Chapter 1-2: Introduction and Preliminaries

Probability Spaces and Stochastic Processes

Fei Lu

Department of Mathematics, Johns Hopkins

$$X_t(\omega) : \Omega \to \mathbb{R}^n$$

Introduction: Applications and an example

Consider the population evolution model:

$$\frac{dN}{dt} = \alpha(t)N(t), \quad N(0) = N_0 \tag{1}$$

Solution: $N(t) = N(0)e^{\int_0^t \alpha(s)ds}$

Introducing Randomness: How do we model uncertainty?

- Random initial condition: $N_0 = N_0(\omega)$
- Random rate: $\alpha(t) = r(t) + X(\omega)$
- Random noise: $\dot{N}(t) = \alpha(t)N(t) + X(\omega, t)$

Q1: How to solve the equation? Integral?
Q2: Properties of $N(t, \omega)$? Limit as $t \to \infty$?
Q3: State-parameter estimation? Control?

$$N(t, \omega) = N_0 \exp\left(\int_0^t \alpha(s)ds + \int_0^t X(\omega, s)ds\right) \tag{2}$$

Section 2: Probability Space

Definition 2.1.1: A probability space is a triple $(\Omega, \mathcal{F}, \mathbb{P})$

- Ω: Sample space (a set)
- $\mathcal{F}$: σ-algebra of events
- $\mathbb{P} : \mathcal{F} \to [0, 1]$: probab. measure, $\mathbb{P}(\Omega) = 1$ and countable additivity.

Properties of σ-algebra $\mathcal{F}$:

1. $\emptyset \in \mathcal{F}$
2. If $F \in \mathcal{F} \Rightarrow F^c \in \mathcal{F}$ (closed under complement)
3. If $A_1, A_2, \cdots \in \mathcal{F} \Rightarrow \bigcup_{k=1}^{\infty} A_k \in \mathcal{F}$ (closed under countable union)

Example: Borel σ-algebra $\mathcal{B}$ for $\Omega = \mathbb{R}^n$: generated by all open subsets.
Example: $\mathcal{H}_X = \sigma$-algebra generated by $X : \Omega \to \mathbb{R}^n$, i.e., the smallest σ-algebra containing all sets of the form $X^{-1}(U)$ for open $U \subset \mathbb{R}^n$.

Random Variables and Independence

Random Variable: A function $X : \Omega \to \mathbb{R}^n$ is $\mathcal{F}$-measurable if $X^{-1}(U) \in \mathcal{F}$ for all open sets $U \in \mathbb{R}^n$.
Doob-Dynkin Lemma: If $X, Y : \Omega \to \mathbb{R}^n$ are given functions, then Y is $\sigma(X)$-measurable IFF there exists a Borel measurable function g such that $Y = g(X)$.

Independence (Def 2.1.3):

- Two sets $A, B \in \mathcal{F}$ are independent if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.
- Two families of sets $\mathcal{A}, \mathcal{B} \subset \mathcal{F}$ are independent if for all $A \in \mathcal{A}, B \in \mathcal{B}$, A and B are independent.
- Two random variables X and Y are independent if $\mathcal{H}_X$ and $\mathcal{H}_Y$ are independent, i.e., for all Borel sets $U, V \subset \mathbb{R}^n$:

$$\mathbb{P}(X^{-1}(U) \cap Y^{-1}(V)) = \mathbb{P}(X^{-1}(U))\mathbb{P}(Y^{-1}(V)).$$

Expectation Property: If X and Y are independent, then $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ when the expectations exist.

Definition of Stochastic Process

Definition 2.1.4: A stochastic process (SP) is a parametrized collection of random variables $\{X_t\}_{t \in T}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$.

- **Parameter space** T: $[0, \infty)$, $[a, b]$, or a discrete set.
- **State space**: Usually $\mathbb{R}^n$.

Three Views:

1. **Random Variable view:** For fixed t, $X_t : \Omega \to \mathbb{R}^n$ is a random variable.
2. **Function view:** $X : T \times \Omega \to \mathbb{R}^n$ (jointly measurable).
3. **Path view:** For each ω, $X(\omega) : T \to \mathbb{R}^n$ is a path/trajectory. Take Ω as the set of paths, $X : \Omega \to (\mathbb{R}^n)^T$.

Finite Dimensional Distributions (FDDs)

Q1: How do we specify a Stochastic Process $(X_t)_{t \in T}$?

Finite Dimensional Distributions (FDDs)

Q1: How do we specify a Stochastic Process $(X_t)_{t \in T}$?
Finite Dimensional Distributions (FDDs). For $t_1, \dots, t_k \in T$, the measure on $\mathbb{R}^{nk}$ is defined by: for Borel sets $F_1, \dots, F_k \subset \mathbb{R}^n$,

$$\mu_{t_1 \dots t_k}(F_1 \times \cdots \times F_k) = \mathbb{P}(X_{t_1} \in F_1, \dots, X_{t_k} \in F_k). \tag{3}$$

Finite Dimensional Distributions (FDDs)

Q1: How do we specify a Stochastic Process $(X_t)_{t \in T}$?
Finite Dimensional Distributions (FDDs). For $t_1, \ldots, t_k \in T$, the measure on $\mathbb{R}^{nk}$ is defined by: for Borel sets $F_1, \ldots, F_k \subset \mathbb{R}^n$,

$$\mu_{t_1 \ldots t_k}(F_1 \times \cdots \times F_k) = \mathbb{P}(X_{t_1} \in F_1, \ldots, X_{t_k} \in F_k). \qquad (3)$$

Q2: Given a family of FDDs, can we construct an SP?

Finite Dimensional Distributions (FDDs)

Q1: How do we specify a Stochastic Process $(X_t)_{t \in T}$?
Finite Dimensional Distributions (FDDs). For $t_1, \ldots, t_k \in T$, the measure on $\mathbb{R}^{nk}$ is defined by: for Borel sets $F_1, \ldots, F_k \subset \mathbb{R}^n$,

$$\mu_{t_1 \ldots t_k}(F_1 \times \cdots \times F_k) = \mathbb{P}(X_{t_1} \in F_1, \ldots, X_{t_k} \in F_k). \tag{3}$$

Q2: Given a family of FDDs, can we construct an SP?
Kolmogorov's Extension Theorem (Thm 2.1.5): Yes.

- Permutation consistency:
  $\mu_{t_{\pi(1)} \ldots t_{\pi(k)}}(F_{\pi(1)} \times \cdots \times F_{\pi(k)}) = \mu_{t_1 \ldots t_k}(F_1 \times \cdots \times F_k)$ for any permutation π.
- Marginal consistency:
  $\mu_{t_1 \ldots t_k}(F_1 \times \cdots \times F_k) = \mu_{t_1 \ldots t_k t_{k+1}}(F_1 \times \cdots \times F_k \times \mathbb{R}^n)$.

Then, $\exists$ a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and an SP $\{X_t\}$ with these FDDs.

FDDs $\{\mu_{t_1 \ldots t_k}\}$ for all $k \in \mathbb{N}$ and $t_1, \ldots, t_k \in T$ completely characterize the distribution of the SP.

Versions and Continuity

Q: Do FDDs determine an SP uniquely?

Versions and Continuity

Q: Do FDDs determine an SP uniquely? Not for path properties.
Two SPs can have the same FDDs but different path properties (e.g., one continuous, one discontinuous).
Example: Let $(\Omega, \mathcal{F}, \mathbb{P}) = ([0, \infty), \mathcal{B}, \mu)$ with μ having no mass on any single point. Let $X_t \equiv 0$ and $Y_t(\omega) = \mathbf{1}_{\{t=\omega\}}$. Then X and Y have the same FDDs

$$\mathbb{P}(X_{t_1} = 0, \ldots, X_{t_k} = 0) = 1 = \mathbb{P}(Y_{t_1} = 0, \ldots, Y_{t_k} = 0)$$

for any $t_1, \ldots, t_k$, but Y is discontinuous (for every path).

Definition 2.2.2 (Version): Y is a version of X if $\mathbb{P}(X_t = Y_t) = 1 \ \forall t$.

Versions and Continuity

Q: Do FDDs determine an SP uniquely? Not for path properties.
Two SPs can have the same FDDs but different path properties (e.g., one continuous, one discontinuous).
Example: Let $(\Omega, \mathcal{F}, \mathbb{P}) = ([0, \infty), \mathcal{B}, \mu)$ with μ having no mass on any single point. Let $X_t \equiv 0$ and $Y_t(\omega) = \mathbf{1}_{\{t=\omega\}}$. Then X and Y have the same FDDs

$$\mathbb{P}(X_{t_1} = 0, \ldots, X_{t_k} = 0) = 1 = \mathbb{P}(Y_{t_1} = 0, \ldots, Y_{t_k} = 0)$$

for any $t_1, \ldots, t_k$, but Y is discontinuous (for every path).

Definition 2.2.2 (Version): Y is a version of X if $\mathbb{P}(X_t = Y_t) = 1 \ \forall t$.

Kolmogorov's Continuity Theorem (Thm 2.2.3): If a process $X = \{X_t\}_{t \in T}$ satisfies the moment condition:

$$\boxed{\mathbb{E}[|X_t - X_s|^\alpha] \leq D|t - s|^{1+\beta}}, \forall t, s \in T \qquad (4)$$

for positive constants α, β, D, then $\exists$ a continuous version of X.

Gaussian Processes (GP)

Definition: A continuous-time GP is an SP whose FDDs are Gaussian.
A GP is characterized by:

1. **Mean:** $m(t) = \mathbb{E}[X_t]$
2. **Covariance:** $C(s,t) = \mathbb{E}[(X_t - m(t)) \otimes (X_s - m(s))]$

Gaussian Processes (GP)

Definition: A continuous-time GP is an SP whose FDDs are Gaussian.
A GP is characterized by:

1. **Mean:** $m(t) = \mathbb{E}[X_t]$
2. **Covariance:** $C(s,t) = \mathbb{E}[(X_t - m(t)) \otimes (X_s - m(s))]$

Examples of GPs: with continuous paths and $m(t) \equiv 0$,

- **Brownian Motion:** covariance $C(s,t) = s \wedge t$.
- **Fractional BM:** $\mathbb{E}[B_t^H B_s^H] = \frac{1}{2}(s^{2H} + t^{2H} - |t-s|^{2H})$.
- **Ornstein-Uhlenbeck (OU) Process:**

$$X_t = e^{-\theta t} X_0 + \int_0^t e^{-\theta(t-s)} dB_s$$

Brownian Motion

History:

- Robert Brown (1828): Pollen grains in water; irregular motion. Diffusion.
- Bachelier (1900): Stock prices fluctuate as BM.
- Einstein (1905): Molecular-kinetic theory.
- Wiener (1923): Rigorous math foundation ($\exists$).
- Levy (1948): construction of BM: interplolation path.
- Kolmogorov (1933): construction: FDDs and Extension Thm.

Definitions of BM

Definition (via FDDs): A Brownian Motion (BM) $\{B_t\}_{t\geq 0}$ starting from $B_0 = x$ is an SP with FDDs satisfying:

$$\mathbb{P}(B_{t_1} \in F_1, \ldots, B_{t_k} \in F_k)$$
$$= \int_{F_1} \cdots \int_{F_k} p(t_1, x, x_1) p(t_2 - t_1, x_1, x_2) \ldots p(t_k - t_{k-1}, x_{k-1}, x_k) dx_1 \ldots dx_k$$

where $p(t, x, y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(y-x)^2}{2t}}$ is the transition density.

- Transition: $B_t | B_s = z \sim \mathcal{N}(z, t - s)$ for $t > s$.
- BM has independent increments: $B_t - B_s \sim \mathcal{N}(0, t - s)$ independent of $\{B_u : u \leq s\}$.
- Green's function of heat equation: $\partial_t u = \frac{1}{2} \Delta u$.

Definitions of BM

Definition (via FDDs): A Brownian Motion (BM) $\{B_t\}_{t\geq 0}$ starting from $B_0 = x$ is an SP with FDDs satisfying:

$$\begin{aligned} &\mathbb{P}(B_{t_1} \in F_1, \dots, B_{t_k} \in F_k) \\ =&\int_{F_1} \cdots \int_{F_k} p(t_1, x, x_1) p(t_2 - t_1, x_1, x_2) \dots p(t_k - t_{k-1}, x_{k-1}, x_k) dx_1 \dots dx_k \end{aligned}$$

where $p(t, x, y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(y-x)^2}{2t}}$ is the transition density.

- Transition: $B_t | B_s = z \sim \mathcal{N}(z, t-s)$ for $t > s$.
- BM has independent increments: $B_t - B_s \sim \mathcal{N}(0, t-s)$ independent of $\{B_u : u \leq s\}$.
- Green's function of heat equation: $\partial_t u = \frac{1}{2} \Delta u$.

Definition (via GP): A BM is a continuous GP with mean $m(t) = 0$ and covariance $C(s, t) = s \wedge t$.

Definitions of BM

Definition (via FDDs): A Brownian Motion (BM) $\{B_t\}_{t\geq 0}$ starting from $B_0 = x$ is an SP with FDDs satisfying:

$$\mathbb{P}(B_{t_1} \in F_1, \ldots, B_{t_k} \in F_k)$$
$$= \int_{F_1} \cdots \int_{F_k} p(t_1, x, x_1) p(t_2 - t_1, x_1, x_2) \ldots p(t_k - t_{k-1}, x_{k-1}, x_k) dx_1 \ldots dx_k$$

where $p(t, x, y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(y-x)^2}{2t}}$ is the transition density.

- Transition: $B_t | B_s = z \sim \mathcal{N}(z, t - s)$ for $t > s$.
- BM has independent increments: $B_t - B_s \sim \mathcal{N}(0, t - s)$ independent of $\{B_u : u \leq s\}$.
- Green's function of heat equation: $\partial_t u = \frac{1}{2} \Delta u$.

Definition (via GP): A BM is a continuous GP with mean $m(t) = 0$ and covariance $C(s, t) = s \wedge t$.

Definition (Lévy characterization): A process $\{B_t\}_{t\geq 0}$ is a martingale with respect to its natural filtration and $B_t^2 - t$ is a martingale.

Properties and Examples

1. BM is a Gaussian Process.
2. Has independent increments.
3. Has a continuous version (by Kolmogorov Continuity Thm).

$$\mathbb{E}[|B_t - B_s|^4] = 3(t-s)^2$$

Properties and Examples

1. BM is a Gaussian Process.
2. Has independent increments.
3. Has a continuous version (by Kolmogorov Continuity Thm).

$$\mathbb{E}[|B_t - B_s|^4] = 3(t-s)^2$$

Transformations of BM:

- **Scaling:** $\hat{B}_t = c^{-1/2} B_{ct}$ is also a BM.
- **Time Reversal:** $B_{T-t} - B_T$ is a BM.
- **Inversion:** $X_t = tB_{1/t}$ (with $X_0 = 0$) is a BM.

OU process

- SDE: $dX_t = -\theta X_t dt + dB_t \Leftrightarrow X_t = e^{-\theta t} X_0 + \int_0^t e^{-\theta(t-s)} dB_s$

OU process

- SDE: $dX_t = -\theta X_t dt + dB_t \Leftrightarrow X_t = e^{-\theta t} X_0 + \int_0^t e^{-\theta(t-s)} dB_s$
- Time change of BM: $X_t = e^{-\theta t} B_{\frac{e^{2\theta t}-1}{2\theta}}$ when $X_0 = 0$.
- GP: Mean $m(t) = e^{-\theta t} m(0)$, Covariance $C(s,t) = \frac{1}{2\theta}(1 - e^{-2\theta s}) e^{-\theta(t-s)}$, $t \geq s$.
- As $t \to \infty$: $X_t \to \mathcal{N}(0, \frac{1}{2\theta})$.

OU process

- SDE: $dX_t = -\theta X_t dt + dB_t \Leftrightarrow X_t = e^{-\theta t} X_0 + \int_0^t e^{-\theta(t-s)} dB_s$
- Time change of BM: $X_t = e^{-\theta t} B_{\frac{e^{2\theta t}-1}{2\theta}}$ when $X_0 = 0$.
- GP: Mean $m(t) = e^{-\theta t} m(0)$, Covariance $C(s,t) = \frac{1}{2\theta}(1 - e^{-2\theta s}) e^{-\theta(t-s)}$, $t \geq s$.
- As $t \to \infty$: $X_t \to \mathcal{N}(0, \frac{1}{2\theta})$.

Quiz (T/F): If $X_0 \sim \mathcal{N}(0, \frac{1}{2\theta})$, then $X_t \sim \mathcal{N}(0, \frac{1}{2\theta})$ for all $t \geq 0$.

Stationary Processes

Strong Stationary: FDDs are invariant under time translation.

$$(X_{t_1}, \ldots, X_{t_k}) \stackrel{d}{=} (X_{t_1+h}, \ldots, X_{t_k+h})$$

Weak (2nd-Order) Stationary:

- $\mathbb{E}[X_t] \equiv \mu$ (constant mean)
- $\mathbb{E}[(X_t - \mu)(X_s - \mu)] = C(t-s)$ (Covariance depends only on lag)

Ergodicity (Prop 1.3, Pav14): If X_t is weak stationary with covariance $C \in L^1$, then:

$$\lim_{T\to\infty} \mathbb{E}\left|\frac{1}{T}\int_0^T X_s ds - \mu\right|^2 = 0$$

1. $X_t \equiv Z$, or (X_t) iid sequence when T is discrete
2. ARMA processes: $X_t = \sum_{i=0}^{p} a_i X_{t-i} + \sum_{j=0}^{q} b_j Z_{t-j}$ with (Z_t) iid Gaussian noise.
3. OU Process: $dX_t = -\theta X_t dt + dB_t$ with $X_0 \sim \mathcal{N}(0, \frac{1}{2\theta})$.