CS476

Basic Statistics

$$\mathrm{Var}(aX+bY)=a^2\ \mathrm{Var}(X)+b^2\ \mathrm{Var}(Y)+2ab\ \mathrm{Cov}(X,Y)$$
$$\mathrm{Var}(aX-bY)=a^2\ \mathrm{Var}(X)+b^2\ \mathrm{Var}(Y)-2ab\ \mathrm{Cov}(X,Y)$$

European Put/Call

$$C_T = \max\{S_T - K, 0\}$$
$$P_T = \max\{K - S_T, 0\}$$

Risk free asset

$$dB(\tau) = r \cdot B(\tau)d\tau$$
$$B(t) = e^{-r(T-t)}B(T)$$

Binomial model

$$S_{T(\mathrm{up})} = uS_0\ ,\ S_{T(\mathrm{down})} = dS_0$$
$$0 < d < 1 + r < u\ \ (\text{Arbitrage free})$$

Trading Strategy is a pair $\Pi = \{\delta_0, \eta_0\}$ representing number of stocks and risk-free bonds. Value at $t = 0$ and $t = T$ is:

$$\Pi_0 = S_0\delta_0 + \eta_0 B_0$$
$$\Pi_T = S_T\delta_0 + \eta_0 B_T$$

An arbitrage strategy is a trading strategy:

$$\Pi_0 = 0 \ \text{ and } \Pi_T > 0 \text{ almost surely}$$
$$\Pi_0 < 0 \ \text{ and } \Pi_T \geq 0 \text{ almost surely}$$

Pricing using replicating portfolio: Given arbitrage free market

$$\text{if } \Pi_T = V_T \text{ almost surely} \Rightarrow \Pi_0 = V_0$$

in general $\Pi_t = V_t$ for $t \in [0, T]$

Replicating portfolio period-1 Binomial model

$$\delta_0 S_T^u + \eta_0 B_T = V_T^u$$
$$\delta_0 S_T^d + \eta_0 B_T = V_T^d$$
$$\Rightarrow \delta_0 = \frac{V_T^u - V_T^d}{S_T^u - S_T^d}$$

Theorem (Put-Call Parity): Assume an arbitrage free market with risk free interest rate $r \geq 0$. Assume S_t doesn't pay dividends. Then at any time $t \in [0, T]$, then European call C_t and put P_t with same strike price K and same expiry T satisfy

$$C_t - P_t = S_t - Ke^{-r(T-t)}$$

Lemma (Risk neutral expected values): Consider an arbitrage free market

$$q^u = \frac{e^{rT} - d}{u - d}$$

and q^u satisfy $q^u \in (0, 1)$. Then under probability measure $\mathbb{Q}$ with $\mathbb{Q}(\mathrm{up}) = q^u$ and $\mathbb{Q}(\mathrm{down}) = 1 - q^u$ we have

$$S_0 = e^{-rT}\mathbb{E}^{\mathbb{Q}}(S_T)$$
$$V_0 = e^{-rT}\mathbb{E}^{\mathbb{Q}}(V_T)$$

Proposition: Under the N −period Binomial model with $0 < d < e^{r\Delta t} < u$. Suppose V_N is a random variable (derivative payout at maturity).

$$V_{n(\omega_1,\omega_2,\ldots,\omega_n)} = e^{-r\Delta t}\mathbb{E}^{\mathbb{Q}}[V_{n+1} \mid \mathcal{F}_{n+1}]$$
$$e^{-r\Delta t}(q^u V_{n+1}(\omega_1,\omega_2,\ldots,\omega_n,\ \mathrm{up}) + (1-q^u)V_{n+1}(\omega_1,\omega_2,\ldots,\omega_n,\ \mathrm{down}))$$

Then V_0 is the fair value of the derivative at time 0

Algorithm: Given N, compute V_0

1. Compute all possible payouts $V_T(\omega_1, \ldots, \omega_N)$
2. Go backwards. `For n = N - 1, ..., 0`
   - For all states $\omega \in \Omega$

$$V_n(\omega_1,\ldots,\omega_N) = e^{-r\Delta t}(q^u V_{n+1}(\mathrm{up}) + (1-q^u)V_{n+1}(\mathrm{down}))$$

3. Return V_0

Definition (Log Normal Returns):

$$X_n = \log(S_n)$$
$$\Delta X_n = X_{n+1} - X_n = \log\left(\frac{S_{n+1}}{S_n}\right)$$

Definition (Standard Brownian Motion):

1. $W_0 = 0$ almost surely,
2. For any $s > t \geq 0$ the increment $W_s - W_t$ satisfies
$$W_s - W_t \sim N(0, s - t)$$
3. For any $0 \leq t_1 < t_2 \leq t_3 < t_4$ the increments $W_{t_2} - W_{t_1}, W_{t_4} - W_{t_3}$ are independent
4. The sample paths (t, W_t) are continuous almost surely.

Definition: Let $(W_t)_{t\geq 0}$ be a Brownian Motion.

1. The process
$$X_t = \mu t + \sigma W_t,\ t \geq 0$$
is called a Brownian Motion with drift μ and volatility σ
2. The process $Y_t = Y_0 \exp(X_t)$ is called a geometric Brownian Motion.

Definition (Black Scholes Model): Under the Black Scholes Model we assume the random stock price process satisfies

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

Proposition (A solution of the BS SDE):

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

is given by

$$S_t = S_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t\right)$$

Definition (Quantile function):

$$F^{\leftarrow}(y) = \inf\{x \in \mathbb{R} : F(x) \geq y\}, y \in [0, 1]$$

Algorithm (Inversion method): Given a CDF F sample $X \sim F$ as follows

1. Sample $U \sim U(0, 1)$
2. Return $X = F^{\leftarrow}(U)$

Monte Carlo estimator

$$\hat{\mu}_n^{\mathrm{MC}} = \frac{1}{n}\sum_{i=0}^{n} g(\boldsymbol{X}_i)$$

$$\sigma^2 = \mathrm{Var}(g(\boldsymbol{X})) \Rightarrow \mathrm{Var}(\hat{\mu}_n^{\mathrm{MC}}) = \frac{\sigma^2}{n}$$

$$1-\alpha \text{ CI: } \hat{\mu}_n^{\mathrm{MC}} \pm Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

$$1-\alpha \text{ CI: (unknown } \sigma) \ : \hat{\mu}_n^{\mathrm{MC}} \pm Z_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}}$$

Where S_n is sample standard deviation.

Antithetic Variates

- Replace n independent observations with $\frac{n}{2}$ pairs of antithetic observations.
- $\frac{g(\boldsymbol{X}_i)+g(\widetilde{\boldsymbol{X}}_i)}{2}$
- $X_i, \widetilde{X}_i$ are negatively correlated.

$$\hat{\mu}_n^{\mathrm{AV}} = \frac{1}{\left(\frac{n}{2}\right)} \sum_{i=1}^{n} \frac{g(\boldsymbol{X}_i) + g\left(\widetilde{\boldsymbol{X}}_i\right)}{2}$$

$$\mathrm{Var}(\hat{\mu}_n^{\mathrm{AV}}) = \frac{\sigma^2}{n} + \frac{1}{n}\,\mathrm{Cov}\left(g(\boldsymbol{X}_i), g\left(\widetilde{\boldsymbol{X}}_i\right)\right)$$

$$= \mathrm{Var}(\hat{\mu}_n^{\mathrm{MC}})\left(1 + \mathrm{Cor}\left(g(\boldsymbol{X}_i) + g\left(\widetilde{\boldsymbol{X}}_i\right)\right)\right)$$

$$\mathrm{Var}(\hat{\mu}_n^{\mathrm{AV}}) \le \mathrm{Var}(\hat{\mu}_n^{\mathrm{MC}}) \Leftrightarrow \mathrm{Cor}\left(g(\boldsymbol{X}_i) + g\left(\widetilde{\boldsymbol{X}}_i\right)\right) \le 0$$

$$\mathrm{res}_i = \frac{g(\boldsymbol{X}_i) + g\left(\widetilde{\boldsymbol{X}}_i\right)}{2}$$

$$1-\alpha \text{ CI: } \hat{\mu}_n^{\mathrm{AV}} \pm Z_{1-\frac{\alpha}{2}} \frac{\mathrm{sd(res)}}{\sqrt{n}}$$

Control Variates

$$\hat{\mu}_n^{\mathrm{CV}} = \frac{1}{n}\sum_{i=1}^{n}(Y_i + \beta(\mu_c - C_i))$$

$$\mathrm{res}_i = Y_i + \beta(\mu_c - C_i)$$

$$(1-\alpha) \text{ CI} : \hat{\mu}_n^{\mathrm{CV}} \pm Z_{1-\frac{\alpha}{2}} \frac{\mathrm{sd(res)}}{\sqrt{n}}$$

$$\mathbb{E}(C) = \mu_C \quad \text{is known}$$

$$\mathrm{Var}\left(\hat{\mu}_n^{\mathrm{CV}}\right) \quad \text{is minimized when}$$

$$\beta^* = \frac{\mathrm{Cov}(Y,C)}{\mathrm{Var}(C)}$$

To estimate β^* we can:

1. Use the same sample, resulting estimate $\hat{\mu}_n^{\mathrm{CV}}$ not necessarily unbiased. For large n bias is negligible.
2. Use pilot study. For n^{pilot} sample (Y_i, C_i) and estimate $\hat{\beta}^*$

- The more correlated C and Y are, the better the improvement over crude MC.
- $\hat{\mu}_n^{\mathrm{CV}}$ is asymptotically normal so CI can be estimated in normal way.

Brownian Bridge

- We can sample the path out of order using conditional distributions. This saves time since we may not need to sample the full path

Theorem (BM Conditional Distribution): If $(W_t)_{t\ge0}$ is a standard BM, then for any $u < v < w$ we have

$$X = W_v \mid (W_u = a, W_v = b)$$

$$X \sim N\left(\frac{w-v}{w-u}a + \frac{v-u}{w-u}b, \frac{(v-u)(w-v)}{w-u}\right)$$

- We can sample the stock price $S_{t_1}, ..., S_{t_N}$ in any order using the conditional distributions.

Multivariate Normal Distribution

Definition (Multivariate Normal):

$$\boldsymbol{X} \sim N_d(\mu, \Sigma)$$

if $\boldsymbol{X} = \mu + \boldsymbol{A}Z$, where $Z = (z_1, z_2, ..., z_d)$ with $Z_j \sim N(0,1)$ and $AA^T = \Sigma$

Sampling from the multivariate normal

- Let A be the e Cholesky factor of Σ, which is a lower triangular

matrix so that $AA^T = \Sigma$

1. For $i = 1...n$
   - Sample $Z_1, ..., Z_d \sim N(0,1)$
   - $\boldsymbol{X_i} = \mu + AZ$
2. Return $\boldsymbol{X}$

Correlated assets

- Let $Z_1, Z_2 \sim N(0,1), Z = (Z_1, Z_2)$
- $A = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix}$ and $\rho \in (-1,1)$
- $\boldsymbol{X} = AZ$, then
- $X_1 = Z_1 + 0 \cdot Z_2$, $X_2 = \rho Z_1 + \sqrt{1-\rho^2} Z_2$
- $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$
- $\mathrm{Cov}(X_1, X_2) = \mathrm{Cor}(X_1, X_2) = \rho$

We can sample two correlated standard Brownian Motions $W^{(1)}, W^{(2)}$, $\mathrm{Cor}(W^{(1)}, W^{(2)}) = \rho$

1. Sample $Z_1, Z_2 \sim N(0,1)$
2. Set $W_{t_j}^{(1)} = W_{t_{j-1}}^{(1)} + \sqrt{\Delta t} Z_1$
3. Set $W_{t_j}^{(2)} = W_{t_{j-1}}^{(2)} + \sqrt{\Delta t}\left(\rho Z_1 + \sqrt{1-\rho^2} Z_2\right)$

Common Random numbers

- Estimating $\mu_1 - \mu_2 = \mathbb{E}(g_1(\boldsymbol{X})) - \mathbb{E}(g_2(\boldsymbol{X}))$

1. **Method 1**: Estimate μ_1, μ_2 using two independent MC estimators $\mu_{n,1}^{\mathrm{MC}}, \mu_{n,2}^{\mathrm{MC}}$

$$\mathrm{Var}(\mu_n^{\mathrm{MC}}) = \frac{1}{n}(\sigma_1^2 + \sigma_2^2)$$

2. **Method 2** Estimate using the *same* random numbers μ_n^{CRN}

$$\mathrm{Var}\left(\mu_n^{\mathrm{CRN}}\right) = \frac{1}{n}(\sigma_1^2 + \sigma_2^2 - 2\sigma_{1,2})$$

$$\sigma_1 = \mathrm{Var}(g_1(\boldsymbol{U}))\ ,$$

$$\sigma_2 = \mathrm{Var}(g_2(\boldsymbol{U}))\ ,$$

$$\sigma_{1,2} = \mathrm{Cov}(g_1(\boldsymbol{U}), g_2(\boldsymbol{U}))$$

- CRN outperform the independent estimator $\iff \sigma_{1,2} > 0$

Lebesgue Integral

$$FV(G) = \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} |G(t_{j+1}) - G(t_j)|$$

Definition (Lebesgue Stieltjes Integral): If $FV(G) < \infty$ then

$$\int_0^T f(t)dG(t) = \lim_{n\to\infty} \sum_{j=0}^{n-1} f(t_j)(G(t_{j+1}) - G(t_j))$$

If G has a first derivative $G' = g(t)$ then

$$\int_0^T f(t)dG(t) = \int_0^T f(t)g(t)dt$$

Definition (Expected Value): Suppose F is a distribution function of some r.v X then

$$\mathbb{E}(X) = \int x dF(X)$$

Definition (Quadratic Variation):

$$[f,f](T) = \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} \left[f(t_{j+1}) - f(t_j)\right]^2$$

Proposition: If f has a continuous first derivative then by Taylor $[f,f](T) = 0$

Theorem: The Brownian Motion $W(t)$ for $f \in [0,T]$ does not have finite first order variation almost surely

$$FV(W) = \infty$$

$$[W,W](T) = T$$

Theorem (Ito Formula for a BM): Let $W(t)$ be a BM and $f(t,w)$ be a function for which the partial derivatives $f_t, f_{w(t,w)}, f_{ww}(t,w)$ defined and continuous. Then for any $T \geq 0$

$$f(T, W(T)) = f(0, W(0)) + \int_0^T f_t(t, W(t))dt$$

$$+ \int_0^T f_w(t, W(t))dW(t) + \frac{1}{2}\int_0^T f_{ww}(t, W(t))dt$$

Definition (Ito Process):

$$X(t) = X(0) + \int_0^t a(u)du + \int_0^t b(u)dW(u)$$

also

$$dX(t) = a(t)dt + b(t)dW(t)$$

Theorem (Ito's formula for Ito processes): Let $X(t)$ be an ito process, (suppose all appearing derivatives are continuous) then

$$f(T,X(T)) = f(0,X(0)) + \int_0^T f_t(t,X(t))dt +$$

$$\int_0^T f_x(t,X(t))dX(t) + \frac{1}{2}\int_0^t f_{xx}(t,X(t))d[X,X](t)$$

After simplification:

$$df(t,X(t)) = \left(f_t + a(t)f_x + \frac{1}{2}f_{xx}b^2(t)\right)dt + f_x b(t)dW(t)$$

Definition (Euler Approximation):

$$dX(t) = a(X(t))dt + b(X(t))dW(t)$$

$\hat{X}(0) = X(0)$ and

$$\hat{X}(t_{i+1}) = \hat{X}(t_i) + a\left(\hat{X}(t_i)\right)\cdot\Delta + b\left(\hat{X}(t_i)\right)\cdot\sqrt{\Delta}\cdot Z_{i+1}$$

Where $Z_1, Z_2, ..., Z_N \sim N(0,1)$ iid.

Definition (Value At Risk):

$$\mathbb{P}\left(X \leq \mathrm{VaR}_{\beta(X)}\right) = \int_{-\infty}^{\mathrm{VaR}_{\beta(X)}} f(x)dx = 1 - \beta$$

Value-at-risk is the $1-\beta$ quantile of the distribution.

Definition (Expected Shortfall):

$$\mathrm{ES}_\beta(X) = \mathbb{E}\left(X \mid X \leq \mathrm{VaR}_\beta\right)$$

$$\mathrm{ES}_\beta(X) = \frac{1}{1-\beta}\int_\beta^1 \mathrm{VaR}_u du$$

Note: The expected shortfall is sub-additive

$$\mathrm{ES}_\beta(X+Y) \leq \mathrm{ES}_\beta(X) + \mathrm{ES}_\beta(Y)$$

Algorithm (VaR and ES from sample): $X_1, X_2, ..., X_n$ is the sample data.

- Let $X_{(1)} < X_{(2)} < ... < X_{(n)}$ be the sorted data.
- $i_\beta = \lfloor (1-\beta)n \rfloor$
- $\widehat{\mathrm{VaR}}_\beta = X_{(i_\beta)}$
- $\widehat{\mathrm{ES}}_\beta = \frac{1}{i_\beta}\left(X_{(1)} + X_{(2)} + ... + X_{(i_\beta)}\right)$

Definition (Jump proccess):

$$dS_t = \mu S_t dt + \sigma S_t dW_t + (J-1)dN(t)$$

Where where $N(t)$ is a poisson process with rate λ. in a small interval of length Δt we have

$$dN(t) = \begin{cases} 1 \text{ with prob } \delta dt \\ 0 \text{ with prob } 1 - \delta dt \end{cases}$$