First introduce myself. He asked about my experience on the resume (intern, stat arb).

Technical question:
Linear regression, how to formulate the optimization problem in different ways (least square, maximum likelihood)

Given n data points on a line, find the point such that the sum of absolute distance is the smallest.

A call and put, both have same parameters, under black scholes model, which have higher implied vol?"
Introduction about their team.
Time Series:
What is a stationary time series?
AR(1) Model? How to determine order?
Integrated AR model, how do you deal with it?
Linear Regression:
How to find the coefficient
How do you know whether the model is good in multivariate linear regression?
If I increase the sample size by 2, how will the t-statistic, R-square and coefficient change?
Monte Carlo:
How do you calculate pi?
How do you estimate the standard deviation?
Suppose I run the simulation 1000 times, what is the order of std?
Programming:
The performance of Hash Table and Binary search tree?
If I want to access all the elements in an ordered way, what is their performance?
Table of N numbers, want to find the largest 10 numbers. How will you do that?
Design patterns. What is a Singleton pattern, and how do you implement it?
What do you know about Multi-thread programming?
How would you calculate the integration of a continuous function on [0,1]?
Financial Products:
If I want to buy a futures contract on SPX, what should be the fair value?
What data inputs do I need?
On the back of the envelop, what is the rate and what is the SPX?
Programming and Finance

1. difference between stack and heap
2. difference between pointer and reference
3. difference between STL map and hash table
4. difference between hash table and binary search tree
5. to implement a hash table. Keys are strings. How to design the hash function, what property do we want from the hash function? Then he emphasize on the fact that the time devoted to calculating the hash function is quite large.

6. Hypothetical option. S0=10, Vol = 1$/share/year, T=1Y, K =12, European Call. Assuming the zero interest rate. Estimate the price of the option.
7. Estimate the price of a digital call option on the same stock, paying $1 if S(T)>K.
8. compare the price of these two options. And estimate again the vanilla option price.

1. Your interest area in finance.
2. Do you know Fannie Mae? Its function?
3. What do you know about 2007 subprime crisis?
4. Bond convexity; when it is negative? (when refinancing is allowed)
5. Difference of mortgage from bond
6. integral of ln(x)dx from 1 to 2
7. X, Y are iid random variables with mean 0 and variance 1. Calculate variance of (X-Y)
8. what is t-squared distribution? How do you generate it?
9. Linux: how to search what files contain some string? how to count # of lines in .txt file?
10. C++: when would you use template and when object?
11. answer to my question: KDB and Q are popular in MS; MATLAB and R are also good; some also use python or C++

1. tell me about your research
2. problems with Black-Scholes formula
3. compare value of two 10-year bonds: coupon 4%, rf 3% and coupon 5%, rf 4%
4. 10-year bond duration estimate: 5yr, 9yr, 10yr, or 12yr?
5. talk about mortgage: financial crisis
6. C++: throw exception in destructor and constructor
7. have you used database?

1. (resume) numerical methods: compare tree and monte carlo; what if derivative is on two assets or more
2. variance reduction in monte carlo; compare performance?
3. derivative pricing: basket of ten stocks, if any has return >5% next year, MS pay $1M. How will the correlation between the stocks in the basket change?
4. (resume) talk about your research experience
5. Object oriented programming: compare inheritance and composition
6. answer to my question: MS is probably the largest credit market maker, so we can do interesting things that other companies cannot handle.

1. Resume questions: how you did time series analysis;
what is linear regression;
what are the coefficients;
R-squarer to measure how good is the regression
2. BS: what is BS equation; its assumptions
3. Heston model: difference from BS model
4. How to use Monte Carlo to simulate a derivative based on Heston model (vert detailed)
5. sorting algorithms: what you know; write bubble sort
6. ATM call option price
7. apply delta hedge to a call option (ATM and ITM), if implied vol is smaller than realized vol, what can you say about P&L?

1. C++: size of an empty class instance (1 byte)
2. C++: virtual function, how to implement
3. C++: define an empty class, what are the default functions?
4. C++: is virtual constructor / destructor allowed?
5. C++: is throw exception allowed in constructor / destructor?
6. write an algorithm to reverse a C-string
7. stack and queue data structures
8. stack_1=(a,b,c), stack_2=(x,y), use a queue to get the result stack_1=(y,x), stack_2=(c,b,a)

1. introduce me to the team and ask my interests in finance
2. time series: know about R-squared, t-distribution, T-squared-distribution
3. If you have time series of a stock price, how to determine if it is integrated ADF - test : for unit root
Assume it follows AR model, how to determine the order of AR? (I don't know)
What is the condition that AR(1) can be stable? What about AR(n)?
4. linear regression: y=ax, what is a? y=ax+b, what is a?
5. Monte Carlo: simulate pi
6. How to get multivaraite normal random variables with some covariance matrix?
7. What design patterns you know? talk about singleton pattern implementation
8. brainteaser: roll 100 dice, sum the values up, what is the 95% confidence interval for the result?
9. How can you get fair future price for S&P500 in three months?

1. BS model: explain intuitively why there the drift of underlying does not appear Read Bob's
2. European debt crisis: what you know about it
3. Monte Carlo: how to simulate uniform distribution in a unit disk
1st way: simulate uniform in a square and abandon those outside the disk
2nd way: simulate radius and angle separately
4. I asked what is the difference between strats and modeling:
- modeling: larger scale, longer term (more than a few days)
- strats: focus on daily activity

1, How to find square root of 2 numerically
2, 100 fair coin, what's the probability of the number of head larger than 60, why using this formula(CLT)
3, How to find 10 largest number from n numbers(n>10)

1. conflicts between Modern Portfolio Theorem and CAPM
2. ARMA and ARIMA model
3. What the p, d, q means, and the intuition, not the definition
4. How you take into account the seasonality (algorithm used) use dummy varible or include lag-n
4. Numerical methods used, and how you implement it
5. Cubic-spline

1. stochastic calculus, random walk
2. how to calculate the covariance matrix
3. how you model CDS
4. Is the hazard rate and the value of the CDS convex or concave?
5. If the hazard rate is not constant but with a volatility, what will the price go? (increase or decrease)
6. How to find the kth largest number of an array with size of 100 million
7. write the algorithm of merge sort

1, A line cut into three pieces, the prob of forming a triangle.
2, Singleton Pattern
3, Strike 95, 100, 105, call price 5,3,2. How to arbitrage

Q1: What's the difference between heap and stack?
Q2: What's the difference between pointer and reference?
Q3: What's the difference between hash table and STL map?
Q4: What's the difference between system call and a function call?
Q5: What's stack frame?
Q6: You mentioned you play poker. Design a computer that plays the game smartly. (No need to write code, just explain some ideas).
Q7: S0=10, K = 12, Vol = 1$/year/share. What's the option price?
What's the corresponding digital call? Which one should be larger?(Answer is the digital call is larger because the Stock has to move almost 3 standard deviation away from mean at the end for the european option to pay more than digital call.

Q1: Answer the following three questions:
What are you looking for in this internship?
What can you tell me about structured finance?
Introduce yourself.

Q2: Duration Question
$100 treasury bond, 5% coupon, if interest increase by 1%, what's the value of the bond? (It's a problem about estimating the duration.) (No formula needed, just a numerical number.) I messed up this question a little bit.
I used the weighted average time approach.

Q3: X year bond duration vs. 2X year bond duration.

Q4: Subprime deal, 100 mortgages same size same interest rate. 50 of these bonds are 15 year mortgages. 50 of these are 30 year mortgages. The public expectation is that none of these are gonna default, but say there is a natural disaster, suddenly 10 of each loan types default, and lose all the money. Originally, there were 2 classes. 1st class takes bottom 50% of risk, 2nd class takes top 50% of risk. First Question: How much principal does the more risky bond lose? 40%
Second Question: If we expect no more natural disaster or defaults for the other 80 bonds, how does it impact the less risky?

1. 2 coins HH and HT, probability the same. First throw is H, what's the probability the second throw is also a H? (5/6)

2. 2 Calls, one with longer maturity than the other. Other terms the same. The one with shorter maturity happens to be priced more than the one with longer maturity. How to arbitrage? (Answer should be, only thing to do at the beginning is to long one and short the other, receive proceeds, then when the earlier maturity arrives, if S goes down, nothing will happen, you win money. If S goes up, you short the underlying at the earlier maturity and lock in the profit for the option you longed.

3. Projects I did. Both C++ and Stats
4. Poker Probability:
If the quality of a hand is uniformly distributed on [0,1]. The betting rule is as follows:
No folds or raises allowed.
If A bets one dollar, B has to bet 1 dollar.
If A checks, B can choose to check or bet 1 dollar. Again if B bets, A also have to bet.

You start, what's your strategy?
Knowledge about the mortgage market in US.

Number of different combinations to choose four numbers out of (1, 2, ..., 10) that are not adjacent to each other.

P(U1 + U2 + U3 < 1).

Three options with different prices, is there an arbitrage. (price-strike relationship.)
3 by 3 grids, you have number 1 - 9, requirements: each row and col sum up to 15. how many different combinations of solutions and algorithm to output a solution.

three toss, you can stop anytime with the money indicated by the toss, what is the expected payoff

three ways of fibonacci and complexity

1. Derive BS equation; while doing delta-hedging, why don't you take the derivative of the units of underlying?
2. write down BS price for put; what is the relationship between gamma and theta?
3. Why binomial tree and finite difference give convergent answers? (no idea...) Why do Monte Carlo methods converge to correct value?
4. What is the convergence speed of finite difference? What is alpha?
5. Explain cubic spline; is there any problem with cubic spline? what if it gives negative rate sometimes?
6. What are the coefficients of linear regression? What is the standard error of this estimate?
7. What are the assumptions of linear regression? what if the error term have different variances? what if the variances of the error term are functions of x with some coefficient to be estimated?

1. What do you know about securitized products? MBS, agent-CDO, non-agent-CDO
2. What are the differences between a treasury bond and a mortgage? (amortization and prepayment) Why is the maturity of mortgage much shorter than that of bond? (most important is prepayment)
3. Why splitting a group of MBS into tranches will reduce the risk? He showed me a concrete example: when the value of the collateral decreases and the borrower defaults, the investor gets part of his money back.

1. A boy wants to go up N stairs and he can go by one stair or two stairs for each step. How many different ways to get there? (Fibonacci number)
2. Write C++ code to compute Fibonacci. (I wrote a recursion) What is the problem with this recursion algorithm? (computationally inefficient) Could you write a better algorithm? (create a look-up table, compute F(1), then F(2)...till F(N).)
3. Write C++ code to output the result of a game: N people sit around a table, go from the 1st one and kick out every other person, which one will be the last person? For example, N=5, kick out 2, then 4, then 1, then 5, so 3 will be left.

Any quantitative projects?
Derive BS PDE
Any numerical methods to solve this?
How volatility will affect delta(positively related for ITM call, and negatively related for OTM call)
Have you done any binomial trees methods?

1. What are the assumptions for Black Scholes formula?
2. If the volatility is not constant but a function of time, what is the formula for the option price?
3. Roll a dice, if you roll 1 to 5, you get the money as the amount shows; if you roll 6, you get $7. What is the fair price to play the game? If you can choose to play a second round, what is the fair price for the game?
Answer: first question 11/3, second 9/2
4. Consider any point on the surface of the earth which is more than 1km away from the south pole. From this point, you go 1km south, then 1km east, then 1km north. What are the initial points that are invariant to this movement?
Answer: Besides north pole, all points that are 1+1/(2*pi*k) away from south pole. (k is any positive integer)
5. (I am still confused what he meant) Assume you and I each have an envelop with money in it. We know one envelop contains twice as much as the other. If you open your envelop and find $10, do you think it would be advantageous to switch your envelop with mine?

1. Can you derive Black Scholes equation?
2. Price this option using Black Scholes equation: Assume no dividend, zero interest rate. A stock is currently $60. If it ever goes above $100, you get $1. There is no maturity (perpetual option).
Answer: the price is not explicitly a function of time, so the partial differential term to t is zero. r is zero, so two terms with r are zero. Only the second order derivative is left and it has to be zero. Price is a linear function of spot price.

1. X and Y are iid normal(0,1) with correlation rho, what's E[X|Y] ?
2. Variance of [1,2,3,4], weighted variance of [1,2,3,4] where the four data points are weighted as 1,2,3,4.

3. Write code to find three elements from an array that adds up to 0. What's the complexity?
Write code for Quick Sort. What other sorting algo do you know?

4. Tell me about a project.

1. Linear Regression.
- The formula for multi-dimensional linear regression in the matrix multiplication form.
- Heteroskedasticity; how do you deal with it; if the betas are still an unbiased estimator and if it is still the best unbiased estimator.

2. Brownian Motion
- Definition & Levy's characteristic for BM.
- Martingale - if B^2(t) is still a martingale and is there a way to map the B^2(t) to another probability space such that the new process is a martingale. (Answers are no to both.)

1. Black-Scholes assumptions, close form formula, what N(d1) and N(d2) means.
2. Monte Carlo
- what kind of options have you priced
- convergence rate and why is that?
3. Implied Volatility
- how does Newton's method work and where does it come from
- what if you dont know the first derivative
4. Random Walk
- starting from 0, the prob of up and down are both 1/2, expected number of steps to hit 1?
- starting from 0, the prob of up is 60%, down is 40%, expected number of steps to hit 1?
5. two normal rv X and Y, what is the distribution of X+Y?

1. If X and Y are standard normal, with correlation rho, what is the expectation of X given Y. What if X and Y are uniform - corpula.

2. You have 1 to 10 ten numbers, what is the probability of picking 4 numbers such that no two numbers are adjacent to each other. (compute the total number of combination of picking 4 numbers, and the number of combination to choose 4 non-consecutive numbers.)

3. maximum subarray problem
http://en.wikipedia.org/wiki/Maximum_subarray_problem

4. You have a class, you want to design it such that:
- There are at most two instances allowed to be created.
- If the user tried to create a third instance or assign another instances, the pointer to the second instance will be returned.
- How do you implement it?
Singleton pattern + reference count + static variable

5. What is agency mbs and non-agency mbs?
1. difference between stack and heap.
2. difference between pointer and reference.
3. difference between STL map and hash table.
4. difference between system call and function call.
5. If you were to design a hash function (NOTE: NOT HASH TABLE), what property you want it to have? Is it fundamentally impossible to have no collision? (Yes). If there is collision, what behavior do you want the function to have? (uniform distribution of collisions)
6. s = 10, k = 12. vol = 1/year. approximate the price of european call after 1 year. I used normal distribution to approximate the value.
- follow up: if it was a lognormal instead of normal you use to price the option, what is going to happen? (higher price because of heavier tail)
- in what condition is normal a good approximation of a lognormal? (small vol, short time to expiration - idea: vol is much smaller than mean)

2. Programming
What is OOP? What is virtual function? examples.
base class and derived class both have two functions, nvf and vf. (non-virtual and virtual)
*pb = class base;
*pc = class derived;
derived D;
*pb = D;
*pc = C;
What are the results to call pb.nvf? pb.vf? pc.nvf? pc.vf?

3. Statistics
Survey among 1000 ppl on a selection whose winner is the one with more votes. 520 and 480 ppl will vote for A and B respectively. One statement: A has 90% probability to be the winner? Comment on this statement.

4. resume related questions

wondershare™

PDF Editor