Stochastic Calculus, Courant Institute, Fall 2020
http://www.math.nyu.edu/faculty/goodman/teaching/StochCalc2020/index.html

Week 5

Jonathan Goodman, November, 2020
Tentative

1 Active strategies for diffusions

This class discusses using dynamic stochastic models to design investment and trading strategies. If you're a scientist making a stochastic model of a system, you ask what the model predicts about the system. This involves asking questions about paths of the stochastic process. An engineer or finance person might instead use the stochastic model to design strategies. *Optimal stochastic control* is one systematic way to use a stochastic model to design strategies. You give an *objective function* (or *merit function*) which involves expectations of quantities related to the stochastic process and your input controls, then you look for the control strategy that minimizes or maximizes this objective function.

2 Utility and choice theory

Utility theory is a philosophy of how people make or should make financial decisions under uncertainty. Today's class is about optimal financial strategies, so there has to be a criterion for optimality. In my opinion, the only sensible approach is to optimize expected utility. The *von Neumann Morgenstern theorem* is the argument for using expected utility. It starts with some natural axioms about choice. The conclusion is that any choice system that satisfies the *von Neumann Morgenstern axioms* is given by a utility function.

The von Neumann Morgenstern framework involves choice under uncertainty. Here, the uncertainty is about how much money you will have at a certain time. Different financial strategies lead to different probability distributions for this. Suppose X and Y are random variables that represent the amount of money (your "wealth") you might have. The axioms say that there are three mutually exclusive possibilities

- $X \prec Y$, you prefer Y to X, as random variables
- $X \sim Y$, you are *indifferent* between the two
- $X \succ Y$, you prefer X to Y.

For example, if $X \sim \mathcal{N}(0,1)$ and $Y \sim \mathcal{N}(5,1)$, then most people would have $X \prec Y$.

Here are the von Neumann Morgenstern axioms

1. If $X \prec Y$ and $Y \prec Z$, then $X \prec Z$. If $X \prec Y$ and $Y \sim Z$, then $X \prec Z$ (transitivity).

The next axiom involves a coin-toss interpolation between two random variables If X and Y are two random variables and p is a probability ($0 < p < 1$), then there is a random variable W_p where we toss a coin and with probability p take $W_p = X$ and with probability $1 - p$ we take $W_p = Y$. To motivate the next axiom, remember that if $f(t)$ is a continuous function of t with $f(0) < 0$ and $f(1) > 0$, then there is a t with $0 < t < 1$ so that $f(t) = 0$. This is the *intermediate value theorem.*

2. Suppose $X \prec Z \prec Y$, then there is p with $0 \leq p \leq 1$ so that $Z \sim W_p$ (continuity)

3. If $Z \geq 0$ and $\Pr(Z > 0) > 0$, then $X \prec X + Z$ (arbitrage)

4. If $\mu = \mathrm{E}[X]$, and if we also use μ to denote the random variable that is always equal to μ, then $X \prec \mu$ (risk aversion)

These axioms are meant to be about probability distributions, but they are stated here in terms of random variables. It is possible that $X \neq Y$ as random variables but have the same distribution. Two random variables have the same distribution if for every number a, $\Pr(X < a) = \Pr(Y < a)$.

5. If X and Y have the same distribution, then $X \sim Y$.

All of these seem (to me) completely natural except the risk aversion axiom. The opposite of risk aversion is *risk seeking.* Some financial situations, such as stock investment competitions encourage risk seeking. An investment or trading competition uses simulated trades with virtual money. Each competitor starts with the same initial wealth and makes decides on simulated trades using reported market prices during the period of the competition. The competitor with the most simulated wealth at the end wins the competition. According to a version of the efficient market hypothesis, none of the players has the knowledge to have an expected return larger than the others. Therefore, your chance of winning is greatest if your variance is the largest – more of your return distribution is above the distributions of your competitors. If the distribution is Gaussian (say), then your probability of winning approaches 50% in the limit where your variance goes to infinity while your competitors' variances stay finite. The optimal strategy is to bet as wildly as possible.

My view is that risk seeking like this indicates that the incentive scheme is broken. The person who wins a stock trading competition is not someone you want investing your money. You can find other views expressed on the web. I will take risk aversion as an axiom today.

A *utility function* $U(x)$ tells you how much x money benefits you, or how happy it makes you. It is a philosophical interpretation that cannot be made completely precise. A utility function should be monotone increasing. If $y > x$, then y makes you happier than x (as amounts of money). A utility function should be *concave*, which means $u''(x) < 0$. The derivative of the utility tells you how much more happy an increment of money makes you $U(x + dx) = U(x) + U'(x)dx$. Economists call $dU = U'(x)dx$ the *marginal utility* corresponding to the increment dx. Concavity represents the fact that the marginal utility of dx decreases as x increases. If you have $x = \$100$ and someone offers you $dx = \$1$, you might be pleased, but if you have $x = \$1,000,0000$ and someone offers you \$1, your marginal pleasure would be less, still positive, but less. Therefore $U'(x)$ should be a decreasing function of x, which is $U''(x) < 0$. I will say that a function $U(x)$ is a *utility function* if it is increasing and concave.

If X is a random variable, a random amount of money, then the *expected utility* is

$$u_X = \mathrm{E}[\,U(X)]\ .$$

Part of the von Neumann Morgenstern theory is that choosing on the basis of expected utility satisfies all of the von Neumann Morgenstern axioms. We prefer Y to X if Y has more expected utility:

$$u_X < u_y \implies X \prec Y\ ,\ \ u_X = u_Y \implies X \sim Y\ .$$

We check that this satisfies the axioms $1-5$.

1. If $X \prec Y \prec Z$, then $u_X < u_Y < u_Z$, so (because $<$ is transitive) $u_X < u_Z$ and $X \prec Z$. The other part is similar. If $u_X < u_Y$ and $u_Y = u_Z$, then $u_X < u_Z$, so $X \prec Z$.

2. The expected utility u_{W_p} is a continuous function of p (math people, a concave function is continuous, and expectations with respect to continuously changing probability distributions are a continuous function of the parameter). We are given that $u_{W_0} = u_X < u_Z < u_Y = u_{W_1}$, therefore there is a p with $u_{W_p} = u_Z$. This makes $W_p \sim Z$, the agent indifferent between W_p and Z.

3. The utility U is strictly increasing. Therefore, if $Z > 0$ then $U(X + Z) > U(X)$. If $\Pr(U(Z + X) > U(X)) > 0$, and $\Pr(U(X + Z) < U(X)) = 0$, then $\mathrm{E}[\,U(X + Z)] > \mathrm{E}[\,U(X)]$. This is like saying that if $f(t) \geq 0$ for all t and if f is continuous and if $f(t) > 0$ for some t, then $\int f(t)dt > 0$.

A mathematical fact called *Jensen's inequality* is the link between concavity $U(x)$ and risk aversion. It says that if U is a strictly concave function and if X is a random variable with $\mathrm{E}[X] = \mu$ and $\mathrm{var}(X) > 0$, then

$$\mathrm{E}[\,U(X)] < U(\mu)\ . \tag{1}$$

There is a geometrical proof that shows how concavity comes in. Consider the line that it tangent to the graph of $U(x)$ at $x = \mu$. The line is the graph of the function

$$V(x) = U(\mu) + U'(\mu)(x - \mu)\ . \tag{2}$$

The graph of U is below the graph of V because U is concave. This means $U(x) < V(x)$ if $x \neq \mu$. Therefore

$$\mathrm{E}[\,U(X)] < \mathrm{E}[\,V(X)]\ , \quad \text{if } \mathrm{var}(X) > 0\ .$$

If $\mathrm{var}(X) = 0$, then X is always equal to the constant μ and $\mathrm{E}[\,U(X)] = U(\mu)$. Look back at the definition (2) of V and you will see that

$$\mathrm{E}[\,V(X)] = U(\mu) + U'(\mu)\,\mathrm{E}[\,X - \mu] = U(\mu)\ .$$

This is a proof of Jensen's inequality (1).

4. This is Jensen's inequality. The expected utility of the constant μ is $U(\mu)$. The expected utility of $U(X)$ is less than that.

5. If X and Y have the same distribution, then they have the same expected utility.

This shows that a choice based on expected utility satisfies the von Neumann Morgenstern axioms.

The harder theorem, which is the *von Neumann Morgenstern theorem*, goes the other way. If $\prec$, $\succ$, $\sim$ is a preference system that satisfies the von Neumann Morgenstern axioms, then there is a utility function $U(x)$ so that the preference system is the same as the one determined by expectation of this utility. A choice system that satisfies the von Neumann Morgenstern axioms is called *rational* The theorem does not say that you have to use expected utility to be rational. It just says that there is a utility function (increasing, concave) that gives the same choices.

An investment advisor (an *investor*) can use expected utility to avoid preferences that are not rational. An increasing and concave utility will guarantee that the investor accepts arbitrages (axiom 3) and is risk averse (axiom 4). A choice system that is not equivalent to an expected utility criterion is guaranteed either to be risk seeking or to decline arbitrage or both, under some circumstances.

The technical meaning of the theorem is this: the decision criterion is a linear function of the outcome probabilities. You can see this simply if X has a discrete probability distribution. Suppose X can take only one of the n values s_j and $\mathrm{Pr}(X = s_j) = p_j$. Then the expected utility is

$$u_X = \sum_{j=1}^{n} U(s_j)\,p_j\ .$$

This is a linear function of the probabilities p_j.

Here is a consequence of the von Neumann Morgenstern theorem. Any portfolio criterion that is not equivalent to a linear function of the probabilities must violate one or more of the von Neumann Morgenstern axioms. One example is *variance penalized expected return*

$$M(X) = \mathrm{E}]X] - \lambda\,\mathrm{var}(X)\ . \tag{3}$$

This is a nonlinear (linear plus quadratic) function of the probabilities:

$$M(X) =; \; \mathrm{E}[X] \; - \lambda \Big[\; \mathrm{E}\big[\, X^2 \big] - \quad \left(\mathrm{E}[X]\right)^2 \Big]$$

$$M(X) = \sum_{j=1}^{n} s_j p_j - \lambda \left[\sum_{j=1}^{n} s_j^2 p_j - \left(\sum_{j=1}^{n} s_j p_j \right)^2 \right] \; .$$

The last term on the right leads to p_j^2 appearing in the formula for $M(X)$. Exercise 1 asks you to construct an example of this.

Here is some non-mathematical commentary. Investors face a tradeoff between return and risk. You can increase your expected return only by taking more risk. Utility theory (above) and mean-variance analysis (exercises 1 and 2) are different approaches to this. The utility theory approach seems indirect. Risk aversion comes from Jensen's inequality, which comes from decreasing marginal utility. Mean variance analysis seems more straightforward, subtracting a penalty for variance. Investment advisors prefer mean variance analysis partly because it is easier to explain to clients

The field of *behavioral finance* is in part about decisions people make that are *irrational* in the sense that they violate the von Neumann Morgenstern axioms. Some people use this as an argument against using expected utility to guide investment decisions. I think this is a mistake. People may wish to be "rational" but lack the internal computing power to determine the expected utility of every financial decision. Using computers and utility theory may help investors avoid irrational decisions.

A drawback of utility theory for practical investing is that nobody knows their utility function. To be fair, this drawback is shared with other ways to take into account risk aversion, such as the parameter λ in the variance penalized criterion (3). It is common to use power law utility

$$U(x) = x^\gamma \; , \quad 0 < \gamma < 1 \; . \tag{4}$$

This is sometimes called *CRRA*, for *constant relative risk aversion.*

3 Optimal dynamic investment

Here is an example of a dynamic investment problem. At any time t, the wealth is divided between two investments, called *cash* and *stock*. More formally, "cash" is called the *risk-free asset* and "stock" is the *risky asset*. Let Z_t denote the wealth at time t. The investment strategy is $Z_t = X_t + Y_t$, where X_t is the investment in stock and Y_t is cash. In a time dt, the investments change according to

$$X_t \to (1 + \mu dt + \sigma dW_t) X_t \tag{5}$$

$$Y_t \to (1 + r dt) Y_t \tag{6}$$

The parameters are r, the risk free rate of return, μ, the expected rate of return of the stock, σ, the *volatility* of the stock. If there is no "trading" (re-allocation of assets), then X_t will be a geometric Brownian motion that satisfies

$$dX_t = \mu X_t dt + \sigma X_t dW_t \ .$$

The solution, we saw in Week 4, is

$$X_t = X_0 e^{\sigma W_t + (\mu - \frac{\sigma^2}{2})t} \ .$$

Without trading, the cash component would satisfy

$$dY_t = rY_t dt \ , \quad Y_t = Y_0 e^{rt} \ .$$

Cash is *risk free* because the risk free rate r is assumed known at the beginning and is constant, and because there is nothing random in the evolution of Y_t (without trading). The dW term in (5) makes stock a *risky asset*. In this model, μ and σ are taken to be known constants.

With dynamic trading, the investor chooses X_t and Y_t at time t subject only to the constraint that $X_t + Y_t = Z_t$. Once this allocation is made, we "watch" the markets for time dt. The two assets change value according to (5) and (6). The result is

$$\begin{aligned} dZ_t &= \mu X_t dt + \sigma X_t dW_t + rY_t dt \\ dZ_t &= rZ_t dt + (\mu - r)X_t dt + \sigma X_t dW_t \end{aligned} \tag{7}$$

The quantity $(\mu - r)$ is the *excess return*.

A non-anticipating *strategy* is a function X_t that is determined by the path from time 0 to time t. Any strategy (assumed to be non-anticipating) leads to a tine T wealth Z_T that is a function of the strategy and the Brownian motion path $W_{[0,T]}$. The best strategy for a given agent ("agent" = "utility function") is the one that maximizes the expected utility. This leads to the problem

$$\max_{\text{strat}} \mathrm{E}[\, U(Z_T)]$$

The final wealth is given by integrating dZ from (7) from $t = 0$ to $t = T$:

$$Z_T = Z_0 + r\int_0^T Z_t dt + (\mu - r)\int_0^T X_t dt + \sigma \int_0^T X_t dW_t \ .$$

This is one of the Merton optimal dynamic investment problems.

Optimal policy problems like this may be solved using value functions. The value function for this problem is

$$f(z,t) = \max_{\text{strat}} \mathrm{E}[\, U(Z_T) \mid Z_t = z] \ . \tag{8}$$

This is the maximum over strategies X_s that are defined for $t \leq s \leq T$. As before, the value function satisfies a PDE for $t < T$ and a final condition

$$f(z,T) = U(z) \ . \tag{9}$$

This formulation implicitly includes the assumption that trading is arbitrary and free. That's why the value function $f(z,t)$ does not depend on the value of X_t. The agent can choose any "stock position" (value of X_t) at time t regardless of X_s for $s < t$ as long as the money comes from cash (the risk free asset). The agent can buy or sell dX value of risky asset at the same price (buying = selling) and using dX amount of cash (no transaction cost). Economists call an ideal market like this *frictionless*. An agent who buys and sells without influencing the price is a *price taker*. All these assumptions are approximations. It is possible to model transaction cost and *market impact* (the effect of dX on the market price).

The PDE is the *Hamilton Jacobi Bellman* equation. It is based on the *dynamic programming principle*, which is the idea that when you make a decision at time t, you can assume that all decisions after that will be optimal. In the conditional expectation that defines the value function (8), the wealth at t is z. Suppose the agent chooses to allocate x to the risky asset at this time. Then at time $t + dt$, the wealth would be given by (7)

$$Z_{t+dt} = z + rZ_t dt + (\mu - r)X_t dt + \sigma X_t dW_t \ .$$

The optimal expected utility starting from $t + dt$ would be

$$f(z+dZ_t, t+dt) = f(z,t) + \partial_z f(z,t) dZ_t + \frac{1}{2}\partial_z^2 f(z,t)(dZ_t)^2 + \partial_t f(z,t)\, dt \ . \quad (10)$$

We should use the "Ito rule" and replace $(dZ_t)^2$ by its expected value, which is $\sigma^2 x^2 dt$, according to (7) with $X_t = x$. The optimal x is the one that maximizes $\mathrm{E}[\, f(z + dZ, t + dt)]$. The only random term on the right (the only thing that goes inside the expectation) is dZ_t. The only quantities that depend on x are dZ_t and $(dZ_t)^2$. Therefore, we can maximize over them to find the optimal x, which we call x_*.

$$x_* = \arg\max_x \left(\partial_z f(z,t)\, \mathrm{E}[\, dZ_t] + \frac{1}{2}\sigma^2 x^2 dt \partial_z^2 f(z,t) \right) \ .$$

In (7) we see that

$$\mathrm{E}[\, dZ_t] = rzdt + (\mu - r)xdt \ .$$

Since z is independent of x, we may leave it out of the maximization. The result is

$$x_* = \arg\max_x \left((\mu - r)\, x\, \partial_z f(z,t) \ + \frac{1}{2}\sigma^2 x^2 \partial_z^2 f(z,t) \right) dt \ .$$

The solution to this maximization problem is

$$x_* = -\frac{(\mu - r)\, \partial_z f(z,t)}{\sigma^2 \partial_z^2 f(z,t)} \ . \quad (11)$$

The dynamic programming principle (*Bellman's principle*) is that $f(z,t)$ is the expected value function if we use the optimal investment at time t. This is x_*

given by (11). Using this value in (10) makes both sides equal to $f(z,t)$. This cancels from both sides of (10), which leaves (with the Ito rule)

$$0 = [\, rz\, dt + (\mu - r)\, x_*\, dt\,]\, \partial_z f(z,t) + \frac{1}{2}\sigma^2 x_*^2\, \partial_z^2 f(z,t)\, dt + \partial_t f(z,t)\, dt\ .$$

After the algebra, the resulting equation is

$$0 = \partial_t f(z,t) + rz\partial_z f(z,t) - \frac{(\mu - r)^2}{2\sigma^2}\, \frac{\left(\,\partial_z f(z,t)\right)^2}{\partial_z^2 f(z,t)}\ . \tag{12}$$

This equation was derived in this way (more or less) by Merton as part of his theory of optimal dynamic investment and consumption. Exercise 4 adds the consumption piece.

Exercise 3 works out this Merton theory more explicitly for the case of a power law utility. You can see in that example that $f(z,t)$ is a utility function as a function of z for each t. You might wonder whether it's true in general that $\partial_z f(z,t) > 0$ and $\partial_z^2 f(z,t) < 0$ as long as the final utility $U(z)$ has these properties. Here are two ways to see that $f(\cdot,t)$ is a utility. You can see it is true by differentiating the equation with respect to z and see that the signs of $\partial_z f$ and $\partial_z^2 f$ do not change. This is "straightforward" but might take a long time to get the details right.

4 Option hedging in continuous time

A *stock option* it the right to buy or sell a specific stock at a specific price at a specific time or until a specific time. The right to buy is a *call* option. The right to sell is a *put* option. If the right exists only at time T, it is a *European style* option. If the right exists at any time up to time T, it is an *American style* option. Options are traded in public exchanges and their price is determined in the market. However, the *Black Scholes* theory of option pricing says what the option price should be, in an economic model. Market prices disagree with this simple theory, but the theory nevertheless provides an important way to think about buying and selling options.

The terminology of this section is that T is the *expiration time* of the option. An option has a *strike price*, written K, that is the price at which the stock will be bought or sold. An option is the right to buy or sell, but not a requirement. Consider a European style option. Suppose you own a put option (option to sell) at price K and the price is S_T. If $S_T > K$, then you can sell a share of stock for S_T on the market or for price K to the *counterparty* (the person who sold you the option). If you have a share of stock, you get more by selling on the market, so you don't exercise the option. We say the option is *out o the money*. If $S_T < K$, then you can buy a share for S_T and sell for K. This gives you a profit of $K - S_T$. For European options that are traded on exchanges, the option is *settled in cash*, which means that the exchange gives you the cash value of the option. For a put, this is

$$V(S_T) = \max\{K - S_T, 0\} = (K - S_T)_+\ .$$

For a call, similar reasoning gives

$$V(S_T) = \max\{S_T - K, 0\} = (S_T - K)_+ \ .$$

American style options with the *early exercise* feature present the owner with a dynamic optimization problem – finding the optimal strategy for exercising the option.

The Black Scholes pricing theory was first developed by Fisher Black and Miron Scholes using reasoning similar to that of Section 3. This section explains the reasoning, as I have come to understand it. Later the *binomial tree* model was invented as a way to explain option pricing to people who are not familiar with stochastic calculus. This is explained in Section 5. The explanations here may be a little quick, because the course *Derivative Securities* covers that material more deeply.

The Black Scholes model of the trading world is this. There is a risk free asset, cash, with rate of return r. There is a risky asset, the stock, whose price is S_t that is a geometric Brownian motion with parameters μ (expected rate of return) and σ (volatility). The market is full of "agents" (traders) who can buy or sell the option or the stock without market "frictions". The amount of cash or stock can be positive or negative. For cash, this is written as "borrowing = lending"; the interest rate you get for your cash is the same as the interest you pay if you borrow. It might seem surprising that this is approximately true for big agents. Owning a negative amount of stock is called having a *short position.* Selling stock you don't own (to get a negative amount of stock) is *short selling.* The Black Scholes theory allows all this, and with zero transaction cost. For European style options, the theory assumes that if you own the option at time T, then

You can get into the Black Scholes theory by asking about *dynamic replication* of an option without buying or selling the option itself. The time t, the agent has a wealth Z_t. The agent allocates X_t to the stock and the rest to cash. The agent seeks a trading strategy so that $Z_T = V(S_T)$. We say the trading strategy *replicates* the option. The strategy satisfies eqrefdZ, which means that it is *self financing.* The trader starts with some wealth and then only takes that wealth to the market. The equivalent of the value function is the wealth you need at time t with stock price s to replicate the option:

$$f(s,t) = Z_t \ \text{ so that } Z_T = V(S_T) \ . \tag{13}$$

This f is the *Black Scholes arbitrage price* of the option. The idea is that if the option price is different from the arbitrage price, then you can make a guaranteed profit by replicating the option. For example, if the option price is $P > f(S_t, t)$, then the trader can sell one option and receive P. The trader then uses $f < P$ of the money to replicate the option and keeps the rest. At time T, the trader has $Z_T = V(S_T)$, so he/she can satisfy the person he/she sold the option to. The rest is risk free profit. Basic finance theory (economic philosophy) is that arbitrage opportunities like this cannot exist. If they did exist, smart traders would jump on them and they would quickly be sold out.

The technical argument of Black and Scholes is ingenious no matter how you say it. In this version, the trading strategy will have the effect that

$$Z_t = f(S_t, t) \; , \quad \text{for all } t \leq T \; .$$

This means that if you follow the strategy (details in the next paragraph) and if you start with the right wealth at time t_0, then at all later times up to time T, you still have exactly the wealth to replicate the option.

For the calculation, we use the terminology of Black and Scholes by writing the stock component of the portfolio as $X_t = \Delta_t S_t$. That means that Δ_t is the number of shares of the stock that you own. As in the Merton theory, this can be any real number. The cash position (amount of wealth in cash) is $Y_t = Z_t - \Delta_t S_t$. We calculate $dZ = f(S_t, t)$ using the market formula (7) and using Ito's lemma. The resulting equation gives a PDE for f, which is the *Black Scholes equation*. First,

$$dZ = rY_t dt + \Delta_t dS_t \; .$$

This is the "tricky" part of this approach to Black Scholes theory. You imaging that you trade (choose Δ_t and Y_t), and then keep them for time dt while the market moves. We used the same idea in the Merton theory. We then use the "budget constraint" to eliminate Y_t and write

$$dZ_t = r(Z_t - \Delta_t S_t) + \Delta_t dS_t \; .$$

The desired replication formula $Z_t = f(S_t, t)$ allows this to be written as

$$dZ_t = r(f(S_t, t) - \Delta_t S_t) + \Delta_t dS_t \; .$$

We compare this to what you get from Ito's lemma, which is (using the Ito rule, $(dS)^2 = \sigma S_t^2 dt$)

$$dZ_t = df(S_t, t) = \partial_s f(S_t, t) dS_t + \partial_t f(S_t, dt) dt + \frac{1}{2} \partial_s^2 f(S_t, t) \sigma S_t^2 dt \; .$$

We compare these expressions and see that we can eliminate the dS term (the term with dW) if we take

$$\Delta_t = \partial_t f(S_t, t) \; . \tag{14}$$

Finally, we equate the remaining terms and drop the dt from both sides. We get

$$r \left[f(S_t, t) - S_t \partial_s (f(S_t, t) \right] = \partial_t f(S_t, t) + \frac{1}{2} \sigma_2 S_t^2 \partial_s^2 f(S_t, t) \; .$$

Some algebra puts this into a more standard form

$$0 = \partial_t f + rs \partial_s f + \frac{1}{2} \sigma^2 s^2 \partial_s^2 f - rf \; . \tag{15}$$

This is the Black Scholes equation.

5 Hedging in discrete time

6 Black Scholes formula

The *Black Scholes formula* is a formula for the solution of the Black Scholes equation with final condition $f(s,T) = V(s) = (s-K)_+$ or $f(s,T) = (K-s)_+$. One way to find the Black Scholes formula is to use the fact that the Black Scholes equation is a backward equation for a geometric Brownian motion

$$dS_t = rS_t dt + \sigma S_t dW_t \ . \tag{16}$$

This is different from the geometric Brownian motion model used to derive the PDE (15) in that the expected rate of return is r instead of μ. More precisely, f is the value function

$$f(s,t) = \mathrm{E}\Big[\, V(S_T) e^{-r(T-t)} \mid S_t = s \Big] \ . \tag{17}$$

To be clear, (17) with process (16) is a formula for the solution of (15), but it is not the derivation. Nevertheless, (17) says that the option price is the expected payout if S is the *risk-free* process (16). An investor is *risk free* or *risk neutral* (as we said before) if he/she makes the price of a risky asset equal to its discounted expected value. The conclusion of the Black Scholes theory may be stated as giving the price as the discounted expected value using the risk free process.

The Black Scholes formula may be derived using the "risk free representation" above. The solution of the SDE is

$$S_T = S_0 e^{\sigma W_T + (r - \frac{\sigma^2}{2})T} \ .$$

We can get the distribution of W_T using $\sqrt{T} Z$, where $Z \sim \mathcal{N}(0,1)$. For a put, we get

$$f(S_0, 0) = e^{-rT} \mathrm{E}\Big[\, (K - S_0 e^{\sigma\sqrt{T}Z + (r - \frac{\sigma^2}{2})T})_+ \Big] \ .$$

As an integral, this is

$$f(S_0, 0) = \frac{1}{\sqrt{2\pi}} e^{-rT} \int_{-\infty}^{z_0} \left(K - S_0 e^{\sigma\sqrt{T}Z + (r - \frac{\sigma^2}{2})T} \right) e^{-\frac{1}{2}z^2} \, dz \ .$$

The endpoint of integration is the value of z that makes $S_T = K$. The result is

$$z_0 = \frac{\log(K/S_0) - \left(r - \frac{\sigma^2}{2} \right)}{\sigma\sqrt{T}} \ .$$

The part of this formula involving just K is

$$\frac{1}{\sqrt{2\pi}} e^{-rT} K \int_{-\infty}^{z_0} e^{-\frac{1}{2}z^2} dz = e^{-rT} N(z_0) \ .$$

7 Exercises

1. Show that the variance penalized expected value violates axiom 3 for any $\lambda > 0$. Find random variables X and Z with $Z \geq 0$ and $Z > 0$ with positive probability so that $M(X + Z) < M(X)$. For this example, you can use discrete probability, such as $X = 2$ with probability .3, etc. It may be simpler to think in terms of X and Y with $Y \geq X$.

2. The *Markowitz mean variance allocation* theory involves random variables R_j, which represent the return (profit) from investing one unit of money on asset j. These have expected returns

$$\mu_j = \mathrm{E}[R_j] \; .$$

The covariance is

$$C_{ij} = \mathrm{cov}(X_i, X_j) \; .$$

Suppose you have a unit amount of money and invest w_j of that in asset j. Then your total return is

$$X = \sum_{j=1}^{n} w_j R_j = w^t R \; .$$

The vector notation in the last version on the right is $w \in \mathbb{R}^n$ with components w_j and $R \in \mathbb{R}^n$ with components R_j. The expected return is

$$r = \mathrm{E}[X] = \sum_{j=1}^{n} w_j \mu_j = w^t \mu \; . \tag{18}$$

The covariance matrix C has entries C_{jk}. The variance of the return is

$$\sigma^2 = w^t C w \; . \tag{19}$$

The *budget constraint* involves the vector $\mathbf{1} \in \mathbb{R}^n$ with all components equal to one:

$$\sum_{j=1}^{n} w_j = w^t \mathbf{1} = 1 \; . \tag{20}$$

An *allocation* (or *portfolio*) is w that satisfies the budget constraint. An allocation is *efficient* if it maximizes r with a fixed σ^2 and budget constraint (20). An allocation is *inefficient* if it is not efficient. Suppose $X = w^t R$ is an efficient allocation and $\widetilde{X} = \widetilde{w}^t R$ is an inefficient allocation with the same variance. Show that $U(X) > U(\widetilde{X})$ for any utility function, if R is Gaussian. According to von Neumann Morgenstern choice theory, any rational investor would prefer an efficient allocation to an inefficient allocation with the same variance. *Harder, attempt only after the rest is finished*: Show that this may not be true for non-Gaussian returns. *Hint.*

If Y and Z are Gaussian with the same variance, then you can think of Z as larger than Y in the sense of the arbitrage axiom if the mean of Z is larger. However, there are non-Gaussian random variables Y and Z have $\mu_Y < \mu_Z$ and $\sigma^2_Y = \sigma^2 Z$ but Z is not an arbitrage from Y in the sense that $\Pr(Z < a) > \Pr(Y < a)$ for some a. This can happen if Z has *fatter tails* than Y. [*My opinion.* Mean variance analysis is popular even though it can lead to "irrational" allocations. You might excuse this by saying it's only supposed to apply to Gaussian returns. Yet, nobody thinks returns are anything like Gaussian.]

3. Suppose the utility is a power law (4). Take the ansatz $f(z,t) = A(t)z^\gamma$.

   (a) Substitute the ansatz into the Merton PDE (12) to show that the ansatz works. Find the differential equation and then the formula for $A(t)$.

   (b) Show that the optimal allocation has the form $x_* = mz$ and find a formula for the *Merton proportion* as a function of the parameters γ, σ, and r.

   (c) An investor is *risk neutral* if they maximize expected wealth rather than expected utility. How does the Merton strategy break down in the risk neutral limit $\gamma \to 1$?

   (d) We saw that in geometric Brownian motion, it can happen that the expected value grows exponentially but the median value goes to zero exponentially. Can this happen for this Merton problem? Can the expected utility grow exponentially while the median utility decays? What does this say about how the utility function z^γ captures risk aversion?

4. (*Extra credit, do only after everything else is done, and if you're interested in economics.*) Here is the optimal policy problem that includes consumption. The rate of consumption at time t will be C_t. You "consume" money, so the wealth dynamics with consumption are

$$dZ_t = rZ_t dt + (\mu - r)X_t dt + \sigma X_t dW_t - C_t dt \ .$$

   As with wealth, we use the utility of consumption rather than consumption itself. The reasoning is similar. You might be very happy to consume two cookies rather than one cookie, but you may not care as much for cookie 101 if you already have 100 of them. There is a *discount rate*, ρ, in addition to the risk free rate. If you consume c at time t, the utility "today" (time $t = 0$) is reduced by $e^{-\rho t}$. The agent chooses X_t and C_t at time t in a way that seeks to maximize

$$H = \mathrm{E}\left[\int_0^T e^{-\rho t} U(C_t)\, dt \right] \ .$$

The constraint is $Z_T \geq 0$. Formulate a value function, the dynamic programming principle, and the Hamilton Jacobi Bellman equation appropriate for this problem. Describe the solution when the utility function has the form $U(c) = c^\gamma$.

5. Most people remember the approximate form of the Black Scholes equation without getting it exactly right. Here are two checks that help you get the terms (the signs, etc) right.

   (a) Suppose the payout is $V(s) = 1$ for all s. Then the option is not random and is not risky. A risk free asset is supposed to increase in value according to the risk free rate (otherwise, there is an arbitrage). Use this reasoning to determine the value of the option and show that this satisfies the Black Scholes equation (15).

   (b) Suppose the payout is $V(s) = s$. Then the option is the same as the stock. What is f in that case? Show that this satisfies (15)