Continuous time finance

Abhinav Anand, IIMB

2022/08/02

Laws of Large Numbers

Weak law

For a very long sequence of iid random variables X_i such that $\mathbb{E}|X| = \mu < \infty$, the sample average converges in probability (and hence in distribution) to the expected value

$$\lim_{n \to \infty} \Pr \left[\left| \frac{\sum_{i=1}^{n} X_i}{n} - \mu \right| \geq \frac{1}{n} \right] = 0$$

Strong law

Almost surely, (i.e., for almost all states of the world $\omega \in \Omega$) the sample average converges to the expected value

$$\Pr \left[\{\omega \in \Omega : \lim_{n \to \infty} X_n(\omega) = X(\omega)\} \right] = 1$$

Strong law implies the weak law. It also implies convergence in rth mean.

$$\lim_{n \to \infty} \sum_{i=1}^{n} \frac{X_i^r}{n} = \mathbb{E}(X^r)$$

In particular, for $r = 1$, this implies that the sample averages converge to their theoretical expectation:

$$\lim_{n \to \infty} \sum_{i=1}^{n} \frac{X_i}{n} = \mathbb{E}(X)$$

Central Limit Theorem

Further, the central limit theorem applies to the *distribution* of the sample average and suggests that when the first two moments of X_i are finite

$$\frac{\left(\sum_{i=1}^{n} \frac{X_i}{n} \right) - \mu}{\frac{\sigma}{\sqrt{n}}} \xrightarrow{D} \mathcal{N}(0, 1)$$

where 'D' denotes convergence in distribution.

The central limit theorem does not need X_i to have normal distribution but suggests that any iid sequence howsoever distributed will converge to the standard normal distribution in large enough samples. The key assumption is that of finite second moments—for example, Cauchy random variables with infinite means or stable random variables with infinite variance do not obey the central limit theorem.

Functional Central Limit Theorem (Donsker's Theorem)

Let X_i be random variables that take value +1 and -1 with equal probability. Clearly, the mean is 0 and variance is 1. The sum of these random variables

$S_k = X_1 + \ldots + X_k$ is a random walk. We consider the position of S_k in time $[0,1]$. We divide $[0, 1]$ in n equal parts of length $1/n$. We then construct a rescaled random walk:

$$U_{k,\frac{1}{n}} = S_k \cdot \sqrt{\frac{1}{n}} = \sqrt{\frac{1}{n}} \sum_{i=1}^{k} X_i$$

$U_{k,\frac{1}{n}}$ is defined for discrete points $k = 1, 2, \ldots$. Hence to extend it to a real $t \in [\frac{k}{n}, \frac{k+1}{n}]$, we interpolate as follows:

$$U_{t,\frac{1}{n}} = \lambda U_{k,\frac{1}{n}} + (1-\lambda)U_{k+1,\frac{1}{n}}$$

Some observations:

1. $\lim_{n\to\infty} U_{t,\frac{1}{n}} = \lim_{n\to\infty} \sqrt{\frac{k}{n}} \frac{S_k}{\sqrt{k}} \xrightarrow{D} \mathcal{N}(0,t)$

2. For $t > s$, assuming $\frac{k}{n} < t < \frac{k+1}{n}$ and $\frac{l}{n} < s < \frac{l+1}{n}$, as $n \to \infty$, $t \approx \frac{k}{n}$ and $s \approx \frac{l}{n}$,

$$U_{k,\frac{1}{n}} - U_{l,\frac{1}{n}} = \sqrt{\frac{1}{n}} \sum_{i=l+1}^{k} X_i = \sqrt{\frac{1}{n}} \sum_{i=1}^{k-l} X_i \xrightarrow{D} \mathcal{N}(0,t-s)$$

3. For $t_1 < t_2 < t_3 < t_4$, $cov(U_{t_2,\frac{1}{n}} - U_{t_1,\frac{1}{n}}; U_{t_4,\frac{1}{n}} - U_{t_3,\frac{1}{n}}) = 0$

Hence the rescaled random walk has the follwing key properties:

1. $U_t \sim \mathcal{N}(0,t)$
2. Increments are independent and stationary.
3. It is continuous

Hence as $n \to \infty$ the rescaled random walk converges to a Wiener process.

Ito integral

Standard phenomena often manifest themselves in the guise of differential equations which may be solved somewhat straightforwardly:

$$dX = g(X,t)dt$$

$$\int_0^T dX = \int_0^T g(X,t)dt$$

However, when we suspect some random noise to be acting additively then we may propose the following modification:

$$dX = g(X,t)dt + \sigma(X,t)dW_t$$

$$\int_0^T dX = \int_0^T g(X,t)dt + \int_0^T \sigma(X,t)dW_t$$

Ito integral tries to make sense of the stochastic integral $\int_0^T \sigma(X,t)dW_t$. Note that the stochastic integral always takes the form of one stochastic process (here $\sigma(X,t)$) being integrated *against* another stochastic process (here dW_t) to yield yet another stochastic process. All other integrals are deterministic and can be handled by standard integration techniques (Riemann or Lebesgue).

We can try to approximate the above integral by means of sums where the time interval $[0,T]$ has been partitioned into n parts of length $\delta t = \frac{T}{n}$ each.

$$X_T - X_0 = \sum_{i=1}^n g(X,t)\delta t + \sum_{i=1}^n \sigma(X,t)\delta W_t$$

Informally, the stochastic integral is a *random variable* defined as a limit of a sequence of random variables. The paths of the Wiener Process are continuous but nowhere differentiable and hence classical integration techniques fail. The main insight is that the stochastic integral can be defined as long as the integrand only depends on information available in the past. Roughly, we choose

a sequence of partitions of the time interval and construct a special type of Riemann sum using particular instantiations of the integrator. Further, it is crucial which point in each of the small intervals is used to compute the value of the function. The limit then is taken *in probability* as the length of the partition becomes smaller and smaller.

Ito integral of a Wiener process

In this section we discuss the Wiener process integrated against itself.

Consider a Wiener process W_t with $t \in [0, 1]$ and $0 = t_1 < t_2 < \ldots < t_n = 1$. Pick $\tau_i \in [t_{i-1}, t_i]$ and evaluate the integral $\int_0^T W_t dW_t$ by a Riemann sum as follows:

$$I_n = \sum_{i=1}^{n} W_{\tau_i}(W_{t_i} - W_{t_{i-1}})$$

In general, $\tau_i \in [t_{i-1}, t_i]$ and hence for some $\lambda \in (0, 1)$ can be represented as

$$\tau_i = t_{i-1} + \lambda \cdot (t_i - t_{i-1})$$

However, it can be shown that the classical Riemann formula that gives $\int_0^T W_t dW_t = 1/2 W_t^2$ holds only when $\lambda = 1/2$—dubbed the *Stratanovich* stochastic integral. The *Ito* integral is obtained at $\lambda = 0$—i.e., when $\tau_i = t_{i-1}$—when the left end point of the partition is chosen. In this case:

$$\int_0^T W_t dW_t = \lim_{n \to \infty} \sum_{i=1}^{n} W_{t_{i-1}}(W_{t_i} - W_{t_{i-1}})$$

Kolmogorov equations

Suppose a given Wiener process starts at $W_0 = 0$. How do we describe the probability that the process will be at time T given that it was located at

x at time $t < T$. In this case, it not hard to see that $W_t \sim \mathcal{N}(x, T - t)$. (Why?) Further, this also helps us compute the probability that the process lies between two points W_l, W_h:

$$\Pr(W_l < W_T < W_h) = \Phi\left(\frac{W_h - x}{T - t}\right) - \Phi\left(\frac{W_l - x}{T - t}\right)$$

where $\Phi(\cdot)$ is the Normal distribution function. If we denote by $p(y, T|x, t)$ the transition probability of the Wiener process reaching y at time T when it is located at x at time t, it has the following density:

$$p(y, T|x, t) = \frac{1}{\sqrt{2\pi(T - t)}} \exp\left(-\frac{1}{2}\frac{(y - x)^2}{T - t}\right)$$

Considering a general diffusion $dX = \mu(X_t, t)dt + \sigma(X_t, t)dW_t$ and its conditional expectation:

$$f(y, t) = \mathbb{E}\left(h(X_T)|y, t\right) = \int_z h(z)p(z, T|y, t)dz$$

where $h(\cdot)$ is smooth. In the same way, we consider $f(x, s)$ where $s < t$:

$$f(x, s) = \mathbb{E}\left(h(X_T)|x, s\right) = \int_z h(z)p(z, T|x, s)dz$$
$$f(x, s) = \int_z h(z) \int_y p(z, T|y, t) \cdot p(y, t|x, s)dy$$
$$f(x, s) = \int_y \left(\int_z h(z)p(z, T|y, t)dz\right) p(y, t|x, s)dy$$
$$f(x, s) = \mathbb{E}(f(y, t)|x, s)$$

This implies that the conditional expected value for the future time T based on information available till time $s < t$ is simply $f(x, s)$. This implies that $f(X_t, t)$ is a martingale and hence its drift must be 0. From Ito's lemma, since

$$df = \left(1/2\sigma^2\frac{\partial^2 f}{\partial x^2} + \mu\frac{\partial f}{\partial x} + \frac{\partial f}{\partial t}\right)dt + \sigma\frac{\partial f}{\partial x}dX_t$$

Since the drift must be 0,

$$1/2\sigma^2 \frac{\partial^2 f}{\partial x^2} + \mu \frac{\partial f}{\partial x} + \frac{\partial f}{\partial t} = 0$$

Since $f(x,t) = \int_z h(z)p(z,T|x,t)dz$, we substitute above to get:

$$\int_z h(z)\left(1/2\sigma^2 \frac{\partial^2 p}{\partial x^2} + \mu \frac{\partial p}{\partial x} + \frac{\partial p}{\partial t}\right)dz = 0$$

This yields the celebrated Kolmogorov backward equation.

$$1/2\sigma^2 \frac{\partial^2 p}{\partial x^2} + \mu \frac{\partial p}{\partial x} + \frac{\partial p}{\partial t} = 0$$

Fundamental theorems of asset pricing

Martingales

A stochastic process M_s is a martingale if for time index $s < t$, the conditional mean of the process equals its level.

$$\mathbb{E}(M_t|\mathcal{F}_s) = \mathbb{E}_s(M_t) = M_s$$

where $\mathcal{F}_s$ denotes 'filtration' or available information till time s.

Examples:

1. Wiener process: Since $\mathbb{E}(W_t|W_s) = W_s$ for $s < t$.
2. Aggregated Wiener process: Since $\mathbb{E}(\int_0^t dW_t| \int_0^s dW_s) = \int_0^s dW_s$.
3. In general, any aggregated function $h(\cdot)$ of a Wiener process—$\int_0^t h(u)dW_u$—is also a martingale as long as $h(\cdot)$ is well-behaved.

Change of drift

Consider a normal random variable $X = \mathcal{N}(\mu, \sigma^2)$. It has a moment generating function $M_X(\lambda) = \mathbb{E}(\exp(\lambda x))$:

$$\mathbb{E}_{\mathbb{P}}(\exp(\lambda x)) = \int \exp(\lambda x) \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx$$

$$\mathbb{E}_{\mathbb{P}}(\exp(\lambda x)) = \exp(\mu t + 1/2\sigma^2 t^2)$$

$$\mathbb{E}_{\mathbb{P}}\left((x-\mu)\lambda - 1/2\sigma^2 t^2\right) = 1$$

$$\int \exp\left((x-\mu)\lambda - 1/2\sigma^2 t^2\right) p(x) dx = 1$$

But the above implies that the integrand is another probability density:

$$q(x) := \exp\left((x-\mu)\lambda - 1/2\sigma^2 t^2\right) p(x)$$

If we expand the normal density of $p(x)$ above we get:

$$q(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{x-(\mu+\lambda\sigma^2)}{\sigma}\right)^2\right]$$

This implies that $q(x) \sim \mathcal{N}(\mu + \lambda\sigma^2, \sigma^2)$

Hence for each normal density $p(x)$, there is a normal density $q(x)$ obtained from its moment generating function that has the same variance but a different mean. Thus, we can change the drift of a stochastic process by changing the measure from $p(\cdot)$ to $q(\cdot)$. Any particularly favored value of drift may be chosen:

$$a = \mu + \lambda\sigma^2 \Rightarrow \lambda = \frac{a-\mu}{\sigma^2}$$

In particular, if a process is a martingale, its drift will be 0 and in that case $a = 0$ implies:

$$\lambda = -\frac{\mu}{\sigma^2}$$

Radon-Nikodym derivative

Note that the above two densities forma a ratio:

$$\frac{q(x)}{p(x)} = \exp\left[\lambda(x - \mu) - \frac{1}{2}\sigma^2\lambda^2\right]$$

If we interpret densities to be the derivatives of their respective distributions, $q(\cdot) = d\mathbb{Q}(\cdot)$ and $p(\cdot) = d\mathbb{P}(\cdot)$. This leads to the well-known *Radon-Nikodym* derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$ $\frac{d\mathbb{Q}}{d\mathbb{P}}$:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left[\lambda(x - \mu) - \frac{1}{2}\sigma^2\lambda^2\right]$$

When $\mathbb{P}$ and $\mathbb{Q}$ are finite and equivalent (measure-0 sets are identical under both) then the Radon-Nikodym derivative exists and is unique and non-negative.

We can apply this change of drift idea to diffusion processes more generally:

$$dX_t = \mu dt + \sigma dW_t \Rightarrow X_t \sim \mathcal{N}(\mu t, \sigma^2 t)$$

This process has mean μt but it can be made 0 under a change of probability measure by choosing $\lambda = -\frac{\mu}{\sigma^2}$. Since $q(x) \sim \mathcal{N}(\mu + \lambda\sigma^2, \sigma^2)$, we get $q(x_t) \sim \mathcal{N}(0, \sigma^2 t)$. Hence change of measure can be used to generate a new diffusion model which is equivalent to a Wiener process with drift 0.

First fundamental theorem of asset pricing (No-arbitrage)

A market is free from arbitrage if and only if there is at least one risk-neutral probability measure $\mathbb{Q}$ equivalent to the original probability measure $\mathbb{P}$.

Second fundamental theorem of asset pricing (Completeness)

An arbitrage-free market with risky stocks and risk-free bonds is *complete* if and only if there is a *unique* risk-neutral measure $\mathbb{Q}$ equivalent to the original probability measure $\mathbb{P}$.

Applications: Option pricing

Options are contracts between a buyer and a seller which stipulate that the buying or selling of an object (called the *underlying*) at a future date will occur at a pre-determined price (called 'strike price'). Buyer of options are said to have 'long' positions while sellers have 'short' positions. In a European *call* option, the buyer has the right (but not the obligation) to buy the underlying at some future date at a pre-determined price. In the same way, buyer of a *put* option gives the agent the right (but not the obligation) to sell the underlying at some future date at a pre-determined price.

Clearly, if the date of expiration of the contract is T and the strike price is K then payoff of the option is:

$$C = \max\{S_T - K, 0\}$$

It is often assumed that stock prices follow a geometric Brownian motion:

$$S_t = \mu S_t dt + \sigma S_t dW_t$$

Assume that the price of the option is a function of the stock price: $f(S_t, t)$. Denoting partial derivatives by subscripts and applying Ito's lemma:

$$df = \left(1/2\sigma^2 S_t^2 f_{SS} + \mu S_t f_S + f_t\right) dt + \sigma S_t f_S dW_t$$

The main idea now is to create a portfolio of 1 unit of the derivative and a short position of f_S in the underlying. The portfolio value then becomes:

$$V_t = f - f_S S_t$$
$$dV_t = df - f_S dS_t$$
$$dV_t = df - f_S(\mu S_t dt + \sigma S_t dW_t)$$
$$dV_t = \left(1/2\sigma^2 S_t^2 f_{SS} + \mu S_t f_S + f_t\right) dt + \sigma S_t f_S dW_t - f_S(\mu S_t dt + \sigma S_t dW_t)$$
$$dV_t = (1/2\sigma^2 S_t^2 f_{SS} + f_t)dt + 0 \cdot dW_t$$

This is a deterministic, non-risky portfolio and hence if the market is free from arbitrage this portfolio's return must equal that of the risk-free interest rate r.

Hence

$$\frac{dV_t}{V_t} = rdt \Rightarrow dV_t = r(f - f_S S_t)dt$$

Equating the two forms for dV_t we get:

$$\frac{1}{2}\sigma^2 S_t^2 f_{SS} + rS_t f_S + f_t = rf$$

This is the celebrated Black-Scholes-Merton fundamental partial differential equation. Its solution depends on the boundary conditions. For example, if a European call option's price is $C(S_t, t)$ then it has terminal condition given by $C(S_T, T) = \max\{S_T - K, 0\}$; and the fundamental PDE assumes the form:

$$\frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 C}{\partial S^2} + rS_t \frac{\partial C}{\partial S} + \frac{\partial C}{\partial t} = rC$$

11

References

Jondeau, Eric, Ser-Huang Poon, and Michael Rockinger. 2007. *Financial Modeling Under Non-Gaussian Distributions.* Springer Finance.

Tsay, Ruey S. 2010. *Analysis of Financial Time Series.* Third Edition. John Wiley; Sons.