CHAPTER 4
Random Walks, Risk and Arbitrage

A. <u>Market Efficiency and Random Walks</u>

Market efficiency was characterized earlier as existing when market prices reflect information. Since information dissemination (news) occurs randomly, security price changes might be expected to occur randomly. Testing for market efficiency (the extent to which prices reflect information) generally involves testing for randomness in price changes. Tests of market efficiency are concerned with which types of information are reflected in security prices and the length of time required for new information to be reflected in security prices. Tests and studies of market efficiency are classified into three types:

1. Weak form efficiency: concerned with whether market prices reflect historical price sequences
2. semi-strong form efficiency: concerned with whether market prices reflect all publicly available information
3. strong form efficiency: concerned with whether security prices reflect all information, including publicly available and not publicly available information.

Such tests are important for purposes of this class for several reasons:

1. It is important to determine whether apparent cycles (for example, booms and busts) in security prices are the result of predictable causes or the result of purely random behavior.
2. It is important to determine whether apparently unusual price behavior is the result of illicit trading behavior or the result of purely random behavior.
3. Understanding whether investors respond rationally to information requires that we be able to distinguish random price behavior from actual price behavior.

Perhaps the most important reason to study the efficiency of markets is that inefficiencies are the key to making stock market profits. That is, one of the primary objectives of investors is to find unpriced or underpriced qualities in stocks. However, this search is likely to be complicated by the searches of thousands or millions of other investors, all looking for an advantage. In his presidential address to the American Finance Association, Richard Roll [1988] discussed the ability of academics to explain financial phenomena:

> *The maturity of a science is often gauged by its success in predicting important phenomena. Astronomy, the oldest science, is able to predict the positions of planets and the reappearance of comets with a high degree of accuracy... The immaturity of our science [finance] is illustrated by the conspicuous lack of predictive content about some of its most intensely interesting phenomena, particularly changes in asset prices. General stock price movements are*

notoriously unpredictable and financial economists have even developed a coherent theory (the theory of efficient markets) to explain why they should be unpredictable.

The theories of capital market efficiency quite powerful in that they are quite simple and do explain much of the behavior that is observed in capital markets. However, much empirical evidence exists which refutes capital markets efficiency. It should be noted that much of this evidence is contradictory and often reflects investors' desire to discover money making strategies, investment advisors' needs to sell their services and academicians' needs for publications. Furthermore, some evidence of market inefficiencies may simply be the result of data mining; one can always find or "demonstrate" an interesting pattern or relationship given enough data. An even greater difficulty for the opponents of capital market efficiency is that they are not able to offer a reasonably coherent and robust set of competing theories of capital markets behavior. Thus, it is reasonable to conclude based on prevailing theoretical and empirical research that the verdict is still out regarding the extent to which capital markets are efficient. Nonetheless, it is still important to discuss some of the many tests which have been conducted and suggest how further testing can still be accomplished.

Random Walks and Submartingales[1]

A *stochastic process* is a sequence of random variables x_t defined on a common probability space (Ω, Φ, P) and indexed by time t.[2] The values of $x_t(\omega)$ define the sample path of the process leading to state $\omega \in \Omega$. The terms $x(\omega,t)$, $x_t(\omega)$ and $x(t)$ are synonymous. A *discrete time process* is defined for a finite or countable set of time periods. This is distinguished from a *continuous time process* that is defined over an interval of an infinite number of infinitesimal time periods. The *state space* is the set of values in process $\{x_t\}$:

$$S = \{x \in \Re : X_t(\omega) \text{ for } \omega \in \Omega \text{ and some } t\}$$

The state space can be discrete or continuous. For example, if stock prices change in increments of eighths or sixteenths, the state space for stock prices is said to be discrete. The state space for prices is continuous if prices can assume any real value.

Consider an example of a particular stochastic process, a discrete time *random walk*, also known as a discrete time *Markov process*.[3] A random walk is a process whose future behavior, given by the sum of independent random variables, is independent of its past. Let z_i be a random variable associated with time i and let S_t be a state variable at time t such that $S_t = S_0 + z_1 + z_2 + \ldots + z_t$. Assume that random variables z_i are independent. The discrete time random walk is described as follows:

---

[1] The formal mathematical notation and definiti0ons offered here are not essential for purposes of this course. Skip them if you find it convenient to do so.
[2] In other words, a stochastic process is a random process through time.
[3] A Markov Process or random walk is a stochastic process whose increments are independent over time; that is, the Markov Process is without memory.

$$E[S_t | S_0, z_1, z_2, ..., z_{t-1}] = S_{t-1} + E[z_t]$$

It is important to note that $E[S_t]$ is a function only of S_{t-1} and z_t; the ordering of z_i where $i < t$ (the price change history) is irrelevant to the determination of the expected value of S_t. A specific type of Markov process, the discrete *martingale process* with $E[z_i] = 0$, is defined with respect to probability measure **P** and history or *filtration* $\Im_{t-1} = \{S_0, z_1, z_2, ..., z_{t-1}\}$ as follows:

$$E_{\mathbf{P}}[S_t | S_0, z_1, z_2, ..., z_{t-1}] = E_{\mathbf{P}}[S_t | \Im_{t-1}] = S_{t-1}$$

which implies:

$$E_{\mathbf{P}}[S_t | S_0, z_1, z_2, ..., z_i] = E_{\mathbf{P}}[S_t | \Im_i] = S_i \quad \forall\, i < t$$

Note that $E[z_i] = E[z_t] = 0$. Thus, a martingale is a process whose future variations cannot be predicted with respect to direction given the process history.[4] A martingale is said to have no memory and will not exhibit consistent trends. A *submartingale* is defined as:

$$E[S_t | S_0, z_1, z_2, ..., z_{t-1}] \geq S_{t-1}$$

A submartingale will trend upward over time such that $E[z_i] > 0$, and a *supermartingale* will trend downward over time.[5]

<u>Weiner Processes</u>

One particular version of a continuous time/space random walk is a *Wiener process*. A Weiner process is a generalized form of a *Brownian motion process*. The Weiner process may allow for drift; the standard Brownian motion process does not. A process z is a standard Brownian motion process if:

1. changes in z over time are independent; $COV(dz_t, dz_{t-i}) = 0$
2. changes in z are normally distributed with $E[dz] = 0$ and $E[(dz)^2] = 1$; $dz \sim N(0, 1)$
3. z is a continuous function of t
4. the process begins at zero, $z_0 = 0$

Brownian motion has a number of unique and very interesting traits. First, it is continuous everywhere and differentiable nowhere under Newtonian calculus; the Brownian motion process is not smooth and does not become smooth as time intervals decrease. We see in Figure 1 that Brownian motion is a *fractal*, meaning that regardless of the length of the observation time period, the process will still be Brownian motion. Consider the Brownian motion process represented by the top graph in Figure 1. If a short segment of is cut out and magnified as in the middle graph in Figure 1, the segment itself is a Brownian motion process; it does not smooth. Further magnifications of cutouts as in the bottom graph continue to result in Brownian motion processes. Many other processes

---

[4] A martingale has increments whose expected values equal zero.

[5] A submartingale has increments whose expected values exceed zero; expected values of increments of a supermartingale are less than zero.

smooth as segments covering shorter intervals are magnified and examined such that they can be differentiated under Newtonian calculus. Once a Brownian motion hits a given value, it will return to that value infinitely often over any finite time period, no matter how short. Over a small finite interval, we can express the change in z (Δz) over a finite period as follows:

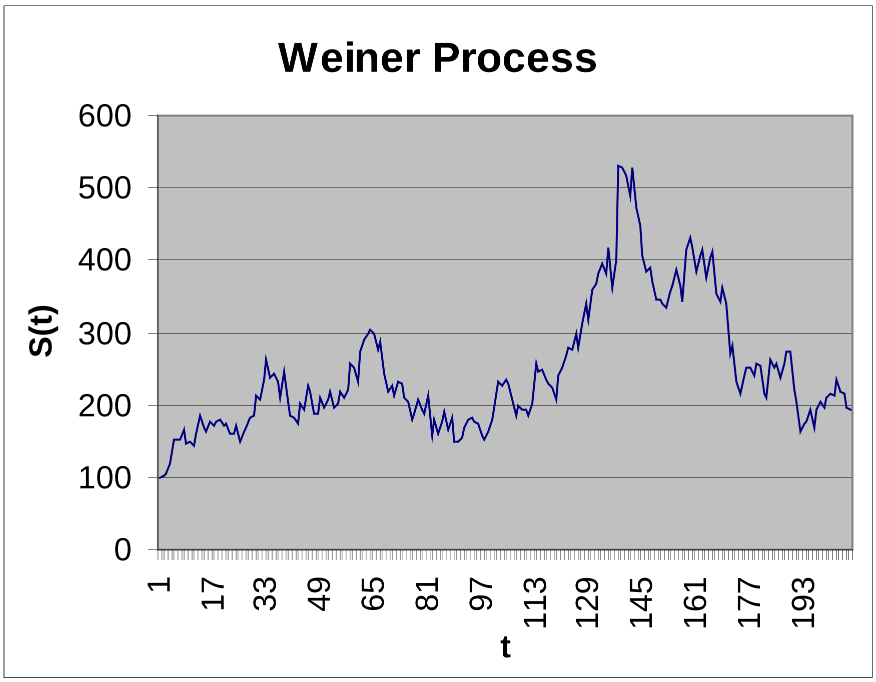
$$\Delta Z = dZ\sqrt{\Delta t} \sim N(0, \Delta t)$$

A *generalized Wiener process* is defined as follows:

$$dS_t = adt + bdz$$

where a represents the drift in the value of S_t and dz is a standard Brownian motion process. Because prices of many securities such as stocks tend to have a predictable drift component in addition to randomness, generalized Wiener processes may be more applicable than standard Brownian motion, which only includes a random element.[6]

Figure 1: Weiner Process



Random Walks and Market Efficiency

In a perfectly efficient market, security prices always reflect all available information. Price changes result from the arrival of new information. New information arrives randomly; otherwise, it is not new. Since news arrives in a random manner, prices in a perfectly efficient market should fluctuate randomly. Thus, one type of test for market efficiency is to determine whether prices can be predicted or whether their movement is entirely random.

[6] A Weiner Process is a a continuous time-space Markov Process with normally distributed increments.

B. Risk

In recent years, increased attention has been focused on forecasting security risk, as its measurement and computation is problematic. Generally speaking, the risk of an investment is simply the uncertainty associated with its returns or cash flows. Analysts typically use absolute risk measures such as variance and relative risk measures such as beta to quantify security risk. Although return variance is a quite simple mathematical construct with many desirable characteristics, its estimation is hampered by lack of ideal data. Suppose that we wish to estimate the risk or variance associated with a stock's returns over the next year. Consider the following discrete expression for ex-ante variance σ_F^2 that considers all potential return outcomes R_i and associated probabilities P_i:

(1)
$$\sigma_F^2 = \sum_{i=1}^{n} (R_i - E[R_i])^2 P_i$$

While this expression for variance is, by definition correct, its computation requires that we delineate all potential returns for the security (which might range from minus infinity to positive infinity). This is actually practical when the list of potential returns is small or when we can ascertain a specific return generating process. However, associating probabilities with these returns is a greater problem. For example, what is the probability that the return for a given stock will range between five and six percent? In many instances, we will be forced to either make probability assignments of a somewhat subjective nature or define a joint return and probability generating process for the security. Availability of historical price data typically makes risk estimation based historical variances more practical.[7]

Historical Volatility Indicators

Because it is frequently difficult to estimate the inputs necessary to estimate security ex-ante variance, analysts often use the volatility of historical returns as a surrogate for ex-ante risk:

(2)
$$\sigma_H^2 = \sum_{t=1}^{n} \frac{\left(R_t - \overline{R_t}\right)^2}{n-1}$$

where R_t represents the return realized during period t $[(P_t - P_{t-1})/P_{t-1}]$ in this n time period framework. Table 1 presents sample daily historical price data for a stock whose returns are given in the third column. The traditional sample daily variance estimator for this stock based on these returns equals .003172; the monthly variance (assuming 30 trading days per month; with weekends there are normally 20-21 days), if we were able assume that returns follow a Brownian motion process is .095. Use of the traditional sample estimator to forecast variance requires the assumption that stock return variances are constant over time, or more specifically, that historical return variance is an appropriate indicator of future return uncertainty. While this can often be a reasonable assumption, firm risk conditions can change and it is well documented that market volatility does fluctuate over time (See for example Officer (1971)). In addition, note that

[7] See the end-of-chapter Appendices A and B for a review of elementary statistics.

the sample variance estimator rather than the population estimator is proposed in Equation (2). This difference becomes more significant with smaller samples. Smaller samples intensify the need for a reliable mean.

Table 1: <u>Traditional Sample Estimators</u>

Time	$Price_t$	$Return_t$
0	30.000	N.A.
1	30.125	0.00417
2	30.250	0.00415
3	30.125	-0.00413
4	32.000	0.06224
5	34.000	0.06250
6	31.000	-0.08824
7	32.000	0.03226
8	30.500	-0.04688
9	30.750	0.00820
10	30.875	0.00407
11	31.000	0.00405
12	30.875	-0.00403
13	31.000	0.00405
14	31.125	0.00403
15	30.250	-0.02811
16	33.000	0.09091
17	30.000	-0.09091
18	35.125	0.17083
19	33.000	-0.06050
20	32.125	-0.02652
21	32.250	0.00389
22	32.375	0.00388
23	32.125	-0.00772
24	32.250	0.00389
25	34.250	0.06202
26	36.375	0.06204
27	38.500	0.05842
28	34.375	-0.10714
29	33.875	-0.01455
30	33.625	-0.00738

$$\sigma_H^2 = .095154$$

Using Equation (2) to estimate security variance requires that the analyst choose a sample series of prices (and dividends, if relevant) at n regular intervals from which to compute returns. Two problems arise in this process:

1. Which prices should be selected and at what intervals?
2. How many prices should be selected?

First, since each of the major stock markets tend to close at regular times on a daily basis, closing prices will usually reflect reasonably comparable intervals, whether selected on a daily, weekly, monthly, annual or other basis. Those prices closer to the date of computation will probably better reflect security risk (e.g., Price volatility of a security thirty years ago hardly seems relevant today). On the other hand, longer-term returns such as those computed on a monthly or annual basis will more closely follow a normal distribution than returns computed on a daily or shorter-term basis. This is a highly desirable quality, since many of the statistical estimation procedures used by analysts assume normal (or lognormal) distribution of inputs; the characteristics of most non-normal distributions are not well known.

Generally speaking, more prices or data points used in the computation process will increase the statistical significance of variance estimates. However, this leads to a dilemma: more data points, particularly pertaining to longer term returns will require prices from the more distant, but less relevant past. On the other hand, shorter estimation intervals may result in non-normal return distributions as well as autocorrelation issues. Hence, the analyst must balance the needs for a large sample to ensure statistical significance, recent data for relevance and longer-term data for independence and normality of distribution. These conflicting needs call for compromise. The convention that has developed over the years both in academia and in the industry is based on computations of five years of monthly returns.

Nonetheless, numerous difficulties still remain with this estimation procedure. For example, as we discussed above, variances are not necessarily stable over time. In our numerical illustration, higher-volatility periods are clustered as are lower volatility periods, in a manner similar to actual return variances. Second, returns themselves may not be independently distributed. In our numerical illustration, returns are inversely correlated, leading to significant differences in our variance estimates. Both problems arise in our numerical illustration data in Table 1. In addition, non-trading may omit returns data for computations.

<u>Extreme Value Estimator</u>

Two difficulties associated with the traditional sample estimator procedure, time required for computation and arbitrary selection of returns from which to compute volatilities may be dealt with by using extreme value estimators. Extreme value estimators are based on high and low values (and sometimes other parameters) realized by the security's price over a given period.

For example, consider the Parkinson Extreme Value Estimator (Parkinson (1980)). This estimating procedure is based on the assumption that underlying stock returns are log-normally distributed without drift. Given this distribution assumption, the underlying stock's realized high and low prices over a given period provide information regarding the stock's variance. Thus, if we are willing to assume that the return distribution is to be the same during the future period, Parkinson's estimate for the underlying stock return variance is determined as follows:

(3) $$\sigma_p^2 = .361 \cdot \left[\ln\left(\frac{HI}{LO} \right) \right]^2$$

where *HI* designates the stock's realized high price for the given period and *LO* designates the low price over the same period.[8]

[8] Accuracy of the Parkinson measure can be improved if the sample period can be subdivided into n equal sub-periods such that variance is estimated as follows:

$$\sigma_p^2 = \frac{.361}{n} \cdot \left[\sum_{t=1}^{n} \ln\left(\frac{HI_t}{LO_t} \right) \right]^2$$

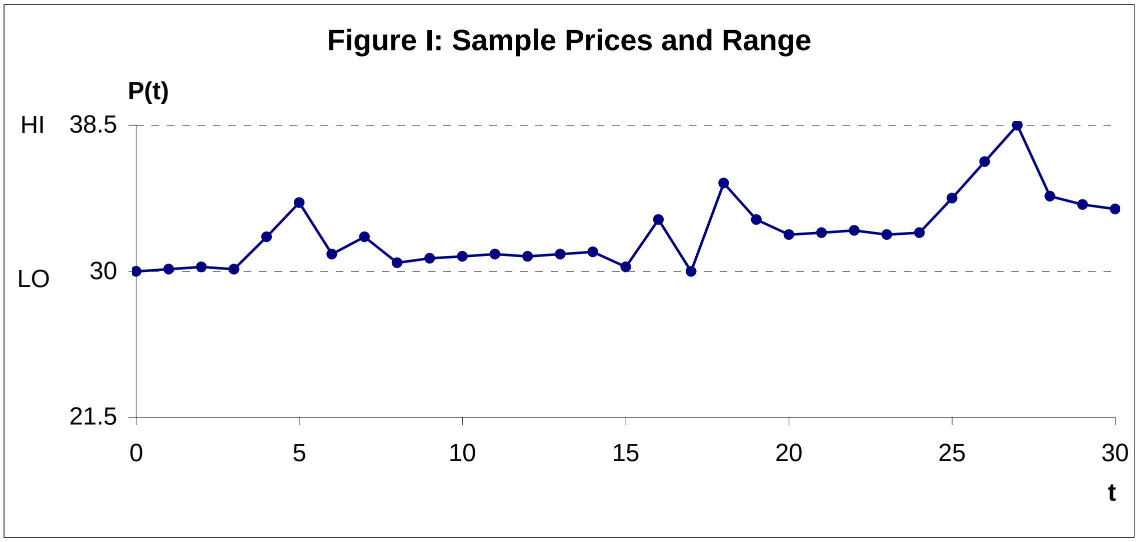


Figure I: Sample Prices and Range

The Parkinson measure results in a variance estimate equal to .022465 for the stock whose historical prices are listed Table 1:

$$\sigma_p^2 = .361 \cdot \left[\ln\left(\frac{HI}{LO} \right) \right]^2 = .361 \cdot \left[\ln\left(\frac{38.5}{30} \right) \right]^2 = .022465$$

Clearly, this is likely to be a simple estimate to obtain when periodic high and low prices for a stock are regularly published as they are for NYSE and many other stocks listed in the *Wall Street Journal*. Furthermore, the efficiency of the Parkinson procedure is several times higher than the traditional sample estimation procedure. To understand why this might be the case, consider Figure I. Several sample prices from which periodic returns might be computed are plotted for a second security, along with the high and low prices of the distribution. One might expect that high and low prices over the life of a process will tell us more about the variance of the distribution than would open and close prices alone. Using the extreme value estimator might be as simple as inserting into Equation (3) the 52-week high and low prices from the *Wall Street Journal*. Note that the Parkinson estimate is significantly smaller than the monthly historical estimate. This difference draws largely from the negative autocorrelation in the returns series.

Implied Volatilities[9]

A problem shared by both the traditional sample estimating procedures and the extreme value estimators is that they require the assumption of stable variance estimates over time; more specifically, that historical variances equal future variances. A third procedure first suggested by Latane and Rendleman (1976) is based on market prices of options that may be used to imply variance estimates. For example, the Black-Scholes

---

This is the more general form of the Parkinson Estimator. Each of the other extreme value estimators discussed in this paper can be generalized in a similar manner. The constant, .361 is the normal density function constant, $1/\sqrt{2\pi}$.

[9] See Appendix C to this chapter for a review of options and the Black-Scholes Option Pricing Model.

Option Pricing Model and its extensions provide an excellent means to estimate underlying stock variances if call prices are known. Essentially, this procedure determines market estimates for underlying stock variance based on known market prices for options on the underlying securities. Consider our stock example from Table 1 on day 30 where the stock is currently trading for \$33.625. Suppose that a one month ($t = 1$) call on this stock with a striking price equal to \$30 is currently trading for $c_0 = \$4.50$:

$$t = 1 \qquad r_f = .005 \qquad c_0 = \$4.50 \qquad X = \$30 \qquad S_0 = \$33.625$$

where r_f equals the monthly riskless return rate and X is the option striking price. If investors use the Black-Scholes Options Pricing Model to evaluate calls, the following must hold:

(7)
$$4.50 = 33.625 \cdot N(d_1) - 30e^{-rT} \cdot N(d_2)$$
$$d_2 = \frac{\ln\left(\frac{33.625}{30}\right) + \left(.005 + .5\sigma^2\right)\cdot 1}{\sigma\sqrt{1}}$$
$$d_2 = d_1 - \sigma\sqrt{1}$$

We find that this system of equations holds when $\sigma^2 = .027459$. Thus, the market prices this call as though it expects that the variance of anticipated returns for the underlying stock is .027459.

Unfortunately, the system of equations required to obtain an implied variance has no closed form solution. That is, we will be unable to solve explicitly for variance; we must search or substitute for a solution. One can substitute values for σ^2 until she finds one that solves the system. One may save a significant amount of time by using one of several well-known numerical search procedures such as the Method of Bisection or the Newton-Raphson Method.

Measure	Best used when
Ex-Ante Measure Based on Probabilities	(1) Ex-ante or future-oriented measure is needed such as when: a. The asset's historical volatility does not properly indicate its future risk b. The asset's risk characteristics have recently changed c. The asset has no price or returns history (2) All potential future return or cash flow outcomes can be specified (3) Probabilities can be associated with each potential return or cash flow outcome (4) Instead of (2) & (3), there is a specific return generating process with known parameters
Traditional Sample Estimator	(1) Variances are expected to be constant between historical and future time periods (2) There are an appropriate number of sampling intervals where: a. More periods increase statistical significance b. More periods increase reliance on older, less relevant historical data (3) Appropriate interval lengths can be determined; longer periods approach normality
Parkinson Extreme Value Estimator	(1) The computationally simplest measure based on a minimum of data is desired (2) Asset returns are log-normally distributed without drift (3) Historical volatility is a good indicator of future risk

Implied Volatility: Analytical Procedures	(1) Option prices on asset are readily available (2) Option pricing model assumptions hold in the relevant market (3) Can be used when historical volatility does not indicate future risk (4) User is able to use the appropriate analytical procedures (5) The market can be assumed capable of assessing risk

Table 2: Basic Risk Measures

C. Arbitrage

Arbitrage, perhaps the single most important pricing tool in modern finance, is defined as the simultaneous purchase and sale of assets or portfolios yielding identical cash flows. Assets generating identical cash flows (certain or risky cash flows) should be worth the same amount. This is known as the *Law of One Price*. If assets generating identical cash flows sell at different prices, opportunities exist to create a profit by buying the cheaper asset and selling the more expensive asset. The ability to realize a profit from this type of transaction is known as an arbitrage opportunity. Rational investors in such a scenario will seek to purchase the underpriced asset, financing its purchase by simultaneously selling the overpriced asset. The arbitrageur will execute such arbitrage transactions, continuing to earn arbitrage profits until the arbitrage opportunity is eliminated. If markets are competitive, the arbitrageur's purchases of the underpriced asset will bid its price up while the arbitrageur's selling transactions will force down the price of the overpriced asset. These arbitrage transactions should continue until no assets are over- or under- priced. Hence, arbitrageurs should force assets that produce identical cash flow structures to have identical prices.

Classic arbitrage is the simultaneous purchase and sale of the same asset at a profit. For example, if gold is selling in London markets for $600 per ounce and in New York markets for $610 per ounce, a classic arbitrage opportunity exists. An investor could purchase gold in London for $600 per ounce and simultaneously sell it in New York for $610. This results in a $10 profit per "round trip" transaction. The transactions involve no risk since both the selling and purchase prices are known. Furthermore, no initial net investment is required because the transactions offset each other; the proceeds of the sale are used to finance the purchase. Thus, if a classic arbitrage opportunity exists, an investor will have the opportunity to make a riskless profit without investing any of his own money. If the laws of supply and demand are not impeded by market inefficiencies, investors will flock to exploit this opportunity. Their buying pressure in London markets will force the London price to rise; their selling pressure in New York markets will force the New York price down. Buying and selling pressure will persist until the prices in the two markets are equal. Thus, classic arbitrage opportunities are not likely to persist long in unimpeded free markets. More generally, arbitrage might refer to the near simultaneous purchase and sale of portfolios generating similar cash flow structures.

The principle of arbitrage is the foundation underlying relative stock valuation. That is, we are able to price securities relative to one another when arbitrageurs are able to exploit violations of the Law of One Price. When The Law of One Price does not hold, one (or both) of the following will hold:

1. There exist opportunities to secure riskless arbitrage opportunities by buying underpriced while selling overpriced assets.
2. There exists some sort of market imperfection such as high transactions preventing arbitrageurs from exploiting arbitrage opportunities.

D. Limits to Arbitrage

Prospects of arbitrage will go far in making many markets respond more efficiently to information. However, despite potential profit opportunities, some markets are slow to react to arbitrage opportunities and some do not react at all. If trading is difficult or expensive, perhaps due to high transactions costs, price adjustments to arbitrage may be delayed or prevented. In recent years, hedge funds have led markets engaging in arbitrage strategies. However, they have not been entirely successful.

Consider the case of Long Term Capital Management (LTCM), which was in the early 1990s by John Meriwether, formerly head of the fixed income unit of Salomon Brothers, Inc. He brought in famed academics Myron Scholes and Robert Merton, developers of some of the most important derivatives and arbitrage strategies of the time, and the fund focused much of its trading activity on derivatives and fixed income. LTCM realized enormous returns and growth until 1998, when it lost $4 billion in one spectacularly disastrous quarter, forcing the fund into liquidation and threatening contagion with potential market-meltdown.

One of LTCM's equity markets arbitrage ventures involved the Royal Dutch/ Shell Group, a dual-listed company independently incorporated in the Netherlands and the U.K. This arrangement originated from a 1907 alliance agreement between Royal Dutch and Shell Transport in which the two companies agreed to merge their interests on a 60/40 basis. Royal Dutch trades primarily in the U.S. and the Netherlands while Shell trades primarily in London. After adjusting for foreign exchange rates, shares of the two firms' stock should trade at a 1.5–1 ratio. However, deviations from this anticipated trading ratio have deviated from this ratio by more than 35%, well beyond levels of tax differentials and transactions costs. In the summer of 1997, Royal Dutch traded at an 8-10% premium over its 1.5 expected level relative to Shell. To exploit this differential, LTCM had taken significant arbitrage positions on the two stocks. As the differential widened to 20% in 1998, LTCM increased its positions, until financial distress caused by trading activities elsewhere in the fund (related to the economic crisis in Russia) forced the firm to liquidate. The positions taken in Royal Dutch/Shell by LTCM were ultimately proven correct, but not until 2001 when the fund was no longer in existence. Rosenthal and Young [1990] argue that significant mispricing in dual-listed companies has prevailed over a long periods of time without satisfactory explanations. Nevertheless, instances, arbitrage can sometimes take significant amounts of time to equalize prices, causing arbitrageurs to maintain positions longer than they are able.

In another well-publicized failure of arbitrage dating from a March 2000 equity carve-out, 3Com spun off its Palm division, a maker of handheld computers. 3Com retained 95% of the shares of Palm and announced that each 3Com shareholder would ultimately receive 1.5 shares of Palm for each share of 3Com. The remaining 5% of Palm shares were issued at $38 per share, increasing to $165 by its first day of trading before closing at $95.06 (See Lamont [2002]). Remember, that ownership of one 3Com share implied ownership of 1.5 shares of Palm shares. The stocks of the two companies should have moved in tandem, but on the date of the IPO, 3Com actually decreased by 21% to

$81.81 as Palm increased. This $81.81 is substantially less than the $142.59 price implied by the 1.5 shares of Palm stock due to each 3Com shareholder (1.5 * $95.06 = $142.59), implying that the remainder of 3Com, on a per share basis, was worth -$60.78. This negative stub value (the whole is worth less than the sum of the parts; in particular, the parent and the subsidiary is worth less than the subsidiary alone) seems particularly unlikely, since 3Com had about $10 per share in cash and marketable securities alone. In other words, what happened wasn't rational, given the numbers, or just could not have been sustainable had investors been able to arbitrage. However, prospective arbitrageurs found themselves unable to short sell shares as the two stocks were under different national regulatory authorities. Thus, arbitrage and price correction could not be implemented because the short selling mechanism was not available for the Palm IPO.

Such negative stub values are not uncommon. For example, in 1923, Benjamin Graham chronicled his purchase of shares of stock in Du Pont, a well-established firm that had negative stub value given its investment in the new company General Motors. Lamont and Thaler (2001) identified five other 1990s technology equity carve-outs with negative stub values: UBID, Retek, PFSWeb, Xpedior, and Stratos Lightwave. Arbitrage in each of these cases was impeded by the inability to short sell. Mitchell, Pulvino and Stafford [2002] found 82 instances in U.S. markets between 1985 and 2000. But in most cases, arbitrage was impeded by inability to short sell, high transactions costs and difficulty in getting reliable price quotes or other information. But, Mitchell, Pulvino and Stafford found that approximately 30% of negative stub values were never eliminated through arbitrage. Some of the spin-offs failed, and others may have faced this risk. But, this probably cannot explain particularly large negative stub values.

References

Lamont, Owen (2002): The Curious Case of Palm and 3 Com; from *Mastering Investment*, James Pickford, ed., pp. 261-266.

Lamont, Owen and Richard Thaler (2001): “Can the market add and subtract? Mispricing in tech stock carve-outs,” Unpublished working paper, University of Chicago.

Latane, Henry and Richard Rendleman.(1976): “Standard Deviations of Stock Price Ratios Implied by Option Prices,” *Journal of Finance* 31: 369-381.

Mitchell, Mark, Todd Pulvino, and Erik Stafford (2002): “Limited Arbitrage in Equity Markets,” *Journal of Finance*, Vol. 57, No. 2: pp. 551-584.

Officer, Robert A. (1971). “A Time Series Examination of the Market Factor of the New York Stock Exchange,” Ph.D. Dissertation, University of Chicago.

Parkinson, Michael. (1980): “The Extreme Value Method for Estimating the variance of the Rate of Return,” *Journal of Business* 57: 61-65.

Ritter, Jay (2003): “Behavioral Finance,” Pacific Basin Finance Journal, September.

Roll, Richard (1988): "R^2," *Journal of Finance*, 43(3), pp. 541-566.

Rosenthal, Leonard, and Colin Young, (1990): “The seemingly anomalous price behavior of Royal Dutch/Shell and Unilever N.V./PLC,” *Journal of Financial Economics* 26, 123-141.

Exercises

1. Mack Products management is considering the investment in one of two projects available to the company. The returns on the two projects (A) and (B) are dependent on the sales outcome of the company. Mack management has determined three potential sales outcomes (1), (2) and (3) for the company. The highest potential sales outcome for Mack is outcome (1) or $800,000. If this sales outcome were realized, Project (A) would realize a return outcome of 30%; Project (B) would realize a return of 20%. If outcome (2) were realized, the company's sales level would be $500,000. In this case, project (A) would yield 15%, and Project (B) would yield 13%. The worst outcome (3) will result in a sales level of $400,000, and return levels for Projects (A) and (B) of 1% and 9% respectively. If each sales outcome has an equal probability of occurring, determine the following for the Mack Company:

a. the probabilities of outcomes (1), (2) and (3).
b. its expected sales level.
c. the variance associated with potential sales levels.
d. the expected return of Project (A).
e. the variance of potential returns for Project (A).
f. the expected return and variance for Project (B).
g. standard deviations associated with company sales, returns on Project (A) and returns on Project (B).

2. Which of the projects in Problem 1 represents the better investment for Mack Products?

3. Historical *percentage* stock returns for the McCarthy and Alston Companies are listed in the following chart along with percentage returns on the market portfolio:

Year	McCarthy	Alston	Market
1988	4	19	15
1989	7	4	10
1990	11	-4	3
1991	4	21	12
1992	5	13	9

Calculate the following based on the preceding diagram:

a. mean historical returns for the two companies and the market portfolio.
b. variances associated with McCarthy Company returns and Alston Company returns as well as returns on the market portfolio.

4. Forecast a variance and a standard deviation of returns for both the McCarthy and Alston Companies based on your calculations in Problem 3.

5. The following table represents outcome numbers, probabilities and associated returns for stock A:

outcome (i) return (R_i) Probability (P_i)

1	.05	.10
2	.15	.10
3	.05	.05
4	.15	.10
5	.15	.10
6	.10	.10
7	.15	.10
8	.05	.10
9	.15	?
10	.10	.10

Thus, there are ten possible return outcomes for Stock A.

a. What is the probability associated with Outcome 9?
b. What is the standard deviation of returns associated with Stock A?

6. The Durocher Company management projects a return level of 15% for the upcoming year. Management is uncertain as to what the actual sales level will be; therefore, it associates a standard deviation of 10% with this sales level. Managers assume that sales will be normally distributed. What is the probability that the actual return level will:

a. fall between 5% and 25%?
b. fall between 15% and 25%?
c. exceed 25%?
d. exceed 30%?

7. What would be each of the probabilities in Problem 6 if Durocher Company management were certain enough of its forecast to associate a 5% standard deviation with its sales projection?

8. Under what circumstances can the coefficient of determination (r-square) between returns on two securities be negative? How would you interpret a negative coefficient of determination? If there are no circumstances where the coefficient of determination can be negative, describe why.

9. Stock A will generate a return of 10% if and only if Stock B yields a return of 15%; Stock B will generate a return of 10% if and only if Stock A yields a return of 20%. There is a 50% probability that Stock A will generate a return of 10% and a 50% probability that it will yield 20%.

a. What is the standard deviation of returns for Stock A?
b. What is the covariance of returns between Stocks A and B?

10. An investor has the opportunity to purchase a risk-free treasury bill yielding a return of 10%. He also has the opportunity to purchase a stock that will yield either 7% or 17%. Either outcome is equally likely to occur. Compute the following:

a. the variance of returns on the stock.
b. the coefficient of correlation between returns on the stock and returns on the treasury bill.

11. The following daily prices were collected for each of three stocks over a twelve-day period.

Company X		Company Y		Company Z	
DATE	PRICE	DATE	PRICE	DATE	PRICE
1/09	50.125	1/09	20.000	1/09	60.375
1/10	50.125	1/10	20.000	1/10	60.500
1/11	50.250	1/11	20.125	1/11	60.250
1/12	50.250	1/12	20.250	1/12	60.125
1/13	50.375	1/13	20.375	1/13	60.000
1/14	50.250	1/14	20.375	1/14	60.125
1/15	52.250	1/15	21.375	1/15	62.625
1/16	52.375	1/16	21.250	1/16	60.750
1/17	52.250	1/17	21.375	1/17	60.750
1/18	52.375	1/18	21.500	1/18	60.875
1/19	52.500	1/19	21.375	1/19	60.875
1/20	52.375	1/20	21.500	2/20	60.875

Based on the data given above, calculate the following:

a. Returns for each day on each of the three stocks. There should be a total of ten returns for each stock - beginning with the date 1/10.
b. Average daily returns for each of the three stocks.
c. Daily return standard deviations for each of the three stocks.

12. Harlow Company stock realized a 52-week high of $50 per share and a 52-week low of $25. What is the Parkinson Extreme Value estimate for variance for this stock? What would be the corresponding standard deviation estimate?

13. Suppose that there is a six-month call currently trading for $8.20 while its underlying stock is currently trading for $75. Other details for this example are as follows:

$t = .5 \quad r_f = .10 \quad c_0 = 8.20 \qquad X = 80 \quad S_0 = 75$

What is the volatility (standard deviation) implied by this call?

14. Emu Company stock currently trades for $50 per share. The current riskless return rate is .06. Under the Black-Scholes framework, what would be the standard deviations implied by six-month (.5 year) European calls with current market values based on each of the following striking prices:

a. $X = 40;\ c_0 = 11.50$
b. $X = 45;\ c_0 = 8.25$
c. $X = 50;\ c_0 = 4.75$
d. $X = 55;\ c_0 = 2.50$
e. $X = 60;\ c_0 = 1.25$

Solutions

1.a. Each outcome has a one-third or .333 probability of being realized since the probabilities are equal and must sum to one.

b. $E[SALES] = (800{,}000 \cdot .333) + (500{,}000 \cdot .333) + (400{,}000 \cdot .333)$
$E[SALES] = 566{,}667$

c. $var[sales] = [\,(800{,}000 - 566{,}667)^2 \times .333 + (500{,}000 - 566{,}667)^2 \times .333 + (400{,}000 - 566{,}667) \times .333\,] = 28{,}888{,}000{,}000 = \sigma^2_{SALES}$

d. Expected return of Project A = $(.3\times.333)+(.15\times.333)+(.01\times.333) = .15333$

e. Variance of A's Returns = $[\,(.3-.1533)^2 \times.333 + (.15-.1533)^2 \times.333 + (.01-.1533)^2 \times.333\,] = .0140222 = \sigma_A^2$

f. Expected Return of Project B = $(.2\times.333)+(.13\times.333)+(.09\times.333) = .14$.

Variance of B's Returns = $[\,(.2-.14)^2 \times.333 + (.13-.14)^2 \times.333 + (.09-.14)^2 \times.333\,] = .0020666 = \sigma_B^2$.

g. Standard deviations are square roots of variances.

$\sigma_{SALES} = 169{,}964$

$\sigma_A = .1184154$

$\sigma_B = .0454606$

2. Project A has a higher expected return; however, it is riskier. Therefore, it does not clearly dominate Project B. Similarly, B does not dominate A. Therefore, we have insufficient evidence to determine which of the projects are better.

3. a. $\overline{R}_{Mc} = .062$

$\overline{R}_A = .106$

$\overline{R}_M = .098$

b. $\sigma^2_{Mc} = .000696$ (Remember to convert returns to percentages.)

$\sigma^2_A = .008824$ (Square roots of these variances are standard

$\sigma^2_M = .001576$ deviations.)

4. Assuming variance and correlation stability, the forecasted values would be the same as the historical values in Problem 3.

5.a. Since probabilities must sum to one, the probability must equal .15.
b. First, note that there is a .25 probability that the return will be .05 (.10+.05+.10) and .20 and .55 probabilities that the return will be .15. Thus, the expected return is .05×.25 + .10×.20 + .15×.55 = .115. The variance is $.25\times(.05-.115)^2 + .20\times(.10-.115)^2 + .55\times(.15-.115)^2 = .001775$, which implies a standard deviation equal to .04213.

6. Standardize returns by standard deviations and consult "z" tables:

$$\frac{R_i - E[R]}{\text{Standard Deviation}} = z.$$ Only use positive values for z.

a. $\frac{.05 - .15}{.10} = z(\text{low}) = 1$ $\frac{.25 - .15}{.10} = z(\text{high}) = 1$

From the "z" table (See Appendix C to this chapter), we see that the probability that the security's return will fall between .05 and .15 is .34. The value .34 is also the probability that the security's return will fall between .15 and .25. Therefore, the probability that the security's return will fall between .05 and .25 is .68.

b. From (4.16.a.), we see that the probability is .34.
c. .16
d. .0668

7. Simply reduce the standard deviations in the z scores in Problem (6) to .05.

a. .95
b. .47
c. .0228
d. .0013

8. Never, because coefficient of determination is always a positive squared value.

9.a. var = .0025; std.dev. = .05
b. -.00125

10.a. VAR = .0025
b. 0 : The coefficient of correlation between returns on any asset and returns on a riskless asset must be zero. Riskless asset returns do not vary.

11.
a.

	Company X	Company Y	Company Z
Date	Return	Return	Return
1/09	-	-	-
1/10	0	0	.00207
1/11	.00249	.00625	-.00413

1/12	0	.00621	-.00207
1/13	.00248	.00617	-.00207
1/14	-.00248	0	.00208
1/15	.03980	.04907	.04158
1/16	.00239	-.00584	-.02994
1/17	-.00238	.00588	0
1/18	.00239	.00584	.00205
1/19	.00238	-.00581	0
1/20	-.00238	.00584	0

b., c.

Stock	Average Return	Standard Deviation
X	.004064	.011479
Y	.006693	.014150
Z	.000869	.015537

12. The variance estimate is computed as follows:

$$\sigma_p^2 = .361 \cdot \left[\ln\left(\frac{HI}{LO}\right) \right]^2 = .361 \cdot \left[\ln\left(\frac{50}{25}\right) \right]^2 = .173444$$

The standard deviation is the square root of this value, or .416466.

13. If investors have used the Black-Scholes Options Pricing Model to evaluate this call, the following should hold:

$$8.20 = 75 \times N(d1) - 80 \times e^{-.1 \times .5} \times N(d2)$$
$$d1 = \{\ln(75/80) + (.1 + .5\sigma^2) \times .5\} \div \sigma\sqrt{.5}$$
$$d2 = d1 - \sigma\sqrt{.5}$$

Thus, we wish to solve the above system of equations for σ. There exists no closed form solution for σ. Thus, we will substitute and iterate to search for a solution. We first arbitrarily select $\sigma_1 = .35$. We find that this estimate for sigma results in a value of 6.90 for c_0. Since this call price is less than the market value 8.20, we know that σ is larger than .35. Thus, we try a larger value for σ, repeating the process until finding that $\sigma = .411466$. We have estimated the implied value with a greater degree of accuracy than is needed for most applications.

14. Implied volatilities are given as follows:

a. $X = 40$; $\sigma = .2579$
b. $X = 45$; $\sigma = .3312$
c. $X = 50$; $\sigma = .2851$
d. $X = 55$; $\sigma = .2715$
e. $X = 60$; $\sigma = .2704$

These values are obtained through a process of substitution and iteration. That is, readers should select trial values for σ to substitute into the Blck-Scholes formula, then compute the trial call value. A closer trial value for the call to the actual market price leads to a closer computed volatility to its Black-Scholes implied value. Each reader will probably use a process that will differ at least slightly from those used by others.

Appendix A: <u>Return And Risk Spreadsheet Applications</u>

Table A.1 contains spreadsheet entries for computing stock variances, standard deviations and covariances. The table (See Problem 11) lists daily closing prices for Stocks X, Y and Z from January 9 to January 20 in Cells B3:B14, E3:E14 and H3:H14. From these prices, we compute returns in Cells B19:B29, E19:E29 and H19:H29. Variance, standard deviation and covariance statistics in Rows 30 to 38 are computed from formulas displayed in Table A.2.

Table A.1: Stock Prices, Returns, Risk and Co-movement

	A	B	C	D	E	F	G	H
1	**CORP. X**				**CORP. Y**		**CORP.**	**Z**
2	**DATE**	**PRICE**		**DATE**	**PRICE**		**DATE**	**PRICE**
3	9-Jan	50.125		9-Jan	20		9-Jan	60.375
4	10-Jan	50.125		10-Jan	20		10-Jan	60.5
5	11-Jan	50.25		11-Jan	20.125		11-Jan	60.25
6	12-Jan	50.25		12-Jan	20.25		12-Jan	60.125
7	13-Jan	50.375		13-Jan	20.375		13-Jan	60
8	14-Jan	50.25		14-Jan	20.375		14-Jan	60.125
9	15-Jan	52.25		15-Jan	21.375		15-Jan	62.625
10	16-Jan	52.375		16-Jan	21.25		16-Jan	60.75
11	17-Jan	52.25		17-Jan	21.375		17-Jan	60.75
12	18-Jan	52.375		18-Jan	21.5		18-Jan	60.875
13	19-Jan	52.5		19-Jan	21.375		19-Jan	60.875
14	20-Jan	52.375		20-Jan	21.5		20-Jan	60.875
15								
16	**CORP. X**				**CORP. Y**		**CORP.**	**Z**
17	**DATE**	**RETURN**		**DATE**	**RETURN**		**DATE**	**RETURN**
18	9-Jan	N/A		9-Jan	N/A		9-Jan	N/A
19	10-Jan	0		10-Jan	0		10-Jan	0.00207
20	11-Jan	0.002494		11-Jan	0.00625		11-Jan	-0.00413
21	12-Jan	0		12-Jan	0.006211		12-Jan	-0.00207
22	13-Jan	0.002488		13-Jan	0.006173		13-Jan	-0.00208
23	14-Jan	-0.00248		14-Jan	0		14-Jan	0.002083
24	15-Jan	0.039801		15-Jan	0.04908		15-Jan	0.04158
25	16-Jan	0.002392		16-Jan	-0.00585		16-Jan	-0.02994
26	17-Jan	-0.00239		17-Jan	0.005882		17-Jan	0
27	18-Jan	0.002392		18-Jan	0.005848		18-Jan	0.002058
28	19-Jan	0.002387		19-Jan	-0.00581		19-Jan	0
29	20-Jan	-0.00238		20-Jan	0.005848		20-Jan	0
30	Mean	0.004064		Mean	0.006694		Mean	0.00087
31	Variance	0.000145		Variance	0.00022		Variance	0.000266
32	Variance (P)	0.000132		Variance (P)	0.0002		Variance (P)	0.000241
33	St.D.	0.01204		St.D.	0.014842		St.D.	0.016296
34	St.D. (P)	0.011479		St.D. (P)	0.014151		St.D. (P)	0.015538
35			COV(X,Y)=	0.0001494		COV(Y,Z)=	0.000192	
36			COV(X,Z)=	0.000139				
37			CORR(X,Y)=	0.9196541		CORR(Y,Z)=	0.8733657	
38			CORR(X,Z)=	0.7791748				

Formulas for computing returns are given in Rows 19 to 29 in Table A.2. Means, variances, standard deviations, covariances and correlation coefficients are computed in Rows 30 to 38. Row 30 computes the arithmetic mean return for each of the three stocks. Table A.2 lists formulas associated with the values in cells A30:H38. The =(Average) function may be typed in directly as listed in Table A.2 Row 30 or obtained from the Paste Function button (***f***$_x$) menu under the Statistical sub-menu. Entry instructions are given in the dialogue box obtained when the Average function is selected. The variance formulas in Row 31 are based on the Sample formula; the Variance (P) formulas in Row 32 are based on the population formula. Standard deviation sample and population results are given in Rows 33 and 34. Covariances and correlation coefficients are given in Rows 35 to 38.

Table A.2: Stock Returns, Risk and Co-movement: Formula Entries

	A	B	C	D	E	F	G	H
16	**CORP. X**			**CORP. Y**			**CORP. Z**	
17	**DATE**	**RETURN**		**DATE**	**RETURN**		**DATE**	**RETURN**
18	9-Jan	N/A		9-Jan	N/A		9-Jan	N/A
19	10-Jan	=B4/B3-1		10-Jan	=E4/E3-1		10-Jan	=H4/H3-1
20	11-Jan	=B5/B4-1		11-Jan	=E5/E4-1		11-Jan	=H5/H4-1
21	12-Jan	=B6/B5-1		12-Jan	=E6/E5-1		12-Jan	=H6/H5-1
22	13-Jan	=B7/B6-1		13-Jan	=E7/E6-1		13-Jan	=H7/H6-1
23	14-Jan	=B8/B7-1		14-Jan	=E8/E7-1		14-Jan	=H8/H7-1
24	15-Jan	=B9/B8-1		15-Jan	=E9/E8-1		15-Jan	=H9/H8-1
25	16-Jan	=B10/B9-1		16-Jan	=E10/E9-1		16-Jan	=H10/H9-1
26	17-Jan	=B11/B10-1		17-Jan	=E11/E10-1		17-Jan	=H11/H10-1
27	18-Jan	=B12/B11-1		18-Jan	=E12/E11-1		18-Jan	=H12/H11-1
28	19-Jan	=B13/B12-1		19-Jan	=E13/E12-1		19-Jan	=H13/H12-1
29	20-Jan	=B14/B13-1		20-Jan	=E14/E13-1		20-Jan	=H14/H13-1
30	Mean	=AVERAGE(B19:B29)		Mean	=AVERAGE(E19:E29)		Mean	=AVERAGE(H19:H29)
31	Variance	=VAR(B19:B29)		Variance	=VAR(E19:E29)		Variance	=VAR(H19:H29)
32	Variance (P)	=VARP(B19:B29)		Variance (P)	=VARP(E19:E29)		Variance (P)	=VARP(H19:H29)
33	St.D.	=STDEV(B19:B29)		St.D.	=STDEV(E19:E29)		St.D.	=STDEV(H19:H29)
34	St.D. (P)	=STDEVP(B19:B29)		St.D. (P)	=STDEVP(E19:E29)		St.D. (P)	=STDEVP(H19:H29)
35		COV(X,Y)=		=COVAR(B19:B29,E19:E29)	COV(Y,Z)=			=COVAR(E19:E29,H19:H29)
36		COV(X,Z)=		=COVAR(B19:B29,H19:H29)				
37			CORR(X,Y)=	=CORREL(B19:B29,E19:E29)	CORR(Y,Z)=			=CORREL(E19:E29,H19:H29)
38			CORR(X,Z)=	=CORREL(B19:B29,H19:H29)				

Appendix B: A Brief Review Of Elementary Statistics Measures

Mean, Variance and Standard Deviation

The purpose of this appendix is to introduce the reader to several important, though elementary concepts from statistics. To begin with, suppose that we wish to describe or summarize the characteristics or distribution of a single population of values (or sample drawn from a population). Two important characteristics include central location (measured by average, mean, median, expected value or mode) and dispersion (measured by range, variance or standard deviation).

In many instances, we will be most interested in the typical value (if it exists) drawn from a population or sample; that is, we are interested in the "location" of the data set. Mean (often referred to as average) or expected values (sometimes referred to as weighted average) are frequently used as measures of location (or central tendency) because they account for all relevant data points and the frequency with which they occur. The arithmetic mean value of a population μ is computed by adding the values x_i associated with each observation i and dividing the result by the number of observations n in the population:

$$\mu = \sum_{i=1}^{n} x_i \div n \tag{A.1}$$

Variance is a measure of the dispersion (variability and sometimes volatility or uncertainty) of values within a data set. In a finance setting, variance is also used as an indicator of risk. Variance is defined as the mean of squared deviations of actual data points from the mean or expected value of a data set. Deviations are squared to ensure that negative deviations do not cancel positive deviations, resulting in zero variances. High variances imply high dispersion of data. This indicates that certain or perhaps many data points are significantly different from mean or expected values. Population and sample variances are computed as follows:

$$\sigma^2 = \sum_{i=1}^{n} (x_i - \mu)^2 \div n \tag{A.2}$$

$$s^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n-1} \tag{A.3}$$

Standard deviation is simply the square root of variance. It is also used as a measure of dispersion, risk or uncertainty. Standard deviation is sometimes easier to interpret than variance because its value is expressed in terms of the same units as the data points themselves rather than their squared values. High standard deviations and high variances imply high dispersion of data. Standard deviations for populations and samples are computed as follows:

(A.4) $$\sigma = \sqrt{\sum_{i=1}^{n} (x_i - \mu)^2 \div n}$$

(A.5) $$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$$

Co-movement Statistics

A joint probability distribution is concerned with probabilities associated with each possible combination of outcomes drawn from two sets of data. Covariance measures the mutual variability of outcomes selected from each set; that is, covariance measures the relationship between variability in one data set relative to variability in the second data set, where variables are selected one at a time from each data set and paired. If large values in one data set seem to be associated with large values in the second data set, covariance is positive; if large values in the first data set seem to be associated with small values in the second data set, covariance is negative. If data sets are unrelated, covariance is zero. Covariance between data set x and data set y may be measured as follows, depending on whether one is interested in covariance of a population, of a sample or expected covariance:

(A.6) $$\sigma_{x,y} = \sum_{i=1}^{n} (x_i - \mu_x)(y_i - \mu_y) \div n$$

(A.7) $$\sigma_{x,y} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

The sign associated with covariance indicates whether the relationship associated with the data in the sets are direct (positive sign), inverse (negative sign) or independent (covariance is zero). The absolute value of covariance measures the strength of the relationship between the two data sets. However, the absolute value of covariance is more easily interpreted when it is expressed relative to the standard deviations of each of the two data sets. That is, when we divide covariance by the product of the standard deviations of each of the data sets, we obtain the coefficient of correlation $\rho_{x,y}$ as follows:

(A.8) $$\rho_{x,y} = \frac{\sigma_{x,y}}{\sigma_x \sigma_y} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}} \cdot \sqrt{\frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{n-1}}}$$

A correlation coefficient equal to 1 indicates that the two data sets are perfectly positively correlated; that is, their changes are always in the same direction, by the same proportions, with 100 percent consistency. Correlation coefficients will always range between -1 and +1. A correlation coefficient of -1 indicates that the two data sets are perfectly inversely correlated; that is, their changes are always in the opposite direction, by the same proportions with 100 percent consistency. The closer a correlation coefficient is to -1 or +1, the stronger is the relationship between the two data sets. A correlation coefficient equal to zero implies independence (no relationship) between the two sets of data.

The correlation coefficient may be squared to obtain the coefficient of determination (also referred to as r^2 in some statistics texts and here as ρ^2). The coefficient of determination is the proportion of variability in one data set that is explained by or associated with variability in the second data set. For example, ρ^2 equal to .35 indicates that 35 percent of the variability in one data set is explained in a statistical sense by variability in the second data set.

Appendix C: A Primer on Option Pricing

First, we will introduce a few option basics. A *stock option* is a legal contract that grants its owner the right (though, not the obligation) to either buy or sell a given stock. There are two types of stock options: puts and calls. A *call* grants its owner to purchase stock (called underlying shares) for a specified exercise price (also known as a striking price) on or before the expiration date of the contract. In a sense, a call is similar to a coupon that one might find in a newspaper enabling its owner to, for example, purchase a roll of paper towels for one dollar. If the coupon represents a bargain, it will be exercised and the consumer will purchase the paper towels. If the coupon is not worth exercising, it will simply be allowed to expire. The value of the coupon when exercised would be the amount by which value of the paper towels exceeds one dollar (or zero if the paper towels are worth less than one dollar). Similarly, the value of a call option at exercise equals the difference between the underlying market price of the stock and the exercise price of the call.

Suppose, for example, that a call option with an exercise price of $90 currently exists on one share of stock. The option expires in one year. This share of stock is expected to be worth either $80 or $120 in one year, but we do not know which at the present time. If the stock were to be worth $80 when the call expires, its owner should decline to exercise the call. It would simply not be practical to use the call to purchase stock for $90 (the exercise price) when it can be purchased in the market for $80. The call would expire worthless in this case. If, instead, the stock were to be worth $120 when the call expires, its owner should exercise the call. Its owner would then be able to pay $90 for a share that has a market value of $120, representing a $30 profit. In this case, the call would be worth $30 when it expires. Let T designate the options term to expiry, S_T the stock value at option expiry and c_T be the value of the call option at expiry. The value of this call at expiry is determined as follows:

(1)

$$c_T = MAX[0, S_T - X]$$

When $S_T = 80$, $C_T = MAX[0, 80 - 90] = 0$
When $S_T=120$, $C_T = MAX[0, 120 - 90] = 30$

A *put* grants its owner the right to sell the underlying stock at a specified exercise price on or before its expiration date. A put contract is similar to an insurance contract. For example, an owner of stock may purchase a put contract ensuring that he can sell his stock for the exercise price given by the put contract. The value of the put when exercised is equal to the amount by which the put exercise price exceeds the underlying stock price (or zero if the put is never exercised).

To continue the above example, suppose that a put option with an exercise price of $90 currently exists on one share of stock. The put option expires in one year. Again, this share of stock is expected to be worth either $80 or $120 in one year, but we do not know which at the present time. If the stock were to be worth $80 when the put expires, its owner should exercise the put. In this case, its owner could use the put to sell stock for \$90 (the exercise price) when it can be purchased in the market for \$80. The put would be worth \$10 in this case. If, instead, the stock were to be worth \$120 when the put expires, its owner should not exercise the put. Its owner should sell for \$90 for a share that has a market value of \$120. In this case, the call would be worth nothing when it expires. Let p_T be the value of the put option at expiry. The value of this put at expiry is determined as follows:

$$p_T = MAX[0, X - S_T] \tag{2}$$

$$\text{When } S_T=80,\ \ p_T = MAX[0, 90 - 80] = 10$$
$$\text{When } S_T=120,\ \ p_T = MAX[0, 120 - 80] = 0$$

The owner of the option contract may exercise his right to buy or sell; however, he is not obligated to do so. Stock options are simply contracts between two investors issued with the aid of a clearing corporation, exchange and broker that ensure that investors honor their obligations to each other. The corporation whose stock options are traded will probably not issue and does not necessarily trade these options. Investors, typically through a clearing corporation, exchange and brokerage firm, create and trade option contracts amongst themselves.

For each owner of an option contract, there is a seller or "writer" who creates the contract, sells it to a buyer and must satisfy an obligation to the owner of the option contract. The option writer sells (in the case of a call exercise) or buys (in the case of a put exercise) the stock when the option owner exercises. The owner of a call is likely to profit if the stock underlying the option increases in value over the exercise price of the option (he can buy the stock for less than its market value); the owner of a put is likely to profit if the underlying stock declines in value below the exercise price (he can sell stock for more than its market value). Since the option owner's right to exercise represents an obligation to the option writer, the option owner's profits are equal to the option writer's losses. Therefore, an option must be purchased from the option writer; the option writer receives a "premium" from the option purchaser for assuming the risk of loss associated with enabling the option owner to exercise.

Most stock options in the United States and Europe are traded on exchanges. The largest U.S. options exchange is the Chicago Board Options Exchange. The American and Philadelphia Exchanges also maintain stock options trading facilities. Options are also traded on several different commodities, currencies and other financial instruments.

Options may be classified into either the European variety or the American variety. European options may be exercised only at the time of their expiration; American options may be exercised any time before and including the date of expiration. Most option contracts traded in the United States (and Europe as well) are of the American variety. We will demonstrate in the next section that American options can never be worth less than their otherwise identical European counterparts.

The simple terminal value examples we discussed above were based on a *Binomial Distribution* where there are two possible outcomes for a given future point in time. If we add more time periods and more trials, we would increase the number of possible terminal outcomes. As the number of trials in a binomial distribution approach infinity, the binomial distribution approaches the *Normal Distribution*. Black and Scholes provide a derivation for an option-pricing model based on the assumption that the natural log of stock price relatives will be normally distributed.[10][11] The assumptions on which the *Black-Scholes Options Pricing Model* and its derivation are based are as follows:

1. There exist no restrictions on short sales of stock or writing of call options.
2. There are no taxes or transactions costs.
3. There exists continuous trading of stocks and options.
4. There exists a constant riskless interest rate that applies for both borrowing and lending.
5. The range of potential stock prices is continuous.
6. The underlying stock will pay no dividends during the life of the option.
7. The option can be exercised only on its expiration date; that is, it is a European Option.
8. Shares of stock and option contracts are infinitely divisible.
9. Stock prices follow an Îto process; that is, they follow a continuous time random walk in two dimensional continuous space. This simply means that stock prices are randomly distributed (in a manner somewhat similar to a normal distribution) and can take on any positive value at any time.

From an applications perspective, one of the most useful aspects of the Black-Scholes Model is that it only requires five inputs. All of these inputs with the exception of the variance of underlying stock returns are normally quite easily obtained:[12]

1. The current stock price (S_0): Use the most recent quote.
2. The variance of returns on the stock ($\mathbf{F}^2$): Several methods will be discussed later.
3. The exercise price of the option (X): Given by the contract
4. The time to maturity of the option (T): Given by the contract
5. The risk-free return rate (r_f): Use a treasury issue rate with an appropriate term to maturity.

It is important to note that the following less easily obtained factors are not required as model inputs:

1. The expected or required return on the stock or option and

[11]The stock price relative for a given period t is defined as $(P_t-P_{t-1})\div P_t$. Thus, the log of the stock price relative is defined as $\ln[(P_t-P_{t-1})\div P_t]$.

[12]These five inputs are the only that are necessary if the assumptions underlying the model hold. The sample sources for deriving input values may or may not be the most appropriate for a given contract.

2. Investor attitudes toward risk

If the assumptions given above hold, the Black-Scholes model specifies that the value of a call option is given as follows:

(3)

$$c_0 = S_0 N(d_1) - \frac{X}{e^{r_f T}} N(d_2)$$

(4)

$$d_1 = \frac{\ln(S_0 / X) + (r_f + \sigma^2 / 2)T}{\sigma\sqrt{T}}$$

(5)

$$d_2 = d_1 - \sigma\sqrt{T}$$

where N(d*) is the cumulative normal distribution function for (d*). This is function frequently referred to in a statistics setting as the "z" value for (d*). From a computational perspective, one would first work through Equation (4), then Equation (5) before valuing the call with Equation (3).

$N(d_1)$ and $N(d_2)$ are areas under the standard normal distribution curves (z-values). Simply locate the z-value on an appropriate table (see Table C.1) corresponding to the $N(d_1)$ and $N(d_2)$ values. Consider the following simple example of a Black-Scholes Model application: An investor has the opportunity to purchase a six month call option for \$7.00 on a stock which is currently selling for \$75. The exercise price of the call is \$80 and the current riskless rate of return is 10% per annum. The variance of annual returns on the underlying stock is 16%.

$$d_1 = \frac{\ln\left(75/80\right) + (.1 + .5 \cdot .16) \cdot .5}{.4 \cdot \sqrt{.5}} = \frac{\ln(.9375) + .09}{.2828} = .09$$

At its current price of \$7.00, does this option represent a good investment? First, we note the model inputs in symbolic form:

t = .5 r_f = .10 **F** = .4 S_0 = 75
X = 80 **F**^{2} = .16 e . 2.71828

Our first steps are to find d_1 from Equation (4) and d_2 from Equation (5):

$$d_2 = d_1 - .4 \cdot \sqrt{.5} = .09 - .2828 = -.1928$$

Next, by either using a z-table (see Table C.1) or by using an appropriate estimation function from a statistics manual, we find normal density functions for d_1 and d_2:

$$N(d_1) = N(.09) = .536 \qquad N(d_2) = N(-.1928) = .420$$

Finally, we use $N(d_1)$ and $N(d_1)$ in Equation (3) to value the call:

$$c_0 = 75 \cdot (.536) - (80 \cdot .9512) \cdot (.420) = 8.23$$

Since the 8.23 estimated value of the call exceeds its 7.00 market price, the call should be a worthy purchase.

Table C.1:
The z-Table

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.0000	.0040	.0080	.0120	.0159	.0199	.0239	.0279	.0319	.0358
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0909	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2356	.2389	.2421	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2703	.2734	.2764	.2793	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3437	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3906	.3925	.3943	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.492	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4986	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

Put-Call Parity

Before proceeding with pricing models applicable to the valuation of call options, we will first discuss a simple model concerning the relationship between put and call values. When this relationship holds, one is able to value a put based on knowledge of a call with exactly the same terms. First, assume that there exists a European put (with a current value of p_0) and a European call (with a value of c_0) written on the same underlying stock that currently has a value equal to X. Both options expire at time T and the riskless return rate is r_f. The basic Put-Call Equivalence Formula is as follows:

(1)

$$c_0 + Xe^{-r_f T} = S_0 + p_0$$

That is, a portfolio consisting of one call with an exercise price equal to X and a pure discount riskless note with a face value equal to X must have the same value as a second portfolio consisting of a put with exercise price equal to X and one share of the stock underlying both options.

A very useful implication of the put call parity relation, we can easily derive the price of a put given a stock price, call price, exercise price and riskless return:

(2)

$$p_0 = c_0 + Xe^{-r_f T} - S_0$$

Appendix Exercises

1. Call and put options with an exercise price of $30 are traded on one share of Company X stock.
   a. What is the value of the call and the put if the stock is worth $33 when the options expire?
   b. What is the value of the call and the put if the stock is worth $22 when the options expire?
   c. What is the value of the call writer's obligation stock is worth $33 when the options expire? What is the value of the put writer's obligation stock is worth $33 when the options expire?
   d. What is the value of the call writer's obligation stock is worth $22 when the options expire? What is the value of the put writer's obligation stock is worth $22 when the options expire?
   e. Suppose that the purchaser of a call in part a paid $1.75 for his option. What was his profit on his investment?
   f. Suppose that the purchaser of a call in part b paid $1.75 for his option. What was his profit on his investment?

2. Evaluate calls and puts for each of the following European stock option series:

Option 1	Option 2	Option 3	Option 4
T = 1	T = 1	T = 1	T = 2
S = 30	S = 30	S = 30	S = 30
F = .3	**F** = .3	**F** = .5	**F** = .3
r = .06	r = .06	r = .06	r = .06
X = 25	X = 35	X = 35	X = 35

Appendix Exercise Solutions

1. a. c_T = \$33 - \$30 = \$3; p_T = 0
   b. c_T = 0; p_T = \$30 - \$22 = \$8
   c. c_T = -\$3; p_T = 0
   d. c_T = 0; p_T = -\$8
   e. \$3 - \$1.75 = \$1.25
   f. \$0 - \$1.75 = -\$1.75

2. The options are valued with the Black-Scholes Model in a step-by-step format in the following table:

	OPTION 1	OPTION 2	OPTION 3	OPTION 4
d(1)	.957739	-.163836	.061699	.131638
d(2)	.657739	-.463836	-.438301	-.292626
N[d(1)]	.830903	.434930	.524599	.552365
N[d(2)]	.744647	.321383	.330584	.384904
Call	7.395	2.455	4.841	4.623
Put	0.939	5.416	7.803	5.665