The data for this problem are closing prices of 50 stocks from the S&P 500, as recorded on 757 days. The problem is to investigate some statistical parameters of the return distributions. These parameters describe the effects of trading strategies. The data is in the class web page.

You are to work individually. Organize and describe carefully the computer output you hand in. We will want to see printouts of some matlab functions you have written, as well as some numerical values and plots.

Even simple looking programming tasks can soak up more time than one estimates. Begin early. Turn in what you can in class by the deadline. Late work can be brought to the TAs (and only to the TAs) any time from the end of the class where it is due, to the end of the following lecture, but not after that. A discount factor of 50% will be applied to late work. Make sure your name is clearly written on all your work. Late work should be identified as such. If you’re expecting to have late work, indicate this on whatever you hand in on the due date.

1. To begin with, build some tools:

   (a) Write a short Matlab function that takes a vector of \( n \) consecutive stock prices and produces the (logarithmic) returns at a given lag \( L \). That is \( \log(P_t/P_{t-L}) \). If the lag is not specified then take it to be one. That is, your function can be called as \( f(\text{prices,lag}) \) or as \( f(\text{prices}) \) with \( \text{lag}=1 \) understood. (It doesn’t have to be called \( f \).)

   (b) Extend your function to handle a matrix of stock prices, one row for each time period and one column for each stock.

   (c) Write a short Matlab function that takes a vector of consecutive logarithmic returns and generates a price series. It should have an optional parameter equal to the starting value. If there are \( n - 1 \) returns, then there should be \( n \) prices.

   (d) Your function from part (c) should handle matrices of logarithmic returns.

2. Now peek at the data a bit:

   (a) Plot a histogram of the returns to Microsoft (symbol MSFT) and report the variance of those returns.
(b) Which of the 50 stocks has the highest variance, and what is that variance?

(c) Which has the lowest, and what is that lowest variance?

(d) Report the stocks having the highest and lowest mean returns, and report the values of those returns.

(e) Produce a scatter plot of the mean returns versus the standard deviations of the returns for these 50 stocks.

3. Now define some statistics of pragmatic interest. These are the mean and variance of the returns of trading strategies applied to these stocks. To keep the programming relatively simple, we will ignore trading costs, interest rates, taxes, and assume that any fractional amount of a stock can be purchased.

(a) Strategy 1 is **Buy and hold**. Write a trivial matlab function to give the value at the end of the period of $1000 invested in a stock at the beginning of the period. To be compatible with the other strategies, this function should take a vector of returns as input.

(b) Strategy 2 is **Rebalance**. In this strategy one starts with $500 invested in one stock and $500 invested in another stock. At the end of each day all the stock is sold and half is reinvested in the first stock, half in the second stock. (In practice if one stock is ahead by $x, then $x/2 of that stock would be sold to buy $x/2 of the other.) The rebalance function computes the value at the end of the period of $1000 invested at the beginning. It takes as arguments an n by 2 matrix of returns to both stocks at n periods.

(c) Strategy 3 is **Buy on weakness**. It is often recommended that technology shares should be bought on the dips. The downside is that they can go up a lot while you’re waiting for a dip. Here is a simplified strategy for buying on weakness. Start with $1000 in cash. Whenever a stock falls 3% or more from day to day, buy $100 worth of the stock at the second day’s price. The cash balance can go negative if there are more than 10 buys. At the end of the period report the sum of the cash balance and the stock value.

(d) Strategy 4 is **Momentum**. This is the same as strategy 3, except that you buy when (and only when) the stock goes up 3% or more on a day.
Programming hints

(a) To rebalance any pair of stocks, you can work with the return to the rebalanced portfolio, expressed in terms of the returns of the individual stocks. Use some care: the function takes logarithmic returns as inputs.

(b) One way to do strategies 3 and 4 is to loop over the days tracking the cash and stock values of the holding. Another way is to identify from the record of returns, on which days if any $100 of stock was purchased, add up the ending value of $100 purchased that day, add in the starting $1000, and subtract out $100 for each purchase event.

(c) If strategy 3 is too hard to program, then partial credit will be given a simpler version that rebalances to cash. For this purpose cash is a second stock with return equal to 0 in all periods. That is, we are ignoring interest.

4. Investigate strategies

(a) What is the final value of $1000 invested in CSCO, IBM and PG, under strategies 1,3,4?

(b) What is the final value of $1000 invested in ORCL and JNJ using strategy 2?

(c) Which stocks give the best (and worst) final value to the buy on weakness strategy, and what are those values?

(d) Which stock gives the best (and worst) final value to the buy on strength strategy, and what are those values?

5. Inference. The mean return to a strategy can be viewed as a parameter of the distribution of returns to that strategy. Of course that parameter is not perfectly estimated by any one time window. This is part of the reason why “Past performance is no gaurantee of future results.”

(a) Generate, but do not turn in, a matrix of size 100 by 756 random numbers. Each random number is uniformly distributed on the values 1 to 755 inclusive and they are all independent. Show the command you used to generate these values.
(b) Interpret each row as a bootstrap index into observed market returns. For each row in the matrix compute the bootstrap values of strategies 1, 3, 4 on CSCO, IBM and PG and report the bootstrap mean and standard deviations of the ending positions.

(c) For each bootstrap realization, subtract the buy and hold value of IBM stock from the strategy 3 value and present a histogram of these 100 results. Repeat for strategy 4.