Lecture IV: General Probability Theory, Brownian Motion, Information, and Conditioning

Dangxing Chen

Duke Kunshan University

Infinite Probability Spaces

Random Variables and Distributions

Random Walk and Brownian Motion

Conditional Expectation

Motivation

- Go from discrete to continuous
- **Remark:** Not rigorous. More serious math will be in measure theory

Difference between discrete and continuous cases

- Two experiments keep in mind:
  - Choose a number randomly from the unit interval $[0, 1]$
  - Toss a coin infinitely many times
- Discrete probability space $(\Omega, \mathbb{P})$
  - Ω is finite
  - $\mathbb{P}(\omega)$ is defined
- Continuous probability space, say experiment 1
  - $\Omega = [0, 1]$ contains infinite points
  - $\mathbb{P}(\omega) =$

Difference between discrete and continuous cases

- Two experiments keep in mind:
  - Choose a number randomly from the unit interval $[0,1]$
  - Toss a coin infinitely many times
- Discrete probability space $(\Omega, \mathbb{P})$
  - Ω is finite
  - $\mathbb{P}(\omega)$ is defined
- Continuous probability space, say experiment 1
  - $\Omega = [0,1]$ contains infinite points
  - $\mathbb{P}(\omega) = 0$
  - Not well defined
- For continuous case, consider events instead
  - Discrete case: First toss is head
  - Continuous case: $[0, 0.5]$
- **Question**: What is the space of events?

Sigma algebra

- **Def:** Let $\mathcal{F}$ be a collection of subsets of Ω. We say that $\mathcal{F}$ is a ***σ*-algebra** provided that:
  1. $\emptyset \in \mathcal{F}$
  2. $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
  3. $A_1, A_2, \cdots \in \mathcal{F} \Rightarrow \cup_{n=1}^{\infty} A_n \in \mathcal{F}$
- E.g., flip coin twice

Sigma algebra

- **Def:** Let $\mathcal{F}$ be a collection of subsets of Ω. We say that $\mathcal{F}$ is a **σ-algebra** provided that:
  1. $\emptyset \in \mathcal{F}$
  2. $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
  3. $A_1, A_2, \cdots \in \mathcal{F} \Rightarrow \cup_{n=1}^{\infty} A_n \in \mathcal{F}$
- E.g., flip coin twice
  - $\Omega = \{HH, HT, TH, TT\}$
  - A_{HH} = The set of all sequences beginning with HH
  - $\mathcal{F} = \{\emptyset, \Omega, A_H, A_T, A_{HH}, A_{HT}, A_{TH}, A_{TT}, A_{HH}^c, A_{HT}^c, A_{TH}^c, A_{TT}^c, A_{HH} \cup A_{TH}, A_{HH} \cup A_{TT}, A_{HT} \cup A_{TH}, A_{HT} \cup A_{TT}\}$
- E.g., continuous case in $[0, 1]$
  - **Borel σ-algebra**: beginning with closed intervals and adding everything else necessary to have a σ-algebra
  - E.g., $(a, b), [a, b], (a, b), [a, b)$
  - E.g., $\left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right], \{a\}$

Probability space

- **Def:** Let Ω be a nonempty set, and let $\mathcal{F}$ be a σ-algebra of subsets of Ω. A probability measure $\mathbb{P}$ is a function that, to every set $A \in \mathcal{F}$, assigns a number in $[0, 1]$, call the probability of A and written $\mathbb{P}(A)$. We require:
  - $\mathbb{P}(\Omega) = 1$
  - Whenever $A_1, A_2, \ldots$ is a sequence of disjoint sets in $\mathcal{F}$, then

$$\mathbb{P}\left(\cup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n).$$

  The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a **probability space**.
- E.g., Uniform measure

$$\mathbb{P}([a, b]) = b - a, 0 \leq a \leq b \leq 1.$$

Infinite Probability Spaces

Random Variables and Distributions

Random Walk and Brownian Motion

Conditional Expectation

- **Def:** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A random variable is a real-valued function X defined on Ω with the property that for for every Borel subset B of $\mathbb{R}$, the subset of Ω given by

$$\{X \in B\} = \{\omega \in \Omega; X(\omega) \in B\}$$

  is in the σ-algebra $\mathcal{F}$.
- Ways to describe the probability
  - Continuous case:
    - Cumulative distribution function (CDF): $F(x) = \mathbb{P}(X \leq x)$.
    - Probability density function (PDF): $\mathbb{P}(a \leq X \leq b) = \int_a^b f(x)\ dx$
  - Discrete case:
    - Probability mass function: $p_i = \mathbb{P}(X = x_i)$.

Example

- X: Standard normal variable
- PDF: $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$
- CDF: $N(x) = \int_{-\infty}^{x} \varphi(z)\ dz$
- Y be a uniformly distributed random variable

Example

- X: Standard normal variable
- PDF: $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$
- CDF: $N(x) = \int_{-\infty}^{x} \varphi(z)\, dz$
- Y be a uniformly distributed random variable
- $X = N^{-1}(Y)$

$$\begin{aligned}
\mathbb{P}(\omega \in \Omega; a \leq X(\omega) \leq b) &= \mathbb{P}(\omega \in \Omega; a \leq N^{-1}(Y(\omega)) \leq b) \\
&= \mathbb{P}(\omega \in \Omega; a \leq N^{-1}(Y(\omega)) \leq b) \\
&= \mathbb{P}(\omega \in \Omega; N(a) \leq Y(\omega) \leq N(b)) \\
&= N(b) - N(a) = \int_a^b \varphi(x)\, dx
\end{aligned}$$

Expectation

- Discrete case: $\mathbb{E}[g(X)] = \sum_{i=1}^{n} g(x_i)\mathbb{P}(X = x_i)$
- Continuous case: $\mathbb{E}[g(X)] = \lim_{n\to\infty} \sum_{i=1}^{n} g(x_i) f(x_i)\Delta x$
- **THM**: For continuous random variable X with the density function f, we have

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\, dx.$$

- **Remark:** Textbook has a more rigorous argument.

Jensen's inequality

- **THM (Jensen's inequality):** If f is a convex, real-valued function defined on $\mathbb{R}$ and $\mathbb{E}[|X|] < \infty$, then

$$f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)].$$

- E.g., is $\mathrm{Var}(X) \geq 0$?

Jensen's inequality

- **THM (Jensen's inequality):** If f is a convex, real-valued function defined on $\mathbb{R}$ and $\mathbb{E}[|X|] < \infty$, then

$$f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)].$$

- E.g., is $\text{Var}(X) \geq 0$?
- $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}^2[X]$
- $f(x) = x^2$ is convex
- Jensen's inequality: $f(\mathbb{E}[X]) = \mathbb{E}^2[X] \leq \mathbb{E}[f(X)] = \mathbb{E}[X^2]$
- $\text{Var}(X) \geq 0$

Infinite Probability Spaces

Random Variables and Distributions

Random Walk and Brownian Motion

Conditional Expectation

Symmetric random walk

- **Question**: Limitation of Binomial model?

Symmetric random walk

- **Question**: Limitation of Binomial model?
- Discrete time and space
- In real life, both time and space are continuous
- Denote successive outcomes of the tosses by $\omega = \omega_1\omega_2\omega_3\dots$
- $X_j = \begin{cases} 1, & \text{if } \omega_j = H, \\ -1, & \text{if } \omega_j = T. \end{cases}$
- Assume $p = \frac{1}{2}$
- Symmetric random walk: $M_0 = 0$ and

$$M_k = \sum_{j=1}^{k} X_j, k = 1, 2, \dots$$

Some properties

- $M_l - M_k$ is called increments
- Nonoverlap increments are independent
- $\mathbb{E}[M_l - M_k] = 0$, $\text{Var}(M_l - M_k) = \sum_{j=k}^{l} \text{Var}(X_j) = l - k$
- **Martingale:** conditional expectation of the future value is equal to the present value

Some properties

- $M_l - M_k$ is called increments
- Nonoverlap increments are independent
- $\mathbb{E}[M_l - M_k] = 0$, $\text{Var}(M_l - M_k) = \sum_{j=k}^{l} \text{Var}(X_j) = l - k$
- **Martingale:** conditional expectation of the future value is equal to the present value

$$
\begin{aligned}
\mathbb{E}[M_l|M_k] &= \mathbb{E}\left[(M_l - M_k) + M_k|M_k\right] \\
&= \mathbb{E}[M_l - M_k|M_k] + \mathbb{E}[M_k|M_k] \\
&= \mathbb{E}[M_l - M_k|M_k] + M_k \\
&= \mathbb{E}[M_l - M_k] + M_k = M_k.
\end{aligned}
$$

Scaled symmetric random walk

- Scaled symmetric random walk (all indices integers):

$$W^{(n)}(t) = \frac{1}{\sqrt{n}} M_{nt}.$$

- $\mathbb{E}[W^{(n)}(t) - W^{(n)}(s)] = 0$, $\mathrm{Var}(W^{(n)}(t) - W^{(n)}(s)) = t - s$
- Martingale: $\mathbb{E}[W^{(n)}(t) | W^{(n)}(s)] = W^{(n)}(s)$
- Quadratic variation:

$$\begin{aligned}
[W^{(n)}, W^{(n)}](t) &= \sum_{j=1}^{nt} \left[W^{(n)}\left(\frac{j}{n}\right) - W^{(n)}\left(\frac{j-1}{n}\right) \right]^2 \\
&= \sum_{j=1}^{nt} \left[\frac{1}{\sqrt{n}} X_j \right]^2 = \sum_{j=1}^{nt} \frac{1}{n} = t.
\end{aligned}$$

- **THM**: Fix $t \geq 0$. As $n \to \infty$, $W^{(n)}(t) \to \mathcal{N}(0, t)$.

Verify

- Check moments
- Suppose $X \sim \mathcal{N}(0, t)$

$$\mathbb{E}[X] = 0,$$
$$\mathrm{Var}(X) = t.$$

- $W^{(n)}(t) = \frac{\sum_{j=1}^{nt} X_j}{\sqrt{n}}$

Verify

- Check moments
- Suppose $X \sim \mathcal{N}(0, t)$

$$\mathbb{E}[X] = 0,$$
$$\mathrm{Var}(X) = t.$$

- $W^{(n)}(t) = \frac{\sum_{j=1}^{nt} X_j}{\sqrt{n}}$

$$\mathbb{E}[W^{(n)}(t)] = \frac{1}{\sqrt{n}} \sum_{j=1}^{nt} \mathbb{E}[X_j] = 0,$$

$$\mathbb{E}[(W^{(n)}(t))^2] = \frac{1}{n} \mathbb{E}\left[\left(\sum_{j=1}^{nt} X_j\right)^2\right] = \frac{1}{n} \mathbb{E}\left[\sum_{j=1}^{nt} X_j^2 + \sum_{j \neq k} X_j X_k\right]$$
$$= \frac{1}{n} \sum_{j=1}^{nt} \mathbb{E}[X_j^2] + \frac{1}{n} \sum_{j \neq k} \mathbb{E}[X_j]\mathbb{E}[X_k] = t.$$

Moment-generating function

- **Q:** How to find a way to incorporate all moments information?
- **Def:** The **moment-generating function** for a random variable X is

$$M_X(s) = \mathbb{E}[e^{sX}].$$

- **Q:** Why moment-generating function contains all moments information?

Moment-generating function

- **Q:** How to find a way to incorporate all moments information?
- **Def:** The **moment-generating function** for a random variable X is

$$M_X(s) = \mathbb{E}[e^{sX}].$$

- **Q:** Why moment-generating function contains all moments information?

$$M_X^{(n)}(0) = \mathbb{E}[X^n].$$

- E.g., $X \sim \mathcal{N}(0, t)$

Moment-generating function

- **Q:** How to find a way to incorporate all moments information?
- **Def:** The **moment-generating function** for a random variable X is

$$M_X(s) = \mathbb{E}[e^{sX}].$$

- **Q:** Why moment-generating function contains all moments information?

$$M_X^{(n)}(0) = \mathbb{E}[X^n].$$

- E.g., $X \sim \mathcal{N}(0, t)$

$$\mathbb{E}[e^{sX}] = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} \ e^{sx} \ dx = \int_{\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-st)^2}{2t}} \cdot e^{\frac{s^2 t}{2}} \ dx$$
$$= e^{\frac{s^2 t}{2}}.$$

Binomial distribution

- $W^{(n)}(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^{nt} X_j$

Binomial distribution

- $W^{(n)}(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^{nt} X_j$

$$\mathbb{E}\left[e^{sW^{(n)}(t)}\right] = \mathbb{E}\left[e^{\frac{s}{\sqrt{n}}\sum_{j=1}^{nt} X_j}\right] = \mathbb{E}\left[\prod_{j=1}^{nt} e^{\frac{sX_j}{\sqrt{n}}}\right] = \prod_{j=1}^{nt} \mathbb{E}\left[e^{\frac{sX_j}{\sqrt{n}}}\right]$$
$$= \prod_{j=1}^{nt}\left(\frac{1}{2}e^{\frac{s}{\sqrt{n}}} + \frac{1}{2}e^{-\frac{s}{\sqrt{n}}}\right)$$
$$= \left(\frac{1}{2}e^{\frac{s}{\sqrt{n}}} + \frac{1}{2}e^{-\frac{s}{\sqrt{n}}}\right)^{nt}.$$

- HW: $\lim_{n\to\infty} \mathbb{E}[e^{sW^{(n)}(t)}] = e^{\frac{s^2 t}{2}}$

Brownian motion

- Obtain Brownian Motion (BM) as the limit of the scaled random walks $W^{(n)}(t)$
- **Definition:** Suppose there is a continuous function $W(t)$ of $t \geq 0$ that satisfies $W(0) = 0$. Then $W(t), t \geq 0$, is a **Brownian motion** if for all $0 = t_0 < t_1 < \cdots < t_m$, the increments

  $$W(t_1) - W(t_0), W(t_2) - W(t_1), \ldots, W(t_m) - W(t_{m-1})$$

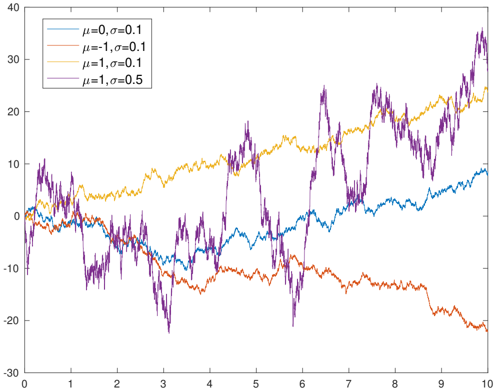
  are independent and each of these increments is normally distributed with

  $$\mathbb{E}[W(t_{i+1}) - W(t_i)] = 0,$$
  $$\text{Var}(W(t_{i+1}) - W(t_i)) = t_{i+1} - t_i.$$

- Key difference: Continuous time and space

Drifted Brownian motion

- $X_t = \mu t + \sigma W_t$
- Stochastic differential equation: $dX_t = \mu dt + \sigma dW_t$
- Discrete version: $X_{\Delta t} = \mu \Delta t + \sigma dW_{\Delta t}$
- Simulation: $X_{t+\Delta t} - X_t \sim \mu \Delta t + \sigma \mathcal{N}(0, \Delta t)$

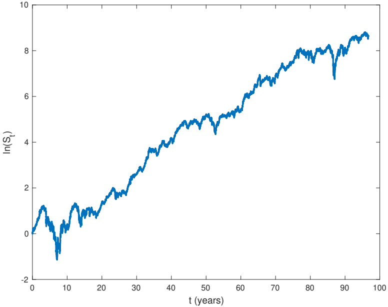


More on drifted Brownian motion

- **Q**: Why linear drift term μt?

More on drifted Brownian motion

- **Q**: Why linear drift term μt?
- Log return increases linearly

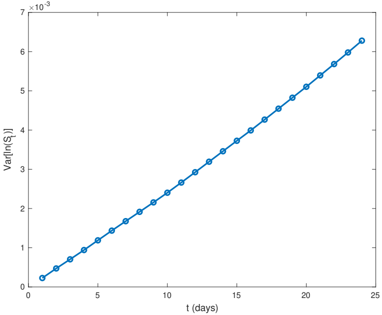


More on drifted Brownian motion

- **Q**: Var$(\sigma W_t) = \sigma^2 t$ reasonable?

More on drifted Brownian motion

- **Q**: Var$(\sigma W_t) = \sigma^2 t$ reasonable?



Properties of Brownian motion

- Covariance $(t > s)$:

Properties of Brownian motion

- Covariance ($t > s$):

$$\begin{aligned}\mathbb{E}[W(s)W(t)] &= \mathbb{E}[W(s)(W(t) - W(s)) + W^2(s)] \\ &= \mathbb{E}[W(s)] \cdot \mathbb{E}[W(t) - W(s)] + \mathbb{E}[W^2(s)] \\ &= 0 + \mathsf{Var}(W(s)) = s.\end{aligned}$$

- **THM:** Brownian motion is a martingale.

Properties of Brownian motion

- Covariance ($t > s$):

$$\begin{aligned}\mathbb{E}[W(s)W(t)] &= \mathbb{E}[W(s)(W(t) - W(s)) + W^2(s)] \\ &= \mathbb{E}[W(s)] \cdot \mathbb{E}[W(t) - W(s)] + \mathbb{E}[W^2(s)] \\ &= 0 + \mathrm{Var}(W(s)) = s.\end{aligned}$$

- **THM:** Brownian motion is a martingale.

$$\begin{aligned}\mathbb{E}[W(t)|\mathcal{F}(s)] &= \mathbb{E}[(W(t) - W(s)) + W(s)|W(s)] \\ &= \mathbb{E}[W(t) - W(s)|W(s)] + \mathbb{E}[W(s)|W(s)] \\ &= \mathbb{E}[W(t) - W(s)] + W(s) = W(s).\end{aligned}$$

Derivative

- **Question**: Is BM $W(t)$ differentiable w.r.t t?

Derivative

- **Question**: Is BM $W(t)$ differentiable w.r.t t?

$$\lim_{\Delta t \to 0} \frac{W(t+\Delta t) - W(t)}{\Delta t} = \lim_{\Delta t \to 0} \frac{\mathcal{N}(0, \Delta t)}{\Delta t}$$
$$= \lim_{\Delta t \to 0} \frac{\mathcal{N}(0,1)}{\sqrt{\Delta t}} = \text{DNE}.$$

- Trajectory of the Brownian motion is zigzag

Transition density function

- **Q:** What is the transition density function of $p(x, y, \tau)$?
- Given that $W_t = y$, the density function of $W_{t+\tau} = x$

Transition density function

- **Q:** What is the transition density function of $p(x, y, \tau)$?
- Given that $W_t = y$, the density function of $W_{t+\tau} = x$
- $W_{t+\tau}|W_t = (W_{t+\tau} - W_t)|W_t + W_t|W_t = \mathcal{N}(0, \tau) + y$

$$p(x, y, \tau) = \frac{1}{\sqrt{2\pi\tau}} e^{-\frac{(x-y)^2}{2\tau}} \ dx$$

Conditional Expectation

- **THM:** Let X and Y be independent random variables, then $f(X)$ and $g(Y)$ are independent random variables.
- **THM:** Let X_t and Y_t be continuous-time random variables with $0 \le s \le t \le T$
  - **Linearity:** $\forall c_1, c_2 \in \mathbb{R}$,
  $$\mathbb{E}[c_1 X_t + c_2 Y_t | X_s, Y_s] = c_1 \mathbb{E}[X_t | X_s] + c_2 \mathbb{E}[Y_t | Y_s].$$
  - **Taking out what is known:** $\mathbb{E}[X_s X_t | X_s] = X_s \mathbb{E}[X_t | X_s]$
  - **Iterated conditioning:** $\mathbb{E}[\mathbb{E}[X_T | X_t] | X_s] = \mathbb{E}[X_T | X_s]$
  - **Independence:** $\mathbb{E}[W_t - W_s | W_s] = \mathbb{E}[W_t - W_s]$
  - **Conditional Jensen's inequality:** If f is a convex function, then
  $$f(\mathbb{E}[X_t | X_s]) \le \mathbb{E}[f(X_t) | X_s].$$
- **Def:** Consider an adapted stochastic process $X(t), 0 \le t \le T$. $\forall 0 \le s \le t \le T$, $\forall f$, $\exists g$ s.t.
$$\mathbb{E}[f(X(t)) | \mathcal{F}_s] = g(X(s)).$$