covariance function $K(s,t)$.

The covariance function K is obviously symmetric, i.e. $K(t,s) = K(s,t)$, by definition. In addition, we have the semi-positivity of K in the following sense.

Theorem 3.2. *Assume the Gaussian process $(X_t)_{t\in[0,T]}$ possesses the regularity $X \in L^2_\omega L^2_t$ in the sense that $X \in L^2(\Omega; L^2[0,T])$, i.e.*

$$\mathbb{E}\int_0^T X_t^2 dt < \infty.$$

We have $m \in L^2_t$ and the operator

$$\mathcal{K}f(s) := \int_0^T K(s,t)f(t)dt, \quad s \in [0,T]$$

is a positive compact operator on L^2_t.

Proof. The mean function $m \in L^2_t$ is obvious since

$$\int_0^T m^2(t)dt = \int_0^T (\mathbb{E}X_t)^2 dt \le \int_0^T \mathbb{E}X_t^2 dt < \infty.$$

In addition, we have

$$\begin{aligned}\int_0^T\int_0^T K^2(s,t)dsdt &= \int_0^T\int_0^T \Big(\mathbb{E}(X_t - m(t))(X_s - m(s))\Big)^2 dsdt \\ &\le \int_0^T\int_0^T \mathbb{E}(X_t - m(t))^2\mathbb{E}(X_s - m(s))^2 dsdt \le \Big(\int_0^T \mathbb{E}X_t^2 dt\Big)^2,\end{aligned}$$

which means $K \in L^2([0,T]\times[0,T])$. Thus $\mathcal{K}$ is a compact operator on L^2_t (c.f. [6]).

It is easy to find that the adjoint operator of $\mathcal{K}$ is

$$\mathcal{K}^* f(s) := \int_0^T K(t,s)f(t)dt, \quad s \in [0,T].$$

From the symmetry of $K(s,t)$, we know that $\mathcal{K}$ is self-adjoint.

To show the positivity of $\mathcal{K}$, we have

$$\begin{aligned}(\mathcal{K}f, f) &= \int_0^T\int_0^T \mathbb{E}(X_t - m(t))(X_s - m(s))f(t)f(s)dsdt \\ &= \mathbb{E}\Big(\int_0^T (X_t - m(t))f(t)dt\Big)^2 \ge 0.\end{aligned}$$

The proof is completed. □

The following important closure property for a collection of Gaussian random variables will be used frequently in this chapter.

Theorem 3.3. *Suppose $X_1, X_2, \ldots$ are a sequence of Gaussian random variables and X_n converges to X in probability. Then X is also Gaussian.*

Proof. Let us denote

$$m_k = \mathbb{E}X_k, \quad \sigma_k^2 = \mathrm{var}X_k.$$

Then by dominated convergence theorem we have

$$e^{i\xi m_k - \frac{1}{2}\sigma_k^2\xi^2} = \mathbb{E}e^{i\xi X_k} \to \mathbb{E}e^{i\xi X} \quad \text{for any } \xi \in \boldsymbol{R}.$$

From the existence of the limit of the above equation, there are numbers m and σ^2 such that

$$m = \lim m_k, \quad \sigma^2 = \lim \sigma_k^2$$

and $\mathbb{E}e^{i\xi X} = e^{i\xi m - \frac{1}{2}\sigma^2\xi^2}$. The proof is completed. □

4 Wiener Process

The rigorous mathematical definition of the Brownian motion, or the Wiener Process, is defined as follows.

Definition 4.1. *(Brownian motion) The one dimensional Brownian motion W_t is defined as*

1. *It is a Gaussian process.*
2. *It has mean function $m(t) = 0$, and covariance function $K(s,t) = s \wedge t = \min(s,t)$.*
3. *With probability one, $t \mapsto W_t$ is continuous.*

The m-dimensional Brownian motion $\boldsymbol{W}_t$ has the form $\boldsymbol{W}_t = (W_t^1, W_t^2, \ldots, W_t^m)$, where each component W_t^j is a Brownian motion and they are independent each other. The Brownian motion (m-dimensional Brownian motion) is usually denoted as W_t or B_t ($\boldsymbol{W}_t$ or $\boldsymbol{B}_t$).

It is not difficult to prove that the above three conditions are equivalent to the following definition.

1′. For any $t_0 < t_1 < \cdots < t_n$, the random variables $W_{t_0}, W_{t_1} - W_{t_0}, \ldots, W_{t_n} - W_{t_{n-1}}$ are independent.

2′. For any $s, t \geq 0$, $W_{s+t} - W_s \sim N(0, t)$.

3. With probability one, $t \mapsto W_t$ is continuous.

One straightforward implication of the second equivalent definition is that we can immediately write down the joint probability distribution density for $(W_{t_1}, W_{t_2}, \ldots, W_{t_n})$ $(t_1 < t_2 < \cdots < t_n)$ as

$$p_n(w_1, w_2, \ldots, w_n) = \frac{1}{\sqrt{2\pi t_1}} e^{-\frac{w_1^2}{2t_1}} \frac{1}{\sqrt{2\pi(t_2 - t_1)}} e^{-\frac{(w_2 - w_1)^2}{2(t_2 - t_1)}} \cdots \frac{1}{\sqrt{2\pi(t_n - t_{n-1})}} e^{-\frac{(w_n - w_{n-1})^2}{2(t_n - t_{n-1})}}.$$

More compactly

$$p_n(w_1, w_2, \ldots, w_n) = \frac{1}{Z_n} \exp(-I_n(w)),$$

where

$$I_n(w) = \frac{1}{2} \sum_{j=1}^{n} \Big(\frac{w_j - w_{j-1}}{t_j - t_{j-1}} \Big)^2 (t_j - t_{j-1}), \quad t_0 := 0, w_0 := 0,$$

$$Z_n = (2\pi)^{\frac{n}{2}} \big[t_1(t_2 - t_1) \cdots (t_n - t_{n-1}) \big]^{\frac{1}{2}}.$$

This also explicitly shows the stationarity and Markovianity of the Brownian motion with transition kernel function $p(x, t|y, s)$

$$\mathbb{P}(W_t \in B | W_s = y) = \int_B \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{(x-y)^2}{2(t-s)}} dx = \int_B p(x, t|y, s) dx$$

where $s < t$ and B is a Borel set on $\mathbf{R}$. The transition probability density $p(x, t|y, s)$ satisfies the stationarity $p(x, t|y, s) = p(x - y, t - s|0, 0)$ and $p(x, t|0, 0)$ satisfies the PDE

$$\frac{\partial p}{\partial t} = \frac{1}{2} \frac{\partial^2 p}{\partial x^2}, \quad p(x, 0|0, 0) = \delta(x).$$

Now mathematically the first question is "Is there a process with these properties?". Though from Kolmogorov's extension theorem we can construct a probability space on $(\boldsymbol{R}^{[0,\infty)}, \mathcal{R}^{[0,\infty)})$ by the consistency of the finite dimensional distributions, it is not straightforward that the condition 3 in Definition 4.1 must be satisfied automatically. In fact, if we define the set

$$C = \{\omega | \omega \in \boldsymbol{R}^T, \omega \text{ is continuous on } T\}, \tag{4.1}$$

we will show that C is not a measurable set in $\mathcal{R}^T$! To understand this, one needs the following theorem

Theorem 4.2. *For any family of real functions $X_t : \Omega \to \boldsymbol{R}$, $t \in T$.*

(i) If $A \in \sigma\{X_t, t \in T\}$ and $\omega \in A$, and if $X_t(\omega') = X_t(\omega)$ for all $t \in T$, then we have $\omega' \in A$.

(ii) If $A \in \sigma\{X_t, t \in T\}$, then $A \in \sigma\{X_t, t \in S\}$ for some countable subset $S \subset T$.

The proof of this theorem may be referred to [2]. To apply the above theorem, we take $T = [0, \infty)$ and S a countable dense subset of T. We will have $C \in \mathcal{R}^S$ if $C \in \mathcal{R}^T$ by the second statement. From the first statement, C should contain all functions which have the same value with some $f \in C$ on S. This should contain lots of discontinuous functions. This contradicts with that C is the set of continuous functions.

To handle this issue, we need the concept "modification" of a process.

Definition 4.3 (Modification)**.** *Two processes X and X' defined on the same probability space are said to be modifications of each other if for each t,*

$$X_t = X'_t \quad a.s.$$

They are called indistinguishable if for almost all ω

$$X_t(\omega) = X'_t(\omega) \quad \text{for every } t.$$

It is clear that if X and X' are modifications of each other, they have the same finite dimensional distribution. If X and X' are modifications of each other and are almost surely continuous, they are indistinguishable.

Theorem 4.4 (Kolmogorov's continuity theorem)**.** *A real-valued process X for which there exist three strictly positive constants γ, β, C such that*

$$\mathbb{E}(|X_t - X_s|^\alpha) \leq C|t - s|^{1+\beta}$$

for any $s, t \geq 0$, then there is a modification $\tilde{X}$ of X which is almost-surely continuous.

For Brownian motion, one has $\alpha = 4, \beta = 1$, thus the condition of the above theorem is satisfied and the continuity of Brownian motion can be ensured in the sense of modifications.

5 Homeworks

- HW1. Let $\{\xi_n\}_{n\in\mathbb{N}}$ be a sequence of *i.i.d.* random variables taking values $+1$ with probability $\frac{2}{3}$ and -1 with probability $\frac{1}{3}$. Consider the (asymmetric) random walk on $\mathbb{Z}$ defined as

  $$S_n = \sum_{j=1}^{n} \xi_j.$$

  We wish to construct a stochastic process Z_t defined for $t \in [0, 1]$ by appropriate rescaling of S_n (similar to what we did to construct the Wiener process). That is we

want to show that there exists $\alpha \in \mathbb{R}$ such that the sequence of piecewise constant functions (here $\lfloor z \rfloor$ denotes the biggest integer smaller or equal to $z \in \mathbb{R}$)

$$Z_t^N = \frac{S_{\lfloor Nt \rfloor}}{N^\alpha}$$

converge as $N \to +\infty$ to some nontrivial Z_t. What is the α you need to chose for this to be the case? And what is Z_t?

- HW2. Let W_t be a Wiener process. Compute

  (a) $\mathbb{E}W_t^4$.

  (b) $\mathbb{E}(W_t - W_s + W_z)^2 \quad (t, s, z \in [0, 1])$.

- HW3. Let $X \in R^n$ be a n-dimensional Gaussian R.V. with mean zero and covariance matrix A (i.e. $\mathbb{E}X_i X_j = A_{ij}$). Suppose B is another strictly positive definite symmetric $n \times n$ matrix. Compute

$$\mathbb{E} \exp(-\frac{1}{2} X^T B X).$$

- HW4. Prove the equivalence of the two definitions for Brownian motion.

References

[1] R.J. Adler. *The Geometry of random fields*. John Wiley & Sons, New York, 1981.

[2] Billingsley. *Probability and measure*. John Wiley and Sons, New York, 1979.

[3] R.H. Cameron and W.T. Martin. The orthogonal development of nonlinear functionals in series of fourier-hermite functionals. *Ann. Math.*, pages 385–392, 1947.

[4] I. Karatzas and S.E. Shreve. *Brownian motion and stochastic calculus*. Springer-Verlag, Berlin, Heidelberg and New York, 1991.

[5] J. Lamperti. *Stochastic processes: A survey of the mathematical theory*. Springer-Verlag, New York, 1977.

[6] P.D. Lax. *Functional analysis*. John Wiley & Sons, 2002.

[7] P. Möters and Y. Peres. *Brownian motion*. Cambridge University Press, Cambridge, UK, 2010.

[8] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*. Springer-Verlag, Berlin and Heidelberg, 3rd edition edition, 2005.

[9] F. Riesz and B. Sz.-Nagy. *Functional analysis*. Blackie & Son limited, London and Glasgow, 1956.

[10] D.F. Walnut. *An introduction to wavelet analysis*. Birkhauser, Boston, Basel and Berlin, 2002.

[11] N. Wiener. Differential space. *J. Math. and Phys.*, 2:131–174, 1923.

[12] D. Xiu and G. Karniadakis. The wiener-askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comp.*, pages 619–644, 2002.