The Epilogue: Brownian Motion and Beyond.

1. Central limit theorem and the universality of Gaussian distribution.

That the number of heads of a coin tossing follows an approximate normal distribution was first discovered and mathematically proved by Abraham de Moivre in 1733 for fair coins, and later for biased coins by Laplace in 1812. In mathematical terms, let $\xi_i = 1$ or 0 if the i-th toss is a head or tail. Then, as $n \to \infty$,

$$\frac{\sum_{i=1}^n \xi_i - np}{\sqrt{np(1-p)}} \to N(0,1).$$

where $N(0,1)$ is the so called standard normal distribution with density $f(x) = e^{-x^2/2}/\sqrt{2\pi}$, with the shape of the famous *bell curve.* A generalization around the end of the nineteenth century and thereafter involved many fundamental works and developments in probability theory. A standard, but not the most general form may be stated as follows. If $X_1, ..., X_n, ...$ are independent and identically distributed random variables with mean μ and variance σ^2. Then, as $n \to \infty$,

$$\frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} \to N(0,1).$$

This is the *central limit theorem* or *de Moivre-Laplace Theorem.*

The most important implication of the central limit theorem is that, regardless of the distribution of X_i, its sample mean can always be approximated by a normal distribution. Here lies the universality and importance of the normal distribution.

2. Definition of Brownian motion.

In 1827, Robert Brown, a Scottish botanist, observed the irregular motion of pollen grains under microscope, caused by constant knocking in all directions by the water molecules. This is the first discovery of Brownian motion (BM). Bachelier (1900) applied BM to model the fluctuation of stock prices. Albert Einstein in 1905 derived the transition density of BM and found the *Einstein relation.* Norbert Wiener (1923, 1924) was the first mathematician to provide rigorous mathematical formulation of BM. In mathematical/literature, BM is often called the *Wiener process.*

Definition. A continuous time stochastic process $\{X(t) : t \in [0, \infty)\}$ with continuous path satisfying

(i) (independent increment) For any $0 \le t_0 < t_1 < ...t_n$,

$$X(t_1) - X(t_0),\ X(t_2) - X(t_1),\ ...,\ X(t_n) - X(t_{n-1}) \qquad \text{are all independent.}$$

(ii) (marginal normal distribution) $X(t) - X(s)$ follows normal distribution with mean 0 and variance $\sigma^2(t-s)$.

(iii) $X(0) = 0$.

Then $\{X(\cdot)\}$ is a BM. If $\sigma = 1$, it is called standard BM.

You may notice the formal resemblance between the definitions of BM and Poisson process, both requiring independent increment. Indeed, they are special cases of *Levy processes* which require independent and stationary increments.

Recall that an understanding of Poisson process, through binomial approximation, is accumulation of independent teeny-tiny Bernoulli head-counting over time. Here, a basic understanding of BM is that it is accumulation of independent itsy-bitsy normal random variables along time.

By some historical misunderstanding, the normal distribution is also called Gaussian distribution, and a stochastic process with continuous path and marginal (finite dimension) normal distributions is called Gaussian process. BM is a special case of Gaussian process.

3. Some basic properties.

BM has some very interesting path properties, some of which are listed below. Let $W_t : t \in [0, \infty)$ be a standard BM.

(a). $\{W_t\}$ is Markov process.

(b). Transformation invariant:

(Scaling) $(1/c)W_{ct}$ is still a BM on $[0, \infty)$ for any $c > 0$;

(Time inversion) $tW_{1/t}$ is still a BM on $[0, \infty)$;

(Time reversal) $W_T - W_{T-t}$ is a BM on $[0, T]$;

(Symmetry) $-W_t$ is a BM on $[0, \infty)$.

(c). With probability 1, $W_\cdot$ is monotone in no interval.

(d). With probability 1, $W_\cdot$ is nowhere differentiable.

(e). The random zero set $\mathcal{L} = \{t \in [0, \infty) : W_t = 0\}$ is, with probability 1, Lebesgue measure 0, closed and unbounded, without isolated point and therefore dense in itself.

(f). (Law of large numbers:) $W_t/t \to 0$ as $t \to \infty$ with probability 1.

(g). (Law of iterated logarithm:) with probability 1,

$$\limsup_{t\to\infty} W_t/\sqrt{2t \log\log t} \to 1 \quad \text{and} \quad \liminf_{t\to\infty} W_t/\sqrt{2t \log\log t} \to -1.$$

4. Donsker's invariance principle.

As with the universality of the normal distribution among all distributions of random variables, the same is true with BM for all stochastic processes. This is demonstrated as Donsker's invariance principle.

Let $\xi_i, i = 1, 2, ...$ be iid with mean 0 and variance σ^2. Set $S_0 = 0$ and $S_n = \sum_{i=1}^n \xi_i$, called partial sums. Define

$$X_t = S_{[t]} + (t - [t])\xi_{[t]+1}, \qquad t \in [0, \infty)$$

where $[t]$ is the integer part of t. $X_t, t \in [0, \infty)$ is an interpolation of $S_0, S_1, S_2, ...$ Then, as $n \to \infty$,

$$\{\frac{X_{nt}}{\sqrt{n}\sigma} : t \in [0, \infty)\} \qquad \text{converges weakly to a standard BM.}$$

If t is fixed, the convergence of $X_{nt}/\sqrt{n\sigma^2}$ to $N(0, t)$ is just the central limit theorem. As a process, its (weak) convergence is somewhat more technical, even for the definition. Ignoring the technicalities, the point here is that, regardless of the distribution of ξ_i, the (weak) limit is (the invariant) the same BM.

5. Stochastic integrals and the Itô formula.

Stochastic integrals and the Itô rule to probability theory play the role of Riemann or Lebesgue integral and the chain rule to calculus. Recall that the Riemann-Stieltjes integral of $\int_0^1 f(x)dG(x)$ with integrand f with respect to G monotone (or, more generally, of bounded variation) is the limit, if exists, of

$$\sum_{j=0}^{n-1} f(t_j)[G(t_{j+1}) - G(t_j)], \qquad \text{as} \quad \max\{t_{j+1} - t_j : 0 \le j \le n-1\} \to 0,$$

where $0 = t_0 < t_1 < ... < t_n = 1$ is any partition of $[0, 1]$. The integrand f does not have to continuous but needs to be good enough to be well approximated by step function.

The Itô integral $\int_0^1 X(t)dW_t$ with integrand $X(\cdot)$, which can be random, and with respect to BM W_t is defined as the limit, if exists, of

$$\sum_{j=0}^{n-1} X(t_j)[W_{t_{j+1}} - W_{t_j}], \qquad \text{as} \quad \max\{t_{j+1} - t_j : 0 \le j \le n-1\} \to 0,$$

where $0 = t_0 < t_1 < ... < t_n = 1$ is any partition of $[0,1]$.

The definition of the Itô integral applies more generally to martingales, in which BM is a special case. Although the definition of the Itô integral and that of Riemann-Stieltjes integral appear to be formally same, the true difficulty lies in the existence of the limit, and requirements on the integrands.

The chain rule in calculus states that, for differentiable functions f and g,

$$df(g(x)) = f'(g(x))dg(x) = f'(g(x))g'(x)dx.$$

This is because, for small Δx, ignoring higher order smaller terms,

$$f(g(x+\Delta x)) - f(g(x)) \approx f'(g(x)(g(x+\Delta x) - g(x)) \approx f'(g(x))g'(x)\Delta x.$$

In the stochastic calculus, the Itô states that, for twice continuously differentiable function f,

$$df(W_t) = f'(W_t)dW_t + \frac{1}{2}f''(W_t)dt$$

or equivalently,

$$f(W_b) - f(W_a) = \int_a^b f(W_t)dW_t + \frac{1}{2}\int_a^b f''(W_t)dt.$$

A naive understanding based on the Taylor expansion is simply, for small Δt,

$$\begin{aligned} f(W_{t+\Delta t}) - f(W_t) &\approx f'(W_t)[W_{t+\Delta t} - W_t] + \frac{1}{2}f''(W_t)[W_{t+\Delta t} - W_t]^2 \\ &\approx f'(W_t)[W_{t+\Delta t} - W_t] + \frac{1}{2}f''(W_t)\Delta t. \end{aligned}$$

The last approximation seems to rest on the fact that $E(W_{t+\Delta t} - W_t)^2 = \Delta t$, but may still be somewhat unnatural. A slightly better way to understand it is through integration. Let $a \le t_0 < ... < t_n = b$ with $t_{i+1} - t_i$ being very small. Then,

$$\begin{aligned} & f(W_b) - f(W_a) \\ = & \sum_{j=0}^{n-1}[f(W_{t_{j+1}}) - f(W_{t_j})] \\ \approx & \sum_{j=0}^{n-1}\Big\{f'(W_{t_j})[W_{t_{j+1}} - W_{t_j}] + \frac{1}{2}f''(W_{t_j})[W_{t_{j+1}} - W_{t_j}]^2\Big\} \\ \approx & \sum_{j=0}^{n-1}\Big\{f'(W_{t_j})[W_{t_{j+1}} - W_{t_j}] + \frac{1}{2}f''(W_{t_j})[t_{j+1} - t_j]\Big\} \quad \text{since } E[W_{t_{j+1}} - W_{t_j}]^2 = t_{j+1} - t_j \\ \approx & \int_a^b f(W_t)dW_t + \frac{1}{2}\int_a^b f''(W_t)dt. \end{aligned}$$

As examples, one can easily see that

$$dW_t^2 = 2W_t dW_t + dt \qquad \text{and} \qquad W_t^2 = \int_0^t 2W_s dW_s + t.$$

6. The Camero-Martin-Girsanov Theorem.

Suppose Y is a random variable (possibly of high dimension) with true density f_1. Let f_0 be another density. Suppose we change the probability on the probability space to, for any event A,

$$\tilde{P}(A) = E[1_A f_0(Y)/f_1(Y)].$$

Then,

$$\tilde{P}(Y \in [a,b]) = E(1_{\{Y\in[a,b]\}} f_0(Y)/f_1(Y)) = \int_a^b f_1(y) f_0(y)/f_1(y)dy = \int_a^b f_0(y)dy.$$

It implies, under $\tilde{P}$, Y has density f_0. This is in relation with the test of hypothesis with $H_1 = f_1$ and $H_0 = f_0$.

Now consider a BM with a drift $W_t - at$ on $[0,T]$, where T is a fixed constant. This is certainly not a BM unless $a = 0$. Define

$$Z_t = e^{aW_t - (1/2)a^2 t}, \qquad t \in [0,T].$$

Suppose we change the probability measure to

$$\tilde{P}(A) = E[1_A Z_T]$$

for all proper sets A. Then, under this new probability measure, $W_t - at$ in a BM on $[0,T]$. This is the Camero-Martin-Girsanov Theorem, proved in Camero-Martin (1944) for nonrandom drift and in Girsanov (1960) for random drifts.

7. Application: Black-Scholes pricing of European option.

Consider a wealth process $X(t)$, which can only be distributed in a risk-free financial instrument, e.g., bank deposit, which grows at compound rate $r > 0$, and a security with a price process $S(t)$ satisfying

$$dS(t) = S(t)[bdt + \sigma dW_t]$$

where $\sigma > 0$ and b are constants, and $S(0) = s_0 > 0$. Then, using the Ito rule,

$$S(t) = s_0 e^{\sigma W_t + (b-\sigma^2/2)t}.$$

Let $\pi(t)$ be the amount wealth in the security at time t. Then $X(t) - \pi(t)$ is the amount deposited in bank. Again, by the Ito rule,

$$dX(t) = rX(t)dt + (b-r)\pi(t)dt + \sigma\pi(t)dW_t$$

with $X(0) = x > 0$. Then,

$$X(t) = e^{rt}(x - \int_0^t e^{-rs}\pi(s)d\tilde{W}_s)$$

where $\tilde{W}_t = W_t + \theta t$ and $\theta = (b-r)/\sigma$.

Define a measure $\tilde{P}$ such that

$$\tilde{P}(A) = E(Z_T 1_A)$$

where $Z_t = e^{-\theta W_t - (1/2)\theta^2 t}$. It follows from the Girsanov Theorem that $\tilde{W}_t$ is a BM under probability measure $\tilde{P}$.

A European call option on security S with strike price K and expiration date T is a payoff of $f_T \equiv (S(T) - K)^+ = \max\{S(T) - K, 0\}$ at time T. A natural question then is: what is the fair price of the option at the outset time 0?

A hedging strategy against the European call option is a portfolio strategy $\pi(\cdot), X(\dot{)} - \pi(\cdot)$ with $X(0) = x > 0$ such that $X(t) \geq 0$ for $t \in [0,T]$ and $X(T) = f_T$. In other words, a hedging strategy is to duplicate the payoff of the option by owning and switching between the security and cash. Should such duplicate strategies exist, then the lowest beginning wealth x should be the fair price of the option. Otherwise, the so-called arbitrage opportunity exists.

The fair price for the call option is

$$e^{-rT}\tilde{E}(f_T) = e^{-rT}E(f_T e^{-\theta W_T - (1/2)\theta^2 T})$$

A heuristic understanding is that, suppose $X(T) = f(T)$ Since $\tilde{W}_t$ is a BM under $\tilde{P}$, $\tilde{E}(\int_0^t \pi(s)\tilde{W}_s) = 0$. As a result, $\tilde{E}(X(t)) = e^{rt}x$. Therefore, $x = e^{-rT}\tilde{E}(X(T)) = e^{-rT}E(f(T))$. A rigorous proof is somewhat technical and is omitted.

It's interesting to observe that since

$$S(t) = s_0 e^{\sigma W_t + (b-\sigma^2/2)t} = s_0 e^{\sigma \tilde{W}_t - \sigma\theta t + (b-\sigma^2/2)t} = s_0 e^{\sigma \tilde{W}_t + (r-\sigma^2/2)t}$$

which does not involve b under measure $\tilde{P}$. Therefore the fair price for the call option,

$$e^{-rT}\tilde{E}(f_T) = e^{-rT}\tilde{E}[(S(T) - K)^+]$$

is also irrelevant with b! In fact, it can be computed in a straightforward fashion. In fact, slightly more generally the fair price for the call option at time $t \in [0, T)$ is $v(t, S(t))$, where

$$v(t, y) = \begin{cases} y\Phi(\rho_+(T-t, y)) - Ke^{-r(T-t)}\Phi(\rho_-(T-t, y)) & \text{for } 0 \le t < T, y > 0 \\ (y-K)^+ & \text{for } t = T, y \ge 0. \end{cases}$$

Here $\Phi(\cdot)$ is the cdf of the standard normal distribution, and

$$\rho_\pm(s, y) = 1/(\sigma\sqrt{t})[\log(y/K) + s(r \pm \sigma^2/2)].$$

Notice that $v(t, y)$ satisfy the partial differential equation

$$\begin{cases} -\frac{\partial v}{\partial t} + rv = \frac{1}{2}\sigma^2 y^2 \frac{\partial^2 v}{\partial y^2} + ry\frac{\partial v}{\partial y} \\ v(T, y) = (y-K)^+ \end{cases}$$

which is the so called Black-Scholes partial differential equation.

Remark. Much of the BM related theory is niced presented in the framework of continuous time martingales. Without turning to martingales, especially the concept of filtration and conditional expectation with respect to σ-algebra, any introduction of BM shall be inevitably shallow. Unfortunately, a deeper exploration exceeds the level/requirement of this course. Students interested in the subject may consider to take an advanced course in probability theory or stochastic processes.