# Reinforcement Learning for Finance:

# Week 1: Lesson 2-4: Discrete-time Black-Scholes model

Igor Halperin

NYU Tandon School of Engineering

Brooklyn NY 2017

# Discrete-time BSM model

We start with a discrete-time version of the BSM model. The problem of option hedging and pricing in this formulation amounts to a **sequential risk minimization**. To define risk in an option, we follow a local risk minimization approach ( Föllmer and Schweizer (1994), Potters and Bouchaud (2001), Kapoor et. al. (2010), Grau (2007)).

We take the view of a seller of a European option (e.g. a put option) with maturity $T$ and the terminal payoff of $H_T(S_T)$.

To hedge the option, the seller use the proceeds of the sale to set up a replicating (hedge) portfolio $\Pi_t$ made of the stock $S_t$ and a risk-free bank deposit $B_t$. The value of hedge portfolio at any time $t \leq T$ is

$$\Pi_t = u_t S_t + B_t \tag{1}$$

where $u_t$ is a stock position at time $t$, taken to hedge risk in the option.

# Hedge portfolio evaluation

The replicating portfolio tries to exactly match the option price in all possible future states of the world. Start at $t = T$:

$$\Pi_T = H_T(S_T) \tag{2}$$

This sets a terminal condition for $H_T$ at $t = T$.
To find $B_t$ for previous times $t < T$, we impose the *self-financing constraint* which requires that all future changes in the hedge portfolio should be funded from an initially set bank account.

$$u_t S_{t+1} + e^{r\Delta t} B_t = u_{t+1} S_{t+1} + B_{t+1} \tag{3}$$

This can be expressed as a recursive relation for $B_t$ at any time $t < T$ using its value at the next time instance:

$$B_t = e^{-r\Delta t} \left[ B_{t+1} + \left( u_{t+1} - u_t \right) S_{t+1} \right], \quad t = T-1, \ldots, 0 \tag{4}$$

# Hedge portfolio evaluation

Plugging this into Eq.(1) produces a recursive relation for $\Pi_t$ in terms of its values at later times, which can therefore be solved backward in time, starting from $t = T$ with the terminal condition (2), and continued all the way to the current time $t = 0$:

$$\Pi_t = e^{-r\Delta t}\left[\Pi_{t+1} - u_t \Delta S_t\right], \quad \Delta S_t = S_{t+1} - e^{r\Delta t} S_t \qquad (5)$$

Eqs.(4) and (5) imply that both $B_t$ and $\Pi_t$ are not measurable at any $t < T$, as they depend on the future. Respectively, their values today $B_0$ and $\Pi_0$ will be random quantities with some distributions. We can compute them using Monte Carlo!

# Hedge portfolio evaluation with Monte Carlo

For any given hedging strategy $\{u_t\}_{t=0}^{T}$, these distributions can be estimated using Monte Carlo simulation:

- **Forward pass** : Simulate $N$ paths of the underlying $S_1 \to S_2 \to \ldots \to S_N$,
- **Backward pass**: Evaluate $\Pi_t$ going backward on each path.

As the choice of a hedge strategy does not affect the evolution of the underlying, such simulation of forward paths should only be performed once. Alternatively, we could use real historical data for stock prices, together with a pre-determined hedging strategy $\{u_t\}_{t=0}^{T}$ and a terminal condition (2).
But first, we need a hedge strategy $u_t$ to implement this Monte Carlo!

# Control question

Select all correct answers:

1. The value of the hedge portfolio at time $T$ is equal to the stock price at time $T$
2. The value of the hedge portfolio at time $T$ is equal to the option price at time $T$, i.e. the option payoff
3. A forward pass should be re-computed every time we recompute hedges.

4 A forward pass should be done only once, as for a small investor assumed in the option pricing problem, a hedge strategy does not impact the market.

Correct answers: 2, 4.