Financial Markets I

Lecture 4: Statistical Facts on Stock Returns

Master Finance & Strategy

Spring 2018

SciencesPo

General Overview

In Lectures 2 and 3 we studied bonds. Now we turn to stocks.

- Lecture 4: *Statistical Facts on Stock Returns*
  - Introduction, and basic facts on stock returns.
- Lecture 5: *Portfolio Theory*
  - How to choose a portfolio of stocks.
- Lecture 6: *The Capital Asset Pricing Model (CAPM)*
  - Equilibrium risk-return tradeoff, and risk-adjusted discount rates.
- Lecture 7: *Valuation of Stocks*
  - How the market values stocks. The PV formula strikes back.
- Lecture 8: *Market Efficiency*
  - Does the market value stocks "efficiently"?

Overview of Lecture 4

1. What Stocks Are.
2. Computing Stock Returns.
3. Stocks vs. Bonds: Historical Performance.
4. Sample Mean, Variance, and Standard Deviation.
5. Covariance and Correlation.
6. Probability and Statistics.
7. Summary: Key Concepts and Basic Facts.

Data sources for this lecture:

- *Stocks, Bonds, Bills, and Inflation, 2014 Yearbook,* Ibbotson Associates.
- Center for Research in Security Prices (CRSP).

1. What Is a Stock?

- Different types of stocks: common stock, non-voting stock, preferred stock.

- Main features of **common stock**:
  - Voting rights
  - Rights to dividends
  - Limited liability

- Dividends are the cashflows of common stock. They are:
  - discretionary. By contrast, debt payments are mandatory.
  - not tax-deductible (because they are not business expenses). By contrast, debt payments are tax-deductible.
  - generally paid quarterly.

2. Computing Stock Returns

Return between two dates, 0 and 1:

- Invest X dollars in the stock at date 0.
- If there are any dividend payments between dates 0 and 1, reinvest them in the stock.
- Suppose that value of the stock investment at date 1 is V dollars.
- The rate of return on the stock R is defined as

$$X(1+R) = V \quad \Rightarrow \quad R = \frac{V}{X} - 1 = \frac{V - X}{X}.$$

A Simple Formula

Assume that the only dividend payment between dates 0 and 1 is at $t = 1$. Suppose you invest X dollars. What is your rate of return over the period?

- With X dollars, buy $n = X/P_0$ shares of the stock at date 0, where P_0 is the date-0 price per share.
- At $t = 1$, these n shares are worth nP_1, where P_1 is the date-1 price. Moreover, they pay nD_1, where D_1 is the date-1 dividend per share.
- Value of the stock investment at $t = 1$ is $V = n(P_1 + D_1)$.
- Return on the stock R between dates 0 and 1 is

$$R = \frac{n(P_1 + D_1) - X}{X} = \frac{D_1 + (P_1 - P_0)}{P_0}.$$

  Return is due to **dividend** payment D_1 and **capital gain** $P_1 - P_0$.

- The rate of return does not depend on the amount X invested.

3. Stocks vs. Bonds: Historical Performance

Annual returns on indices in %

Year	Large Stocks	Small Stocks	Long-Term Govt Bonds	T-Bills	Inflation
1999	21.04	29.79	-8.96	4.68	2.68
2000	-9.11	-3.59	21.48	5.89	3.39
2001	-11.88	22.77	3.70	3.83	1.55
2002	-22.10	-13.28	17.84	1.65	2.38
2003	28.70	60.70	1.45	1.02	1.88
2006	15.80	16.17	1.19	4.80	2.54
2007	5.49	-5.22	9.88	4.66	4.08
2008	-37.00	-36.72	25.87	1.60	0.09
2009	26.46	28.09	-14.90	0.10	2.72
2010	15.06	31.26	10.14	0.12	1.50
2011	2.11	-3.26	28.23	0.04	2.96
2012	16.00	18.24	3.31	0.06	1.74
2013	32.39	45.07	-11.36	0.02	1.50

Stocks vs. Bonds (cont'd)

Annual returns on indices in % before and during the Great Depression

Year	Large Stocks	Small Stocks	Long-Term Govt Bonds	T-Bills	Inflation
1926	11.62	0.28	7.77	3.27	-1.49
1927	37.49	22.10	8.93	3.12	-2.08
1928	43.61	39.69	0.10	3.56	-0.97
1929	-8.42	-51.36	3.42	4.75	0.20
1930	-24.90	-38.15	4.66	2.41	-6.03
1931	-43.34	-49.75	-5.31	1.07	-9.52

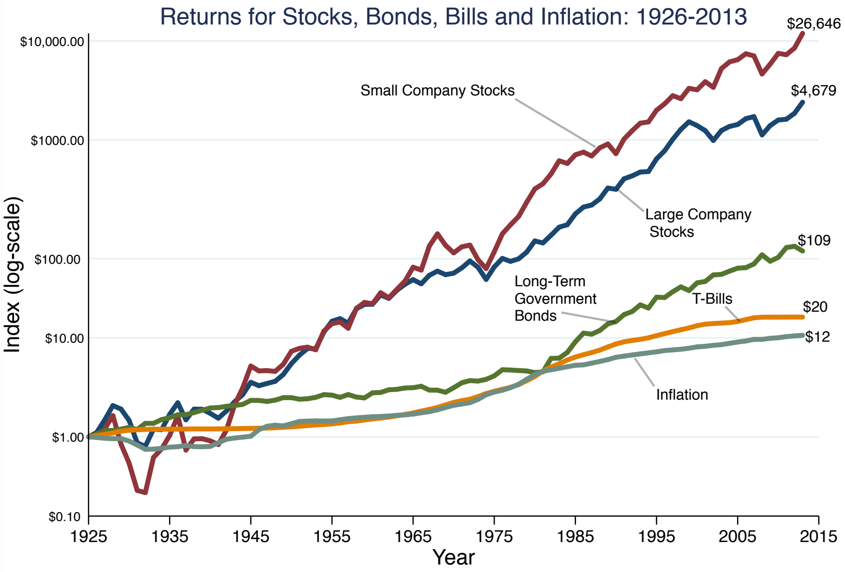
Compounded Returns

- Let R_t denote the rate of return in year t.
- Suppose you start with initial wealth W_0 at date 0.
- After T years, your accumulated wealth is

$$W_T = (1 + R_1)(1 + R_2) \dots (1 + R_T) W_0.$$

- The next figure shows, for a $1 initial investment made in Jan 1926, the accumulated wealth up until Dec 2013, for several assets.

Compounded Returns: Stocks vs. Bonds



4. Sample Mean, Variance, Standard Deviation

- We will use basic notions of **statistics** to describe some key aspects of the returns on various financial assets.
- To start with, we recall the definition of the sample mean, sample variance, and standard deviation.
- Say we have a 'sample' of N observations for a given variable X.
  - The observations $X_1, X_2 \ldots, X_N$ constitute a data series.
- For instance, we observe the returns on the S&P500 every year between 1926 and 2013.

Basic Statistics

- A measure of the average value of the series is the **sample average (or sample mean)**
$$\overline{X} = \frac{X_1 + \cdots + X_N}{N}.$$
- Two measures of the dispersion of the series around the average value are the **sample variance**
$$v(X) = \frac{(X_1 - \overline{X})^2 + \cdots + (X_N - \overline{X})^2}{N-1}$$
and the **sample standard deviation**
$$s(X) = \sqrt{v(X)}.$$
- The sample standard deviation is in the same units as the data series. It measures the "typical" distance of the series elements from the average value.

Stocks vs. Bonds

Using historical returns 1926-2013, we can compute:

	Sample Average (%)	Sample St. Dev. (%)
Large Stocks	12.1	20.2
Small Stocks	16.8	32.3
Long-Term Govt Bonds	5.9	9.8
T-Bills	3.5	3.1
Inflation	3.0	4.1

First observation (to be refined later in Lecture 6, see CAPM):

"Riskier" investments tend to have higher returns on average

Individual Stocks vs. Indices

Using annual returns 1967-2017, we can compute:

	Sample Average (%)	Sample St. Dev. (%)
Coca-Cola	15.3	25.2
Disney	19.1	39.8
Ford	19.4	59.7
GE	12.8	26.2
IBM	11.7	28.2
Xerox	10.5	39.8
S&P500	8.4	16.3

Note: The fact that individual stocks beat the index in terms of average returns is partly due to "survivorship bias".

Individual Stocks vs. Indices (cont'd)

- The sample standard deviation of the S&P500 is much smaller than those of the individual stocks.
- By contrast, the average returns are comparable.
- Second important observation (to be refined in Lecture 5):

> There are benefits to diversification.
>
> By holding a diversified portfolio, we can reduce risk without sacrificing expected return.

5. Covariance and Correlation

- Say we want to measure how much two data series X_1, .., X_N, and Y_1, .., Y_N, tend to move together.
- Two related measures of association are the **sample covariance**

$$Cov(X, Y) = \frac{\sum_{i=1}^{N}(X_i - \overline{X})(Y_i - \overline{Y})}{N - 1}$$

and the **sample correlation**

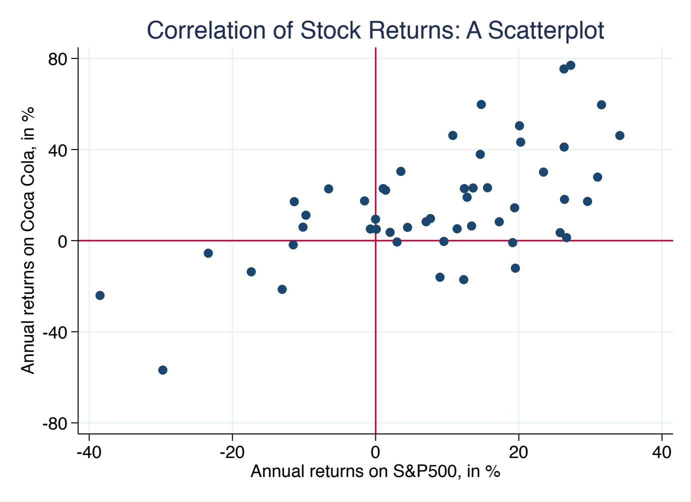
$$\rho(X, Y) = \frac{Cov(X, Y)}{s(X)s(Y)}.$$

Properties of Covariance and Correlation

- The sample covariance and correlation have the same sign. They are:
  - positive, if X and Y are high at the same time;
  - zero, if X and Y are unrelated;
  - negative, if X is high when Y is low.
- The correlation is always a number between -1 and 1. It is:
  - equal to 1 if there is an exact linear relation with positive slope between X and Y (perfect correlation);
  - equal to -1 if there is an exact linear relation with negative slope between X and Y.

Return Correlation: A Scatterplot

Coca-Cola vs. SP500, annual returns 1967-2017. Sample correlation 0.62.



Correlation of Stock Returns

Using annual returns 1967-2017, we can compute the following correlation matrix:

	Coca-Cola	Disney	GE	IBM	Xerox
Coca-Cola	1				
Disney	0.55	1			
GE	0.57	0.43	1		
IBM	0.34	0.06	0.43	1	
Xerox	0.35	0.19	0.35	0.37	1

Note that:

- Usually stocks are positively correlated.
- Correlations can be significant.

Serial Correlation

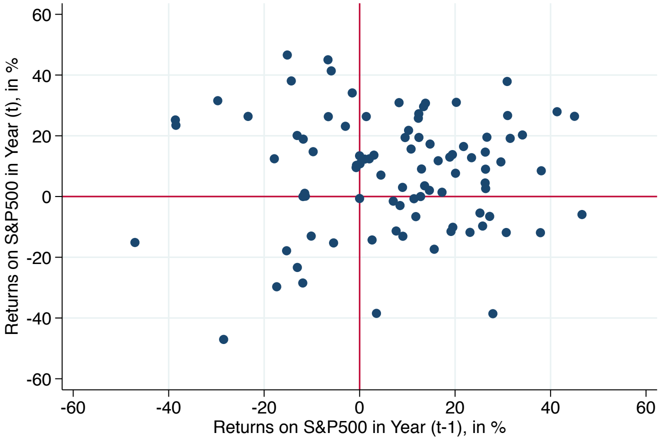
The **serial correlation**, or auto-correlation, of a stock is the correlation between the return on a given year, and the return on the previous year.

Using annual returns 1926-2013, we can compute:

	Serial Correlation
Large Stocks	0.02
Small Stocks	0.06

A Scatterplot

S&P500 vs. S&P500 in previous year: Annual returns 1926-2017.
Sample correlation is 0.016.



Serial Correlation and Predictability

- There is very little serial correlation in stock returns.
  - The sign of return auto-correlation may depend on data frequency.

- Third basic observation (to be refined in Lecture 8):

The evidence suggests that it is hard to predict future stock returns based on past returns.

6. Notions of Probability

- Consider a **random variable** Z which can take any of K values depending on the 'state of the world'.
  - e.g., Z could describe the outcome of rolling a dice ($K = 6$).

- Suppose that Z takes values z_1, ..., z_K with probabilities p_1, ..., p_K.

- The **expectation** of Z is

$$E(Z) = p_1 z_1 + \cdots + p_K z_K = \sum_{k=1}^{K} p_k z_k.$$

- The **variance** of Z is

$$V(Z) = p_1 \left(z_1 - E(Z)\right)^2 + \cdots + p_K \left(z_K - E(Z)\right)^2 = E\left[\left(Z - E(Z)\right)^2\right],$$

  and the **standard deviation** of Z is

$$\sigma(Z) = \sqrt{V(Z)}.$$

Example

- Consider the random variable Z describing the outcome of rolling a dice: Z can take values 1, 2, 3, 4, 5, or 6, each with probability $1/6$.

- The expectation of Z is

$$E(Z) = \frac{1}{6} \times 1 + \frac{1}{6} \times 2 + \frac{1}{6} \times 3 + \frac{1}{6} \times 4 + \frac{1}{6} \times 5 + \frac{1}{6} \times 6 = 3.5.$$

- The variance of Z is

$$V(Z) = \frac{1}{6} \times (1 - 3.5)^2 + \frac{1}{6} \times (2 - 3.5)^2 + \frac{1}{6} \times (3 - 3.5)^2 + \frac{1}{6} \times (4 - 3.5)^2 + \frac{1}{6} \times (5 - 3.5)^2 + \frac{1}{6} \times (6 - 3.5)^2 = 2.9167.$$

- The standard deviation of Z is

$$\sigma(Z) = \sqrt{2.9167} = 1.7078.$$

Statistics and Probability

- Suppose that we roll a dice many times, and record the data series formed by the successive outcomes.
- For this data series, we can compute a sample average, sample variance, and sample standard deviation.
- These will generally be different than the expectation (3.5), variance (2.9167), and standard deviation (1.7078), of Z.
- However, if we roll the dice very many times, they will get very close. If we roll the dice an "infinite" number of times, they will become equal.
- **General result (Law of Large Numbers):** Statistics concepts converge to their probability counterparts as the sample size (N) grows.

Using Statistics in Finance

- The return on a stock over the next year can be viewed as a random variable.
- From past return data, we can **estimate**
  - its expectation (the expected return) by the sample average;
  - its variance by the sample variance;
  - its standard deviation by the sample standard deviation.
- For variance and standard deviation, we can obtain quite precise estimates using a sample of even one year (the most recent), provided that we use daily or weekly returns.
- For expected return, we should use as large a sample as possible.
- One should bear in mind that estimates may vary a lot depending on the sample period being used for estimation.

Estimate of the Market Risk Premium

- We want to know the expected return of large stocks relative to T-bills. We refer to this as the **market risk premium** (MRP).
- We can obtain an estimate of the MRP using sample averages.
- The 1926-2013 sample averages for large stocks and T-bills are 12.1% and 3.5%, respectively. Large stocks outperformed by 8.6%.
- An estimate of the MRP is 8.6%.
- We will come back to the MRP in Lectures 6 and 7.

Realized vs. Expected Returns

- 8.6% is an estimate of the **expected return** of large stocks relative to T-bills.
- i.e., we *expect* that large stocks will, on average, outperform T-bills by 8.6% over one year.
- This does not mean that large stocks will outperform T-bills by 8.6% *for sure* over the next 12 months.
- The actual return of large stocks in excess of T-bills over the next 12 months will only be known one year from now. This will be the **realized return**.

Using Statistics in Finance (cont'd)

- The rates of return on two different stocks over the next year can be viewed as two random variables.
- The tendency of the two stock returns to 'comove' depends on their **joint distribution**.
- Given the joint distribution for two random variables Y and Z, one can define their covariance as

$$Cov(Y, Z) = E\Big[\big(Y - E(Y)\big)\big(Z - E(Z)\big)\Big].$$

  and their correlation as

$$\rho(Y, Z) = \frac{Cov(Y, Z)}{\sigma(Y)\sigma(Z)}.$$

- From past return data for two stocks, we can **estimate**
  - their covariance by their sample covariance;
  - their correlation by their sample correlation.

7. Summary: Key Concepts

- How to compute stock returns.
- Sample average, sample variance, and sample standard deviation.
- Sample covariance and sample correlation.
- Expectation, variance, and standard deviation of a random variable.
- Realized vs. expected returns.
- Using sample statistics as estimates.

Summary: Basic Facts

1. "Riskier" investments tend to have higher returns on average.
   - Lecture 6 will develop a theory of the trade-off between risk and return.
2. There are benefits to diversification. By holding a diversified portfolio, we can reduce risk without sacrificing expected return.
   - Diversification will play a big role in Lecture 5 (portfolio theory).
3. The evidence suggests that it is hard to predict future stock returns based on past returns only.
   - We will examine more refined forms of predictability in Lecture 8.