Consider now a derivative security that pays \( f(\bar{S}(T)) \) at time \( T \) in the future.

\( f \) is a deterministic function

\[
\text{Examples: } \quad \begin{align*}
(S_i(T) - K)^+ & \quad \text{call option on the first asset} \\
\left( \sum_{k=1}^k \alpha_k S_k(T) - K \right)^+ & \quad \text{call option on a weighted average on the first } k \text{ assets (basket option).}
\end{align*}
\]

The main problem of financial engineering: compute the price of that derivative security at time \( t < T \) and understand the risks.

**Main idea (Black–Scholes and Merton 1973):** Find a self-financing trading strategy \( \bar{\mathcal{V}}(t) \) such that \( \bar{\mathcal{V}}(T) \bar{S}(T) = f(\bar{S}(T)) \).

The price at time \( t \) is \( \bar{\mathcal{V}}(t) \bar{S}(t) \) and the risks of the derivative security are hedged away by the trading strategy.

**How do we find \( \bar{\mathcal{V}} \)?**

Let us suppose that \( \bar{\mathcal{V}}(t) \bar{S}(t) = V(t, \bar{S}(t)) \) for some deterministic function \( V \).

**Multi-dimensional version of Itô’s formula:**

\[
dV(t, \bar{S}(t)) = \left[ \frac{\partial V}{\partial t}(t, \bar{S}(t)) + \frac{1}{2} \sum_{i, j=1}^n \Sigma_{i, j} S_i S_j \frac{\partial^2 V}{\partial S_i \partial S_j}(t, \bar{S}(t)) \right] dt + \sum_{i=1}^n \frac{\partial V}{\partial S_i}(t, \bar{S}(t)) dS_i(t)
\]

We want \( V(t, \bar{S}(t)) \) to be the value of a self-financing trading strategy, therefore

\[
dV(t, \bar{S}(t)) = \bar{\mathcal{V}}(t) \cdot d\bar{S}(t)
\]

Therefore

\[
\bar{\mathcal{V}}_i(t) = \frac{\partial V}{\partial S_i}(t, \bar{S}(t))
\]

and

\[
\begin{align*}
\frac{\partial V}{\partial t}(t, \bar{S}(t)) + \frac{1}{2} \sum_{i, j=1}^n \Sigma_{i, j} S_i S_j \frac{\partial^2 V}{\partial S_i \partial S_j}(t, \bar{S}(t)) & = 0 \\
V(T, \bar{S}) & = f(\bar{S})
\end{align*}
\]

If there exists such a function \( V \), then the strategy \( \bar{\mathcal{V}}_i(t) \frac{\partial V}{\partial S_i}(t, \bar{S}(t)) \) is self-financing and replicates the derivative security whose value is

\[
V(t, \bar{S}(t)) = \sum_{i=1}^n \frac{\partial V}{\partial S_i}(t, \bar{S}(t)) S_i(t)
\]

\( \frac{\partial V}{\partial S_i} \) is called the delta (\( \Delta \))
Suppose an investor is selling the derivative security. He can hedge his risks away by entering the self-financing replicating trading strategy. This is often called “delta hedging”.

**Example:** Black-Scholes $n = 2$ assets

\[
\begin{align*}
\frac{dS_t}{S_t} &= \mu_t dt + \sigma_t dW_t, \quad \text{“stock”} \\
\frac{dB_t}{B_t} &= r dt, \quad \text{“bank account”}
\end{align*}
\]

$S_0$ given, $B_0 = 1$

\[
\Sigma = \begin{bmatrix} \sigma^2 & 0 \\ 0 & 0 \end{bmatrix}
\]

\[
V(t, S_t, B_t) \text{ solves } \begin{cases} 
\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} = 0 \\
V(T, S, B) = f(S).
\end{cases}
\]

\[
V(t, S_t, B_t) = \frac{\partial V}{\partial S} S_t + \frac{\partial V}{\partial B} B_t.
\]

Here $B_t = e^{rt}$ so \[V(t, S_t, B_t) = \tilde{V}(t, S_t).\]

\[
\frac{\partial \tilde{V}}{\partial t} = \frac{\partial V}{\partial t} + r B_t \frac{\partial V}{\partial B} \quad \frac{\partial \tilde{V}}{\partial S} = \frac{\partial V}{\partial S} \quad \frac{\partial^2 \tilde{V}}{\partial S^2} = - \frac{\partial V}{\partial S^2}
\]

so \[\tilde{V} \text{ solves } \begin{cases} 
\frac{\partial \tilde{V}}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 \tilde{V}}{\partial S^2} + r S \frac{\partial \tilde{V}}{\partial S} - r \tilde{V} = 0 \\
\tilde{V}(T, S) = f(S)
\end{cases}\]

where $f(s) = (s - K)^+$, explicit solution is known:

\[
\tilde{V}(t, S) = S \Phi \left( d_1 \right) - Ke^{-r(T-t)} \Phi \left( d_2 \right)
\]

\[
d_1 = \frac{\ln(S/e^{-r(T-t)})}{\sigma \sqrt{T-t}} + \frac{1}{2} \sigma \sqrt{T-t} \quad d_2 = d_1 - \sigma \sqrt{T-t}
\]

\[
\frac{\partial \tilde{V}}{\partial S} = \Phi \left( d_1 \right).
\]
We have just seen that to price a derivative security one has to solve a PDE. There are a number of problems with PDE's. (1) They are computationally infeasible when \( n \geq 2 \) or 3 (2) the PDE approach does not work when dynamics \( S(t) \) is more complex (i.e., non Markov) or when derivative securities are path-dependent (i.e., payoffs like barrier options, ...). We are going to develop another approach that will address these issues.

**Arbitrage and Stochastic discount factors.**

Given a selffinancing trading strategy \( \varphi(t) \), an arbitrage is such that for some \( t \) fixed \( \varphi(0) \cdot S(0) = 0 \) and \( P \{ \varphi(t) \cdot S(t) > 0 \} > 0 \) and \( P \{ \varphi(t) \cdot S(t) > 0 \} = 1 \) or \( \varphi(0) \cdot S(0) < 0 \) and \( P \{ \varphi(t) \cdot S(t) > 0 \} = 1 \).

Rolling out arbitrage is a basic requirement but it has far reaching consequences and is very powerful.

A process \( V \) is called an attainable price process if \( V(t) = \varphi(t) S(t) \) for some selffinancing trading strategy \( \varphi \).

Note that each \( S(t) \) is attainable (take \( \varphi(t) = (0, ..., 0, 1, 0, ..., 0) \) with position).

**Definition:** A strictly positive process \( Z \) is called a stochastic discount factor if the ratio \( \frac{V}{Z} \) is a martingale for every attainable price process \( V \), i.e.,

\[
V_A \leq t \quad E \left\{ \frac{V_t}{Z_t} \bigg| \mathcal{F}_s \right\} = \frac{V_s}{Z_s}.
\]

(or equivalently, \( V_A = E \left\{ \frac{Z_s V_t}{Z_t} \bigg| \mathcal{F}_s \right\} \) \( \frac{Z_s}{Z_t} \) is a discount factor.

Note that if \( Z \) is a sdf then \( \alpha Z \) is another one for any \( \alpha > 0 \), so can assume \( Z_0 = 1 \).

The point of that definition is the following Fundamental Theorem of Asset Pricing:

**Existence of a sdf \( \iff \) no arbitrage.**

Of course, some assumptions are needed but we won't need them.
One implication is easy: assume there exists a self-financing trading strategy \( \mathcal{V} \) such that \( \mathcal{V}(t) S(t) \) is an attainable price process, so

\[
\mathcal{V}(0) \cdot S(0) = E \left\{ \frac{Z_0}{Z_t} \mathcal{V}(t) \frac{S(t)}{S_T} \right\} \tag{by the martingale property}
\]

Then \( \mathcal{V} \) can never be an arbitrage: if \( P \{ \mathcal{V}(t) \frac{S(t)}{S_T} > 0 \} > 0 \) and \( P \{ \mathcal{V}(t) \frac{S(t)}{S_T} \geq 0 \} = 1 \)

then \( \mathcal{V}(0) \cdot S(0) > 0 \)

\[
\text{if } P \{ \mathcal{V}(t) \frac{S(t)}{S_T} > 0 \} = 1 \text{ then } \mathcal{V}(0) \cdot S(0) > 0.
\]

The other implication is very deep and much harder to prove.

**Risk Neutral pricing**

Assume one of the asset is risk free (usually bank account) meaning for now that \( \sigma_{i,k} = 0 \) for \( k = 1, \ldots, d \), with constant growth rate \( p_{o} = r \). We usually call it \( B \) assume \( B_0 = 1 \) so that \( B_t = e^{rt} \).

Assume there is no arbitrage (or equivalently, that there is a self fin.

\[
\frac{B_t}{Z_t} \text{ is a martingale}
\]

\[
E \frac{B_t}{Z_t} = \frac{B_0}{Z_0} = 1, \quad \text{for } B_t > 0,
\]

so we can use that as a Radon-Nikodym density to define a new probability measure \( \Omega \):

\[
\frac{d\Omega}{dP} \bigg|_{\mathcal{F}_t} = \frac{B_t}{Z_t} \tag{Q in called risk-neutral}
\]

meaning that if \( A \in \mathcal{F}_t \)

\[
\Omega(A) = E^P \left\{ 1_A \frac{d\Omega}{dP} \bigg|_{\mathcal{F}_t} \right\} = E^P \left\{ 1_A \frac{B_t}{Z_t} \right\}
\]

for any \( \mathcal{F}_t \)-r.v. \( X \)

\[
E^\Omega \left\{ X \right\} = E^P \left\{ X \frac{d\Omega}{dP} \bigg|_{\mathcal{F}_t} \right\}
\]

Now take \( V \) an attainable price process \( V \) is a \( \Omega \)-martingale:

\[
V(0) = E^P \left\{ \frac{V(T)}{Z_T} \right\} = E^\Omega \left\{ \frac{V(T)}{Z_T} \frac{1}{B_T/Z_T} \right\} = B_T^{-1} E^\Omega \left\{ V(T) \right\} = e^{-rT} E^\Omega \{ V(T) \}
\]

So: the current value of an attainable price process is the expected value of its future value under \( \Omega \) discounted at the risk free rate.

It is much more interesting to work with \( \Omega \) because:

(i) We don't have to find \( Z \), whose dynamics are usually hard to know.

(ii) Dynamics of \( S \) are easier under \( \Omega \) than under \( P \).
Indeed, since $S_t$ is an attainable price process, \( \frac{S_t(t)}{Z(t)} \) is a $P$-martingale and \( \frac{S_t(t)}{Z(t)} = \mathbb{E}^P \left\{ \frac{S_t(T)}{Z(T)} \mid \mathcal{F}_t \right\} = \mathbb{E}^Q \left\{ \frac{S_t(t)}{Z(t)} \mid \mathcal{F}_t \right\} \), which translates into \( \frac{S_t(t)}{Z(t)} \) is a $Q$-martingale.

It means that under $Q$:
\[
\frac{dS_t(t)}{S_t(t)} = r dt + \sum_{k=1}^{d} \sigma_{ik}(t, \mathcal{F}_t) dW_k(t)
\]

indeed by Itô's formula applied to \( f(S_t(t), B_t) = \frac{S_t(t)}{B_t} \),
\[
\frac{d}{S_t(t)/B_t} = \sum_{k=1}^{d} \sigma_{ik}(t, \mathcal{F}_t) dW_k(t).
\]

So under $Q$, drift of traded assets are equal to $r$.

---

**Risk-Neutral pricing recipe to compute the price at time $0$ of a derivative security whose payoff at time $T$ is an $\mathcal{F}_T$-measurable r.v. $X = f(S_T)$:**

- (may be path dependent: $X = (\sup_{t \leq T} S_t - K)^+$)
- (ii) model dynamics of $S$ under the risk-neutral measure $Q$ (essentially mean that drift must be set to $r$)
- (ii) compute $\mathbb{E}^Q \left\{ e^{-rT} f(S_T) \right\}$.

**Example:** Black–Scholes: assume volatility is constant $\sigma$.

Under $Q$:
\[
\frac{dS_t}{S_t} = r dt + \sigma dW_t
\]

Where solution is $S_t = S_0 \exp((r - \frac{\sigma^2}{2}) t + \sigma W_t)$

To price a call option compute:
\[
\mathbb{E}^Q \left\{ e^{-rT} (S_T - K)^+ \right\} = S_0 \Phi(d_1) - Ke^{-rT} \Phi(d_2)
\]

We find back the Black–Scholes formula.
Monte Carlo is the name of a simple idea to compute expectations based on the law of large numbers.

Take a random variable \( X \) and a function \( f \) such that \( E|f(X)| < \infty \)
we want to evaluate \( \alpha = Ef(X) \).

The idea is to use a sequence of independent random variables \( X_1, \ldots, X_n, \ldots \)
having the same law as \( X \), \( X_i \sim X \).

Put \( \hat{\alpha}_n = \frac{1}{n} \sum_{i=1}^{n} f(X_i) \)
\( \hat{\alpha}_n \) is called the sample mean.

The law of large numbers says that
\[
\lim_{n \to \infty} \hat{\alpha}_n = \alpha \quad \text{as.}
\]

As \( n \) tends to infinity, \( \hat{\alpha}_n(x) \) converges to the value \( Ef(X) \) for almost every \( w \). Also, the estimator is said to be unbiased because \( E\hat{\alpha}_n = \alpha \).

Speed of convergence:

Since \( n \) is finite in practice, we need to know the precision, or the speed of convergence. It is given by the Central Limit Theorem.

Assuming \( Ef(x)^2 < \infty \), let \( \sigma_f^2 = \text{Var} f(X) = Ef(x)^2 - (Ef(x))^2 \)

The Central Limit Theorem says that
\[
\sqrt{n} (\hat{\alpha}_n - \alpha) \overset{\text{law}}{\rightarrow} \text{G} \text{m}(0, \sigma_f^2)
\]

In other words, \( \hat{\alpha}_n - \alpha \) is approximately a Gaussian random variable with mean 0 and variance \( \sigma_f^2/n \).

For finite \( n \), the error is typically of size \( \sigma_f / \sqrt{n} \).

\( \sigma_f \) is usually unknown, as is \( \alpha \), but it can be estimated using sample standard deviation

\[
\sigma_f = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (f(X_i) - \hat{\alpha}_n)^2}
\]

Monte Carlo methods typically yield to a convergence rate in \( \sqrt{n} \).
Remark: An expectation is an integral, so why can't we use standard integration technique?

For instance, \( x = \int_0^1 f(x) \, dx = Ef(u) \) when \( U \sim \text{Unif}[0,1] \).

The integral can be computed using trapezoidal rule:

\[
\Delta_h^k = \frac{f(0) + f(1)}{2h} + \frac{1}{h} \sum_{i=1}^{n-1} f(ih)
\]

It can be shown that \( |\Delta_h^k - x| \leq \frac{C(f)}{h} \).

The speed is much higher, but in general, we don't know the exact distribution of \( X \) and the method will fail. Also, in higher dimension (meaning \( E_f(U_1, U_2, \ldots, U_k) \)) the speed will be lower. Monte-Carlo methods always have rate \( \sqrt{n} \) independently of the dimension of the problem.

As in the case of the trapezoidal rule, one can replace random points \( U_i \) by deterministic ones, \( i/h \). This is sometimes called Quasi Monte Carlo.

The rate of convergence can be better in low dimensions but not in high dimensions.

Example 1: Call option in Black-Scholes,

\[
C = e^{-rT} E \left\{ (S_T - K)^+ \right\} \quad \text{where} \quad S_T = S_0 \exp \left( \frac{r - \sigma^2}{2} T + \sigma W_T \right).
\]

Even though, there exists a closed form expression for the price of the option, let us sketch what MC would give:

for \( i = 1: N \)

\[
Z = \text{r Andreas}(1, 1);
S = S_0 \exp \left( (\frac{r}{1} - \frac{\sigma^2}{2}) \times T + Z \sigma \sqrt{T} \right);
C(i) = \exp \left( -rt \right) \times \left( S - K, 0 \right);
\]

\[
\hat{C}_n = \frac{1}{n} \sum_{i=1}^{n} (C(i) - \hat{C}_n)
\]

Denote \( S_c = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (C(i) - \hat{C}_n)^2} \) Let \( z_\delta \) be the \( 1 - \delta \) quantile of standard Gaussian distribution. (If \( \delta = 0.05 \), \( z_{0.05} \approx 1.645 \))
then \( \left[ \hat{C}_n - \frac{s \sqrt{2}}{\sqrt{n}} \frac{S_0}{\sqrt{n}} ; \hat{C}_n + \frac{s \sqrt{2}}{\sqrt{n}} \frac{S_0}{\sqrt{n}} \right] \) is an asymptotically valid
1-\( \delta \) confidence interval for \( C \).

\( \hat{C}_n \) is unbiased \( (E \hat{C}_n = C) \) and strongly consistent \( (\lim_{n \to \infty} \hat{C}_n = C \text{ a.s.}) \).

**Example 2**: An Asian option has payoff \( \left( \frac{1}{T} \int_0^T S_u \ du - K \right)^+ \).

**Risk-Neutral valuation rule**: price = \( E \left\{ e^{-rT} \left( \frac{1}{T} \int_0^T S_u \ du - K \right)^+ \right\} \).

We need to simulate \( (i.e. \text{ get samples from the law of }) \) \( \frac{1}{T} \int_0^T S_u \ du \).

We know how to simulate a discretized sample path \( (S_{iT/N})_{i=0, \ldots, N} \),
so we are going to approximate \( \frac{1}{T} \int_0^T S_u \ du \) by \( \frac{1}{N} \sum_{i=1}^{N} S_{iT/N} \).

Simulate \( p \) paths and form the estimator:

\[ \hat{C}_p = \frac{1}{p} \sum_{k=1}^{p} e^{-rT} \left( \frac{1}{N} \sum_{i=1}^{N} S_{iT/N} - K \right)^+ . \]

Approximating the integral by a discrete sum introduces a bias \( E \hat{C}_p \neq C \).

**Example 3**: Another bias: model discretization error.

So far we have sampled stock price in the Black-Scholes model using the explicit formula

\[ S_{(t+i)T/N} = S_{iT/N} \exp \left( \left( \mu - \frac{\sigma^2}{2} \right) \frac{T}{N} + \sigma \sqrt{\frac{T}{N}} Z_i \right) \]

where \( Z_i \)'s are i.i.d. \( N(0,1) \) random variables.

We could also discretize the SDE \( dS_t = S_t (\mu dt + \sigma dW_t) \)

\[ S_{(t+i)T/N} = S_{iT/N} \left( 1 + \frac{\mu T}{N} + \frac{\sigma \sqrt{T}}{N} Z_i \right) \]

\( S_T \) will not have the lognormal distribution in this case. (for instance \( S_T \) could be negative although with a very very small probability).

This introduces a bias called discretization bias. When there is no explicit formula for the solution of the SDE, we usually introduce a bias when discretizing.