MS&E 245A - Sample Investment Log Entry

Mathematical Overview

Returns

When we first look up stock data, we initially see a series of prices for a stock. One day it might sell for $101 per share, another day for $103, and so on. As investors, however, we are much more interested in how the prices for a stock *change* than what they are at any given point in time. After all, it is these changes in price (either up or down) that will determine whether we make or lose money after acquiring an investment position. One measure of such changes is called the *return* which is basically a representation of the percentage change in an assets value over some period of time.

Formally, let S_i represent the price of a stock at time i. i could be measured in days, weeks, months, years, milliseconds, or any other unit of time. Whatever unit of time we use for i, that is what the returns we calculate will be with respect to - thus, we could have daily returns, representing the percentage change in a stock's price over a day, weekly returns, and so on. Formally, we let r_i represent the return at time i and define:

$$r_i := \frac{S_i - S_{i-1}}{S_{i-1}}$$

In other words:

$$\text{return} := \frac{\text{Value now} - \text{Value in the past}}{\text{Value in the past}}$$

Thus, suppose we have a stock that was worth $100 a week ago and now is worth $103. We would calculate the weekly return as:

$$r_i = \frac{103 - 100}{100} = \frac{3}{100} = 0.03 = 3\%$$

The stock has grown in value by 3% over the past week.

Logarithmic Returns

In much of work for the simulation in this class, we will use logarithmic returns, which are a slight variation on the basic returns discussed above. The logarithmic return is defined as follows:

$$r_i := \log(S_i) - \log(S_{i-1}) = \log\left(\frac{S_i}{S_{i-1}}\right)$$

The logarithmic rate of return does **not** equal the logarithm of the rate of return. Instead, in most situations, it will be pretty close to the rate of return. For instance, taking the example above with the stock that grew from \$100 to \$103 we have:

$$r_i = \log\left(\frac{103}{100}\right) = \log(1.03) = 0.0295588 \approx 0.03$$

Returns Over Multiple Periods

Two Periods

One of the reasons that logarithmic returns are frequently used is because they make calculations involving returns over multiple periods much simpler and easier. For instance, consider three points in time, t_0, t_1, t_2. Suppose you know the returns from t_0 to t_1 (call this r_1) and the returns from t_1 to t_2 (call this r_2). You want to know the total return from t_0 to t_2 (call this $r_{0,2}$).

For the basic returns, we have:

$$(1 + r_{0,2}) = (1 + r_1) * (1 + r_2)$$

E.g. suppose we had a 3% return between t_0 to t_1 and from t_1 to t_2. If you are familiar with compound interest, it may be relatively intuitive for you to then calculate:

$$(1 + 0.03) * (1 + 0.03) = 1.0609$$

So, from this, you may intuitively see that your return over the two periods together is 6.09%. But, note that we're expressing returns as percents, and 1.0609 isn't actually a percent. Instead, to get a percent from it, we need to first subtract 1 from it to get 0.0609 (and then multiply by 100 if we want it to be 6.09 instead of 0.0609). So, the formula for the return $r_{0,2}$ between two periods using basic returns is:

$$r_{0,2} = [(1 + r_1) * (1 + r_2)] - 1$$

That might not seem terrible, but as we'll see, it's a lot simpler with log return.

For the log returns, we simply have:

$$r_{0,2} = r_1 + r_2$$

To see why this is, recall that we use S_0, S_1, S_2 to denote the value of our stock at times $0, 1, 2$. Note that for log returns, we have:

$$r_{0,2} := \log(S_2) - \log(S_0) = \log(S_2) - \log(S_1) + \log(S_1) - \log(S_0)$$

(here, we just subtracted and added $\log(S_1)$ in the middle)

$$= [\log(S_2) - \log(S_1)] + [\log(S_1) - \log(S_0)] = r_1 + r_2$$

Multiple Periods

When considering returns over multiple periods, the advantage of log returns becomes even more stark. Suppose we want to know the yearly return that would be produced by a given daily return. We frequently assume that there are 252 days in a year, since this is roughly the number of days in which financial markets are open for trading (after excluding holidays and weekends). Thus, we have for basic returns:

$$1 + r_{yearly} = (1 + r_{daily})^{252}$$

or, in other words:

$$r_{yearly} = (1 + r_{daily})^{252} - 1$$

For log returns, we have simply:

$$r_{yearly} = 252 * r_{daily}$$

What if we want to go the other direction? We have a yearly return and want to know what daily return would produce it. For basic returns, we must calculate:

$$r_{daily} = (1 + r_{yearly})^{1/252} - 1$$

whereas for log returns, we have simply:

$$r_{daily} = \frac{1}{252} r_{yearly}$$

Sharpe Ratios

One tool that is often used to evaluate an investment is called the Sharpe Ratio. You will learn more about this later in the quarter, but here is a brief intro so that you can calculate it for your selected fund. The goal of the Sharpe ratio is to evaluate the returns that an investment yields in light of (a) the "risk-free" return that an investor could receive from investing in US Treasury securities and (b) the riskiness or variability of the return. It is natural that riskier investments tend to have higher returns than safer investments. Otherwise, if there were say, two stocks with the same average return but one were much riskier than the other, no one would buy the riskier stock! Simply comparing the returns of two investments then doesn't necessarily give a good sense of which is "better" than the other. One investment might have a slightly higher return than another, but if it is much, much riskier, then it probably still isn't a good deal. The Sharpe ratio seeks a standardized way in which investments with different risks and returns can be compared across a common metric. It is calculated from historical data as follows:[1]

- Let r_i be the return of an asset in period i.
  - This can be either the regular or the logarithmic return. Depending on which you use, you will get slightly different values for the Sharpe ratio. Both versions get used in various investing contexts. <u>In this investment simulation we will calculate Sharpe ratios using logarithmic returns</u>.
  - Also, this can be the return of the asset over different intervals, such as daily, weekly, monthly, or yearly. Different intervals of returns yield different Sharpe ratios - e.g. there is a daily Sharpe ratio and a yearly Sharpe ratio. <u>In this investment simulation, unless otherwise stated, we will use daily Sharpe ratios.</u>
  - If you are using logarithmic returns, it is relatively straightforward to convert from, for example, a daily to a yearly Sharpe ratio. If you are using regular returns, however, the calculation is more difficult.[2]
  - If you look up information on various investments online, you may well see Sharpe ratios reported. Note that it can be difficult to compare these to your calculations. The reported ratios may use regular or logarithmic returns, they may be based on annual, monthly, or some other type of returns, they may be calculated over different time periods, and they may use different assumptions about the risk-free interest rate. As long as you do all of your calculations yourself, using the same assumptions, however, then you will get a consistent set of Sharpe ratios that you can use to compare across different investments.
- Let r_i^f be the risk-free return yielded by a US Treasury security during that same period.
- The "excess return" (D_i) of the asset over the risk free asset is then $D_i := r_i - r_i^f$. On average, we would hope that risky stocks will outperform the risk free rate, meaning that excess returns will be positive. But, for any given day, week, month, etc. this might not always be true.
- The Sharpe ratio for an investment is then defined as:

$$\text{Sharpe Ratio} := \frac{\bar{D}_i}{\text{SD}(D_i)}$$

  where $\bar{D}_i$ represents the average excess return and $\text{SD}(D_i)$ is its standard deviation.

[1] See this article here, by Nobel laureate William Sharpe for more details.

[2] For more information on additional complexities that can accompany calculating annual versions of the Sharpe ratio, see the optional reading "The Statistics of Sharpe Ratios" by Andrew Lo, available here

- See the document “Download_Financial_Data.pdf” for information on how you can download financial data on your chosen fund in order to calculate the r_i in this calculation.
- In order to calculate the r_i^f, you can use the Yahoo index symbol ˆIRX which tracks the returns of 13-week treasury bills. You can download this data using the same methods as you use for downloading data on your fund returns.
- **NOTE:** the data that you download for ˆIRX is not given in daily prices, as it is for the fund data that you download. Instead, it is given in annualized returns. Please see the document “Logarithmic and Regular Returns.pdf” for details on how to convert the ˆIRX to the form you'll need it for calculating Sharpe ratios.
- See the segments for Excel and R below for more detailed practical guidance on calculating Sharpe ratios from real data.

Mean / Expected Value

All modern investment theory considers investment returns to be random variables.[3] Thus, finance professionals routinely use descriptions of random variables, such as mean, variance, covariance, and so forth to describe patterns of stock returns. Suppose we have many weeks over which we observe a stock's return, $r_1, r_2, r_3, \ldots$. We define the mean return as:

$$\text{mean} := \sum_{i=1}^{n} \frac{1}{n} r_i$$

That is, we sum all of the returns over n weeks and divide by n.

We frequently will use $\bar{r}$ to refer to the mean return. Another common term for mean is “expected value”, often times written in mathematical form as $\mathbf{E}[r_i]$.

Variance and Standard Deviation

The variance of a return is one way to describe how much it moves around from week to week. We define the variance as:

$$\text{var}(r_i) := \sum_{i=1}^{n} \frac{1}{n-1} (r_i - \bar{r})^2$$

We also will at times refer to the standard deviation of a return, which is simply the square root of its variance.

Covariance

Suppose that we have two stocks whose returns we represent as $r_i^{(1)}$ and $r_i^{(2)}$. The covariance between these stocks is a measure of how closely the movements of the stocks are related to one another. A high covariance means that when one stock goes up, the other stock tends to go up too. A negative covariance means that when one stock goes up in value, the other stock tends to go down in value. A covariance that is zero or close to zero means that in certain ways, there is relatively little relationship between the movements of the stocks. We define covariance as:

$$\text{cov}\left(r_i^{(1)}, r_i^{(2)}\right) := \sum_{i=1}^{n} \frac{1}{n-1} \left(r_i^{(1)} - \bar{r}^{(1)}\right)\left(r_i^{(2)} - \bar{r}^{(2)}\right)$$

Note that we could also consider the covariance of one stock with itself, e.g. $\text{cov}\left(r_i^{(1)}, r_i^{(1)}\right)$. In this case, as you can check on your own, the covariance of a stock's returns with itself is equal to the variance of that stock's returns. Thus, you can think of variance as a special case of the more general notion of covariance.

The below section give information on how to calculate these figures using a variety of software platforms.

[3] Random variables can still exhibit certain predictable patterns, and thus much of investment science is devoted to understanding the predictable patterns within the ultimately random stock returns.

R

The `quantmod` package in R has a built-in function `Delt()` which can be used to directly calculate returns from a list of prices. Note that since we will always have one fewer returns than we have prices (since there is nothing to compare the first price to in order to find a return) the first element of the result of the `Delt()` function will be `NA`. We may thus want to remove this before proceeding with additional calculations on the returns.

To calculate log returns using the `Delt()` function, we specify the optional argument `type = 'log'`

After downloading the data (see the online resource section for "Stock Return Calculations.R" giving the full script used here) we can use:

```
## pull out adjusted closing prices as separate objects
DJI_Adj_Close = data.frame(DJI)$DJI.Adjusted
SP500_Adj_Close = data.frame(SP500)$GSPC.Adjusted

DJI_Return = Delt(DJI_Adj_Close, type = 'log')[2:length(DJI_Return)] ## remove the first element, which is NA
SP500_Return = Delt(SP500_Adj_Close, type = 'log')[2:length(SP500_Adj_Close)]

mean(DJI_Return)
var(DJI_Return)

cov(DJI_Return, SP500_Return)
```

Note, if we want a covariance matrix like we got in the output from Matlab (see discussion above) we can use `cbind()` (for column bind) to combine the two sets of returns as columns in a matrix and then apply the `cov()` function to that:

```
> cov(cbind(DJI_Return, SP500_Return))

              DJI_Return SP500_Return
DJI_Return 2.332940e-05 1.975636e-05
SP500_Return 1.975636e-05 1.884656e-05
```

Sharpe Ratio

First, we want to download interest rate data for our risk free rate of return:

```
IRX = getSymbols(
  "^IRX",
  from = fromdate,
  to = todate,
  auto.assign = FALSE)
```

The interest rate data from Yahoo Finance might contain missing values. For this course, we simply take the intersection of the dates where we have observations for both SP500 and risk free rate.

```
SP500_Dates = rownames(data.frame(SP500))
IRX_Dates = rownames(data.frame(IRX))

SP500_Data = data.table(Date = SP500_Dates, SP500$GSPC.Adjusted)
IRX_Data = data.table(Date = IRX_Dates, IRX$IRX.Adjusted)
```

```
setkey(SP500_Data, Date)
setkey(IRX_Data, Date)

Data = IRX_Data[SP500_Data]
Data = Data[complete.cases(Data)]

SP500_Adj_Close = Data$GSPC.Adjusted
IRX_Adj_Close = Data$IRX.Adjusted
```

Next, we convert that data to daily log returns. See “Logarithmic and Regular Returns.pdf” for details on why we perform these next calculations as we do.

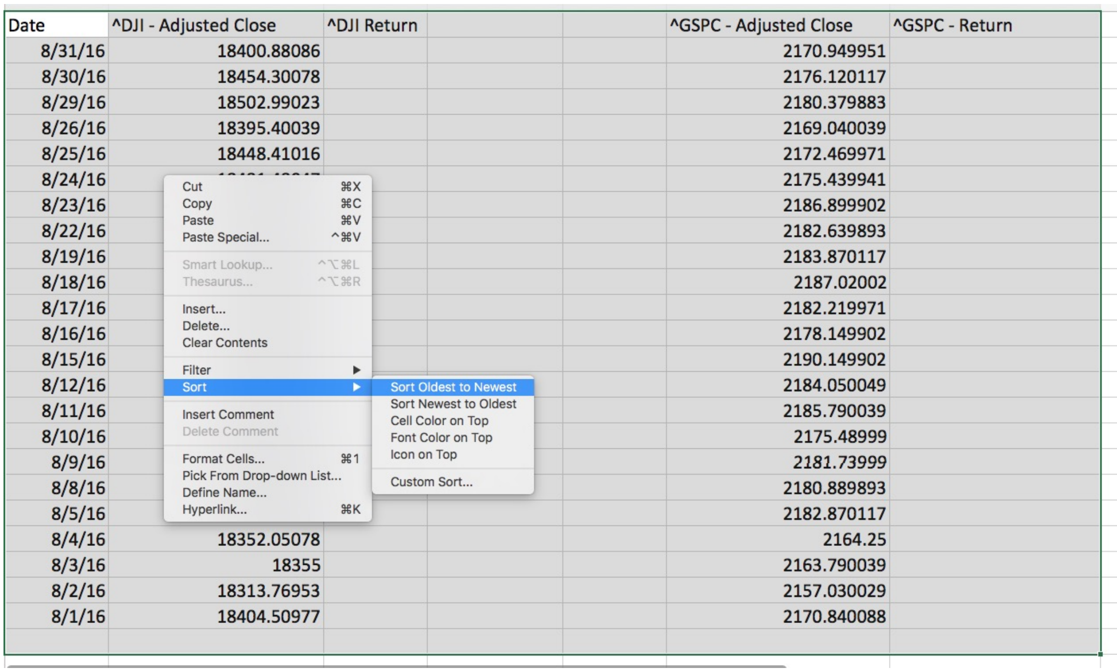
```
SP500_Return = Delt(SP500_Adj_Close, type = 'log')[2:length(SP500_Adj_Close)]

IRX_Annual_Return = (IRX_Adj_Close/100)[2:length(IRX_Adj_Close)]
IRX_Annual_Log_Return = log(1+IRX_Annual_Return)
IRX_Daily_Log_Return = IRX_Annual_Log_Return/252
```

Finally, we perform the specific calculations to find the Sharpe Ratio:

```
Excess_Return = SP500_Return - IRX_Daily_Log_Return

Sharpe_Ratio = mean(Excess_Return, na.rm = TRUE) / sd(Excess_Return, na.rm = TRUE)

Sharpe_Ratio ## daily Sharpe ratio (what we use)

Sharpe_Ratio*sqrt(252) ## annual Sharpe ratio (for those interested)
```

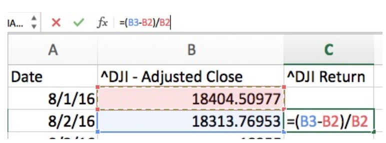
Excel

First, import your data into Excel. **IMPORTANT!** Once imported, you need to sort your data so that it is going forward in time, not backward, which is the default. Otherwise, all of your returns, for example, will be the opposites of what they should be. You can do so by highlighting all of your data (both the columns with closing prices and those with returns, and also highlight the row with your column headings), right clicking, and choosing the option to "Sort Oldest to Newest". E.g.



Returns

Now, go to the column that is one to the right of the column where you have the adjusted closing prices stored. Select the cell that is one cell from the top of this column (it is the cell highlighted in green in this picture). Type into this cell an "=" sign and then enter a formula such as is shown here:



Note that you can either type in the coordinates of the cells (e.g. B3) manually, or you can just click on the cells that you want and their coordinates will be put into your formula. Notice that this is the formula we defined for investment returns above. Once your formula is complete, hit "Enter/Return."

Now, notice that when you have a cell highlighted, there is a small darker square in the bottom right of the cell that appears. Once you have entered the formula to calculate the return between the first and second day's stock prices, double click this small box. It will automatically apply your formula to all of the other cells in the column below it.

Log Returns

To calculate log returns with Excel, use much the same procedure, but just substitute the formula given above for log returns. **Note:** make sure to use the Excel function `LN()` for the natural log, rather than `log()` which gives a base-10 log (which is confusing because in most scientific computing platforms, the function `log()` gives a natural log by default).

Date	^DJI - Adjusted Close	^DJI Log Return
8/1/16	18404.50977	
8/2/16	18313.76953	=LN(B3/B2)

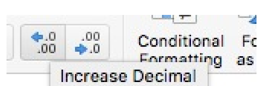
Mean, Variance, Covariance

To calculate the mean / expected value for the returns, go to a new, blank cell. Type into this cell `=AVERAGE(`. Then, select all of the values for the stock returns in your sheet that you calculated previously. Finally, type in the close of the parentheses `)` and hit "Enter/Return."

To find the variance, do the same as for mean, but use the formula `=VAR.S()` instead. The `.S` signifies that this is a *sample covariance* meaning that we are calculating it based on a limited sample of all stock returns out there. When performing calculations on stock data, we will always use the sample versions of variance and covariance.[4] Note, you may well find it helpful to write the label of what you have calculated in a cell above or to the side of your numbers that you calculate so that you can find it easily in the future, e.g.:

Date	^DJI - Adjusted Close	^DJI Log Return	
8/1/16	18404.50977		**^DJI Mean Ret**
8/2/16	18313.76953	-0.00494252	-0.0000090
8/3/16	18355	0.002248807	**^DJI Var Ret**
8/4/16	18352.05078	-0.00016069	0.0000144
8/5/16	18543.5293	0.010379576	

Note also, it is possible that when you first do these calculations, your numbers may appear in scientific notation. If you prefer them to be in decimal notation, then highlight the cells in question, right click on them, select "format" from the options and then select "Number." After this, you may need to increase the number of decimals displayed in order to prevent them from being rounded down to zero. You can do so with these buttons at the top of the screen:



Finally, to calculate covariance, calculate the returns for two different sets of returns and use the formula `=COVARIANCE.S(Returns1,Returns2)` where for `Returns1` and `Returns2` you fill in the highlighted cells for the respective returns. See the spreadsheet "Stock Return Calculations.xlsx" posted with the other project materials, for a full example of this.

Sharpe Ratio

1. Download data for the ˆIRX symbol - this gives annual risk-free returns, expressed as percentages. It is based on the 13-week US Treasury interest rate.

2. See the document "Logarithmic and Regular Returns.pdf" for details on the formula to use when converting the ÎRX data into daily log returns of the form we want to use for the risk-free rate, r_f.

3. Calculate the daily log returns (r_i) for your given asset as described above.

4. Calculate the daily "excess return" (D_i) by subtracting the daily risk free returns (r_f) from your daily asset returns (r_i).

5. Use Excel to find the average and standard deviation of D_i.

[4]See here if you need a refresher on sample vs. population variance and covariance.

6. Use Excel to calculate the ratio of the average of D_i to the standard deviation of D_i. This is your daily Sharpe ratio.

See the spreadsheet “Stock Return Calculations.xlsx” posted with the other project materials, for a fully worked example of this. See also the video for the first Friday’s problem session.